Journal of Sustainability Science and Management *eISSN: 2672-7226*
Volume 14 Number 3, June 2019: 55-66 *© Penerbit UMT*

SPATIAL LOGISTIC REGRESSION MODELS FOR PREDICTING PEATLAND FIRE IN BENGKALIS REGENCY, INDONESIA

SANDHI IMAM MAULANA[1*], LAILAN SYAUFINA[2], LILIK BUDI PRASETYO[3], AND MUHAMMAD NUR AIDI[4]

[1]Study Program of Natural Resources and Environmental Management Science, Bogor Agricultural University (IPB), Bogor 16144, Indonesia
[2]Department of Silviculture, Faculty of Forestry, IPB, Bogor 16680, Indonesia
[3]Department of Forest Resources Conservation and Ecotourism, Faculty of Forestry, IPB, Bogor 16680, Indonesia
[4]Department of Statistics, Faculty of Mathematics and Natural Sciences, IPB, Bogor 16680, Indonesia

**Corresponding email: sandhimaulana2018@gmail.com*

Abstract: Peat land fires have received increasing attentions as a major recurrent environmental problem in Indonesia, particularly across the eastern coast of Sumatera Island, where Bengkalis Regency is located. Although peat land fire prediction analysis has become an essential aspect of fire management, however, it seems that studies are still very limited in Indonesia. Therefore, this study objective is to develop a prediction models for peat land fire particularly in the Bengkalis Regency, Riau Province-Indonesia. This study was conducted based on spatial logistic regression method, in which that a prediction model that can provide preparation time up to 2 months before the beginning of peat land fire season was selected to produce a prediction map based on physiographic variables, peat physical characteristics variables, human activity variables, and climate variables. Performance of the selected peat land fire prediction model has been verified and validated using an independent testing subset, and the results showed are consistently reliable (Chi-square, p <0.001; Nagelkerke R^2=0.314; AUC=0.8309; Overall accuracy of correct predicted value=85.16%). This finding is useful to improve peat land fire management in Bengkalis Regency, and can also be used to help the authorities in the spatial domain to make more appropriate decisions related to fire prevention strategies.

Keywords: Fire-prediction model, fire management, spatial logistic-regression method, Sumatra Island, Riau

Introduction

Nowadays, peat land fires have received considerable attentions since it is also proven to trigger other environmental and health related issues, such as increasing threat to endangered flora and fauna due to significant habitat loss (Curran *et al.*, 2004; Barlow & Silveira, 2009), huge amount of greenhouse gasses emission (Simmonds *et al*., 2005; Page *et al.*, 2011), as well as greater intensity of smoke haze disaster (Heil *et al.*, 2007; Wooster *et al.*, 2012). Specifically, in regard to smoke haze disaster, Koplitz *et al*., (2016) have reported that the smoke haze generated from fire throughout Sumatera in 2015 had caused health problem to more than 100,000 people across Indonesia, Malaysia, and Singapore.

In general, peat land fire in a tropical country like Indonesia could only be triggered by human activities (Tacconi *et al.*, 2007; Murdiyarso & Adiningsih, 2007; Hooijer *et al*., 2012) Nevertheless, many studies also reveal that series of those major burning actions were strongly supported and related to prolonged drought characteristics due to El Nino seasonal phenomenon (van der Werf *et al*., 2008; Chen *et al*., 2011; Wooster et al., 2012; Doblas-Reyes *et al.*, 2013; Spessa *et al*., 2015). During the warm phase of *El Nino Southern Oscillation* (ENSO), *Sea Surface Temperature* (SST) in the western region of Pacific tends to drop from its normal average, causing significant reduction in precipitation rate and prolonging drought period that eventually increasing the risk of fire, particularly across degraded area (Siegert *et al.*, 2001; Aldrian & Susanto, 2003; Zhao & Yang, 2014). Hence, if the relationship between seasonal phase of ENSO and series of peat land fire events be thoroughly studied, it is possible that the risk of incoming peat land fire may be predicted and anticipated accurately several months prior the event took place (Li *et al.*, 2008; Jin *et al.*, 2008; Barnston *et al*., 2010).

Previously, many studies have been conducted to develop prediction models of seasonal fire. Using correlation information between series of forest fire and seasonal sea surface temperature movements. Chen *et al*. (2011) have successfully established prediction model for fire season severity in South America with 3-5 months lead times. Meanwhile, based on logistic regression approach, del

Hoyo *et al.*, (2011) have produced models for human-caused wildfire risk estimation in Spain. Harnessing similar approach, Mohammadi *et al.*, (2013) have developed forest fire risk zone modelling in Iran, whilst Pan *et al.*, (2016) have built probabilistic models of fire occurrences and fire risk zoning in China. Although all of those studies indicated current advances in the development of fire prediction models across the globe, however, it seems that such studies are still very limited in Indonesia. Considering this issue, by using spatial logistic regression approach, this study primarily aims to develop prediction models for peat land fire in Bengkalis Regency, Riau Province-Indonesia. Spatial logistic regression was chosen in this study since this approach is reasonably flexible and is able to accept a mixture of both numerical and categorical variables (Catry *et al.*, 2009; Pan *et al.*, 2016). This approach is also known as the most commonly used to develop fire prediction models which typically contain a binary response variable with "0" represents fire absence and "1" represents fire presence (Bisquert *et al.*, 2012; Mohammadi *et al.*, 2013; Zhang *et al.*, 2014). The development of such prediction model is essentially needed to provide rapid and accurate information related to the probability of incoming peat land fire event, so that it is possible to prepare appropriate fire prevention and suppression strategies.

Materials and Methods

Research Area

This study was conducted in Bengkalis Regency, which is located in Riau Province-Indonesia, as illustrated in Figure 1. This regency was chosen since 65% of the total area is peat land (Syaufina & Hafni, 2018). The regency is also considered as the largest contributor of peat land fires in Riau Province (Rosul, 2015). Geographically, this regency is located at 2°7'37.2"-0°55'33.6" North and 100°57'57.6"-102°30'25.2" East. Administratively, this regency covers eleven districts, including the Bandar Laksamana, Bathin Solapan, Talang Muandau, Kecamatan Mandau, Pinggir, Bukit Batu, and Siak Kecil that are located across eastern coast of Sumatera Island. Rupat and Rupat Utara are located in the Rupat Island, as well as Bengkalis and Bantan are located in the Bengkalis Island. According to the Indonesian Centre for Statistic Agency (BPS, 2017), overall land extent of this regency is about 7.773,93 km^2, which is dominated by lowland plain with the altitude ranging from 2 to 6 meters above sea level.

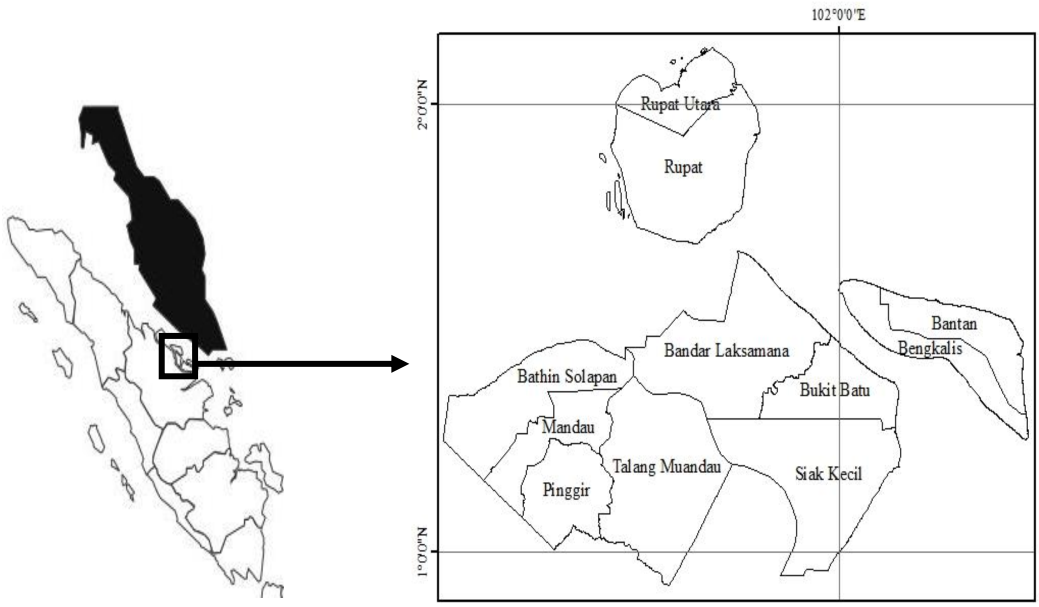

Figure 1: Research area

Data Collections

Dependent variable in this study is the burned area on peat land in Bengkalis Regency, that based on binary logistical approach for each cell sized 100 x 100 meter, value of cell with burned area detected will be defined as 1, whilst cell with no burned area detected will be defined as 0. Meanwhile independent variables in this study are including physiographic variables (e.g. river and canal network density), peat physical characteristics variables (e.g. peat depth and peat decomposition type), human activity variables (e.g. type of peatland use and cover, and road network density), as well as climate variables (e.g. monthly precipitation rate and moisture index).

Data for dependent variable, which is peat land fire occurrences, were extracted from Moderate Resolution Imaging Spectroradiometer (MODIS)-MCD64a1 burned area product images. Since most of major peat land fire in Indonesia had taken place during *El Nino* years (Wooster *et al.*, 2012; Spessa *et al.*, 2015;

Field *et al.*, 2016), data collection of this dependent variable was focused on years with significant indication of El Nino, which are 2002, 2004, 2006, 2009, 2014 and 2015. The trend of peat land fire data is shown in Table 1. The data showed that for each of those years, the initial fire month ws indeed January, while peat land fires were only detected to presence until August.

In regard to independent variables, data for peat depth and peat type were extracted from peat land map from Wetlands International Indonesia (WII) after it was updated using peat land map from the Ministry of Agriculture-Republic of Indonesia that had been accessed through Global Forest Watch-World Resources Institute and the National Peat Ecosystem Function Map from The Ministry of Environment and Forestry-Republic of Indonesia. In the meantime, data related to land use and land cover, road density, river density, and canal density were extracted from Indonesian Topographic Map from the Agency for Geospatial Information-Republic of Indonesia.

In order to obtain monthly precipitation data, the Tropical Rainfall Measuring Mission (TRMM) 3B43v7 satellite images product that were obtained from the GES-DISC Interactive Online Visualization and Analysis Infrastructure (GIOVANNI)-NASA. Meanwhile, to obtain moisture index, the Normalized Difference Moisture Index (NDMI) from Landsat 7 and Landsat 8 satellite images were collected from the United States Geological Survey (USGS)-Earth Explorer. Subsequently, to define the optimum lead time of prediction models in this study, those monthly precipitation and moisture index data were collected and calculated following 4-monthly moving average approach, for the time span of one (t-1) to three (t-3) months prior the initial burning month.

Table 1: Trend of peatland fire in Bengkalis Regency, Indonesia

Year	Burned Area (Hectares)												
	Jan	Feb	Mar	Apr	May	Jun	Jul	Ags	Sep	Oct	Nov	Dec	Total
2002	527	8294	5151	-	-	-	20	-	-	-	-	-	13992
2004	2392	3713	142	-	-	1530	763	1379	-	-	-	-	8540
2006	546	225	474	-	-	-	1105	1616	-	-	-	-	2350
2009	818	3343	-	-	312	2696	3025	2328	-	-	-	-	10194
2014	731	21907	28831	318	191	861	1124	467	-	-	-	-	53963
2015	369	389	2379	898	126	723	1675	-	-	-	-	-	6559
Total	5383	37871	36977	1216	629	5810	7712	5790	-	-	-	-	95598

Source: calculated from MODIS 64a1 Burned Area Product satellite images

Table 1 shows that the initial burning month of peat land fire occurrences were started from January, so that it was defined in this study as the initial burning month, while December, November, and October of the prior year were defined as t-1, t-2, and t-3 respectively. Following the 4-monthly moving average calculation approach, it implies that climate data for October (t-3) was calculated as the average data from July, August, September, and October. While, climate data for November (t-2) was calculated as the average data from August, September, October and November. Then, climate data for December (t-1) was calculated as the average data from September, October, November, and December. To implement this moving average approach, since the trend of peatland fire was always started from January during the above-mentioned years (2002, 2004, 2006, 2009, 2014, 2015; Table 1), the collection of both climate data (precipitation and NDMI) were conducted from the month of December, November, October, September, August, and July from the prior year (2001, 2003, 2005, 2008, 2013, 2014). This kind of moving average calculation approach was previously implemented by Chen *et al.*, (2011), who have successfully developed forest fire prediction models with 3-5 months lead times.

Model Development

Overall process of model development and analysis in this study was conducted using R Studio software, which is principally an open source programming software that is able to support big data computation.

Model development was started by partitioning collected data into two separate data set, namely training data set and testing data set. For this purpose, we applied stratified random data partition approach, in which that collected data were firstly stratified into six strata based on land use/land cover type (e.g. activities and residential area, bare land, farming field, plantation, forest, and shrubs). Then for each stratum, the data were proportionally and randomly divided with proportion of 80% for training data that be used for model development, and the rest of 20% were kept as testing data set that be used for model validation using confusion matrix mechanism. Similar data partition approach was also implemented by several previous studies such as del Hoyo *et al.*, (2011), Quintano *et al.*, (2011), Mohammadi *et al.*, (2013), and Pan *et al.*, (2016). Details of the number of data partitioned per stratum in this study is shown in Table 2, where it can be seen that the total cell population (N) in this study was 521.948 cells. Overall, 80% of that cell population, or about 417.558 cells, were used to develop prediction model; while the rest 20 %, or about 104.390 cells were used to validate the developed models.

Table 2: Data partition per stratum

Strata	Number of cells partitioned		Total Cells (100%)
	Train Set (80%)	Test Set (20%)	
Activities and residential area	2826	707	3533
Bare Land	741	185	926
Farming Field	2580	645	3225
Forest	187854	46964	234818
Plantation	176893	44223	221116
Shrubs	46664	11666	58330
Total cells	417558	104390	521948

Remark: size of each cell is 100x100m

Subsequently, as suggested by several previous studies, such as Paciorek (2006), Zhang *et al.*, (2010), del Hoyo *et al.*, (2011), Mohammadi *et al.*, (2013), and Pan *et al.*, (2016), spatial logistic regression was then applied to develop peatland fire prediction models using training data set that had already been partitioned previously. Basic equation of spatial logistic regression used to develop models in this study can be written as follow:

$$P=\frac{\mathrm{Exp}\,(\beta_0+\beta_1*\mathrm{Cnl\_Den}+\beta_2*\mathrm{Rvr\_Den}+\beta_3*\mathrm{Rd\_Den}+\beta_4*\mathrm{Pt\_LUC}+\beta_5*\mathrm{Pt\_Depth}+\beta_6*\mathrm{Pt\_Type}+\beta_7*\mathrm{NDMI}+\beta_8*\mathrm{Precip})}{1+\mathrm{Exp}\,(\beta_0+\beta_1*\mathrm{Cnl\_Den}+\beta_2*\mathrm{Rvr\_Den}+\beta_3*\mathrm{Rd\_Den}+\beta_4*\mathrm{Pt\_LUC}+\beta_5*\mathrm{Pt\_Depth}+\beta_6*\mathrm{Pt\_Type}+\beta_7*\mathrm{NDMI}+\beta_8*\mathrm{Precip})} \quad (1)$$

Where:

- P is the probability that a fire occurs
- β_0 is the intercept
- β_1, β_2, ... , β_8 are numerical coefficients of each variable
- Cnl_Den is the density of canal network (numerical variable in square map unit)
- Rvr_Den is the density of river network (numerical variable in square map unit)
- Rd_Den is the density of road network (numerical variable in square map unit)
- Pt_LUC is the type of peat land use and land cover (categorical variable covering six classes, namely activities and residential area, bare land, farming field, plantation, forest, and shrubs)
- Pt_Depth is the class of peat depth (categorical variable covering four classes, namely D1 for peat depth ranging from 50-100 centimeters, D2 for peat depth ranging from 100-200 centimeters, D3 for peat depth ranging from 200-400 centimeters, and D4 for peat depth more than 400 centimeters)
- Pt_Type is the type of peat decomposition (categorical variable covering eight classes, namely H1a for peat decomposition type of 60%Hemic/40%Sapric at D1 depth class, H2a for peat decomposition type of 60%Hemic/40%Sapric at D2 depth class, H3a for peat decomposition type of 60%Hemic/40%Sapric at D3 depth class, H4a for peat decomposition type of 60%Hemic/40%Sapric at D4 depth class, S1a for peat decomposition type of 60% Sapric/40%Hemic at D1 depth class, S2a for peat decomposition type of 60% Sapric/40%Hemic at D2 depth class, S2c for peat decomposition type of 50% Sapric/50%Mineral at D2 depth class, and S3a for peat decomposition type of 60% Sapric/40%Hemic at D3 depth class)
- NDMI is the monthly Normalized Difference Moisture Index (numerical variable with value ranging from -1 to 1)
- Precip is the monthly precipitation rate (numerical variable in mm/month)

Model Evaluation

After prediction models were established, we then conducted verification tests using the value of Chi square test, Nagelkerke R^2, and also area under

Receiver Operating Characteristic (ROC) curve, that has been also known as AUC test. In addition to those tests, we also conducted validation test to calculate the accuracy of established models by comparing prediction values with actual values from testing data set that had been partitioned previously. For this purpose, we applied confusion matrix, in which that the accuracy of a prediction model was calculated as a ratio (%) between the sum of true negative (TN) and true positive (TP) divided by total number of testing cells, as illustrated in Table 3.

Table 3: Confusion matrix for model validation

	Predicted: No Fire (0)	Predicted: Fire Occur (1)
Actual: No Fire (0)	*True Negative* (TN)	*False Positive* (FP)
Actual: Fire Occur (1)	*False Negative* (FN)	*True Positive* (TP)

Remark: Model accuracy = (TN+TP)/(Total Cells); where total cells = TP+TN+FP+FN

Results and Discussion

Peatland Fire Prediction Models

Harnessing spatial logistic regression approach, we have established three prediction models of peat land fire for initial fire month of January. Firstly, the t-1 model (Table 4), which we used the 4-monthly moving average of NDMI and precipitation at December. Secondly, the t-2 model (Table 5), which we used 4-monthly moving average of NDMI and precipitation at November. Finally, the t-3 model (Table 6), which we used 4-monthly moving average of NDMI and precipitation at October. In addition to those climate variables, the developed models also incorporated physiographic variables (e.g. river and canal network density), peat physical characteristics variables (e.g. peat depth and peat decomposition type), and human activity variables (e.g. type of peat land use and cover, and road network density). Overall, those explanatory variables illustrate good level of significance to be used as predictors of peat land fire in Bengkalis Regency.

Interestingly, as illustrated in Table 4, 5, and 6, it seems that the first level for each categorical explanatory variable, namely activities and residential area in peatland use and cover, D1 in peat depth, and H1a in peat type are not shown. As explained by Jaccard (2001), this often happen since the coefficients of those first levels of categorical explanatory variables are already absorbed into the models' intercept. Additionally, looking at the coefficient of explanatory variables presented in Table 4, 5, and 6, it seems that if their absolute values were displayed in descending sequence, then the variables that have the greatest portion in defining peatland fire occurrences are the climate variable (NDMI), followed by human activity variable (peatland use and cover), then peat physical characteristics variable (peat type), and lastly physiographic variable (river density). This finding is in line with several previous studies (Tacconi *et al.*, 2007; Hooijer *et al.*, 2012; Wooster *et al.*, 2012; Doblas-Reyes *et al.*, 2013;; Mohammadi *et al.*, 2013; Spessa *et al.*, 2015) that explained that climate dynamics and human activities are the most determining factors in forest and peatland fire ignitions.

Table 4: Explanatory variables and significance levels for prediction model of December (t-1)

	Estimate	Std. Eror	z value	Pr(>|z|)
(Intercept)	-9.891E+00	2.006E-01	-4.930E+01	< 2E-16***
Cnl_Den	-1.598E-03	8.506E-05	-1.879E+01	< 2E-16***
Rvr Den	-1.103E-02	2.444E-04	-4.511E+01	< 2E-16***
Rd_Den	5.074E-03	5.935E-05	8.550E+01	< 2E-16***
Pt_LUC Bare Land	4.998E+00	1.940E-01	2.577E+01	< 2E-16***
Pt_LUC Farming Field	3.742E+00	1.830E-01	2.045E+01	< 2E-16***
Pt_LUC Forest	2.363E+00	1.770E-01	1.335E+01	< 2E-16***
Pt_LUC Plantation	2.994E+00	1.765E-01	1.696E+01	< 2E-16***
Pt_LUC Shrubs	3.311E+00	1.767E-01	1.873E+01	< 2E-16***
Pt_Depth D2	8.242E-02	1.430E-01	5.760E-01	5.64E-01***
Pt_Depth D3	7.892E-01	1.609E-01	4.906E+00	9.28E-07***
Pt_Depth D4	-3.266E-01	1.766E-01	-1.849E+00	6.44E-02 .
Pt_Type H2a	9.861E-01	1.442E-01	6.838E+00	8.02E-12***
Pt_Type H3a	-1.329E+00	1.623E-01	-8.186E+00	2.71E-16***
Pt_Type H4a	8.201E-01	1.774E-01	4.623E+00	3.79E-06***
Pt_Type S1a	1.483E+00	3.775E-02	3.929E+01	< 2E-16***
Pt_Type S2a	2.069E+00	1.448E-01	1.429E+01	< 2E-16***
Pt_Type S2c	1.324E+00	1.496E-01	8.854E+00	< 2E-16***
Pt_Type S3a	2.737E+00	1.624E-01	1.685E+01	< 2E-16***
NDMI_December (t-1)	-7.569E+00	9.140E-02	-8.282E+01	< 2E-16***
Precip_December (t-1)	2.482E-02	3.573E-04	6.948E+01	< 2E-16***

Significant codes: 0 '***'; 0.001 '**'; 0.01'*'; 0.05'.'; 0.1 ' '; 1
Initial fire month: January

Table 5: Explanatory variables and significance levels for prediction model of November (t-2)

	Estimate	Std. Eror	z value	Pr(>|z|)
(Intercept)	-1.230E+01	2.166E-01	-5.679E+01	< 2e-16***
Cnl_Den	-1.433E-03	8.485E-05	-1.689E+01	< 2e-16***
Rvr Den	-1.156E-02	2.456E-04	-4.708E+01	< 2e-16***
Rd_Den	4.998E-03	5.941E-05	8.412E+01	< 2e-16***
Pt_LUC Bare Land	5.092E+00	1.937E-01	2.628E+01	< 2e-16***
Pt_LUC Farming Field	3.637E+00	1.829E-01	1.988E+01	< 2e-16***
Pt_LUC Forest	2.447E+00	1.768E-01	1.384E+01	< 2e-16***
Pt_LUC Plantation	3.034E+00	1.763E-01	1.721E+01	< 2e-16***
Pt_LUC Shrubs	3.373E+00	1.765E-01	1.911E+01	< 2e-16***
Pt_Depth D2	1.182E-01	1.436E-01	8.230E-01	4.10E-01***
Pt_Depth D3	7.783E-01	1.611E-01	4.832E+00	1.35E-06***
Pt_Depth D4	-2.696E-01	1.766E-01	-1.526E+00	1.27E-01***
Pt_Type H2a	9.614E-01	1.448E-01	6.640E+00	3.14E-11***
Pt_Type H3a	-1.206E+00	1.625E-01	-7.419E+00	1.18E-13***
Pt_Type H4a	8.386E-01	1.774E-01	4.727E+00	2.28E-06***
Pt_Type S1a	1.544E+00	3.764E-02	4.101E+01	< 2e-16***
Pt_Type S2a	2.033E+00	1.453E-01	1.399E+01	< 2e-16***
Pt_Type S2c	1.352E+00	1.501E-01	9.008E+00	< 2e-16***
Pt_Type S3a	2.604E+00	1.626E-01	1.602E+01	< 2e-16***
NDMI_November (t-2)	-6.577E+00	8.533E-02	-7.708E+01	< 2e-16***
Precip_November (t-2)	3.393E-02	4.952E-04	6.851E+01	< 2e-16***

Initial fire month: January

Table 6: Explanatory variables and significance levels for prediction model of October (t-3)

	Estimate	Std. Eror	z value	Pr(>|z|)
(Intercept)	-1.211E+01	2.296E-01	-5.272E+01	< 2e-16***
Cnl_Den	-7.066E-04	8.375E-05	-8.437E+00	< 2e-16***
Rvr Den	-1.318E-02	2.453E-04	-5.371E+01	< 2e-16***
Rd_Den	4.747E-03	6.029E-05	7.875E+01	< 2e-16***
Pt_LUC Bare Land	5.352E+00	1.935E-01	2.766E+01	< 2e-16***
Pt_LUC Farming Field	3.586E+00	1.830E-01	1.959E+01	< 2e-16***
Pt_LUC Forest	2.832E+00	1.765E-01	1.605E+01	< 2e-16***
Pt_LUC Plantation	3.274E+00	1.760E-01	1.860E+01	< 2e-16***
Pt_LUC Shrubs	3.743E+00	1.762E-01	2.124E+01	< 2e-16***
Pt_Depth D2	1.263E-01	1.428E-01	8.850E-01	3.76E-01***
Pt_Depth D3	6.566E-01	1.607E-01	4.086E+00	4.38E-05***
Pt_Depth D4	-3.230E-01	1.733E-01	-1.864E+00	6.23E-02**·
Pt_Type H2a	8.133E-01	1.440E-01	5.648E+00	1.62E-08***
Pt_Type H3a	-9.853E-01	1.621E-01	-6.078E+00	1.22E-09***
Pt_Type H4a	7.953E-01	1.740E-01	4.570E+00	4.89E-06***
Pt_Type S1a	1.677E+00	3.739E-02	4.485E+01	< 2e-16***
Pt_Type S2a	2.108E+00	1.445E-01	1.458E+01	< 2e-16***
Pt_Type S2c	1.258E+00	1.493E-01	8.429E+00	< 2e-16***
Pt_Type S3a	2.174E+00	1.619E-01	1.342E+01	< 2e-16***
NDMI_October (t-3)	-6.658E+00	8.193E-02	-8.126E+01	< 2e-16***
Precip_October (t-3)	3.533E-02	6.382E-04	5.536E+01	< 2e-16***

Initial fire month: January

Having established above mentioned peatland fire prediction models, we then conducted several tests to verified and validate them, where details of the results are shown in Table 7. Based on this table, although there is a slight decreasing trend in the result along with the increase in models' lead times, it can be seen that all of those models are able to produce very good values for Chi square test, Nagelkerke R^2, AUC and overall accuracy test. Looking first at Chi-square test result, this test has been commonly harnessed to calculate p-value for logistic regression based models, to indicate the significance of the overall model (Mangiafico, 2015). Based on this notion, overall, it can be said that the three models established in this study are highly significance to predict peatland fire occurrences since their p-value are constantly maintained at <0.001. Meanwhile, their Nagelkerke R^2 range, which is 0.314-0.326, is above the average value of Nagelkerke R^2 produced by previous studies, such as Mohammadi *et al*., (2013) and Pan *et al*., (2016).

Afterwards, further examining the models, we calculated the value of area under the Receiver Operating Characteristic (ROC) curve that has been known as the AUC. Principally, according to Massada *et al.,* (2013), the ROC plot illustrates relationship between the true-positive rate (sensitivity or the proportion of ignitions correctly predicted) and the false-positive error rate (1-specificity; where specificity is the proportion of non-ignitions correctly predicted) for each threshold value to the probability of the presence predicted by the model. The AUC value is ranging from 0.5 - 1, where 0.5 represent worthless prediction model since it equal to a fully random prediction, while 1 implies perfect prediction model (McCune *et al*., 2002; Pan *et al*., 2016). Based on this perspective, it seems that models established in this study are able to conduct reliable performance to predict peatland fire since their AUC value are 0.8333, 0.8331, and 0.8309 for t-1, t-2, and t-3 model respectively (Figure 2).

Moving on to the validation measure to evaluate the models overall accuracy of correct predicted peatland fire occurrences. It can be seen on Table 7 that those models ability to produce true positive cells (correct predicted fire occurrences cell) is decreasing along with the increase in their lead times. In other words, it can be said that the longer the lead time, the lower number of correct predicted fire occurrences cells will be produced by the models. Nevertheless, all those models (t-1, t-2, & t-3) still manage to hold relatively stable overall accuracy at 85%.

In regard to the difference between lead times and available preparation time (Table 7), it should be kept in mind that although our t-3 model is able to predict peatland fire occurrences in January by using 4-monthly moving average climate data at October (3 months prior the event happens), however, climate data for October could only available to be accessed in November. Hence, the available time to prepare before initial fire month are only 2 months.

Similarly, for t-2 model that is able to predict peatland fire occurrences in January by using 4-monthly moving average climate data at November (2 months prior the event happens), however, climate data for November could only available to be accessed in December. Therefore, the available time to prepare before initial fire month are only one month. Subsequently, for t-1 model that is able to predict peatland fire occurrences in January by using 4-monthly moving average climate data at December (a month prior the event happens), however, climate data for December could only available to be accessed in January. This implies that there is no time available to prepare before initial fire month. Since our focus in this study is to establish prediction model with the maximum possible lead time and preparation time, so that we selected t-3 model to produce peatland fire prediction map (Figure 3). The equation (1) of this model can be written mathematically as follow:

$$P = \frac{\text{Exp}\{-12.11+(-0.0007066*\text{Cnl Den})+ -0.01318*\text{Rvr Den}+ \ldots\ +(0.03533*\text{Precip})\}}{1+\text{Exp}\{-12.11+(-0.0007066*\text{Cnl Den})+ -0.01318*\text{Rvr Den}+ \ldots\ +(0.03533*\text{Precip})\}}$$

Table 7: Result of verification and validation tests to developed peatland fire prediction models for initial fire month of January

	Peatland Fire Prediction Models		
	December (t-1)	November (t-2)	October (t-3)
Lead time	1 month	2 months	3 months
Available preparation time	0	1 months	2 months
Verification measures:			
Chi square	88045	86683	84404
	p <0.001	p <0.001	p <0.001
Nagelkerke R^2	0.326	0.321	0.314
Area under ROC curve (AUC)	0.8333	0.8331	0.8309
Validation measure:			
Number of true positive cells (correct predicted fire occurrences)	5063	4792	4309
Number of true negative cells (correct predicted no fire occurrences)	84031	84095	84594
Total number of testing cells	104390	104390	104390
Overall accuracy of correct predicted values	85.35%	85.15%	85.16%

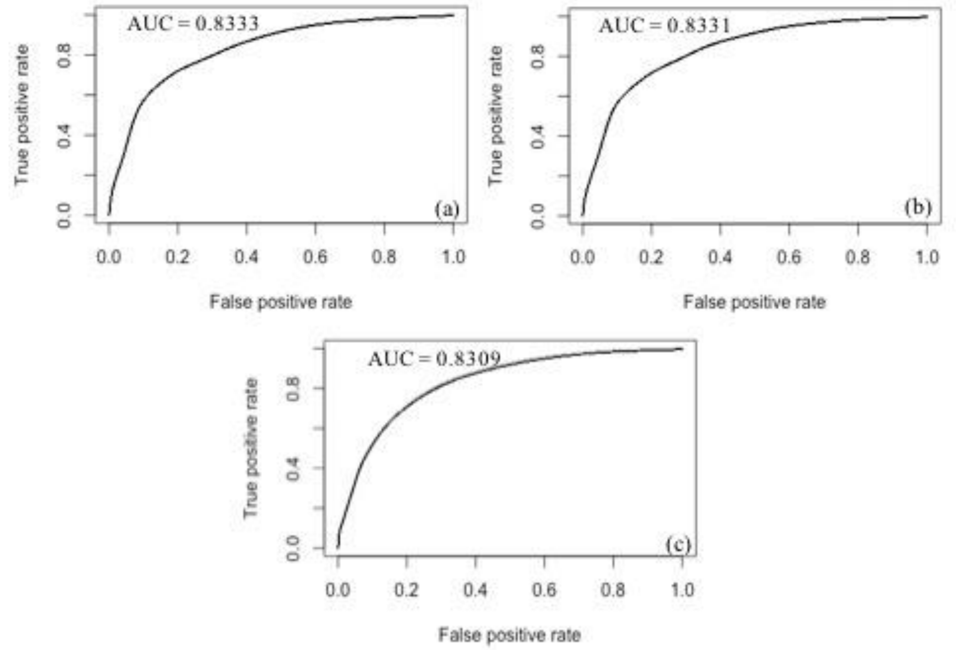

Figure 2: Calculated values of area under the Receiver Operating Characteristic curve (AUC) for (a) t-1 model, (b) t-2 model, and (c) t-3 model

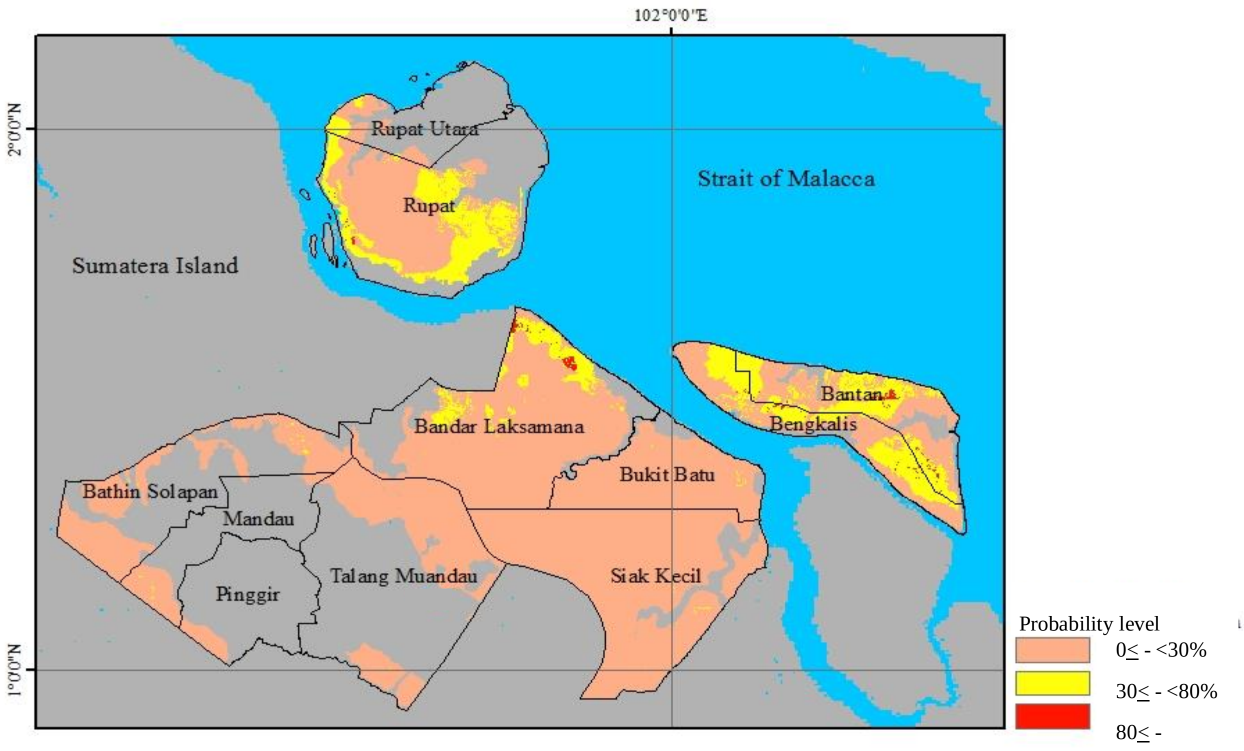

Figure 3: Peatland fire prediction map for initial fire month of January based on prediction model of October (t-3)

Influencing Factors of Peatland Fire in Bengkalis Regency

In order to investigate predicted peatland fire occurrences related to influencing factors, established prediction map (Figure 3) was classified into three levels of probability, as suggested by Giglio (2015). We then analyze the third level, which is predicted fire occurrences with probability $\geq$ 80%, since according to Giglio (2015) areas with this level of probability were seen to be highly prone to fire occurrences. Results of this examination were depicted in Table 8, in which it can be seen that about 96% predicted peatland fire are related to NDMI range of -0.06 – 0.30; whilst 99% of predicted peatland fire are detected in areas with monthly precipitation range of 230-246 mm/month, or about 7.6 – 8.2 mm/day. This finding is consistent with a recent study by Field *et al.*, (2016), who reported that forest and peatland fire in Sumatera are strongly related to climate variables, and tend to emerge when the daily precipitation rate less than 10 mm/day.

Meanwhile, there are three types of peatland use and cover that have been identified with predicted probability of peatland fire $\geq$ 80%, namely, shrubs, plantation, and forest. Among those three, it seems that we need to be more aware to peatland that have been used as plantation, since 92% of predicted peatland fire with probability level $\geq$ 80% are detected in this area. This finding is also in line with findings reported by Tacconi (2003), and Tacconi *et al.*, (2007), who stated that fires were often used during land opening and clearing, particularly in plantations and planted forests in Sumatera. In regard to peat physical characteristics, it can be seen that 90% of peatland fire are predicted to take place at peat depth of above 100 cm, whereas 78% of predicted fire are likely to take place at hemic peat type. Considering canal and river network density, it seems that both of these topographic variables are having a negative relationship with peatland fire prediction, since the lower the density of both variables, the higher the proportion of peatland fire predicted. On the other hand, road density is posing a positive trend towards peatland fire prediction.

Table 8: Predicted peatland fire occurrences related to influencing factors

Variable	Classes	Fire occurrences prediction with probability level above 80%	
		Number of identified cells	Proportion
NDMI (numerical index)	-0.06 - 0.15	902	46%
	0.15 - 0.30	988	50%
	0.30 - 0.45	89	4%
	Total identified cells	1979	
Precipitation (mm/month)	222-230	22	1%
	230-238	1456	74%
	238-246	501	25%
	Total identified cells	1979	
Peatland use and cover (categorical)	Shrubs	64	3%
	Plantation	1830	92%
	Forest	85	4%
	Total identified cells	1979	
Peat depth (categorical)	D1 (50-100 cm)	189	10%
	D2 (100-200 cm)	888	45%
	D3 (20-400 cm)	366	18%
	D4 (>400 cm)	536	27%
	Total identified cells	1979	
Peat type (categorical)	H1a (60%Hemic/40%Sapric at D1)	60	3%
	H2a (60%Hemic/40%Sapric at D2)	650	33%
	H3a (60%Hemic/40%Sapric at D3)	294	15%
	H4a (60%Hemic/40%Sapric at D4)	542	27%
	S1a (60% Sapric/40%Hemic at D1)	126	6%
	S2a (60% Sapric/40%Hemic at D2)	220	11%
	S2c (50% Sapric/50%Mineral at D2)	11	1%
	S3a (60% Sapric/40%Hemic at D3)	76	4%
	Total identified cells	1979	
Canal Density (km/km^2)	0 - 0.5	570	29%
	0.5 - 1.1	1192	60%
	1.1 - 1.6	217	11%
	Total identified cells	1979	
River Density (km/km^2)	0 - 0.1	1806	91%
	0.1 - 0.2	162	8%
	0.2 - 0.3	11	1%
	Total identified cells	1979	
Road Density (km/km^2)	0 - 0.8	317	16%
	0.8 - 1.7	1087	55%
	1.7 - 2.5	575	29%
	Total identified cells	1979	

Model used: t-3 model (Chi-sq: 84404***; Nagelkerke R^2: 0.314; AUC: 0.8309; Overall accuracy: 85.16%)

Conclusion

This study presented the development of comprehensive peat land fire prediction models to be implemented in Bengkalis Regency, Indonesia. The spatial logistic regression method used in this study also shows that this kind of approach evidently can be a reliable way to establish fire prediction models. Such models have been presented in this report, in which that a prediction model that can provide preparation time up to 2 months before the beginning of peatland fire season was selected to produce a prediction map based on physiographic variables, peat physical characteristics variables, human activity variables, and climate variables. Performance of the selected peatland fire prediction model has been verified and validated using an independent testing subset, and the results showed are consistently reliable. The established prediction map can be used to answer important questions, particularly in regard to the causative factors of peatland fire ignition in the spatial domain, so that it can help the authorities to undertake necessary and sufficient prevention measures.

Acknowledgements

The authors are grateful to the Indonesia Endowment Fund for Education (Lembaga Pengelola Dana Pendidikan/LPDP) for providing financial support. Thanks are due to the Center of Statistic Agency (Badan Pusat Statistik/BPS), Geospatial Information Agency (Badan Informasi Geospasial/BIG), the Ministry of Environment and Forestry, the Ministry of Agriculture, Wetlands International Indonesia (WII), World Resources Institute (WRI) and the National Aeronautics and Space Administration (NASA) who have supported us with important data needed in this study. Sincere thanks are also due to all reviewers of this manuscript who have given us valuable suggestions and advices.

References

Aldrian, E., & Susanto, R. D. (2003). Identification of three dominant rainfall regions within Indonesia and their relationship to sea surface temperature. *International Journal of Climatology*, *23*, 1435–1452.

Barlow, J., & Silveira, J. M. (2009). The consequences of fire for the fauna of humid tropical forests. In M. Cochrane (Ed.), *Tropical Fire Ecology* (pp. 543–556). Berlin, Germany: Springer Praxis Books.

Barnston, A. G., Li, S., Mason, S. J., DeWitt, D. G., Goddard, L., & Gong, X. (2010). Verification of the first 11 years of IRI's Seasonal Climate Forecasts. *Journal of Applied Meteorology Climatology*, *49*, 493–520.

Bisquert, M., Caselles, E., Sanchez, J. M., & Caselles, V. (2012). Application of artificial neural networks and logistic regression to the prediction of forest fire danger in Galicia using MODIS data. *International Journal of Wildland Fire*, *21*, 1025–1029.

[BPS] Badan Pusat Statistik Kabupaten Bengkalis. (2017). *Bengkalis Regency in Figures 2017*. Bengkalis: BPS Kabupaten Bengkalis [in Indonesian].

Catry, F. X., Moreira, F., Duarte, I., & Acacio, V. (2009). Factors affecting post-fire crown regeneration in cork oak (Quercus suber L.) trees. *European Journal of Forest Research*, *128*(3), 231–240.

Chen, Y., Randerson, J. T., Morton, D. C., DeFries, R. S., Collatz, G. J., Kasibhatla, P. S., & Malier, M. E. (2011). Forecasting fire season severity in South America using sea surface temperature anomalies. *Science*, *334*, 787–791.

Curran, L. M., Trigg, S. N., & McDonald, A. K. (2004). Lowland forest loss in protected areas of Indonesian Borneo. *Science*, *303*, 1000–1003.

del Hoyo, L. V., Isabel, M. P. M., & Vega, F. J. M. (2011). Logistic regression models for human-caused wildfire risk estimation: analysing the effect of the spatial accuracy in fire occurrence data. *European Journal of Forest Resources*, *130*, 983–996.

Doblas-Reyes, F. J., García-Serrano, J., Lienert, F., Biescas, A. P., & Rodrigues, L. R. (2013). Seasonal climate predictability and forecasting: status and prospects. *Wiley Interdisciplinary Reviews: Climate Change*, *4*, 245–268.

Field, R.D., van der Werf, G.R., Fanin, T., Fetzer, E.J., Fuller, R., Jethva, H., (...) & Worden, H. M. (2016). Indonesian fire activity and smoke pollution in 2015 show persistent nonlinear sensitivity to El Nino-induced drought. *Proceedings of the National Academy of Sciences of the United States of America (PNAS)*, *113*(33), 9204–9209.

Giglio, L. (2015). *MODIS Collection 6 Active Fire Product User's Guide Revision A*. Maryland: Department of Geographical Sciences-University of Maryland.

Heil, A., Langmann, B., & Aldrian, E. (2007). Indonesian peat and vegetation fire emissions:

Study on factors influencing large-scale smoke haze pollution using a regional atmospheric chemistry model. *Mitigation and Adaptation Strategies for Global Change*, *12*(1), 113–133. https://doi.org/10.1007/s11027-006-9045-6

Hooijer, A., Page, S., Jauhiainen, J., Lee, W. A., Lu, X. X., Idris, A., & Anshari, G. (2012). Subsidence and carbon loss in drained tropical peatlands. *Biogeosciences*, *9*, 1053–1071.

Jaccard, J. (2001). *Interaction Effects in Logistic Regression*. Thousand Oaks, California: Sage Publications Inc.

Jin, E. K., Kinter, I. J. L., Wang, B., Park, C. K., Kang, L. S., Kirtman, B. P., & Yamagata, T. (2008). Current status of ENSO prediction skill in coupled ocean-atmosphere models. *Climate Dynamic*, *31*, 647–664.

Koplitz, S. N., Mickley, L. J., Marlier, M. E., Buonocore, J. J., Kim, P. S., Liu, T., … Myers, S. S. (2016). Public health impacts of the severe haze in Equatorial Asia in September–October 2015: demonstration of a new framework for informing fire management strategies to reduce downwind smoke exposure. *Environmental Research Letters*, *11*(9), 094023. https://doi.org/10.1088/1748-9326/11/9/094023

Li, S., Goddard, L., & DeWitt, D. G. (2008). Predictive skill of AGCM seasonal climate forecasts subject to different SST prediction methodologies. *Journal of Climate*, *21*, 2169–2186.

Mangiafico, S. S. (2015). *An R Companion for the Handbook of Biological Statistics*. New Brunswick, New Jersey: Rutgers Cooperative Extension.

Massada, A. B., Syphard, A. D., Stewart, S. I., & Radeloff, V. C. (2013). Wildfire ignition–distribution modelling: a comparative study in the Huron-Manistee National Forest, Michigan, USA. *International Journal of Wildland Fire*, *22*, 174–183.

McCune, B., Grace, J. B., & Urban, D. L. (2002). *Analysis of Ecological Communities*. Gleneden Beach, Oregon: MjM Software Design.

Mohammadi, F., Bavaghar, M. P., & Shabanian, N. (2013). Forest fire risk zone modelling using logistic regression and GIS: an Iranian case study. Small-scale Forestry. *Small-Scale Forestry*. https://doi.org/10.1007/s11842-013-9244-4

Murdiyarso, D., & Adiningsih, E. S. (2007). Climate anomalies, Indonesian vegetation fire and terrestrial carbon emissions. *Mitigation and Adaptation Strategies for Global Change*, *12*, 101–112.

Paciorek, C. J. (2006). Computational techniques for spatial logistic regression with large data sets. *Computational Statistics and Data Analysis*, *52*, 3631–3653.

Page, S. E., Rieley, J. O., & Banks, C. J. (2011). Global and regional importance of the tropical peatland carbon pool. *Global Change Biology*, *17*, 798–818.

Pan, J., Wang, W., & Li, J. (2016). Building probabilistic models of fire occurence and fire risk zoning using logistic regression in Shanxi Province, China. *Natural Hazards*, *81*, 1879–1899.

Quintano, C., Fernandez-Manso, A., Stein, A., & Bijker, W. (2011). Estimation of area burned by forest fire in Mediterranean countries: A remote sensing data mining perspective. *Forest Ecology and Management*, *262*, 1567–1607.

Rosul, P. (2015). *Analysis of Policy Effectiveness on Forest Fires in Riau, Indonesia*. M.Sc. Dissertation. Duke University, North Carolina.

Siegert, F., Ruecker, G., Hinrichs, A., & Hoffmann, A. A. (2001). Increased damage from fire in logged forests during droughts caused by El Nino. *Nature*, *414*, 437–440.

Simmonds, P. G., Manning, A. J., Derwent, R. G., Ciais, P., Ramonet, M., Kazan, V., & D., R. (2005). A burning question: Can recent growth rate anomalies in the greenhouse gases be attributed to large-scale biomas burning events? *Atmospheric Environment*, *39*, 2513–2517.

Spessa, A. C., Field, R. D., Pappenberger, F., Langner, A., Englhart, S., Weber, U., & Moore, J. (2015). Seasonal forecasting of fire over Kalimantan, Indonesia. *Natural Hazards and Earth System Science*, *15*, 429–442.

Syaufina, L., & Hafni, D. A. F. (2018). Variability of Climate and forest and peat fires occurrences in Bengkalis Regency, Riau. *Jurnal Silvikultur Tropika*, *9*(1), 60–68, [in

Indonesian].

Tacconi, L. (2003). Fire in Indonesia: Causes, costs and policy implications. *CIFOR Occasional Paper*, *38*.

Tacconi, L., Moore, P. F., & Kaimowitz, D. (2007). Fire in tropical forests-what is really the problem? Lesson from Indonesia. *Mitigations and Adaptation Strategies for Global Change*, *12*, 55–66.

van der Werf, G. F., Dempewolf, J., Trigg, S. N., Randerson, J. T., Kasibhatla, P. S., Giglio, L., & DeFries, R. S. (2008). Climate regulation of fire emissions and deforestation in equatorial Asia. *Proceedings of the National Academy of Sciences*, *105*(51), 20350–20355.

Wooster, M. J., Perry, G. L., & Zoumas, A. (2012). Fire, drought and El Nino relationships on Borneo (Southeast Asia) in the pre-MODIS