WILDFIRE SUSCEPTIBILITY MAPPING IN BHUTAN USING GEOINFORMATICS TECHNOLOGY

Samdrup Dorji[1,2] and Suwit Ongsomwang[1*]

Received: May 19, 2017; Revised: June 28, 2017; Accepted: July 02, 2017

Abstract

The integration of geoinformatics technology with suitable geospatial models has been widely employed in many wildfire studies to develop and enhance wildfire management systems in different parts of the world. In Bhutan, wildfire is perceived as one of the most prominent causes of forest degradation and a serious threat to national conservation efforts. Thus, wildfire susceptibility analysis is seen as a necessary component of the wildfire management system for Bhutan. The main aim of the study is to apply the innovative approach of geoinformatics technology with the integration of GIS-based logistic regression (LR) and frequency ratio (FR) models to establish a wildfire susceptibility map. Herein, the study collected and prepared various influential wildfire factors, analyzed them, and established probability maps. The efficiency of each of the 2 models was then evaluated and compared with each other to determine an optimal model using the relative operating characteristics method.

The interpretations of the results revealed that the most significant predictor variables that played a major role in determining a wildfire occurrence in the study area are land surface temperature, proximity to roads, elevation, population density, enhanced vegetation index, distance to agricultural land, relative humidity, and aspect. The prediction and success rates of the LR model were 88.3% and 88.1%, while for the FR model they were 85.3% and 85.5%, respectively. The results indicated that both models are good predictors of wildfire with the LR model performing slightly better than the FR model. The predicted probability map from the optimum LR model was further classified into 5 categories of wildfire susceptibility zones: very low, low, moderate, high, and very high. The results from the study demonstrate that the integration of geoinformatics technology with GIS-based LR and FR models is an inevitable component of wildfire mapping that can effectively determine the most significant influential factors of a wildfire and its probability and eventually lead to the development of the wildfire susceptibility map.

Keywords: **Wildfire susceptibility analysis, remote sensing, GIS, logistic regression, frequency ratio, Thimphu and Paro, Bhutan**

Introduction

Wildfires present a substantial threat to precious forest resources and numerous studies have indicated an increasing trend in wildfire occurrences around the world. Besides, wildfires

[1] *School of Remote Sensing, Institute of Science, Suranaree University of Technology, Nakhon Ratchasima 30000, Thailand. E-mail: suwit@sut.ac.th*

[2] *National Land Commission, Kawajangsa, Post Box 142, Thimphu, Bhutan. E-mail: samdrupdorji@nlcs.gov.bt*

* *Corresponding author*

play an important role in global warming and climate change which are expected to increase with variations of climatic parameters (Stocks *et al*., 1998). They also influence vegetation dynamics and land use change at a global scale and contribute to the Earth's deforestation, desertification, and ecological damage.

Bhutan has about 80.9% of its total land as pristine forest areas, where 70.46% is covered by trees (Ministry of Agriculture and Forests, 2015). This has contributed to the country's position of being the first carbon negative country in the world (Energy and Climate Intelligence Unit, 2015), consequently making a significant contribution to the world that is threatened by climate change. It also helps to maintain the geologically fragile mountain ecosystem and environmental balance. Today, Bhutan has its rich natural and pristine forest cover still intact, providing homes to diverse flora and fauna including critical and endangered faunal species. As a result, Bhutan is also recognized as one of the 10 most threatened global biodiversity hotspots (Royal Government of Bhutan, 1999). These rich forest resources also help to sustain the hydropower industry, rural livelihoods, and food subsistence, thus contributing to the overall development of Bhutan. Hence, the future economy of the people and the country depends on the protection, conservation, and scientific management of forest resources (Royal Government of Bhutan, 1999).

However, wildfire is one of the most serious and consistent threats among many natural disasters such as earthquakes, glacial lake outburst floods, flash floods, and windstorms, and poses a potential hazard to the physical, biological, and ecological environments. It is perceived as one of the most prominent causes of forest degradation in the country and is a serious threat to the national conservation efforts (Tshering, 2006). It is estimated that more than 10000 acres of forest cover is lost every year because of wildfires (Gyelmo, 2016). During the period 2010 to 2015, the country recorded 216 wildfire incidences that burned about 950352 acres of forest cover (Department of Forest and Park Services, 2015). Though the majority of fires in Bhutan are related to human activities, the impact of various influential factors of fire still remain unknown, so there is a need for more advanced studies to examine and assess their correlation and degree of influence on wildfires. Thus, wildfire susceptibility analysis is seen as a necessary component of the wildfire management system for Bhutan. Moreover, geoinformatics technology on wildfire studies is rarely applied and is still at a developing stage.

Nowadays, geoinformatics technology provides comprehensive information and valuable tools to develop a wildfire susceptibility map that can be effectively used in wildfire management. Many studies have been conducted to establish wildfire susceptibility maps using geoinformatics technology (Chuvieco and Congalton, 1989; Jaiswal *et al.,* 2002; Pradhan *et al*., 2007; Adab *et al*., 2013) and different fire models have been developed based on various influential factors of wildfire coupled with suitable geospatial models. To determine the impact of different factors, various statistical methods, models, and algorithms have been tested by different researchers (Syphard *et al*., 2008; Oliveira *et al*., 2012; Ghoadi *et al*., 2012; Mohammadi *et al*., 2013) and varying results have been obtained based on the study areas and models being tested. Generally, GIS-based LR and FR models have given promising results at different study sites with a high prediction accuracy (Pradhan *et al*., 2007; Zhang *et al*., 2013; Mohammadi *et al*., 2013; Pourtaghi *et al*., 2014; Guo *et al*., 2015) and they have been proved to be reliable tools for wildfire susceptibility assessment.

The primary objective of the study is to apply the innovative approach of remote sensing and GIS technology in wildfire susceptibility maps using the geospatial models (LR and FR). The specific objectives include:

(1) To apply remote sensing and GIS technology with the integration of the geospatial models and determine the effect of 3 key influential factors (environmental, climatic, and anthropogenic) on wildfire occurrence;

(2) To formulate wildfire probability models and generate probability maps based on identified significant influential factors;

(3) To examine an optimum geospatial

model based on accuracy assessment and validation using the relative operating characteristics (ROC) method and establish a reliable wildfire susceptibility zonation map.

Materials and Methods

Study Area

The study area covers Thimphu and Paro districts in western Bhutan (Figure 1). According to the Asian Disaster Reduction Center statistics of 2015, the 2 districts have recorded one of the highest fire incidences in the country. The study area is characterized by a fragile mountain ecosystem and rugged topographic terrain combined with high ground fuel loads and fluctuating wind conditions that attribute to the high number of fire incidences. The study area is bounded by the geographic coordinates of longitude 89° 07' 20" to 89° 45' 56" E and latitude 27° 8' 41" to 28° 0' 3" N, approximately. It covers a total area of 3084 sq. km with elevations ranging from 1906 to 7092 meters above the mean sea level. The climate varies substantially from one place to another due to variations in topography and its elevation. Most of the developments and settlements are located in the low valleys surrounded by mountains.

Research Methodology

The framework of the research methodology on wildfire susceptibility analysis consists of 3 major components: (1) data collection and preparation, (2) wildfire susceptibility analysis based on the LR and FR models, and (3) accuracy ssessment and validation of the results to determine the optimum model for the final wildfire susceptibility mapping (Figure 2).

Data Collection and Preparation

The basic information of the collected input data for analysis is provided in Table 1, while a description of the prepared input data including dependent and independent variables

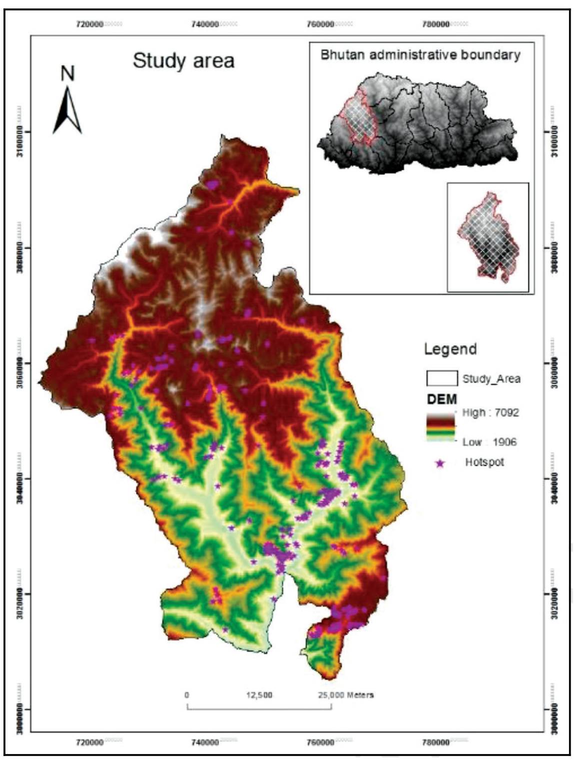

Figure 1. Study area

and the sampling technique is summarized in the following sections.

Dependent Variable (Hotspot)

The wildfire inventory map depicts the spatial location of wildfire points and represents the dependent variable in the analysis. However, the spatial data for the wildfire incidences in the study area were not available and did not exist at all. The spatial locations of the wildfire hotspots were obtained from the MODIS active fire/hotspot data of the Terra and Aqua satellites from NASA Fire Information for Resource Management System (FIRMS) (https://firms.modaps.eosdis.nasa.gov) for 15 years (2002-2016) via E-mail. The hotspot represents the center of a 1 km pixel that is flagged by the MODIS fire detection algorithm as containing 1 or more fires within the pixel (Giglio *et al*., 2010). The acquired hotspot's data are re-projected to a standard coordinate system. According to analysis of the MODIS statistics, the active fire season occurs during winter between October and May, and February was observed as the peak fire period. The obtained information was found to be consistent with the actual fire situation in the study area. This confirmed that the hotspot data can be reliable for the wildfire susceptibility analysis. The extracted hotspot points were further overlaid to high-resolution Google Earth images, analyzed, processed, and

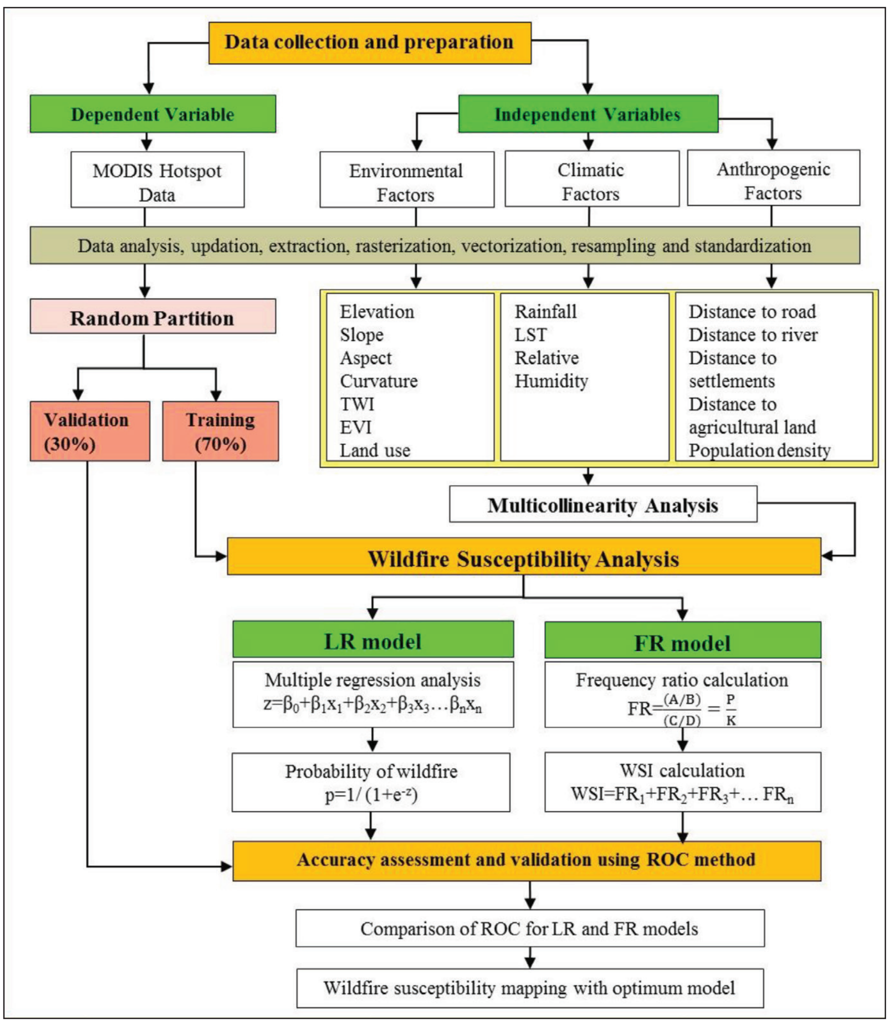

Figure 2. Schematic framework of research methodology

finally converted to the raster format with a 100 m cell size as the dependent variable. The entire study area comprises 2546 hotspot pixels to be used in the analysis.

Independent Variables

To accomplish a reliable wildfire susceptibility map, a comprehensive evaluation of associated influential wildfire factors is very essential and a prerequisite for wildfire analysis. In this study, 15 influential wildfire variables were extracted from the input databases comprising 3 major categories - the environmental, climatic, and anthropogenic variables. These variables were selected and established according to the basic characteristics of wildfires and extensive relevant literatures reviews.

(a) Environmental parameters include topographic features and fuel characteristics. Topography is one of the main factors applied in any fire hazard rating system because it characterizes the landscape features and it is strongly recommended in wildfire studies (Preisler *et al*., 2004; Brown and Davis, 1973; Chuvieco and Congalton, 1989). It also affects the vegetation distribution, composition, and flammability and has an influence on climatic variations (Syphard *et al.*, 2008). Thus, topographic variables including elevation, slope, aspect, and curvature were derived from the 10 m resolution of the ALOS digital elevation model (DEM) using surface analysis tools in the ESRI ArcGIS software. Likewise, the topographic wetness index (TWI) that represents moisture content was deduced from the DEM using the hydrological tools in the ESRI ArcGIS software. The amount of fuel has a significant influence on the rate of combustion and fire behavior. Thus, a 16-day composite of the MODIS enhanced vegetation index (EVI) product with a 250 m resolution was downloaded, re-projected, and extracted prior to the active fire season to represent the fuel characteristic. Land cover, which represents the landscape features of the Earth's surface, has been associated with fire occurrence (Syphard *et al*., 2008). It represents the type of vegetation available for burning and applies as a proxy for fuel types because it reflects the possible interactions with human

Table 1. Basic remote sensing and GIS input data for wildfire susceptibility analysis

No	Input Data	Data Format	Scale/ Resolution	Date	Source
1	MODIS Wildfire hotspot	Vector	1 km	2002-2016	NASA FIRMS (LANCE)
2	ALOS DEM	Raster	10 m	2010	National Land Commission (NLCS), Bhutan
3	ALOS image	Raster	10 m	2010	National Land Commission (NLCS), Bhutan
4	Topographic map	Vector	1:25,000	2002	National Land Commission (NLCS), Bhutan
5	LULC map	Vector	10 m	2010	Ministry of Agriculture and Forest (MoAF)
6	NCRP map	Vector	10-20 cm	2012	National Land Commission (NLCS), Bhutan
7	Meteorological data	Excel	NA	2005-2015	Department of Meteorology, Bhutan
8	Population data	Excel	NA	2010	National Statistical Bureau (PHCB-2010), Bhutan
9	EVI	Raster	250 m	2016	NASA, MODIS vegetation indices
10	LST	Raster	1 km	2016	NASA, MODIS LST product
11	Google satellite images	Raster	65 cm	2016	DigitalGlobe (QuickBird), 2016

activities. Herein, land use data are extracted from the Bhutan Land Cover Assessment 2010 (LCMP-2010) and were reclassified into 9 classes - coniferous forest, broadleaf forest, broadleaf and coniferous forest, shrubs and meadows, agricultural fields, built-up areas, snow cover, water bodies, and miscellaneous land.

(b) Climatic conditions are known to affect fuel accumulation and the moisture content (Syphard *et al*., 2008) and determine the type of vegetation in a region; thus, they play a dominant role in creating fire-prone areas. The drier the climate, the higher the probability of fire igniting and spreading. Considering the temporal scale of the data, climatic variables were derived from the available average weather conditions over a period of 11 years (2005-2015) from the Meteorological Department of Bhutan. Climatic factors which include rainfall and relative humidity were generated by the inverse distance weighted (IDW) interpolation technique using the coordinates of weather stations. Meanwhile, the MODIS land surface temperature (LST) at a 1 km resolution available for an 8-day composite was downloaded from NASA's website to represent temperature. Herein, the 8-day interval data prior to the active fire season was extracted and converted to Celsius using the scale factor of 0.002 provided in the metadata file.

(c) Anthropogenic factors, which include proximity and socio-economic variables, are one of the most significant driving factors of wildfire occurrence, since most of the wildfire incidences are related to human activities. The proximity variables represent the human accessibility to areas where fires can occur. Forest located near roads, settlements, and agricultural land is more prone to fires because of the habitation/cultural practices. Hence, proximity variables including distance to roads, rivers, settlements, and agricultural land are prepared using the Euclidean distance tools in the ERSI ArcGIS software. Roads, rivers, and settlements are extracted from the topographic map obtained from the National Land Commission at a 1:25000 scale. Those missing and new features were here updated according to the NCRP data. Missing rivers and streams are generated from the DEM using the hydrological tools in the ERSI ArcGIS software. New road networks were digitized and extracted using Google Earth images. The agricultural field class was updated according to the National Cadastral Resurvey Program data. In addition, population density is a socio-economic factor that represents the distribution of potential human influence on fire. This was generated at the sub-district level from the Population and Housing Census of Bhutan (PHCB) 2010 data of the National Statistics Bureau. It was interpolated using the IDW method, considering the major towns and cities as the center points of highly populated areas.

In summary, the selected variables in the analysis are the following 15 independent factors explained in Table 2: ELV, SLP, ASP, CRV, TWI, EVI, LU, RF, LST, RH, Dist_Road, Dist_River, Dist_Sett, Dist_AgriL, and Pop_Density. All variables were standardized to a uniform scale and resampled to a uniform cell size of 100 m using the spatial analyst tools in the ERSI ArcGIS software. The MODIS hotspot map as a dependent variable and the factor maps as independent variables are presented in Figure 3.

Sampling Technique

Selecting an appropriate sample for the LR model involves consideration of the sample size and the proportion of hotspot and non-hotspot pixels (Schicker and Moon, 2012). Thus, an appropriate number of samples should be considered to create a dependent variable. Earlier studies have recommended using an equal proportion for the presence (1) and absence (0) of a hotspot (Ayalew and Yamagishi, 2005), because it can eliminate the bias associated with the unequal proportion of samples (Zhu and Huang, 2006) and reduces the volume of data in the analysis. Since the number of hotspot pixels (2546 pixels) was comparatively less than the non-hotspot pixels in the study area, all hotspot pixels are taken into account for the analysis. Then, an equal number of non-hotspot pixels (2546 pixels) was randomly selected from the non-hotspot pixels and then combined with the hotspot pixels. Therefore, the total number of hotspot and non-hotspot pixels is 5092 pixels for the entire study area. These samples are then partitioned into a training and validation dataset by applying the random sampling technique to the proportions of 70% and 30%, respectively, using the geostatistical analysis tools in the ERSI ArcGIS software. In practice, the training and validation dataset was applied for both the LR and FR models. For the LR model, the dependent variable includes randomly sampled hotspot and non-hotspot pixels while the FR model requires only hotspot pixels as the

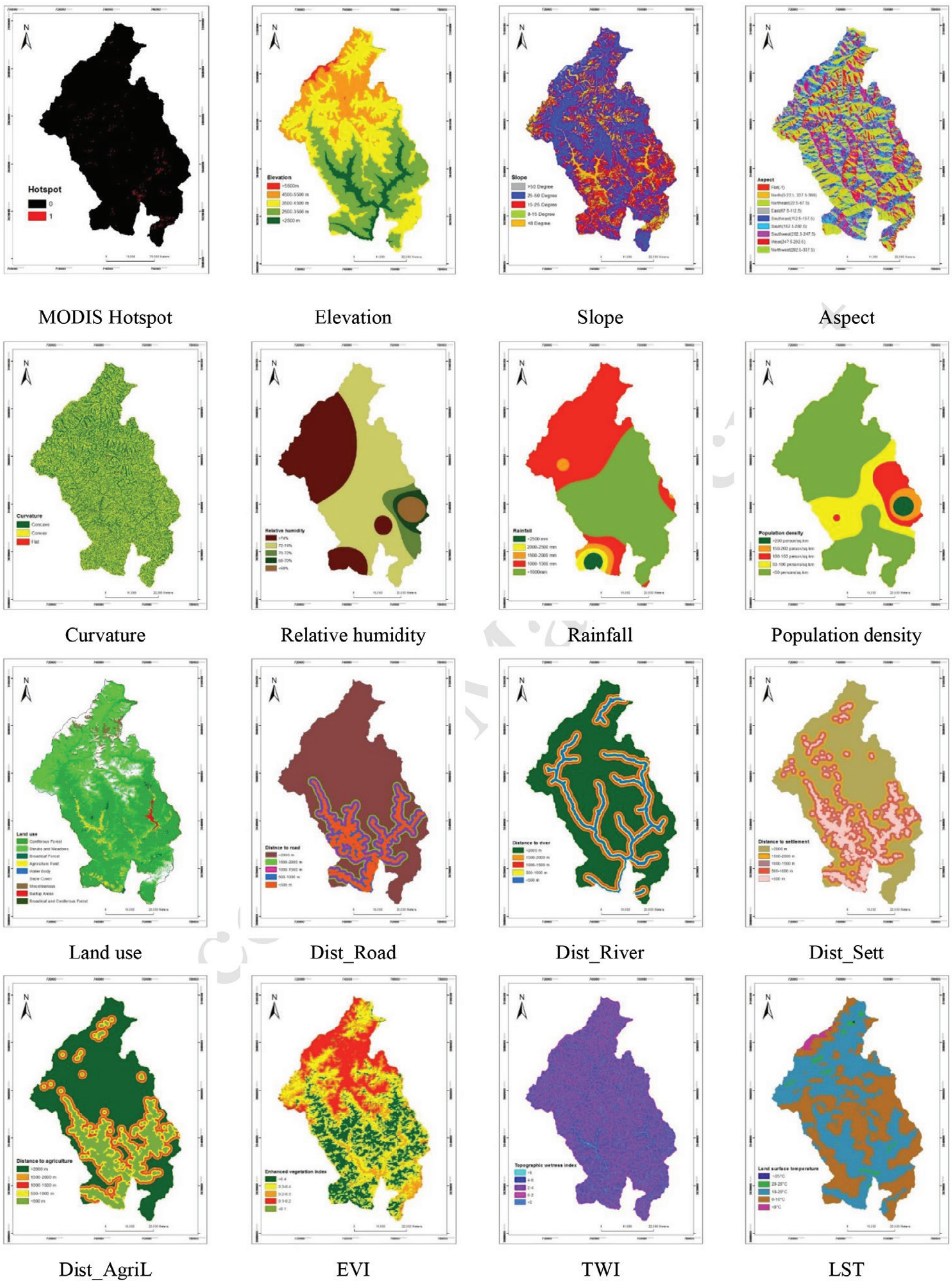

Figure 3. MODIS hotspot and predictor maps used in the analysis

dependent variable.

Wildfire Susceptibility Analysis Based on the LR and FR Models

The LR and FR models are here used to develop a wildfire probability map because the comparison of results from 2 models can both provide more insights on the complicated relationship between fire events and the influential factors in the study area and eventually enable the preparation of a reliable wildfire susceptibility map based on the optimum model through validation.

Logistic Regression (LR) Model

The LR model, which is sometimes known as a logistic or logit model, is a special case of multiple regression analysis for predicting the binary outcome variable (presence or absence) based on the set of predictor variables. It is suitable for modelling where the dependent variable is dichotomous or binary in nature. It gives the freedom to use both categorical and continuous variables in a regression analysis, whereby independent variables can be non-linear (Schicker and Moon, 2012). The main purpose of the LR model is to find the best fitting model to describe the relationship between a dependent and the independent variables (Ayalew and Yamagishi 2005). In this study, the LR model was applied to examine the relative strength and significance of each factor in wildfire prediction. Herein, the MODIS hotspot is considered as the dependent variable while environmental, climatic, and anthropogenic variables are the independent variables.

In the analysis, the presence of a wildfire hotspot is coded as "1" (y = 1), while the absence of a wildfire hotspot is coded as "0" (y = 0) (Atkinson and Massari 1998). Further, the presence of wildfire (y =1) is denoted as P and the absence (y = 0) as (1 - *P*). This allows logistic regression to model the probability of the occurrence of wildfire in association with each variable (Yesilnacar and Topao, 2005). Since the result of the LR model is binary, the probability value cannot be expressed as the linear function of the explanatory variables. Thus, the predicted probability of predictors is transformed to a linear function applying the logit transformation by executing the logarithm of *P*/ (1-*P*), known as odds. Therefore, the LR model is expressed in equation form as:

$$Logit\,(Y) = log\,\frac{P}{(1-P)} = \frac{e^{(\beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 \dots\dots \beta_n x_n)}}{1+e^{(\beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 \dots\dots \beta_n x_n)}} \tag{1}$$

where *P* represents the probability of the presence of wildfire, while 1-*P* represents the absence of wildfire, and *P*/(1-*P*) is the odds ratio. Quantitatively, the relationship between the probability of wildfire and its influential variable is expressed as (Preisler *et al*., 2004):

$$P = \frac{1}{1+e^{-\beta_0+\beta_i X_i}} \approx \frac{1}{1+e^{-z}}\,; \qquad 0 < P < 1 \tag{2}$$

$$Z = \beta_o + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 \dots\dots\dots \beta_n x_n \tag{3}$$

where *P* is the probability that wildfire occurs (Y=1) at a given location and varies from 0 to 1, β_0 is the intercept/constant of the model and β_i are the coefficients associated with the independent (X_i) variables, *Z* is the linear combination of the independent variables (X_i) in use weighted by their regression coefficients, and *e* is the base of the natural log. The coefficients of variables with positive values indicate a positive correlation while those with negative coefficients indicate a negative correlation with wildfire occurrence (Yalcin *et al*., 2011).

To apply the LR analysis, 1,782 pixels of hotspot (70%) were selected for training and another 764 pixels of hotspot (30%) have been retained for accuracy assessment and validation. An equal number of non-hotspot pixels (1782 pixels) was also randomly selected from the non-hotspot pixels and then combined with the hotspot pixels. Therefore, the total number of hotspot and non-hotspot pixels comprises 3564 pixels for the entire study area as a training dataset. Meanwhile, the corresponding values of independent variables for 3564 pixels were extracted and combined in an MS-Excel spreadsheet as the input data for the LR analysis. Before performing the LR analysis, the relative percentage of hotspot density was firstly used to transform nominal variables to numeric variables because this technique avoids the creation of an excessive number of dummy variables and considers the existing hotspot distribution (Yesilnacar and Topal, 2005). After that, all independent variables were normalized in the manner LR requires, because the independent variables are measured at different scales and they do not contribute equally to the analysis, making it difficult to assess relative importance and creating problems during interpretation of the final result (Ayalew and Yamagishi, 2005). Therefore, all variables were normalized using a linear transformation scale method under the spatial analyst tools in the ERSI ArcGIS software. The LR analysis is performed using the SPSS statistical software. Herein, the backward LR with stepwise analysis using the maximum likelihood method was applied to identify significant influential factors and the probability of wildfire occurrence.

Multicollinearity Analysis

Prior to the LR analysis, consideration of the multicollinearity issue is one of the most important steps to detect the correlation among the predictor variables because this will distort the model estimation and may provide an erroneous result (Rogerson, 2006). According to O'Brien (2007), a tolerance (TOL) of less than 0.20 or 0.10 and/or a variance inflation factor (VIF) of 5 or 10 and above presented a multicollinearity problem. In this study, if the TOL value is less than 0.1 and the VIF value is greater than 10, the variables were considered to have high correlation and were excluded in the analysis. The TOL and VIF values can be calculated using the following equation (Rogerson, 2006):

$$TOL = 1 - r^2 \tag{4}$$

$$VIF = \frac{1}{1-r^2} \tag{5}$$

where r^2 is associated with the regression of the independent variable on all other independent variables. The test was carried out in the SPSS statistical software.

Frequency Ratio (FR) Model

The FR model is another approach applied to assess the wildfire susceptibility. It is a simple geospatial assessment tool for computing the probabilistic relationship between dependent and independent variables, including multi-classified maps (Oh *et al*., 2011), and it is defined as the ratio of occurrence probability to nonoccurrence probability for specific attributes. The FR model is based on the observed relationships between the distribution of hotspots and each hotspot-related factor to reveal the level of correlation between hotspot locations and the influential factors (Pradhan *et al*., 2007). The FR model has several advantages of simplicity. More importantly, inputs, outputs, and the calculation process are easy to understand. In addition, a large amount of data can be processed quickly and easily in the GIS environment, whereas the LR model requires the specific statistical package and has difficulty in processing a large amount of data.

For the FR analysis, the thematic maps of all 15 wildfire influential factors were classified according to the objective, accuracy and scale of the data, and literature reviews. The first step of the FR analysis is to calculate the frequency ratio (FR) of each factor using the equation below (Lee and Pradhan, 2007):

$$FR = \frac{Hotspot\ Ratio}{Area\ Ratio} = \frac{(A/B)}{(C/D)} = \frac{P}{K} \tag{6}$$

where A is the number of hotspot pixels in each class of factor, B is the total number of hotspot pixels in the entire study area, P represents the percentage of hotspot pixels of the whole study area, C is the number of pixels (hotspot and non-hotspot) in each class of the factor, D is the total number of pixels (hotspot and non-hotspot) for the entire study area, and K represents the percentage of pixels (hotspot and non-hotspot) in each class of the factor. An FR value of 1 illustrates an average correlation, while a value greater than 1 illustrates a high correlation (indicates a higher chance of having fire in that specific class), and a value less than 1 indicates a lower correlation (Oh *et al*., 2011; Ozdemir and Altural, 2013).

The next step of the FR analysis is to assign those computed FR values of each class of factors using the reclassify function of the spatial analyst tool in the ESRI ArcGIS software. Finally, all the factor maps, with assigned FR values, are added to produce a wildfire susceptibility index (WSI) map using the equation below:

$$WSI = FR_1 + FR_2 + FR_3 + \cdots \ldots \ldots \ldots \ldots \ldots \ldots . FR_n \quad (7)$$

where *WSI* represents the wildfire susceptibility index; it indicates the relative susceptibility to wildfire occurrence, where higher values are associated with high susceptibility and lower values represent low susceptibility. FR_i represents the weighted factor maps of wildfire influential factors.

For the FR analysis, the hotspot pixels of the entire training dataset were overlaid with the classified wildfire factor maps and the number of hotspot pixels in each class was cross-tabulated and examined using the spatial analyst tool in the ESRI ArcGIS software. It was then imported to the MS Excel spreadsheet to calculate the FR values. Herein, the FR of each factor's class is calculated in 3 steps. First, the area ratio of each class of factor is computed followed by the calculation of the hotspot ratio. Finally, the FR is obtained by dividing the hotspot ratio by the area ratio for each factor's class (Equation 6). Next, the wildfire susceptibility index (WSI) was computed based on the FR values of each variable applying Equation 7 and using the spatial analyst tools in the ESRI ArcGIS software.

Accuracy Assessment and Validation

The accuracy assessment and validation of the predicted wildfire probability map is the most important component, otherwise the prediction model has no scientific significance (Chung and Fabbri, 2003). The accuracy assessment of the wildfire probability maps obtained from the LR and FR models was evaluated using the ROC technique based on the independent validation dataset (30%). The model that provides the better ROC value is selected for the final wildfire susceptibility zonation mapping.

Basically, the ROC determines whether the model is fit or not by checking the prediction performance of the model. It determines the accuracy of the prediction model at a user defined threshold value using area under curve (AUC). The AUC is also known as the index of accuracy or concordant index which represents the performance of the ROC curve. The higher the ROC value, the better is the model. The value of the ROC varies from 0.5 to 1. If the ROC value is 1, it indicates a perfect fit and an ROC value of 0.5 indicates a random fit.

Practically, the derived probability maps of the LR and FR models from the GIS environment are firstly exported to ERDAS IMAGINE software in the IMG format. Then, it is imported to the IDRISI software and converted to the RST format. Likewise, the dependent training and validation dataset is also imported to the IDRISI environment. Herein, the probability maps represent the input image while the training and validation map is used as a reference image for the calculation of the ROC. The probability map is compared with the training and the validation dataset to obtain the success rate and prediction rate curves, respectively, and then the AUC of the ROC is obtained. Generally, the ROC graph is plotted with a true positive rate (sensitivity) on the Y-axis against a false positive rate (1-specificity) on the X-axis for possible classification thresholds. The true positive rate (sensitivity) is the proportion of hotspots that are correctly classified, while the true negative rate (specificity) is the proportion of non-hotspots correctly classified. Here, the false positive rate (1-specificity) and false negative rate (1-sensitivity) are the proportions of non-hotspot and hotspot pixels that are erroneously classified. Both the true positive rate (sensitivity) and false positive rate (1-specificity) range from 0 to 1.

Results and Discussion

Multicollinearity Analysis

The multicollinearity analysis is reported in Table 2. The test results confirmed that there is no multicollinearity among the independent variables. In fact, the lowest TOL value was 0.161 and the highest VIF index was 6.218 for elevation which is greater than the TOL threshold (0.1) and less than the VIF threshold (10). Meanwhile, all other variables have TOL and VIF values within the threshold value which indicates there is no multicollinearity problem. Hence, all independent variables are applied for the LR and FR analyses.

LR Analysis on Wildfire Probability

In the LR analysis, the process firstly starts by entering all 15 predictor variables into the model and then sequentially eliminates the predictor variables based on the probability of the likelihood-ratio statistic, based on conditional parameter estimates. The elimination process of variable removal terminated after the 4^{th} step. In the process, 3 insignificant predictor variables, namely CRV, SLP, and Dist_River were removed while 12 significant predictor variables were retained by the model. All twelve variables have the estimated coefficients (β) statistically different from 0 with the given null hypothesis H_0: $\beta = 0$. They have a significance value (Sig.) less than 0.05 and are considered as significant influential factors of wildfire occurrence. The statistical test used a Wald chi-square value at 95% confidence level for the corresponding degree of freedom (df) indicating all 12 variables are significant, because the Wald values are greater than 4 which gives the level of significance value (p-value) less than 0.05 (Table 3). According to the classification summary in Table 3, the model correctly predicted for 2488 pixels out of 3564 pixels for an overall success rate of 70%. The LR goodness of fit measured by the Nagelkerke R^2 statistic of 0.267 is the pseudo-R^2 which indicates that the estimated LR model can approximately explain 27% of the variance in wildfire occurrence. The value of the pseudo R^2 (> 0.2) indicates that the performance of the model is good (Clark and Hosking, 1986) and the model can efficiently explain and interpret the relationship between the independent variables and the occurrence of wildfire.

The coefficient (β) of the LR model indicates the contribution of each factor to wildfire occurrence and its statistical significance. The relative importance of predictor variables is assessed using the corresponding coefficients and it is used in predicting the probability of wildfire occurrence. In principle, the coefficient (β) explains a change in the probability of wildfire occurrence for a unit increase in the corresponding independent variables. The variables with positive coefficients indicate a positive correlation and

Table 2. Multicollinearity diagnostic test of independent variables

No.	Independent variables (Abbreviation)	Collinearity statistics value	
		TOL	VIF
1	Elevation (ELV)	0.161	6.218
2	Slope (SLP)	0.840	1.190
3	Aspect (ASP)	0.855	1.170
4	Curvature (CRV)	0.506	1.974
5	Topographic wetness index (TWI)	0.496	2.016
6	Enhanced vegetation index (EVI)	0.624	1.602
7	Land use (LU)	0.875	1.143
8	Rain fall (RF)	0.684	1.461
9	Land surface temperature (LST)	0.543	1.842
10	Relative humidity (RH)	0.184	5.443
11	Distance to road (Dist_Road)	0.546	1.830
12	Distance to river (Dist_River)	0.515	1.943
13	Distance to settlement (Dist_Sett)	0.282	3.541
14	Distance to agricultural land (Dist_AgriL)	0.270	3.697
15	Population density (Pop_Density)	0.174	5.753

those with negative coefficients indicate a negative correlation to wildfire occurrence.

The probability of wildfire occurrence shows a positive correlation with the LST, ASP, Dist_AgriL, Dist_Sett, and LU variables, whereas Dist_Road, ELV, Pop_Density, EVI, RH, RF, and TWI show a negative correlation. Basically, the variables with positive coefficients have more of an explanatory capability in terms of causing wildfire while the variables with negative coefficients will tend to suppress the probability of wildfire occurrence. The results indicate that the most significant influential factors of wildfire are LST and Dist_Road followed by ELV, Pop_Density, EVI, Dist_AgriL, ASP, and RH. The remaining factors have a relatively low influence.

In this study, the influence of explanatory factors including LST, ASP, LU, Dist_Road, ELV, Pop_Density, EVI, RH, RF, and TWI principally agree with wildfire behavior and they are consistent with the previous works of Zhang et al. (2009); Mohammadi *et al*. (2014); Pourtaghi *et al*. (2014); Guo *et al*. (2015); Zhang *et al*., (2013); and Abdi *et al*. (2016). However, a positive linear relationship of the proximity variables including Dist_AgriL, and Dist_Sett indicates that as the Euclidean distance increases the occurrence of wildfire increases. This is an unexpected result because generally areas closer to agricultural land and settlement areas are more likely to initiate wildfire due to human activities like the burning of debris (agriculture/orchards/waste). To verify the result, the wildfire hotspots were overlaid with the raster maps of the agricultural land and settlements. It was observed that for certain areas closer to the agricultural lands and settlements, the density of hotspots was high while more fires seem scattered farther away from the agricultural and settlement areas in the northern part of Paro and southeastern part of Thimphu. As a result, the overall impact seems to show a positive correlation to the Euclidean distance of agricultural land and settlements. Moreover, all variables do not necessarily have a consistent linear relationship

Table 3. Coefficients of LR model and statistics

Factors	β	S.E.	Wald	df	Sig.	Exp(β)
Land surface temperature (LST)	5.099	0.480	112.995	1.000	0.000	163.778
Distance to agricultural land (Dist_AgriL)	1.769	0.320	30.562	1.000	0.000	5.862
Aspect (ASP)	1.540	0.137	125.692	1.000	0.000	4.663
Distance to settlement (Dist_Sett)	0.997	0.354	7.921	1.000	0.005	2.709
Land use (LU)	0.805	0.225	12.851	1.000	0.000	2.237
Topographic wetness index (TWI)	-0.680	0.231	8.671	1.000	0.003	0.507
Rain fall (RF)	-0.790	0.328	5.792	1.000	0.016	0.454
Relative humidity (RH)	-1.388	0.446	9.668	1.000	0.002	0.250
Enhanced vegetation index (EVI)	-1.798	0.565	10.135	1.000	0.001	0.166
Population density (P_Density)	-1.841	0.436	17.786	1.000	0.000	0.159
Elevation (ELV)	-2.937	0.563	27.238	1.000	0.000	0.053
Distance to road (Dist_Road)	-3.261	0.366	79.312	1.000	0.000	0.038
Constant	-1.785	0.601	8.833	1.000	0.003	0.168

Note: β = logistic coefficient; S.E. = standard error of estimate; Wald = Wald chi-square values; df = degree of freedom; Sig. = Significance; Exp(β) = exponentiated coefficient.

Classification summary and model statistics

Overall Percentage	70%
Nagelkerke R Square	0.267
Cox & Snell R Square	0.200

with wildfire occurrence (Wu *et al*., 2015). Therefore, it is important to focus on the possible reasons for such types of unexpected outcomes.

Using the coefficients of the LR analysis, the equation of the LR model for the probability of wildfire occurrence was formulated. This was achieved using the multiple linear regression equation of the LR model separately for 3 categories of variables: environmental, climatic, and anthropogenic and later combined with the estimated constant of the model. The combined Z values are then applied to calculate the probability of wildfire occurrence as:

$$Z_C = 5.099(LST) - 1.388(RH) - 0.79(RF) \quad (8)$$

$$Z_C = 5.099(LST) - 1.388(RH) - 0.79(RF) \quad (9)$$

$$Z_A = 1.769(Dist\_AgriL) - 3.261(Dist\_Road) + 0.997(Dist\_Sett) - 1.841(Pop\_Density) \quad (10)$$

$$Z = -1.785 + Z_E + Z_C + Z_A \quad (11)$$

$$=> P = \frac{1}{1+e^{-(\beta_o+\beta_1 x_1+\beta_2 x_2+\beta_3 x_3 \ldots\ldots \beta_n x_n)}} \approx \frac{1}{1+e^{-z}} \quad (12)$$

where Z_E, Z_C, and Z_A are the parameters that represent the linear combination of the environmental, climatic, and anthropogenic variables in use weighted by their individual regression coefficients, respectively, and P is the probability of the occurrence of a wildfire hotspot. This whole operation was done in the ESRI ArcGIS environment using the model builder tools and raster calculator of the spatial analyst tools (Figure 4). The result is a raster layer with the cell

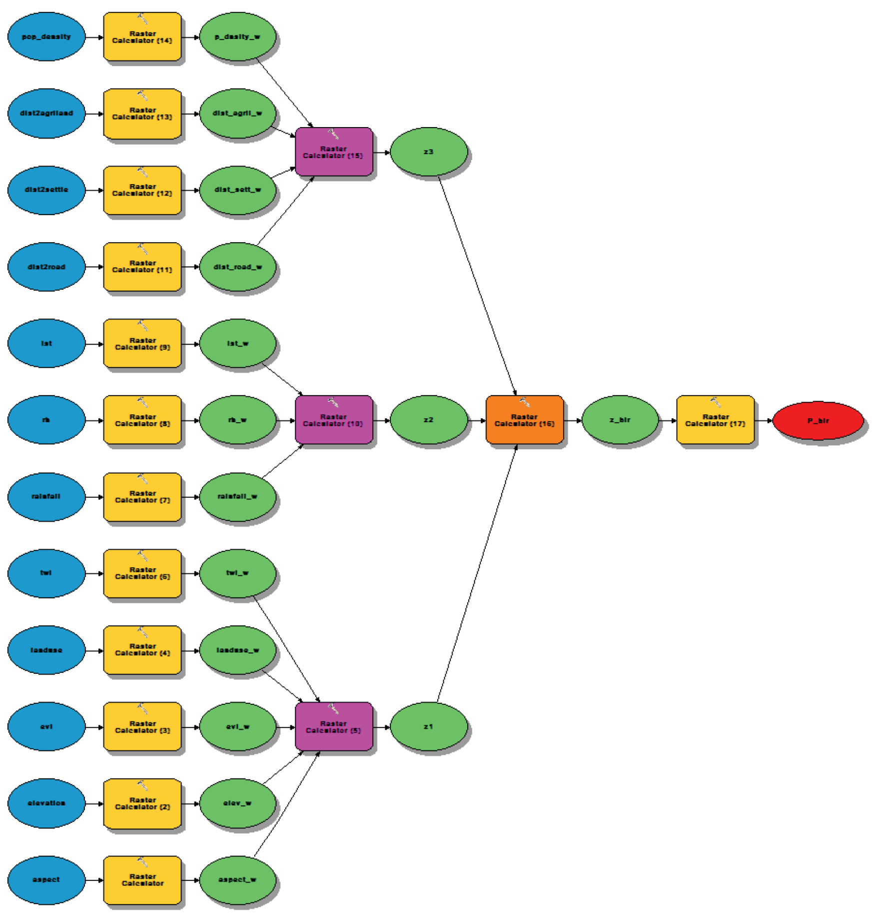

Figure 4. Model structure for probability of wildfire occurrence under LR analysis

values representing the estimated probability of wildfire occurrence, which varies from 0 to 0.945 (Figure 5).

The probability map of the LR model indicates that all the influential factors have a different degree of influence to the occurrence of wildfire. As indicated by the regression coefficients (β), the probability map of wildfire shows a higher probability at lower elevations that corresponds to high land surface temperature with low rainfall and humidity. The probability is very high closer to the roads and in places where the vegetation is predominant with dry grasslands, shrubs, and meadows and are, therefore, more susceptible to wildfire occurrence. In contrast, the probability is rather low at high altitude where humidity and rainfall are quite high and those areas covered by snow, glaciers, and lakes.

In summary, the LR model demonstrates that most of the fires are induced by human activities along the roads while they are also controlled by climatic and environmental conditions, particularly temperature and fuel. However, their degree of influence varies from one place to another due to the variation of topography.

FR Analysis on Wildfire Probability

The result of the FR analysis is presented in Table 4 and the WSI was produced using Equation 7 under the model builder and spatial analyst tools in the ESRI ArcGIS software (Figure 6). The output of the WSI that represents the probability of wildfire occurrence is displayed in Figure 7.

The analysis of the FR model reveals that LST, SLP, ASP, CRV, EVI, LU, Dist_Road, Dist_River, Dist_Sett, Dist_AgriL, and Pop_Density have a positive correlation to the frequency of wildfire occurrence in the study

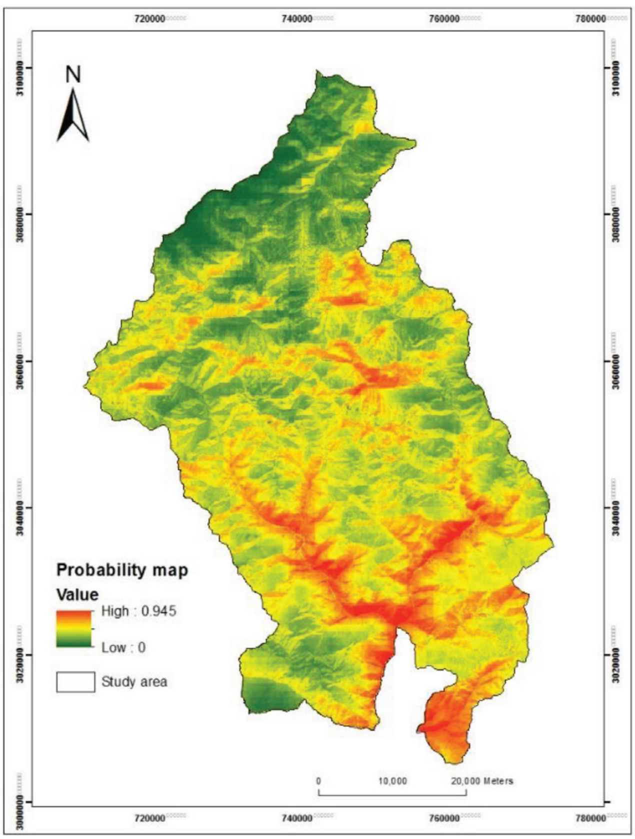

Figure 5. Wildfire probability map of LR model

area, while ELV, RF, RH, and TWI show a negative correlation. Generally, as the value of positively correlated factors increases, the frequency of wildfire tends to increase, while for negatively correlated factors it will tend to decrease their influence on wildfire occurrence. However, the FR values may deviate slightly according to the classification of factors.

Table 4. Frequency ratio value of wildfire factor classes computed from FR model

Factor	Class (Unit)	No. of pixels in each class	% of pixels in each class (B)	No of hotspot pixels	% of hotspot pixels (A)	FR = A/B
Elevation	<2500 m	22507	7%	272	15%	2.092
	2500-3500 m	107241	35%	711	40%	1.148
	3500-4500 m	112812	37%	734	41%	1.126
	4500-5500 m	63253	21%	65	4%	0.178
	>5500 m	2661	1%	0	0%	0
Slope	0-8o	13473	4%	45	3%	0.578
	8–15 o	34880	11%	172	10%	0.854
	15–25 o	104848	34%	592	33%	0.977
	25-50 o	153493	50%	970	54%	1.094
	>50 o	1780	1%	3	0%	0.292
Aspect	Flat (-1)	59	0%	0	0%	0.000
	North (0-22.5; 337.5-360)	36565	12%	103	6%	0.488
	Northeast (22.5-67.5)	43033	14%	131	7%	0.527
	East (67.5-112.5)	40158	13%	217	12%	0.935
	Southeast (112.5-157.5)	39793	13%	421	24%	1.831
	South (157.5-202.5)	39760	13%	376	21%	1.637
	Southwest (202.5-247.5)	37175	12%	300	17%	1.397
	West (247.5-292.5)	36587	12%	147	8%	0.696
	Northwest	35344	11%	87	5%	0.426
Curvature	Concave	158018	51%	862	48%	0.944
	Flat	3001	1%	16	1%	0.923
	Convex	147455	48%	904	51%	1.061
TWI	<0	108949	35%	635	36%	1.009
	0 - 2	101073	33%	667	37%	1.142
	2-4	62128	20%	311	17%	0.867
	4-6	20944	7%	108	6%	0.893
	>6	15380	5%	61	3%	0.687
EVI	<0.1	827	0%	0	0%	0.000
	0.1-0.2	58858	19%	78	4%	0.229
	0.2-0.3	54116	18%	364	20%	1.164
	0.3-0.4	113695	37%	1009	57%	1.536
	>0.4	80978	26%	331	19%	0.708
Land Use	Coniferous Forest	149563	48%	798	45%	0.924
	Shrubs and Meadows	95445	31%	735	41%	1.333
	Broadleaf Forest	1388	0%	1	0%	0.125
	Agriculture Fields	8638	3%	45	3%	0.902
	Water Body	972	0%	0	0%	0.000
	Snow Cover	34686	11%	185	10%	0.923
	Miscellaneous	15397	5%	16	1%	0.180
	Built-up Areas	1975	1%	1	0%	0.088
	Broadleaf and Coniferous Forest	410	0%	1	0%	0.422

Table 4. continued).

Factor	Class (Unit)	No. of pixels in each class	% of pixels in each class (B)	No of hotspot pixels	% of hotspot pixels (A)	FR = A/B
Rainfall	<1000 mm	154940	50%	1271	71%	1.420
	1000-1500 mm	135384	44%	429	24%	0.549
	1500-2000 mm	8150	3%	39	2%	0.828
	2000-2500 mm	5546	2%	34	2%	1.061
	>2500 mm	4454	1%	9	1%	0.350
LST	< 0°C	3081	1%	0	0%	0.000
	0 – 10°C	140505	46%	460	26%	0.567
	10 – 20°C	160574	52%	1234	69%	1.330
	20-25°C	4233	1%	88	5%	3.599
	>25°C	81	0%	0	0%	0.000
Relative humidity	<68%	10226	3%	125	7%	2.116
	68-70%	8719	3%	134	8%	2.660
	70-72%	15076	5%	139	8%	1.596
	72-74%	161191	52%	809	45%	0.869
	>74%	113262	37%	575	32%	0.879
Distance to road	<500 m	41450	13%	351	20%	1.466
	500-1000 m	22578	7%	363	20%	2.783
	1000-1500 m	17514	6%	123	7%	1.216
	1500-2000 m	15186	5%	74	4%	0.844
	>2000 m	211746	69%	871	49%	0.712
Distance to river	<500 m	31233	10%	229	13%	1.269
	500-1000 m	27100	9%	300	17%	1.916
	1000-1500 m	26676	9%	218	12%	1.415
	1500-2000 m	26368	9%	212	12%	1.392
	>2000 m	197097	64%	823	46%	0.723
Distance to settlement	<500 m	50630	16%	487	27%	1.665
	500-1000 m	41601	13%	372	21%	1.548
	1000-1500 m	34438	11%	173	10%	0.870
	1500-2000 m	28846	9%	97	5%	0.582
	>2000 m	152959	50%	653	37%	0.739
Distance to agricultural land	<500 m	54186	18%	582	33%	1.859
	500-1000 m	32447	11%	292	16%	1.558
	1000-1500 m	26166	8%	74	4%	0.490
	1500-2000 m	21770	7%	54	3%	0.429
	>2000 m	173905	56%	780	44%	0.776
Population density	<50 person/sq.km	215810	70%	1255	70%	1.007
	50-100 persons/sq.km	56084	18%	167	9%	0.515
	100-150 persons/sq.km	21856	7%	166	9%	1.315
	150-200 persons/sq.km	9023	3%	105	6%	2.014
	>200 persons/sq.km	5701	2%	89	5%	2.702

According to the FR values, the frequency of wildfire is found to be highly correlated with LST, Dist_Road, ELV, RH, and Pop_Density while other variables have a comparatively lower correlation. Wildfire frequency is very high at elevations lower than 2500 m above mean sea level and low above 2500 m. At elevations above 5500 m no fire incidences were observed. A progressive increase in FR values was noted as the slope angle increases. The FR value is high in the slope class between 25°–50° indicating a high correlation in this class. South-facing aspects experienced the highest number of wildfires compared to other faces as indicated by their high FR values. In the case of the land use factor, the wildfire frequency is very high in classes like shrubs and meadows, coniferous forest, and snow cover compared to other classes. However, due to the mismatch of the temporal scale of hotspot data and land use data, an unexpected high frequency value in snow cover was observed. In addition, snow cover is one of the dynamic variables that characterizes spatial-temporal phenomena. Convex curvature shows high wildfire frequency compared to concave and flat curvatures. The EVI class between 0.2 and 0.4 has a high frequency of wildfires compared to other classes, indicating the presence of more wildfires in shrubs and meadows and grasslands. The classes with the TWI lower than 0.2 have more wildfires, while with the classes with the TWI greater than 0.2 experience fewer fires. In the same way, places that have a mean annual rainfall less than 1000 mm and mean relative humidity less than 70% have a higher wildfire frequency while areas with a mean annual rainfall greater than 1000 mm and relative humidity greater than 70% have lower fire frequencies. The progressive increase in the FR values of the LST shows that the frequency of wildfire increases as the LST increases. The frequency of wildfire occurrences is relatively high in places where the LST is between 20°C and 25°C. The proximity factors indicate that within the Euclidean distance of 1500 m from roads and 100 m from rivers, there is the highest number of wildfire incidents. Likewise, within a proximity of 1000 meters from settlements and

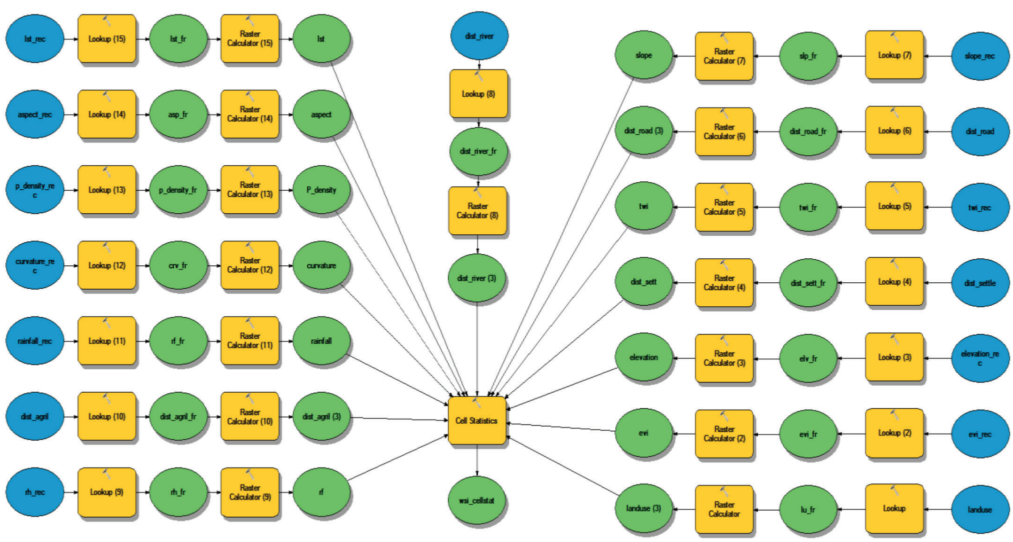

Figure 6. Model structure for wildfire susceptibility index under FR analysis

agricultural land, a high frequency of wildfires was found. As the Euclidean distance from the proximity factors increases, the frequency of wildfires tends to decrease. However, there are few instances where wildfire incidences increase beyond 2000 m from settlements and agricultural land. The FR results also reveal that a population density with more than 100 persons/km^2 has a high correlation with wildfire occurrence.

In addition, the prediction rates computed from the FR value of each factor confirm that the LST factor followed by the ELV, EVI, Dist_Road, and Pop_Density factors have a highly significant influence on wildfire occurrence (Table 5). Factors like Dist_AgriL, RF, ASP, RH, SLP, and Dist_Sett show a moderate influence while CRV, TWI, and Dist_River show a very low influence. CRV and Dist_River were also eliminated by the LR analysis since their contributions to the model were insignificant.

Moreover, the comparative analysis of the wildfire probability map from the LR model and the wildfire susceptibility index map from the FR model, which both represent the probability of wildfire occurrences, show a similar pattern along the low valleys, and present a slightly dissimilar pattern in hilly and mountain areas, because the representation of input data for the LR model as a continuous format and the FR model as a discrete format are different. In addition, the classification system of influential factors plays an important role for the FR analysis. Despite this, no significant difference was observed between the 2 probability maps which signifies that both the LR and FR models are reliably good in predicting the wildfires in the study area.

Furthermore, the comparison of predictive power from the LR model and prediction rate from the FR model demonstrates that LST, ELV, and Dist_Road followed by EVI and Pop_

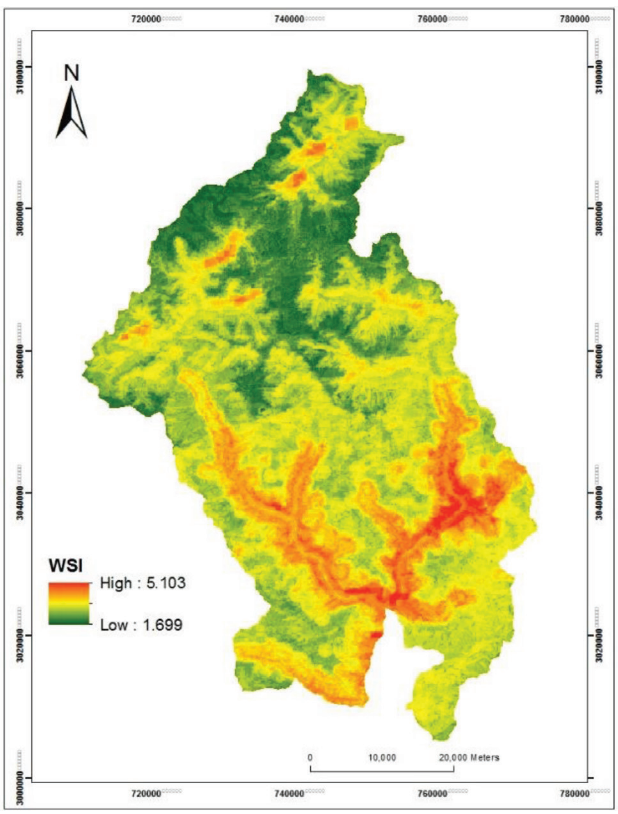

Figure 7. Wildfire susceptibility index map of FR model

Density are the most influential factors of wildfire while ASP, RH, Dist_Sett, RF, LU, and SLP have a moderate influence. The factors like TWI, CRV, and Dist_River have a very minimum role. The degree and pattern of the influence of each factor on wildfire probability show a similar pattern for both models (Figure 8).

In summary, the results from the LR and FR models show similar influences with an acceptable degree of correlation to wildfire occurrence in this study. This further confirms the reliability of both models in predicting wildfire. Overall, the spatial pattern of the areas predicted as having the highest probability of

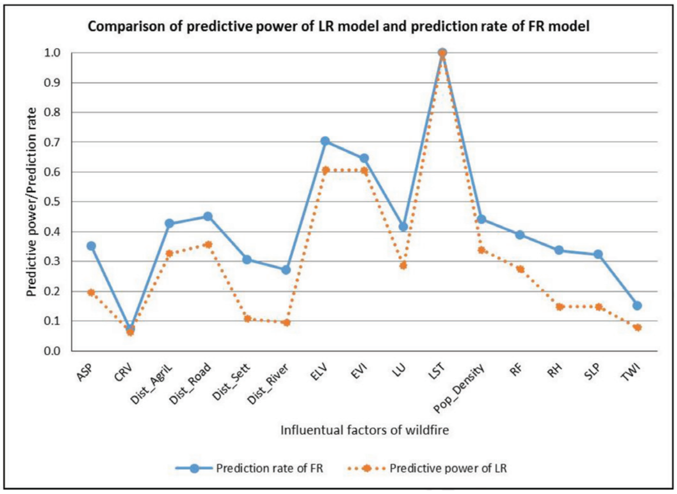

Figure 8. Graphical representation of predictive power of LR model and prediction rate of FR model

Table 5. Prediction rates of wildfire influential factors

No	Factor	Prediction Rate
1	Elevation (ELV)	0.703
2	Slope (SLP)	0.323
3	Aspect (ASP)	0.352
4	Curvature (CRV)	0.070
5	Topographic wetness index (TWI)	0.151
6	Enhanced vegetation index (EVI)	0.645
7	Land use (LU)	0.416
8	Rainfall (RF)	0.388
9	Land surface temperature (LST)	1.000
10	Relative humidity (RH)	0.337
11	Distance to road (Dist_Road)	0.451
12	Distance to river (Dist_River)	0.271
13	Distance to settlement (Dist_Sett)	0.306
14	Distance to agricultural land (Dist_Agril)	0.427
15	Population density (Pop_Density)	0.442

wildfire occurrences from both models reflects the significant influence of the land surface temperature, distance to roads, elevation, EVI, and population density in the study area.

Accuracy Assessment and Validation of LR and FR Models

The study prepared wildfire probability maps using the LR and FR models and the performance of each model was evaluated using the ROC method based on a 30% (764 pixels) independent validation set retained during data sampling.

The results showed success rates with AUC values of 0.881 and 0.855 for the LR and FR models, respectively (Figure 9). Based on Chung and Fabbri (2003) the results indicate that both models have a very good capability of classifying the area, and the models have a high goodness of fit with the training dataset and wildfire variables. Subsequently, the prediction rate curves with AUC values of 0.883 and 0.853 are obtained for the LR and FR models, respectively (Figure 9). The results indicate that both models have a relatively high predictive capability to discriminate the presence and absence of wildfire in the study area. Moreover, compared to other previous studies which employed both models, the results obtained in the present study provide better accuracy in predicting wildfire occurrences (Intarawichian and Dasananda (2010); Zhang *et al*. (2013); Pourtaghi *et al*., 2014; Guo *et al*. (2015).

Although, the LR model performed slightly better than the FR model, as indicated by the higher AUC value, the FR model can also be considered as an equally acceptable model that can be applied for susceptibility mapping in the area. The close similarities of the success and prediction rate curves of the 2 models indicate that both models are reliable and can be used in predicting future wildfires. However,

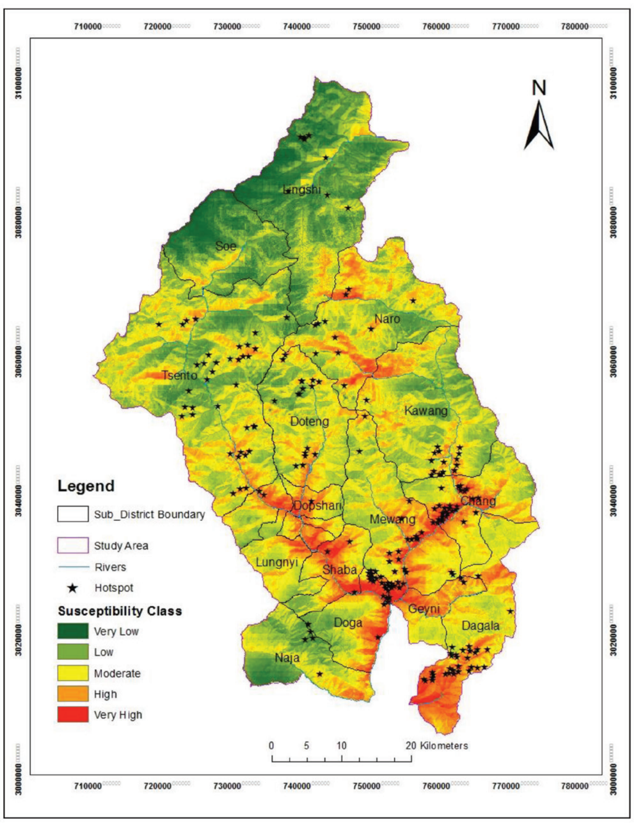

Figure 9. Success rate curves and prediction rate curves of LR and FR models

the LR model is here considered as the optimum model for the final wildfire susceptibility mapping based on the comparative assessment and validation. The LR model shows a slightly higher performance for both training and validation datasets compared to the FR model with a high AUC value for success and prediction rates of 0.881 and 0.883, respectively. Some studies have also found that the LR model has performed better than the FR model (Lee and Evangelista, 2008; Meten *et al*., 2015), while others have found the FR model better than the LR model (Lee and Pradhan, 2007).

Wildfire Susceptibility Mapping

To generate the final wildfire susceptibility map of different zones using the optimum LR model, the method adopted in many previous studies is to divide the histogram of the probability map into different categories based on expert opinions (Dai and Lee, 2002; Ohlmacher and Davis, 2003) and many studies have chosen and applied different classification methods depending on their interest and the type of data. For instance, Ayalew and Yamagishi (2005) applied 4 classification methods, namely quantiles, natural breaks, equal intervals, and standard deviation and selected one that provides the best information according to the scale of investigation. They also found that the standard deviation method was suitable and provided good information. In other studies, Meinhardt *et al*. (2015) and Intarawichian and Dasananda (2010) applied the manual and natural breaks, respectively, for better classification.

The present study examined all the available inbuilt classification methods in the ESRI ArcGIS software and deduced that the standard deviation method provides the best information that is more suitable to the study area compared to other methods. The standard deviation method has a certain advantage of applying the mean to generate the class breaks (Ayalew and Yamagishi, 2005). Moreover, probability values of the final output map are normally distributed according to a histogram report where the standard deviation method is suitable when the samples are normally distributed (Environmental Systems Research Institute, Inc., 2016). Herein, the wildfire susceptibility map comprises 5 zones: very low, low, moderate, high, and very high (Figure 10).

The percentage of hotspots and the area coverage computed for each zone are presented in Table 6. According to the classified zones, 39% of the total hotspots are found in very high susceptibility zones covering about 9% of the total study area, while the high susceptibility zones have 33% of the total hotspots covering 21% of the total area. This indicates that the majority of the total hotspots (72%) are found in high and very high susceptibility zones covering 30% of the total study area. The zones corresponding to moderate, low, and very low susceptibility constitute 17%, 11%, and 0% of the total hotspots, with the corresponding area coverage of 1146.420, 860.260, and 161.530 km^2, respectively.

Upon the visual interpretation, the deduced wildfire susceptibility map conveys useful information and it appears to be highly satisfying and rational. According to the classification zones, most parts of the very high and high zones are located in the sloping valleys at lower elevations where vegetation is mostly dominant with shrubs and meadows/grasslands, and in areas that have

Table 6. Percentage of hotspots and the area and its percentage coverage in each susceptibility zones

Probability Range	Susceptibility class	Hotspot (%)	Area (Sq. Km)	Percent (%)
0.000 - 0.110	Very low	0	161.530	5
0.110 - 0.300	Low	11	860.260	28
0.300 - 0.500	Moderate	17	1146.420	37
0.500 - 0.700	High	33	639.150	21
0.700 - 0.945	Very high	39	277.380	9

a high land surface temperature with low rainfall and humidity. They are also found closer to the roads where most of the daily human activities are involved. In addition, most of the agricultural land also seems to fall into the high susceptibility zone including a few patches in remote areas, especially nearby settlements. The susceptibility to wildfires appears to decrease as the distance from the road increases where there is less human interference. Most of the areas that are covered by coniferous forest in the mid-altitude areas seem to fall into the moderate susceptibility zone and those areas in the high altitudes covered by snow and bare soils fall into either the low or very low susceptibility zones. A few areas of the low susceptibility zone are located in the southwest and center of the study area that lie in higher altitudes while most of the very low susceptibility zone is situated in the northern part. Although, the proportion of the very high and high susceptibility zones are smaller compared to other zones, the resulting map is agreeable with the actual fire situation in the study area. Particularly during the winter, when there is no rainfall, the surrounding air becomes very dry with fluctuating winds and the humidity remains very low. The trees shed their leaves adding more fuel to the ground, and the grasses, shrubs, and meadows become dry. As a result, they become more susceptible to wildfires.

Conclusions

According to the results of the LR analysis, it can be concluded that the probability of wildfire occurrences has a positive correlation with the land surface temperature, aspect, distance to agricultural land, distance to settlement, and land use factors, whereas distance to road, elevation, population density, enhanced vegetation index, relative humidity, rainfall, and topographic wetness index have a negative correlation to the occurrence of wildfire. On the other hand, the factors including curvature, slope, and distance to river were eliminated during the process of the stepwise LR analysis. This suggests that they have a very weak correlation to wildfire occurrences. The variables with positive coefficients have a more explanatory capability than variables with negative coefficients in terms of causing wildfires in the study area. The factors with negative coefficients will tend to suppress the probability of wildfire occurrences, which means that, with a unit increase in the variables with negative coefficients, the probability of wildfire occurrences will decrease. Thus, the findings from the LR model concluded that the probability of wildfire occurrences is higher at lower elevations with a high land surface temperature and closer to the roads that are associated with a high frequency of human activities. The wildfires are more likely to occur

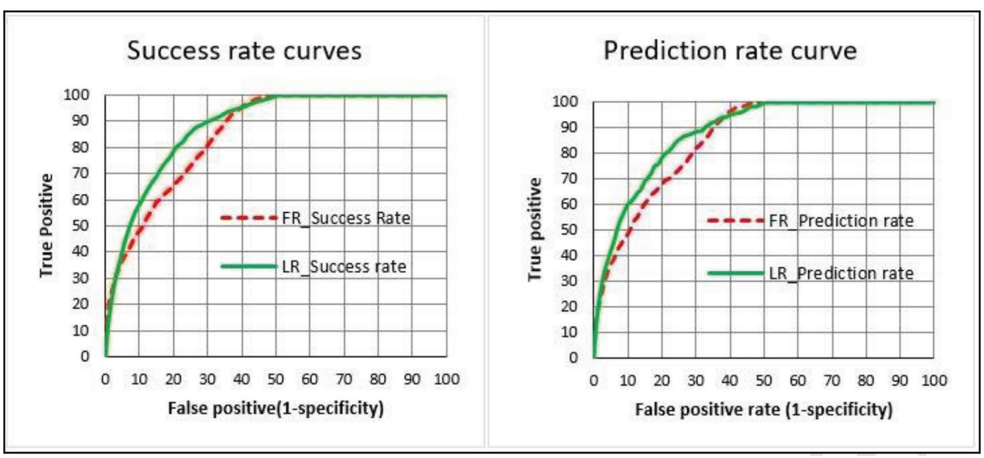

Figure 10. Final wildfire susceptibility map with hotspot from an optimum model

in the sloping valleys where most of the vegetation comprises shrubs and meadows/ grasslands with low humidity and less rainfall.

The analysis of the FR model revealed that the land surface temperature, slope, aspect, curvature, enhanced vegetation index, land use, Euclidean distance to roads, rivers, settlements and agricultural land, and population density have a positive correlation to the occurrence of wildfire in the study area, while elevation, rainfall, relative humidity, and topographic wetness index showed a negative correlation. As the value of positively correlated factors increases, the frequency of wildfires tends to increase, while for negatively correlated variables the frequency of wildfires tends to decrease their influence of wildfire occurrences. In addition, the prediction rates of each factor confirmed that the land surface temperature, followed by elevation, enhanced vegetation index, and distance to road are the highest contributing factors of wildfire occurrences. Factors like population density, distance to agriculture, rainfall, aspect, relative humidity, slope, and distance to settlements showed a moderate influence while other factors including curvature, topographic wetness index, and distance to rivers showed very little influence on the occurrences of wildfire.

The accuracy assessment and validation results showed the success rate with AUC values of 0.881 and 0.855 for the LR and FR models, respectively. This indicates that both models have a very good capability of classifying the wildfire susceptibility areas. Meanwhile, the prediction rate of the LR and FR models were 0.883 and 0.853, respectively. The results indicated that both models had a relatively high predictive capability to discriminate the presence and absence of wildfire in the study area. Nonetheless, the LR model is chosen as the optimum model for the final wildfire susceptibility mapping based on the comparative analysis.

The final wildfire susceptibility map revealed that the high and very high susceptibility zones covered 30% of the total study area and contained the majority (70%) of the total hotspots. These zones fall in sloping valleys in lower elevations associated with a high land surface temperature where the vegetation was dominant with shrubs and meadows, dry grasslands mixed with scattered conifers and blue pines. These zones also correspond to the area closer to the roads within a proximity of 1,500 m where active human activities were involved. This implies that areas closer to the roads were more susceptible to wildfires due to human activities that contribute to starting fires either accidentally or intentionally. In addition, the low rainfall and humidity in the area also contributed to high susceptibility to wildfires.

In a nutshell, it can be concluded that the integration of geoinformatics technology with GIS-based LR and FR models can effectively determine the most significant influential factors of wildfire occurrences and probability, and eventually develop the wildfire susceptibility map. The findings may provide valuable information that can guide and help in the effective wildfire management system of Bhutan. In addition, the methodology adopted in the current study may also have the potential to be implemented in the other areas of Bhutan that have similar environmental, climatic, and anthropogenic influence.

References

Adab, H., Devi, K., and Solaimani, K. (2013). Modeling forest fire risk in the northeast of Iran using remote sensing and GIS techniques. Nat Hazards, 65:1723-1743.

Atkinson, P.M. and Massari, R. (1998). Generalised linear modelling of susceptibility to land sliding in the Central Apennines, Italy. Comput. Geosci., 24:373-385.

Ayalew, L. and Yamagishi, H. (2005). The application of GIS-based logistic regression for landslide susceptibility mapping in the Kakuda-Yahiko Mountains, Central Japan. Geomorphology, 65(1-2):15-31. Brown, A.A. and Davis, K.P. (1973). Forest Fire: Control and Use. 2nd ed. McGraw-Hill, NY, USA, 686p.

Chang, K. (2014) . Introduction to Geographic Information System, 7th ed. McGraw-Hill Education, NY, USA, 425p.

Chung, C.J.F. and Fabbri, A.G. (2003). Validation of spatial prediction models for landslide hazard mapping. Nat. Hazard. Earth Sys., 30(3):451-472.

Chuvieco, E. and Congalton, R.G. (1989). Application of remote sensing and geographic information systems to forest fire hazard mapping. Remote Sens. Environ., 29(2):147-159.

Clark, W. and Hosking, P. (1986). Statistical Methods for Geographers. John Wiley and Sons, Hoboken, NJ, USA, 528p.

Dai, F.C. and Lee, C.F. (2002). Landslide characteristics and slope instability modelling using GIS, Lantau Island, Hong Kong. Geomorphology, 42:213-228.

Department of Forest and Park Services (2015). RNR monthly newsletter. A monthly publication of the Ministry of Agriculture and Forests (MoAF), Information and Communication Services (ICS), Department of Forest and Park Services (DoFPS), Thimphu, Bhutan.

Drukpa, K. (2015). Bhutan RNR Statistics. Royal Government of Bhutan, Ministry of Agriculture and Forests, Thimphu, Bhutan. Energy and Climate Intelligence Unit. (2015). ECIU Comparator Tool Reveals Surprising Results. Available from: http://www.eciu.net/press-releases/2015/comparator-tool. Accessed date:

Environmental Systems Research Institute, Inc. (2016). Available from: http://webhelp.esri.com/arcgis-desktop. Accessed date :

Giglio, L., Randerson, J.T., Van der Werf, GR., Kasibhatla, P.S., Collatz, G.J., Morton, D.C., and DeFries, R.S. (2010). Assessing variability and long-term trends in burned area by merging multiple satellite fire products. Biogeosciences, 7:1171-1186.

Ghobadi, J.G., Gholizadeh, B., and Dashliburun, O.M. (2012). Forest fire risk zone mapping from geographic information system in Northern Forests of Iran. International Journal of Agriculture and Crop Science, 4(12):818-824.Guo, F., Su, Z., Wang, G., Sun, L., Lin, F., and Liu, A. (2015). Wildfire ignition in the forests of southeast China: Identifying drivers and spatial distribution to predict wildfire likelihood. Appl. Geogr. 66:12-21.

Gyelmo, D. (2016). Forest fires burning up Bhutan's wilderness. Available from: https://www.thethirdpole.net/2016/03/21/forest-fires-burning-up-bhutans-wilderness/ Accessed date:

Intarawichian, L.S. and Dasananda, S. (2010). Forest fire susceptibility mapping based on logistic regression and frequency ratio methods: a case study of Chiangmai province, Thailand. Proceedings of the 33rd Asian Conference on Remote Sensing; November 26-30, 2012; Pattaya, Thailand, p. 2038-2045.

International Business Machines, Corp. (2012). IBM-SPSS File Help (Version 21) Armonk, NY, USA.

Jaiswal, R.K., Mukherjee, S., Raju, K.D., and Saxena, R. (2002). Forest fire risk zone mapping from satellite imagery and GIS. Int. J. Appl. Earth Obs., 4:1-10.

Lee, S. and Evangelista, D.G. (2008). Landslide susceptibility mapping using probability and statistics models in Baguio City, Philippines. Proceedings of the 31st International Symposium on Remote Sensing of Environment; Global Monitoring for Sustainability and Security; May 20-24, 2005; St. Petersburg, Russia, p. 396-399.

Lee, S. and Pradhan, B. (2007). Landslide mapping hazard at Selangor, Malaysia using frequency ratio and logistic regression models. Landslides, 4:33-41.

Meinhardt, M., Fink, M., and Tünschel, H. (2015). Landslide susceptibility analysis in central Vietnam based on an incomplete landslide inventory: Comparison of a new method to calculate weighting factors by means of bivariate statistics. Geomorphology 234:80-97.

Meten, M., Bhandary, N.P., and Yatabe, R.J. (2015). GIS based frequency ratio and logistic regression for landslide susceptibility mapping of Sina Debre area in central Ethiopia. J. Mt. Sci., 12(6):1355-1372.

Ministry of Agriculture and Forests. (2015). Bhutan RNR Statistics. Royal Government of Bhutan, Ministry of Agriculture and Forests, Thimphu, Bhutan, 133p.

Mohammadi, F., Bavaghar, M.P., and Shabanian, N. (2013). Forest fire risk zone modelling using logistic regression and GIS: an Iranian case study. Small-scale Forestry, 13:117-125.

O'Brien, R. (2007). A caution regarding rules of thumb for variance inflation factors. Qual. Quant., 41:673-690.

Oh, H.J., Kim, Y.S., Choi, J.K., and Lee, S. (2011). GIS mapping of regional probabilistic ground water potential in the area of Pohang City, Korea. J. Hydrol., 399:158-172.

Ohlmacher, G.C. and Davis, J.C. (2003). Using multiple logistic regression and GIS technology to predict landslide hazard in northeast Kansas, USA. Eng. Geol., 69:331-343.

Oliveira, S., Oehler, F., San-Miguel-Ayanz, J., Camia, A., and Pereira, J.M. (2012). Modeling spatial patterns of fire occurrence in Mediterranean Europe using multiple regression and random forest. Forest Ecol. Manag., 275:117-129.

Ozdemir, A. and Altural, T. (2013). A comparative study of frequency ratio, weights of evidence and logistic regression methods for landslide susceptibility mapping: Sultan Mountains, SW Turkey. J. Asian Earth Sci., 63:180-197.

Pourtaghi, Z.S., Pourghasemi, H.R., and Rossi, M. (2014). Forest fire susceptibility mapping in the Minudasht forests, Golestan province, Iran. Environmental Earth Sciences, 73:1515-1533.

Pradhan, B., bin Suliman, M.D.H., and bin Awang, M.A (2007). Forest fire susceptibility and risk mapping using remote sensing and geographical information systems (GIS). Disaster Prevention and Management, 16(3):344-352.

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

Rogerson, P. (2006). Statistical Methods for Geography. SAGE Publications, Thousand Oaks, CA, USA, 248p.

Royal Government of Bhutan. (1999). Vision 2020. A vision of peace, prosperity and happiness. Planning Commission, Thimphu, Bhutan. Part I, 40p.

Schicker, R. and Moon, V. (2012). Comparison of bivariate and multivariate statistical approaches in landslide susceptibility mapping at a regional scale. Geomorphology, 161-162:40-57.

Stocks, B.J., Fosberg, M.A., Lynham, T.J., Mearns, L., Wotton, B.M., Yang, Q., Jin, J-Z., Lawrence, K., Hartley, G.R., Mason, J.A., and McKenney, D.W. (1998). Climate change and forest fire potential in Russian and Canadian boreal forests. Climatic Change, 38:1-13.

Syphard, A.D., Radeloff, V.C., Keuler, N.S., Taylor, R.S., Hawbaker, T.J., Stewart, S.I., and Clayton, M.K. (2008). Predicting spatial patterns of fire on a southern California landscape. Int. J. Wildland Fire, 17(5):602-613.

Tshering, K. (2006). Development of an effective forest fire management strategy for Bhutan, [MSc. Thesis]. College of Forestry and Conservation, University of Montana. Missoula, MT, USA, 87p.

Yalcin, A., Reis, S., Aydinoglu, A.C., and Yomralioglu, T. (2011). A GIS-based comparative study of frequency ratio, analytical hierarchy process, bivariate statistics and logistic regression methods for landslide susceptibility mapping in Trabzon, NE Turkey. Catena, 85(3):274-287.

Yesilnacar, E. and Topal, T. (2005). Landslide susceptibility mapping: A comparison of logistic regression and neural networks methods in a medium scale study, Hendek region (Turkey). Eng. Geol., 79(3-4): 251-266.

Zhang, H., Han, X., and Dai, S. (2013). Fire occurrence probability mapping of Northeast China with binary logistic regression model. IEEE J. Sel. Top. Appl., 6(1):121-127.

Zhang, Z.X., Zhang, H.Y., and Zhou, D.W. (2009). Using GIS spatial analysis and logistic regression to predict the probabilities of human-caused grassland fires. J. Arid. Environ., 74(3):386-393.

Zhu, L. and Huang, J. (2006). GIS-based logistic regression method for landslide susceptibility mapping in regional scale. J. Zhejiang Univ-SC A., 7(12):2007-2017.