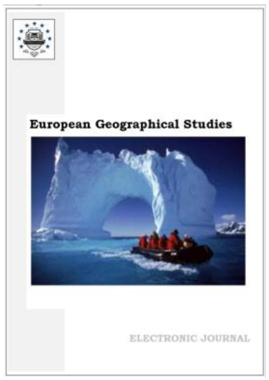
Published in the USA
European Geographical Studies
Has been issued since 2014.
E-ISSN: 2413-7197
2022. 9(1): 3-11

DOI: 10.13187/egs.2022.1.3
https://egs.cherkasgu.press



Articles

Forest Fire Risk Assessment and Mapping Using Support Vector Machine Algorithm, A Case Study in Nghe An Province, Vietnam

Thi Nam Phuong Doan [a, *], Van Trung Nguyen [a], Thi Thanh Hoa Pham [a], Thanh Ha Tran [a], Thi Thu Ha Le [a]

[a] Hanoi University of Mining and Geology, Hanoi, Vietnam

Abstract

In recent years, forest fires have occurred frequently in Vietnam due to the influence of climate change and human activities. This paper presents the results of modeling the risk of forest fires in the west of Nghe An province (north-central Vietnam) from remote sensing and GIS data. 09 factors affect the risk of forest fire, including vegetation cover (NDVI index), soil moisture (NMDI index), elevation, slope, aspect, wind speed, land surface temperature, average monthly precipitation and population density are used to build a model for mapping forest fire risk based on Suppor Vector Machine (SVM) algorithm. The past forest fire data is collected from the database of the Forest Protection Department (Vietnam Ministry of Agriculture and Rural Development) to evaluate the accuracy of the model. Different values of cost parameter (C) are tested to select the value with the highest accuracy in predicting forest fire risk. The results obtained in the study can be used effectively for monitoring and early warning of forest fire risk in the localities, helping to reduce damage caused by forest fires.

Keywords: forest fire risk, remote sensing, GIS, support vector machine algorithm, Nghe An province.

1. Introduction

A forest fire is a complex phenomenon that is difficult to model and manage. There are many factors that contribute to the ignition and spread of forest fires, such as human factors, topographic and meteorological variables. Development of forest fire risk prediction model and classification of forest fire risk is an urgent problem, providing timely information for the protection and development of forest resources. Geospatial technology, including remote sensing technology and geographic information system (GIS), has been widely and effectively used in building forest fire risk prediction models. Remote sensing and GIS allow the collection of data on forest cover for analysis, management, and modeling for early warning of the risk of forest fires.

Jaiswal et al. (2002) used analytical hierarchical method (AHP) to map forest fire risk in the Gorna Subwatershed area (Madhya Pradesh state, India) based on IRS satellite image database and GIS technology. The obtained results show that nearly 30 % of the study area is at "high" to "very

[*] Corresponding author
E-mail addresses: doanthinamphuong@humg.edu.vn (T.N.P. Doan)

high" forest fire risk, consistent with the locations where forest fires occur (Jaiswal et al., 2002). In the study (Wimberly et al., 2008), the authors used multi-temporal Landsat satellite image data to map the severity of forest fires in the southern Appalachians (North Carolina, USA). Normalized Burn Ration Index (NBR) calculated from near-infrared and short-wave infrared bands of Landsat imagery is used to evaluate the land cover change before and after the fire.

Many studies have used artificial intelligence techniques (artificial neural network, random forest, support vector machine) combined with AHP hierarchical analysis method to improve the accuracy of forest fire risk prediction models from remote sensing and GIS data (Vasilakos et al., 2009; Oliveira et al., 2012; Dieu Tien Bui et al., 2016, Dieu Tien Bui et al., 2017). Regression techniques such as multiple regression (Oliveira et al., 2012), logistic regression (Pourghasemi, 2015), geographically weighted regression (GWR) (Fernandez et al., 2012), data mining techniques (Arpaci et al., 2014) are also used to build forest fire risk prediction models based on assessing the relationship between natural-society factors and the possibility of forest fire occurrence.

In Vietnam, several studies have used land surface temperature data calculated from Landsat and MODIS thermal infrared images for early warning of areas with high risk of forest fires (Vuong Van Quynh, 2005; Doan Ha Phong, 2007; Tran Quang Bao et al., 2016). Thermal infrared remote sensing data is also used in the study (Trinh, Zablotskii, 2017) to detect subsurface coal fire in coal mines. Studies (Nguyen Ngoc Thach et al., 2015; Dang Ngo Bao Toan, 2021; Hoang et al., 2020) have also used remote sensing and GIS data for mapping forest fire risk in in different areas in Vietnam based on machine learning techniques. The obtained results show that machine learning techniques allow to classify forest fire risks with higher accuracy than traditional methods using hierarchical analysis technique (AHP). In general, the above studies have demonstrated the effectiveness of remote sensing and GIS technology to develop forest fire risk prediction model for monitoring and minimizing damage caused by forest fires.

2. Materials and methodology

Study area

The study area is located in the west of Nghe An province, in the north-central region of Vietnam, with geographical coordinates from 18°33' to 20°01'N, 103°52' to 105°48'E. This area has a diverse and complex topography, strongly divided by hills, mountains and a system of rivers and streams.

Nghe An has the largest area of forest and forestry land in the country, the forest coverage density in 2020 will reach 58.50 % according to the forest status report in 2020 of the Ministry of Agriculture and Rural Development of Vietnam. Although forest cover is high and tends to increase in recent years, most of the increased forest area in province is planted forest. The natural forest has a marked decrease in both quality and area.

The province is located in the tropical monsoon climate zone, with hot, humid and rainy summers (from May to October) and cold, less rainy winters (from November to April next year). The Feohn wind is most active in the western region of Nghe An province, causing a hot and dry climate, negatively affecting production activities as well as increasing the risk of forest fires.

Materials

In this study, the Sentinel 2 MSI and Landsat 8 multispectral images are used to calculate the land cover, soil moisture and land surface temperature factors. Remote sensing data is collected and processed directly on the Google Earth Engine (GEE) cloud computing platform.

65 Sentinel 2 MSI scenes, including Sentinel 2A and Sentinel 2B images taken from November 15, 2021, to January 16, 2022, are used to create cloudless image, and then calculate the Normalized Difference Vegetation Index (NDVI) and Normalized Multi-band Drought Index (NMDI). 15 Landsat 8 OLI_TIRS multispectral images with path/row of 127/047, 127/046 and 128/046 taken during the period from November 15, 2021, to 16 January 2022 were used to calculate the land surface temperature.

The Landsat 8 and Sentinel 2 MSI images in the western region of Nghe An province used in the study are shown in Figure 2.

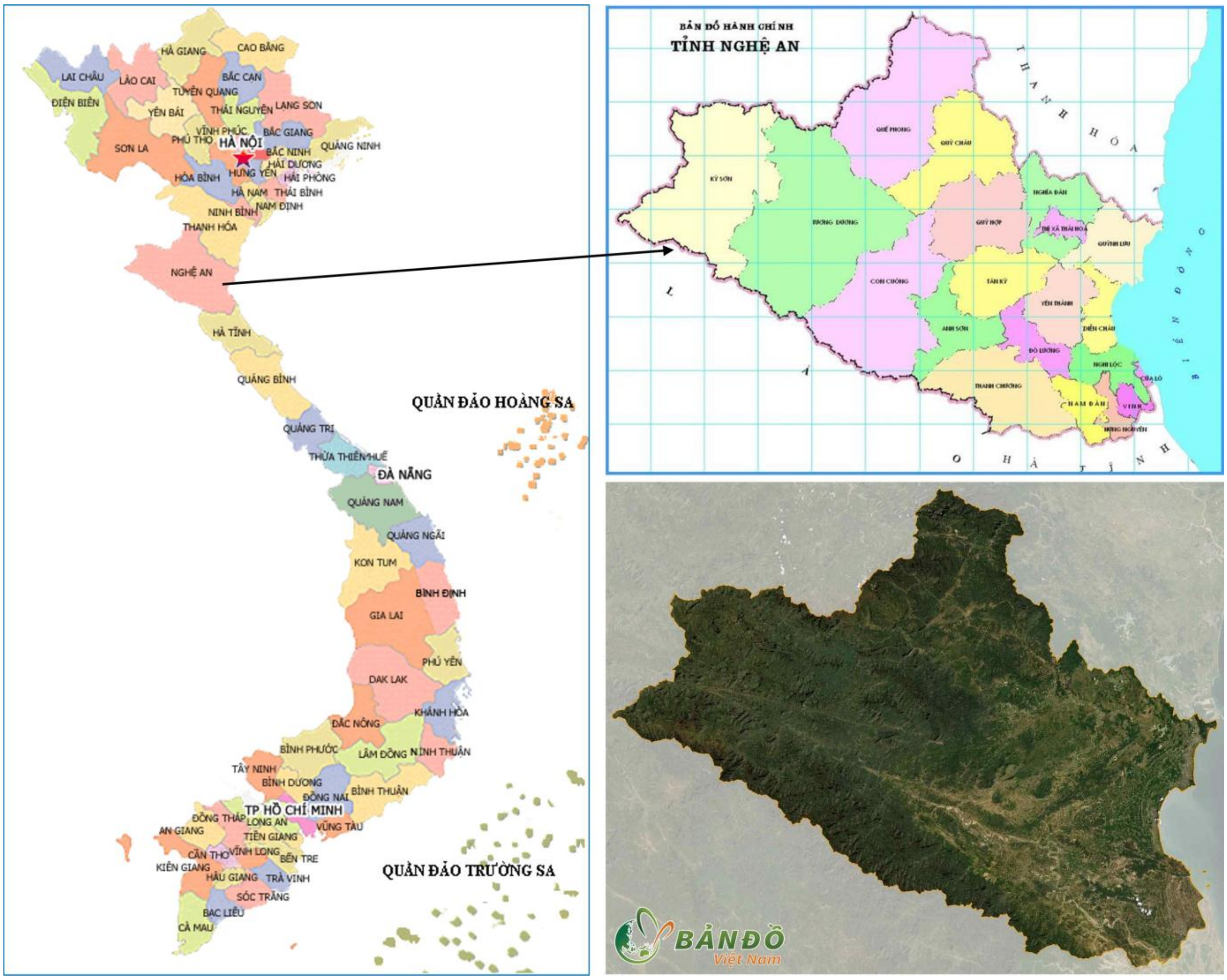


Fig. 1. The study area in the west of Nghe An province, Vietnam

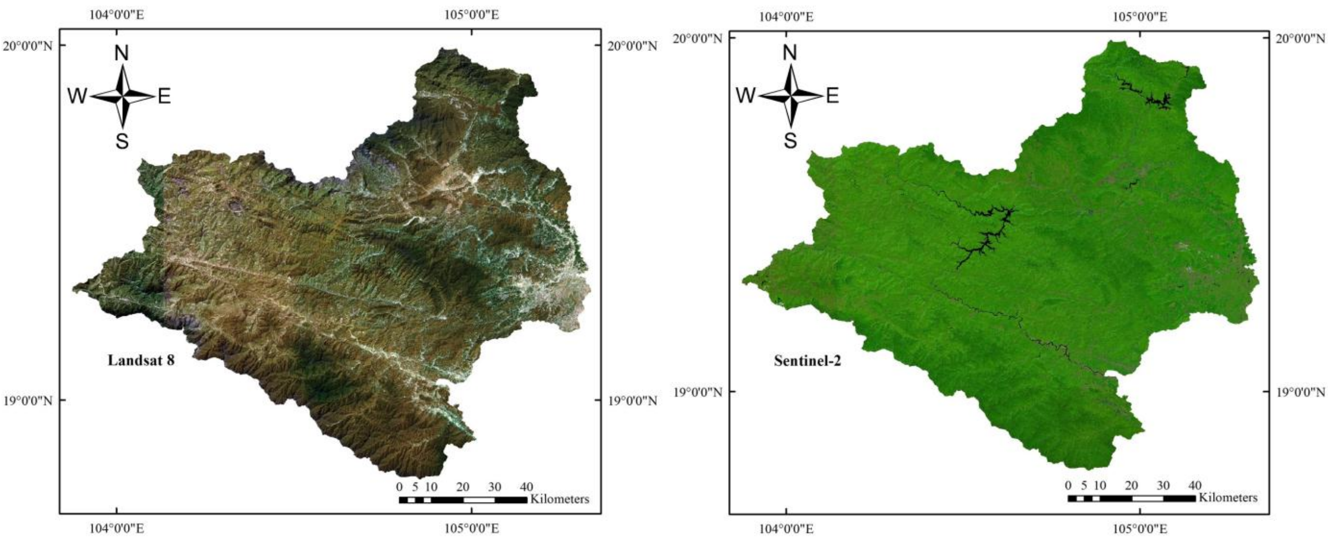


Fig. 2. Sentinel 2 MSI images in Lao Cai and Yen Bai provinces

The Shuttle Radar Topography Mission (SRTM) elevation data with 30 m spatial resolution is collected and processed to build the terrain factors for forest fire risk model, including elevation, aspect and slope.

Population density data is collected from the WorldPop database (Index). Meanwhile, wind speed and average monthly precipitation are collected from the WorldClim database (WorldClim).

Methodology

Based on the natural and social characteristics of the western region of Nghe An province, in this study, 09 factors are used to build the forest fire risk prediction model, including:

+ Population density
+ Vegetation cover (NDVI index)
+ Soil moisture (NMDI index)
+ Aspect
+ Slope
+ Wind speed
+ Elevation
+ Land surface temperature
+ Average monthly precipitation.

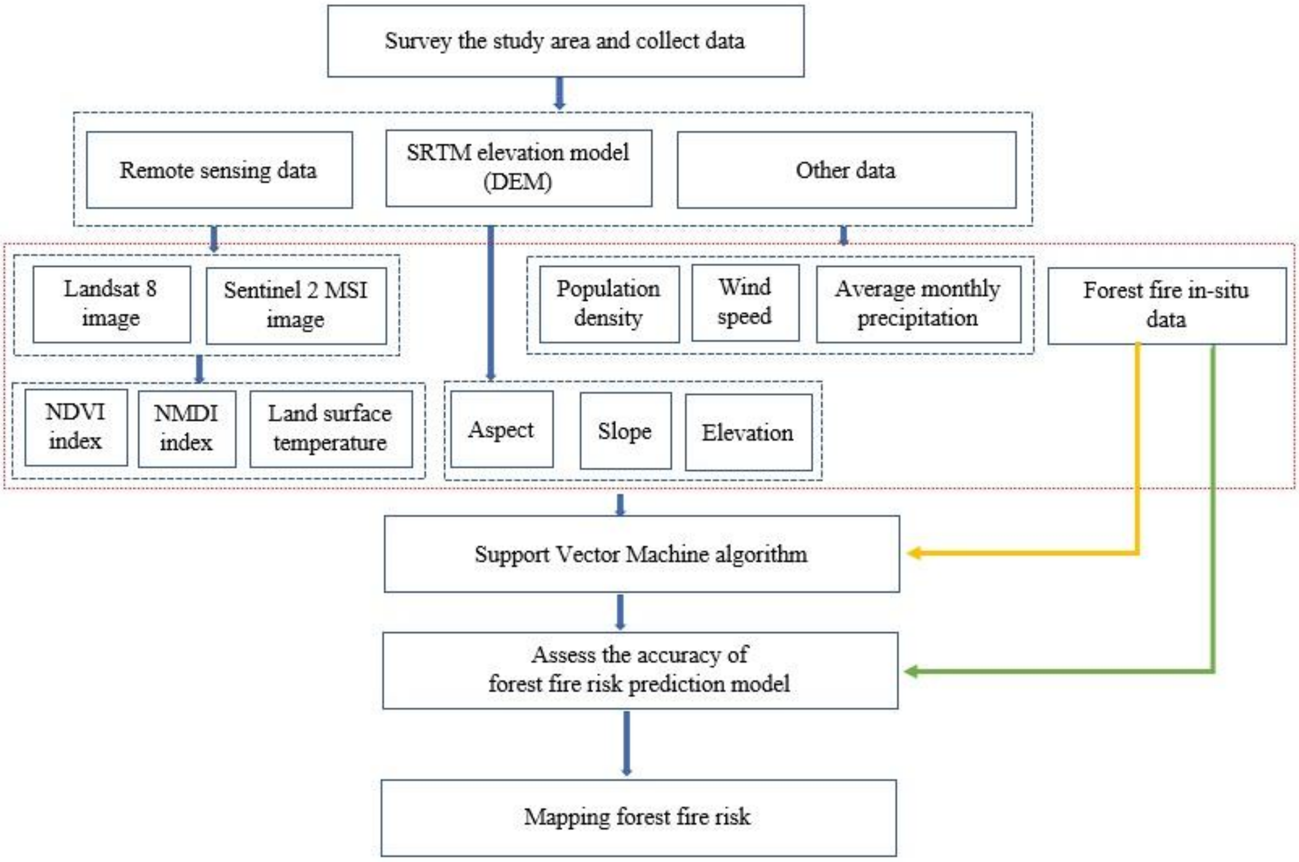


Fig. 3. Flowchart of the methodology for mapping forest fire risk using SVM algorithm from remote sensing and GIS data

The NDVI index is used to represent the vegetation cover factor in the forest fire risk prediction model. NDVI is calculated from the spectral reflectance value in the red and near-infrared channels of the Sentinel 2 MSI image following the equation (Rouse et al., 1973):

$$NDVI = \frac{\rho_{NIR} - \rho_{RED}}{\rho_{NIR} + \rho_{RED}} \quad (1)$$

Where: ρ_{RED} and ρ_{NIR} are reflectance values of red (Band4) and near infrared (Band8) bands of Sentinel 2 MSI image.

In this study, the NMDI moisture index (Wang, Qu, 2007) was used to represent the surface moisture information of the study area. The NMDI index is also determined from optical remote sensing images according to the following formula:

$$NMDI = \frac{R_{860nm} - (R_{1610nm} - R_{2150nm})}{R_{860nm} + (R_{1610nm} - R_{2150nm})} \quad (2)$$

In which, R_{860nm}, R_{1610nm} and R_{2150nm} are spectral reflectances from near-infrared (band 8), short-wave infrared (band 11 and band 12) bands of Sentinel 2 MSI satellite images.

Land surface temperatures were extracted from Landsat 8 thermal infrared remote sensing data using NASA model (Trinh, 2014).

The elevation, slope, and aspect data layers are calculated from the SRTM elevation digital model with a spatial resolution of 30 m. All 9 data layers of the forest fire risk prediction model are interpolated to a spatial resolution of 10m for consistency with Sentinel 2 MSI image data. Finally, the SVM algorithm is used to classify forest fire risk levels from 9 input data layers and forest fire data in the western region of Nghe An province.

The flowchart of the methodology for mapping forest fire risk using SVM algorithm from remote sensing and GIS data is shown in Figure 3.

3. Results and discussion

To build topographic data layers such as slope, elevation, and aspect, the SRTM model with spatial resolution of 30 m was used in this study. The collection and processing are done on the Google Earth Engine (GEE) platform. DEM SRTM data of the study area is presented in Figure 4. The slope, elevation and aspect data layers in the western region of Nghe An province built from the SRTM model are shown in Figure 5.

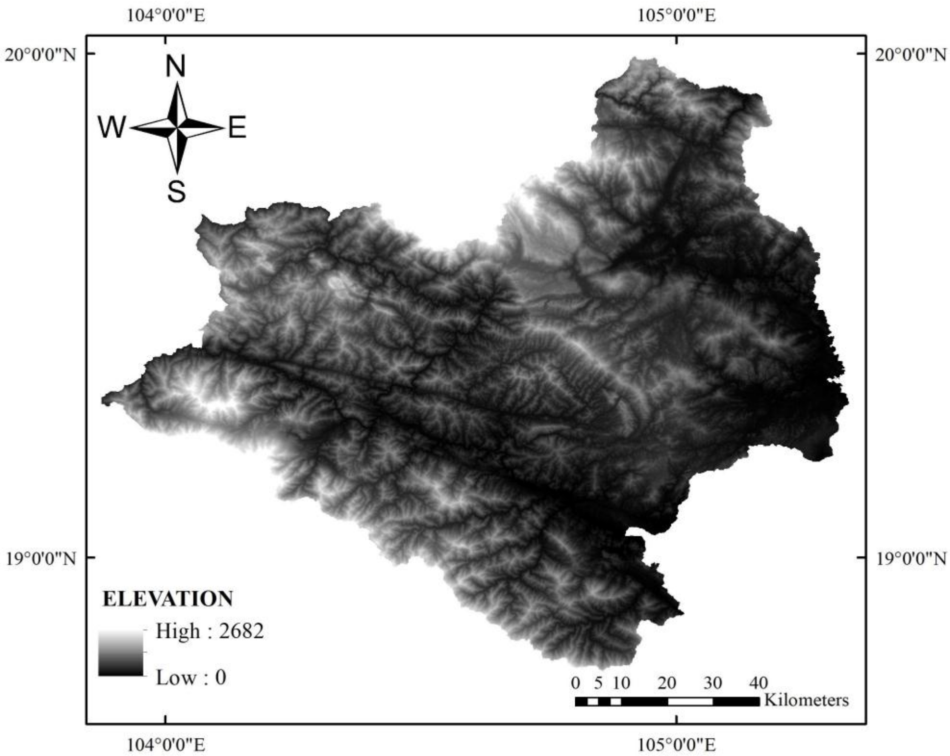


Fig. 4. The SRTM model of study area

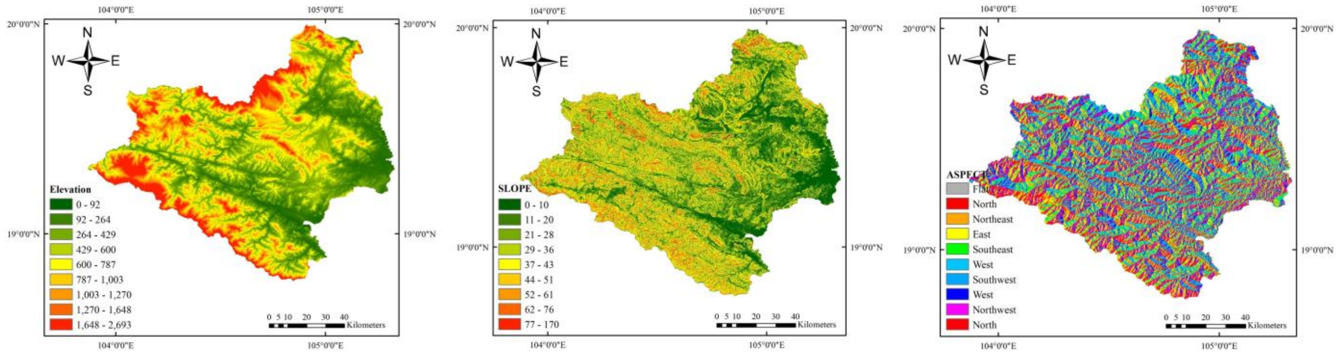


Fig. 5. Elevation, Slope and Aspect data layers for forest fire risk prediction model

Sentinel 2 MSI images after collection and pre-processing were used to calculate the vegetation index NVDI (formula 1) and the drought index NMDI (formula 2). Meanwhile, the Landsat 8 image is used to calculate the land surface temperature according to the NASA model. The results of mapping vegetation cover (NDVI index), soil moisture (NMDI index) and land surface temperature data layers in the western region of Nghe An province from remote sensing data are presented in Figure 6. The surface temperature of the study area ranges from 10,15°C to 35,47°C. Areas with low land surface temperature are concentrated in mountainous areas with dense vegetation cover, while areas with high land surface temperatures are concentrated in residential and bare land. For visualization, in the study, the land surface temperature is divided into 9 ranges: less than 18°C, 18-19°C, 19-20°C, 20-21°C, 21-22°C, 22-23°C, 23-24°C, 24-30°C and greater 30°C (Figure 6). Land surface temperature determined from Landsat 8 image has spatial resolution of 30m, to be consistent with NDVI and NMDI data, in the study the land surface temperature is interpolated to 10 m pixel size.

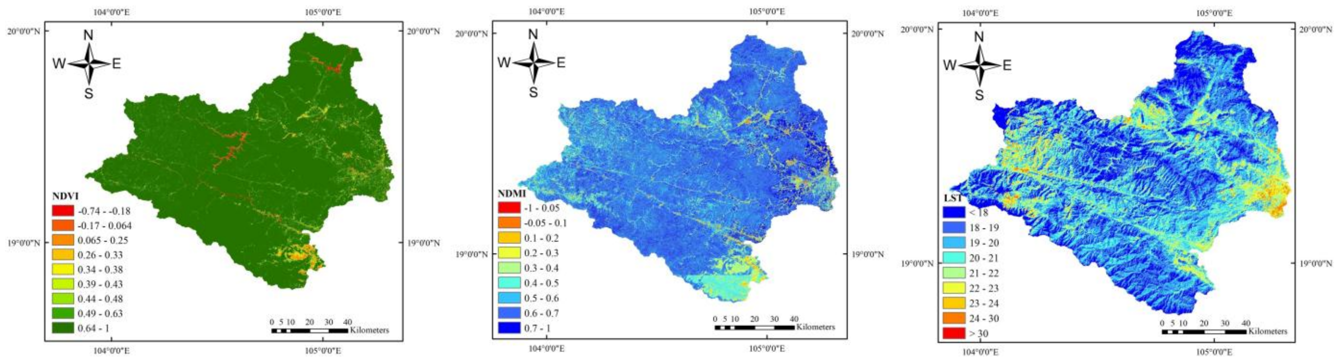


Fig. 6. NDVI, NMDI and Land surface temperature data layers for forest fire risk prediction model

Figure 7 presents the results of building data layers on wind speed, average monthly precipitation and population density in the western region of Nghe An province. Since the resolution of data collected from WorldPop and WorldClime databases is 1000m, to be consistent with other data layers of forest fire risk prediction model, these data layers are interpolated to 10m pixel size.

To evalute the forest fire risk using the SVM algorithm, this study has tested with different values of cost parameter (C). The parameter C tells the SVM algorithm how to balance the two competing objectives which are to maximize the margin between the two classes and to not allow any samples to be misclassified. If C = 0 then the algorithm does not allow any samples to be misclassified. If your data is not linearly separable then the algorithm will not be able to find a separating hyperplane. If C > 0 then the algorithm can trade-off some misclassified samples in-order to find a margin that better separates the remaining points.

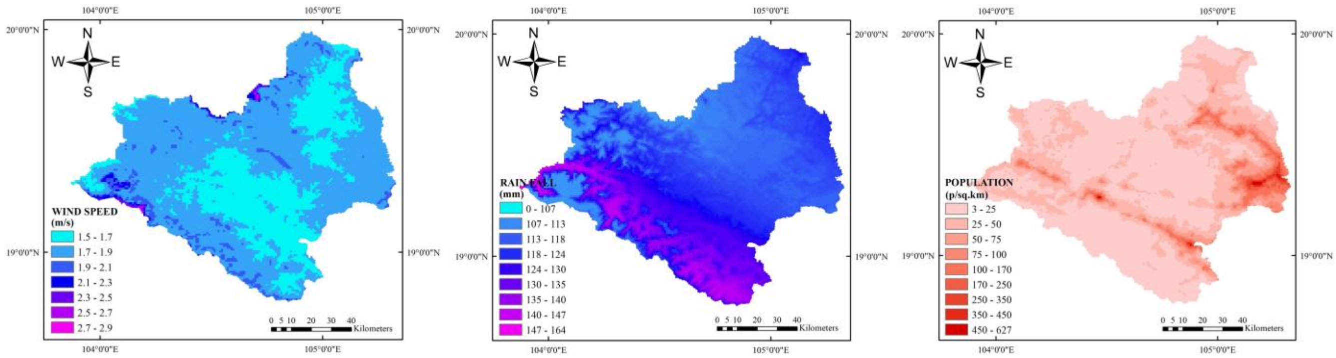


Fig. 7. Wind speed, Average monthly precipitation, and Population density data layers for forest fire risk prediction model

From analyzing the distribution of forest fire data, the parameter C = 25 allows forecasting the risk of forest fire in the study area with the highest accuracy. Out of 36 points where forest fires have occurred, 18/36 fire points are distributed in areas with a «high» level of forest fire risk, 3 points in areas with «very high» level; 4 fire points distributed in the area with «very low» and «low» level. The forest fire hazard zoning map in the western region of Nghe An province based on SVM algorithm is presented in Figure 8, in which the forest fires risk is divided into 5 levels: very high, high, medium, low and very low.

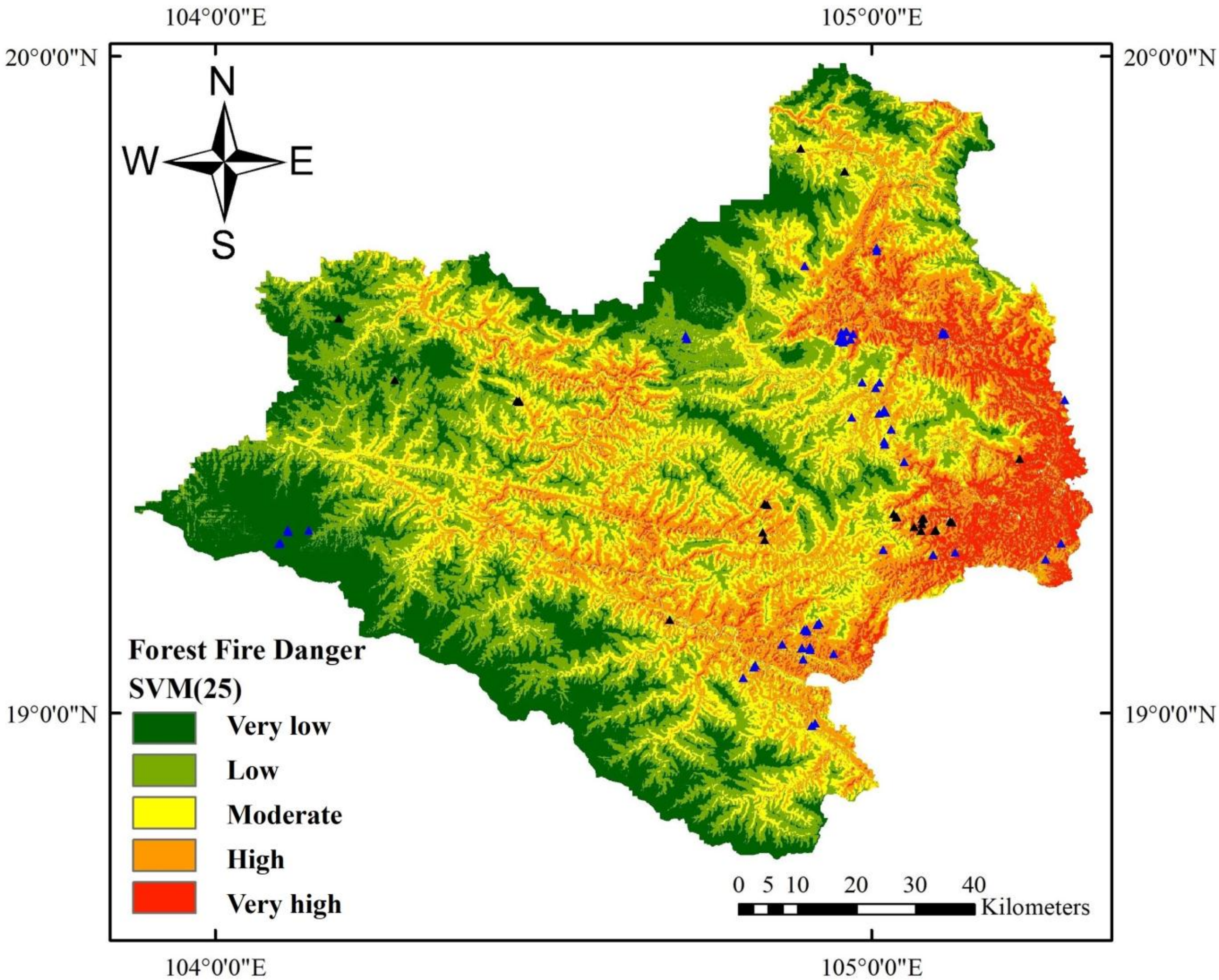


Fig. 8. Result of mapping forest fire risk in the western region of Nghe An province

4. Conclusion

In this study, 09 factors affect the risk of forest fire, including vegetation cover (NDVI index), soil moisture (NMDI index), elevation, slope, aspect, wind speed, land surface temperature, average monthly precipitation and population density are used for mapping forest fire risk based on Suppor Vector Machine (SVM) algorithm. Sentinel 2 MSI and Landsat 8 images data are used to create the vegetation cover, soil moisture and land surface temperature data layer. Topographic data layers, including slope, elevation and aspect are built from the SRTM model with a spatial resolution of 30 m. The population density, average monthly precipitation, and wind speed factors were built from the WorldPop and WorldClim databases. The obtained results show that the SVM algorithm allows forecasting the risk of forest fires with high accuracy. From the analysis of 36 points where fires have occurred in the past, 21/36 points are distributed in areas with a high and very high fire risk level. The results obtained can be used effectively for monitoring and early warning of forest fire risk and helping to reduce damage caused by forest fires.

References

Arpaci et al., 2014 – *Arpaci A., Malowerschnig B., Sass O., Vacik H.* (2014). Using multi variate data mining techniques for estimating fire susceptibility of Tyrolean forests. *Applied Geography*. 53: 258-270.

Dang Ngo Bao Toan, 2021 – Dang Ngo Bao Toan (2021). Study on the risk and warning of forest fire in Son La province based on the application of geoinformation technology. Geography Doctoral Thesis, Hanoi National University.

Dieu et al., 2017 – *Dieu Tien Bui, Quang Thanh Bui, Quoc Phi Nguyen, Pradhan B., Nampak H., Phan Trong Trinh* (2017). A hybrid artificial intelligence approach using GIS-based neural-fuzzy inference system and particle swarm optimization for forest fire susceptibility modeling at a tropical area. *Agricultural and Forest Meteorology*. 233: 32-44.

Dieu et al., 2016 – *Dieu Tien Bui, Kim Thoa Thi Le, Van Cam Nguyen, Hoang Duc Le, Revhaug I.* (2016). Tropical Forest Fire Susceptibility Mapping at the Cat Ba National Park Area, Hai Phong City, Vietnam, Using GIS-Based Kernel Logistic Regression. *Remote Sensing*. 8: 347. DOI: 10.3390/rs8040347

Doan, 2007 – *Doan Ha Phong* (2007). Developing algorithms for determining land surface temperature in forest fire warning monitoring based on MODIS satellite images (TERRA and AQUA) in the territory of Vietnam. Thesis of Doctor of Engineering, Hanoi University of Mining and Geology.

Fernandez et al., 2012 – *Fernandez, J., Chuvieco, E., Koutsias, N.* (2012). Modelling long-term fire occurrence factors in Spain by accounting for local variations with geographically weighted regression. *Natural Hazards Earth System Scieces*. 12: 1-17.

Hoang et al., 2020 – *Hoang, V.T., Chou, T., Fang, Y., Nguyen, N.T., Nguyen, Q.H., Pham, X.C., Dang, N.B.T., Nguyen, X.L., Meadows, M.* (2020). Mapping forest fire risk and development of early warning system for NW Vietnam using AHP and MCA/GIS methods. *Applied Sciences*. 10(12): 4348.

Jaiswal et al., 2002 – *Jaiswal, R., Mukherjee, S., Raju, K., Saxena, R.* (2002). Forest fire risk zone mapping from satellite imagery and GIS. *International Journal of Applied Earth Observation and Geoinformation*. 4(1): 1-10.

Ngoc et al., 2018 – *Ngoc Thach Nguyen, Bao Toan Ngo Dang, Xuan Canh Pham, Hong Thi Nguyen, Hang Thi Bui, Nhat Duc Hoang, Dieu Tien Bui* (2018). Spatial pattern asessment of tropical forest fire danger at Thuan Chau area (Vietnam) using GIS-based advanced machine learning algorithms: A comparative study. *Ecological Informatics*. 46: 74-85.

Olivera et al., 2012 – *Olivera, S., Oehler, F., Ayanz, J., Camia, A., Pereira, J.* (2012). Modeling spatial patterns of fire occurrence in Mediterranean Europe using Multiple Regression and Random Forest. *Forest Ecology and Management*. 275: 117-129.

Pourghasemi, 2015 – *Pourghasemi, H.* (2015). GIS-based forest fire susceptibility mapping in Iran: A comparison between evidential belief function and binary logistic regression models. *Scandinavian Journal of Forest Research*. 40 p. DOI: 10.1080/02827581.2015.1052750

Rouse et al., 1973 – *Rouse, J.W., Haas, R.H., Schell, J.A., Deering, D.W.* (1973). Monitoring Vegetation Systems in the Great Plains with ERTS. *Third ERTS Symposium*. NASA SP-351 I: 309-317.

Tran et al., 2016 – *Tran Quang Bao, Phung Nam Thang, Le Ngoc Hoan* (2016). Research on designing new software for early detection of forest firesfrom MODIS satellite image. *Journal of Agriculture and Rural Development*. 21: 114-120.

Trinh, Zablotskii, 2017 – *Trinh, L.H., Zablotskii, V.R.* (2017). The application of Landsat multi-temporal thermal infrared data to identify coal fire in the Khanh Hoa coal mine, Thai Nguyen province, Vietnam. *Izvestiya, Atmospheric and Oceanic Physics*. 53(9): 1181-1188.

Trinh, 2014 – *Trinh, L.H.* (2014). Studies of land surface temperature distribution using multispectral image Landsat. *Vietnam Journal of Earth Sciences*. 36(1): 82-89.

Wang, Qu, 2007 – *Wang, L., Qu, J.* (2007). NMDI: a normalized multi-band drought index for monitoring soil and vegetation moisture with satellite remote sensing. *Hydrology and Land Surface Studies*. 34(20). DOI: https://doi.org/10.1029/2007GL031021

Wimberly, Reilly, 2007 – *Wimberly, M., Reilly, M.* (2007). Assessment of fire severity and species diversity in the southern Appalachians using Landsat TM and ETM+ imagery. *Remote Sensing of Environment*. 108(2): 189-197.

Vasilakos et al., 2009 – *Vasilakos, C., Kalabokidis, K., Hatzopoulos, J., Matsinos, T.* (2009). Identifying wildland fire ignition factors through sensitivity analysis of a neural network. *Natural Hazards*. 50: 125-143.

Vuong, 2005 – *Vuong Van Quynh* (2005). Research and develop solutions to prevent and overcome the consequences of forest fires for the U Minh and Central Highlands regions. Summary report of topic KC08.24.

WorldClim – WorldClim. [Electronic resource]. URL: https://www.worldclim.org/

Index – Index [Electronic resource]. URL: https://data.worldpop.org/

watch.pcccr.vn – Quản Lý Cháy Rừng, Mất Rừng & Khai Thác Thông Tin [Forest Fire Management, Deforestation and Information Exploitation]. [Electronic resource]. URL: http://watch.pcccr.vn/thongKe/diemChay (date of access: 20.11.2022). [in Vietnamese]

Landsat 8 – Landsat 8 Data Users Handbook. [Electronic resource]. URL: https://www.usgs.gov/landsat-missions/landsat-8-data-users-handbook