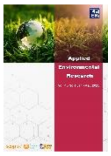
Applied Environmental Research (2026) 48(1), 008
https://doi.org/10.35762/AER.2026008

Applied Environmental Research
- RESEARCH FOR SUSTAINABLE PLANET -



Research Article

Assessment of The Influence of Human Activities on The Occurrence of Forest Fires in Thailand via Multiple Linear Regression (MLR)

Sittipong Ruktamatakul[1], Jirarat Insuk[1], Benjaporn Pinwongpet[1], Sarisa Ruktametakul[2], Pornpis Yimprayoon[1,*]

[1] Department of Computational Science and Digital Technology, Faculty of Liberal Arts and Science, Kasetsart University, Kamphaeng Saen Campus, Nakhon Pathom 73140, Thailand
[2] Triam Udom Suksa School, Bangkok 10330, Thailand

*Corresponding Email: faasppy@ku.ac.th

Abstract

Forest fires represent one of the most critical environmental challenges in Thailand, with impacts varying depending on forest type, fuel characteristics, terrain conditions, fire intensity, and the frequency of fire occurrence on the same landscape. While forest fires can contribute to ecosystem degradation, biodiversity loss, and the depletion of natural resources, such effects are not uniformly severe across all forest ecosystems. Understanding the human-induced factors contributing to forest fire occurrence is crucial for developing effective prevention strategies and promoting sustainable forest management. This study aimed to identify the anthropogenic factors influencing forest fire areas in Thailand via multiple linear regression (MLR) analysis. Eight independent variables related to human activities, including agricultural burning, forest product gathering, hunting, livestock raising, tourism, local conflicts, illegal logging, and accidents or negligence, were analyzed via annual data from 1998--2024 obtained from governmental and environmental agencies. The analysis revealed that forest product gathering, livestock raising, tourism, local conflicts, and negligence were significantly and positively correlated with burned areas. Although agricultural burning, hunting, and illegal logging were not statistically significant in the final regression model, these activities have been reported to contribute to extensive burned areas in Thailand. The lack of statistical significance in this study may reflect limitations related to data aggregation, temporal resolution, or indirect pathways through which these activities influence fire occurrence. The final regression model demonstrated high predictive accuracy, explaining approximately 97.83% of the interannual variation in Thailand's burned area. The findings indicate that human-related activities play a significant role in forest fire occurrence within the scope of the anthropogenic factors analyzed in this study, and the developed statistical model provides an effective tool for predicting fire-prone areas and supporting policy development for sustainable forest management and environmental protection.

ARTICLE HISTORY

Received: 3 Nov. 2025
Revised: 8 Jan. 2026
Accepted: 9 Feb. 2026
Published: 17 Feb. 2026

KEYWORDS

Forest fire;
Human activities;
Multiple linear regression;
Natural resource management;
Sustainable management

Introduction

Forest fires have become one of the most persistent environmental challenges in tropical regions, particularly in Thailand, where seasonal droughts and anthropogenic activities increase the risk of ignition and the potential spread of fire. In tropical countries such as Thailand, forest fires primarily result in forest degradation rather than deforestation, causing biodiversity loss and substantial emissions of carbon dioxide and particulate matter that deteriorate air quality and contribute to climate change (Smith et al., 2014). Effective forecasting of fire occurrence is therefore crucial for sustainable forest and natural resource management.

Previous studies have explored various mechanisms and modeling approaches for fire behavior and prediction. Mathematical and simulation-based studies have provided foundational knowledge for understanding surface fire spread (Sullivan, 2009), whereas stochastic differential equation models have offered insights into fire growth dynamics (Sanni et al., 2017). Statistical approaches such as multiple linear regression (MLR) remain widely used because they provide interpretable quantitative relationships between influencing factors and observed burned areas. Moreover, advanced machine learning and deep learning frameworks, including support vector machines (Yang et al., 2023; Yuan et al., 2020) and transformer-based time series models (Miao et al., 2023), have demonstrated potential for large-scale prediction and pattern recognition in wildfire research.

Environmental and remote sensing studies have further improved fire detection, monitoring, and mapping accuracy. The MODIS Collection 6 active fire detection algorithm (Giglio et al., 2016) and global geospatial platforms such as the Google Earth Engine (Gorelick et al., 2017) have enhanced spatial and temporal data accessibility. Broader analyses have linked wildfire frequency to vegetation structure, topography, and meteorological variables (Doerr and Santín, 2016; Keeley et al., 1999; Kogan, 1997; Rodrigues et al., 2014). In the context of Thailand, satellite-based and statistical investigations have recently focused on fire risk assessment and prediction (Quan et al., 2023; Talukdar et al., 2024) [13–14], whereas media coverage and institutional reports have highlighted human activities as dominant ignition sources in northern regions (Thai PBS, 2025).

However, limited research has quantified the contributions of diverse human activities to forest fire occurrence across Thailand over multiple decades. To address this gap, the present study develops an MLR model to evaluate the influence of socioeconomic and behavioral variables, including agricultural burning, forest product gathering, hunting, livestock grazing, tourism, local conflicts, illegal logging, and neglect, on burned area variation in Thailand from 1998–2024. The analysis is based on annual burned area statistics and human activity-related data compiled from official national sources, including the Department of National Parks, Wildlife and Plant Conservation (DNP) and other relevant governmental agencies. The MLR framework provides interpretable and empirical insights that support data-driven prevention, targeted mitigation, and sustainable management of Thailand's forest ecosystems. By integrating these findings into national policy and planning frameworks, this study contributes to long-term natural resource management, ensuring the protection and resilience of Thailand's forested landscapes for future generations.

Materials and methods

1) Multiple linear regression framework

This study employed an MLR framework to analyze the influence of human-related activities on forest fire occurrence in Thailand from 1998–2024. The MLR method was applied to examine the linear relationships between the dependent variable (Y), which represents the total annual burned area, and multiple independent variables ($x_1, x_2, \ldots, x_k$), which represent anthropogenic factors contributing to fire ignition and spread. The general form of the MLR model is expressed as follows:

$$\hat{Y} = a + b_1x_1 + b_2x_2 + b_3x_3 + \cdots + b_kx_k$$

where $\hat{Y}$ denotes the predicted burned area (hectares), a is the intercept, and b_i is the partial regression coefficient of each independent variable x_i. The coefficients were estimated via the least squares method, which minimizes the sum of squared residuals to obtain the best linear unbiased estimators under the Gauss–Markov assumption.

In matrix notation, the estimation of regression coefficients is represented as:

$$B = (X^TX)^{-1}X^TY$$

where

$$B = \begin{bmatrix} a \\ b_1 \\ b_2 \\ \vdots \\ b_k \end{bmatrix},\quad X = \begin{bmatrix} 1 & x_{11} & x_{21} & \cdots & x_{k1} \\ 1 & x_{12} & x_{22} & \cdots & x_{k2} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & x_{1n} & x_{2n} & \cdots & x_{kn} \end{bmatrix},\quad Y = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}$$

In this study, k = 8 independent variables and n = 27 annual observations were used.

2) Data and variables

The analysis was based on annual data spanning 27 years (1998–2024). The dependent variable (Y) represents the total annual burned area, measured in hectares. The burned area data were obtained from the forest fire situation in Thailand annual reports published by the DNP, which is the official governmental agency responsible for forest fire monitoring and reporting in Thailand (Department of National Parks, Wildlife and Plant Conservation, 2024).

The burned area statistics were compiled via standardized national fire reporting procedures and have been consistently published annually, ensuring temporal comparability across the study period. These datasets have been widely used in previous national assessments and policy reports, reflecting their reliability and institutional validation.

Eight independent variables (x_1– x_8) were defined to represent major human-related activities potentially associated with forest fire ignition and spread: agricultural burning (x_1), forest product gathering (x_2), hunting (x_3), livestock grazing (x_4), tourism (x_5), local conflicts (x_6), illegal logging (x_7), and negligence (x_8). Data for these variables were compiled from official government statistics, national reports, and publicly available datasets released by relevant governmental agencies.

Prior to analysis, all datasets were subjected to consistency checks, unit verification, and temporal alignment to ensure data quality and comparability across years. Because the study relied exclusively on aggregated secondary data obtained from public institutional sources and did not involve human subjects, animals, or biosafety-related experiments, ethical approval was not needed. The raw annual datasets are publicly available through official DNP reports and can be accessed upon request.

3) Statistical analysis

The overall model significance was evaluated via the F statistic, whereas individual coefficients were tested via t statistics at the 0.05 significance level. The hypotheses were $H_0: b_i = 0$ (no effect) and $H_1: b_i \neq 0$ (significant effect), where b_i represents the regression coefficient associated with predictor x_i.

Backward elimination was applied, sequentially removing predictors with nonsignificant t values (p > 0.05) until all remaining variables were statistically significant. This approach ensured a parsimonious and interpretable model while maintaining high explanatory

Results and discussion

1) Descriptive analysis of variables

Descriptive statistics across the 27-year period (1998–2024) revealed pronounced interannual variability (Table 1). The total burned area averaged 20,866.00 ± 14,525.22 hectares, ranging from 4,078.26 to 63,905.38 ha. Among the independent variables, forest product accumulation presented the highest mean and variability, followed by hunting and agricultural burning. Livestock grazing and tourism presented relatively low mean values, whereas local conflicts, illegal logging, and negligence presented high variability relative to their means. These patterns highlight the heterogeneity of anthropogenic influences contributing to nationwide variability in burned areas.

2) Regression results and model performance

Table 2 presents the final MLR coefficients, t values, and significance levels for the five predictors retained after backward elimination. The model explained $R^2 = 0.9783$ of the total variation in the annual burned area, and the overall F test ($p < 0.001$) confirmed strong statistical significance and high reliability.

Table 3 summarizes the stepwise regression process, indicating the sequential elimination of nonsignificant variables. Variables with the smallest nonsignificant t values (p > 0.05) were removed one at a time until only statistically significant predictors remained.

Figure 1 shows the strong linear relationship between the observed and predicted burned areas, indicating an excellent model fit ($R^2 = 0.9783$). Figure 2 displays the temporal variations in the observed and predicted burned areas from 1998–2024, confirming the model's ability to capture interannual fluctuations and the recent surge in 2024.

Table 1 Descriptive statistics of the dependent and independent variables (1998–2024) (unit: ha)

Description	Min	Max	Mean ± SD
Total burned area (*Y*)	4,078.26	63,905.38	20,866.00 ± 14,525.22
Agricultural burning (x_1)	99.42	6,677.28	2,131.98 ± 2,002.39
Forest-product gathering (x_2)	1,929.80	31,656.62	8,196.91 ± 5,874.59
Hunting (x_3)	520.80	13,422.24	3,503.08 ± 3,294.19
Livestock grazing (x_4)	86.32	1,753.12	607.48 ± 517.39
Tourism (x_5)	0.00	398.24	53.39 ± 90.22
Local conflicts (x_6)	4.64	6,595.36	748.74 ± 1,574.12
Illegal logging (x_7)	1.60	811.04	197.89 ± 246.10
Negligence (x_8)	1.70	3,523.04	194.64 ± 667.81

Table 2 Regression coefficients and significance levels for the final MLR model

Predictor	Coefficient	t value	p-value
Intercept	-13,527.1474	-2.1742	0.0413*
x_2	1.9841	24.7734	0.0000**
x_4	5.2503	4.9002	0.0001**
x_5	31.6717	2.3224	0.0303*
x_6	1.7030	2.4774	0.0218*
x_8	3.1414	2.3156	0.0308*

Note: *significant at p < 0.05, **significant at p < 0.001.

Table 3 Stepwise regression results for forest-fire area estimation

Model	Regression equation	F value	Minimum t value of the excluded variable
1	$\hat{Y} = -14748.9378 + 1.0890x_1 + 2.0555x_2 - 0.5840x_3 + 4.2737x_4 + 37.7008x_5 + 1.5614x_6 - 1.4572x_7 + 3.1396x_8$	–	–
2	$\hat{Y} = -14575.2816 + 0.8616x_1 + 2.0437x_2 - 0.5395x_3 + 4.5728x_4 + 35.6454x_5 + 1.5378x_6 + 3.4216x_8$	0.0626	-0.2501 (x_7)
3	$\hat{Y} = -13882.8530 + 0.6850x_1 + 1.9718x_2 + 4.1157x_4 + 20.3127x_5 + 1.5843x_6 + 3.5608x_8$	0.4800	-0.6928 (x_3)
4	$\hat{Y} = -13527.1474 + 1.9841x_2 + 5.2503x_4 + 31.6717x_5 + 1.7030x_6 + 3.1414x_8$	0.5183	0.7199 (x_1)

Note: The table shows successive regression models after each step of variable elimination. The variables with the lowest nonsignificant t values (p > 0.05) were removed sequentially until all remaining predictors were statistically significant.

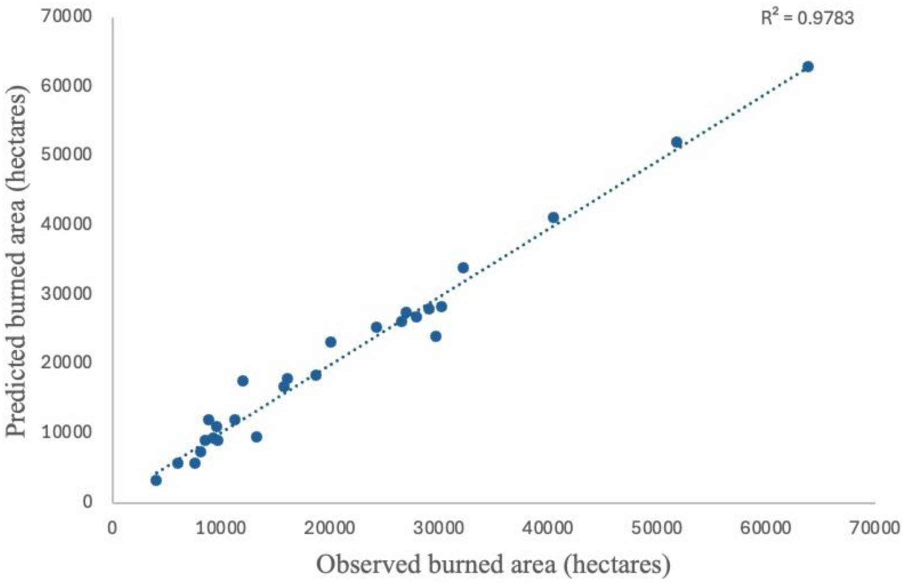


Figure 1 Relationships between observed and predicted burned areas (ha).

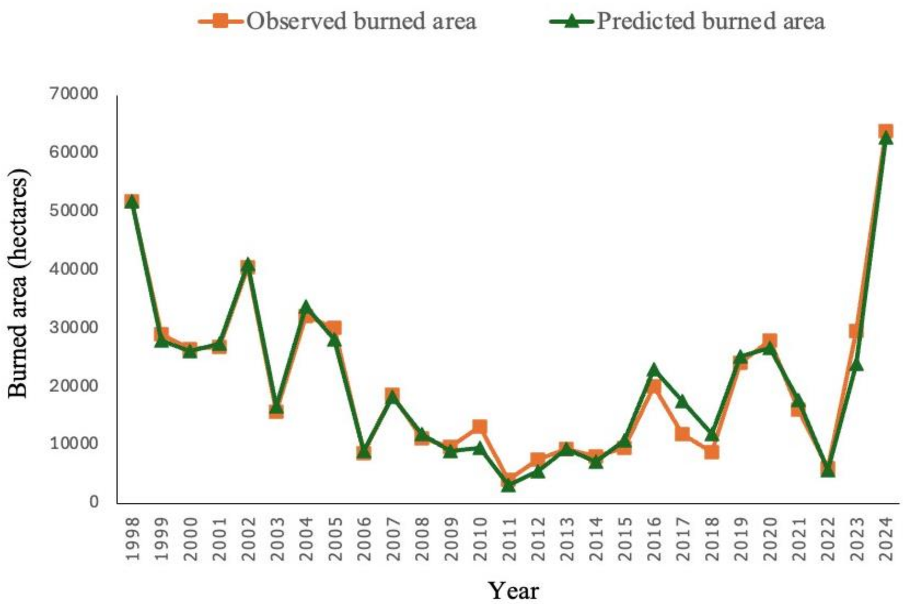


Figure 2 Temporal trends of observed and predicted burned areas (1998–2024).

The final regression equation was as follows:

$$\hat{Y} = -13527.1474 + 1.9841x_2 + 5.2503x_4 + 31.6717x_5 + 11.7030x_6 + 3.1414x_8$$

After backward elimination, five predictors remained statistically significant and positive: forest product gathering (x_2), livestock grazing (x_4), tourism (x_5), local conflicts (x_6), and negligence (x_8). Their significance at the 0.05 level indicates that these human activities are the principal drivers of the total burned area in Thailand. The corresponding *t* values were x_2 = 24.7734, x_4 = 4.9002, x_5 = 2.3224, x_6 = 2.4774, x_8 = 2.3156 and intercept = -2.1742.

The very high coefficient of determination and significant F test result indicate that the MLR model captures nearly all the interannual variability in the burned area, demonstrating a robust and reliable statistical fit.

3) Interpretation of key predictors

Tourism (x_5) exhibited the largest marginal coefficient (31.6717), indicating that years with increased tourism pressure, particularly during dry months, tended to experience substantially larger burned areas. Forest product gathering (x_2) had the strongest statistical signal, which is consistent with widespread foraging and collection activities linked to ignition risk. Livestock grazing (x_4) also contributed markedly, reflecting pasture clearing and understory burning that may escape containment. Local conflicts (x_6) and negligence (x_8) remain significant contributors, underscoring the influence of human disputes and careless behavior in initiating wildfires.

This finding suggests that the positive association between tourism activity and burned area may be attributed to increased human presence in forested and protected areas, particularly during the dry season. Tourism-related activities such as picnicking, roadside stops, and recreational use of forest trails can increase ignition risk through improperly extinguished campfires, discarded cigarette butts, and other careless behaviors. In addition, increased traffic and temporary settlements during peak tourism periods may increase fire susceptibility in surrounding forest landscapes.

Conversely, agricultural burning (x_1), hunting (x_3), and illegal logging (x_7) were excluded during model refinement owing to nonsignificant t values, thereby improving model parsimony without reducing explanatory power.

Although agricultural burning, hunting, and illegal logging were not statistically significant in the final model, this does not necessarily imply the absence of their influence on forest fires. These activities often exhibit strong seasonal or localized characteristics, whereas the present analysis was based on aggregated annual national-level data. As a result, their effects may be diluted when they are examined at broader spatial and temporal scales. Furthermore, the impacts of agricultural burning and illegal logging may be indirectly captured through other correlated variables, such as local conflicts or negligence, which remain statistically significant in the model.

4) Comparative EOL management approach

Regression coefficients quantify the expected change in total burned area per 1,000-hectare increase in each human-activity variable (Department of National Parks, Wildlife and Plant Conservation, 2024). Holding other predictors constant, a 1,000-hectare increase in forest product gathering corresponds to approximately 317.44 ha of additional burned area; an equivalent increase in livestock grazing produces approximately 840.00 ha more; a comparable rise in tourism contributes approximately 5,067.52 ha; and increases in local conflicts and negligence add approximately 272.48 and 502.56 ha, respectively.

The relative strength of the predictors, ranked by t statistics, was as follows: forest product gathering > livestock grazing > local conflicts > tourism > negligence. This ordering highlights the human-activity domains where targeted management could most effectively reduce forest fire risk nationwide.

5) Model adequacy and limitations

The residual analysis in this study focused on model significance testing via F and t statistics. Additional diagnostics, such as the Durbin–Watson test and variance inflation factor, were not performed; therefore, potential autocorrelation or multicollinearity cannot be ruled out. These represent limitations but also opportunities for future refinement, for example, through the incorporation of meteorological or spatially explicit variables.

Nevertheless, the high R^2 value, coherent coefficient magnitudes, and consistent statistical significance confirm the model's robustness and practical reliability for national-scale fire assessment.

6) Management implications

The findings indicate that management efforts should prioritize the following:

i. Regulation and monitoring of forest product gathering zones,
ii. Enhanced grazing management and fire break maintenance, and
iii. Visitor management programmes during the peak dry season.

Complementary strategies, including public awareness campaigns, conflict mediation initiatives, and anti-negligence enforcement, are also justified, given their significant positive associations with total burned area. Together, these measures can substantially reduce the annual extent of forest fires and support sustainable landscape management throughout Thailand.

Conclusions

This study successfully quantified the influence of eight human-related activities on forest fire occurrence across Thailand from 1998–2024 via an MLR approach. The model demonstrated strong explanatory power (R^2

= 0.9783; F test, $p < 0.001$), with five key predictors consisting of forest product gathering, livestock grazing, tourism, local conflicts, and negligence, all of which remained significant after refinement. These findings confirm that human activities are the main drivers of national-scale burned area variation. The results provide a clear foundation for developing management policies that emphasize forest access control, grazing management, and fire prevention awareness. Future studies should incorporate meteorological, vegetation, and spatial variables to improve model accuracy and strengthen long-term fire risk evaluations for sustainable forest management in Thailand.

Importantly, the robustness of the findings is supported by the quality and consistency of the underlying data. The burned area statistics were derived from the forest fire situation in Thailand annual reports published by the DNP, the official governmental authority responsible for national forest fire monitoring in Thailand. These data are compiled via standardized reporting procedures and have been published consistently on an annual basis over multiple decades, with public availability ensuring temporal comparability and institutional reliability. Furthermore, all the variables were subjected to data consistency checks and temporal alignment prior to analysis, reducing potential measurement bias. Therefore, the observed statistical relationships are unlikely to be artifacts of poor data quality and provide a credible empirical basis for national-scale fire management and policy development.

Data availability statement

Information and data used in the study will be disclosed upon request.

Author contributions

Sittipong Ruktamatakul: Conceptualization, Methodology, Writing – Original draft, Visualization
Jirarat Insuk: Software, Investigation, Resources, Data curation
Benjaporn Pinwongpet: Software, Investigation, Resources, Data curation
Sarisa Ruktametakul: Software, Investigation, Resources, Data curation
Pornpis Yimprayoon: Conceptualization, Methodology, Validation, Formal analysis, Writing – Original draft, Writing – Review & editing, Supervision, Project administration

Conflicts of interest

The authors declare that there are no conflicts of interest in competing financial or personal relationships that could have appeared to influence the work reported in this work.

References

Department of National Parks, Wildlife and Plant Conservation. (2024). *Forest fire situation in Thailand Annual report 2024*. Retrieved June 14, 2025, from https://portal.dnp.go.th/Content/Info_section4?contentId=30293

Doerr, S. H., & Santín, C. (2016). Global trends in wildfire and its impacts: Perceptions versus realities in a changing world. *Philosophical Transactions of the Royal Society B*, *371*(1696), 20150345.

Giglio, L., Schroeder, W., & Justice, C.O. (2016). The collection 6 MODIS active fire detection algorithm and fire products. *Remote Sensing of Environment*, *178*, 31–41.

Gorelick, N., Hancher, M., Dixon, M., Ilyushchenko, S., Thau, D., & Moore, R. (1999). Google Earth Engine: Planetary-scale geospatial analysis for everyone. *Remote Sensing of Environment*, *202*, 18–27.

Keeley, J. E., Fotheringham, C. J., & Morais, M. (1999). Reexamining fire-suppression impacts on brushland fire regimes. *Science*, *284*(5421), 1829–1832.

Kogan, F. N. (1997). Global drought watch from space. *Bulletin of the American Meteorological Society*, *78*(4), 621–636.

Miao, X., Li, J., Mu, Y., He, C., Ma, Y., Chen, J., Wei, W., & Gao, D. (2023). Time series forest fire prediction based on improved transformer. *Forests*, *14*(8), 1596.

Quan, X., Wang, W., Xie, Q., He, B., Resco de Dios, V., Yebra, M., Jiao, M., & Chen, R. (2023). Improving wildfire occurrence modelling by integrating time-series features of weather and fuel moisture content. *Environmental Modelling & Software*, *170*, 105840.

Rodrigues, M., de la Riva, J., & Fotheringham, S. (2014). Modeling the spatial variation of explanatory factors of human-caused wildfires in Spain using geographically weighted logistic regression. *Applied Geography*, *48*, 52–63.

Sanni, K., Viegas, D. X., Viegas, M. T., & Ferreira, A. D. (2017). A stochastic differential-equation approach to modelling the growth phase of fire spread. *Proceedings of the 22nd International Congress on Modelling and Simulation (MODSIM2017), Hobart, Australia*, 320–326.

Smith, P., Bustamante, M., Ahammad, H., Clark, H., Dong, H., Elsiddig, E. A., …, &Tubiello, F. N. (2014). Agriculture, forestry and other land use (AFOLU). In Edenhofer, O., Pichs-Madruga, R., Sokona, Y., Farahani, E., Kadner, S., Seyboth, K., Adler, A., Baum, I., Brunner, S., Eickemeier, P., Kriemann, B., Savolainen, J., Schlomer, S., von Stechow, C., Zwickel, T., & Minx, J. C. (Eds.), *Climate change 2014: Mitigation of climate change* (pp. 811–922). Cambridge University Press.

Sullivan, A. L. (2009). Wildland surface fire spread modelling, 1990–2007: Simulation and mathematical analogue models. *International Journal of Wildland Fire*, *18*, 387–403

Talukdar, N. R., Ahmad, F., Goparaju, L., Choudhury, P., Arya, R., Qayum, A., & Rizvi, J. (2024). Forest fire estimation and risk prediction using multispectral satellite images: Case study. *Natural Hazards Research*, *4*(2), 304–319.

Thai PBS. (2025). *Wildfire news and reports in northern Thailand*. Retrieved June 14, 2025, from https://www.thaipbs.or.th/now/content/845

Yang, X., Hua, Z., Zhang, L., Fan, X., Zhang, F., Ye, Q., & Fu, L. (2023). Preferred vector machine for forest-fire detection. *Pattern Recognition*, *143*, 109722.

Yuan, Q., Shen, H., Li, T., Li, Z., Li, S., Jiang, Y., Xu, H., Tan, W., Yang, Q., Wang, J., Gao, J., & Zhang, L. (2020). Deep learning in environmental remote sensing: Achievements and challenges. *Remote Sensing of Environment*, *241*, 111716.