Mugla Journal of Science and Technology

A NEW APPROACH TO DETERMINE THE INFLUENCE OF WEATHER CONDITIONS ON FOREST FIRE RISK IN THE MEDITERRANEAN REGION OF TÜRKİYE

Semra TÜRKAN, Department of Statistics, Hacettepe University, Türkiye, sturkan@hacettepe.edu.tr
(https://orcid.org/0000-0002-4188-184X)
Gamze ÖZEL, Department of Statistics, Hacettepe University, Türkiye, gamzeozl@hacettepe.edu.tr
(https://orcid.org/0000-0003-3886-3074)
Coşkun Okan GÜNEY, Department of Forest Fires, Aegean Forest Research Institute, Izmir, Türkiye, coskunokanguney@ogm.gov.tr
(https://orcid.org/0000-0003-4664-8024)
Ceren ÜNAL*, Department of Statistics, Hacettepe University, Türkiye, cerenunal@hacettepe.edu.tr
(https://orcid.org/0000-0002-9357-1771)
Özdemir ŞENTÜRK, Department of Forestry, Mehmet Akif Ersoy University, Burdur, Türkiye, osenturk@mehmetakif.edu.tr
(https://orcid.org/0000-0002-1841-7213)
Kürşad ÖZKAN, Department of Forestry, Isparta University of Applied University, Isparta, Türkiye.
(https://orcid.org/0000-0002-8526-7243)

Received: 29.03.2023, Accepted: 20.07.2023 Research Article
*Corresponding author
DOI: 10.22531/muglajsci.1273256

Abstract

The risk of forest fires is a major problem in Türkiye's Mediterranean region and has a significant impact on ecosystems and atmospheric conditions. Throughout the previous century, a significant portion of Türkiye's Mediterranean Region has been destroyed by forest fires. This study aims to determine the meteorological covariates, such as relative humidity, maximum temperature, and wind speed, that affect forest fires. We classified forest fires into two groups. The first group (F1) refers to small forest fires, with burned forest areas of less than 10 hectares. The second group (F2), representing rare events, corresponds to burned areas of more than 10 hectares. The data is composed of binary values (F1=0 and F2=1) taken between the years 2015-2019 from different locations in the Mediterranean Region of Türkiye. For binary data modeling, the ordinary logistic regression (LR) has been frequently used. However, such a method tends to give biased results when using rare event data. Therefore, we employed three different modeling techniques specifically designed for rare event data. According to the results obtained from the best model, Firth's Logistic Regression (FLR), wind speed, and maximum temperature are found to be statistically significant variables in the occurrence of forest fires greater than 10 hectares.

Keywords: Climate Change, Humidity, Natural Hazard, Penalized Regression.

TÜRKİYE'NIN AKDENIZ BÖLGESINDE HAVA KOŞULLARININ ORMAN YANGIN RISKINE ETKISINI BELIRLEMEYE YÖNELIK YENI BIR YAKLAŞIM

Özet

Orman yangınları riski, Türkiye'nin Akdeniz bölgesinde önemli bir sorundur ve ekosistemlere ile atmosferik özelliklere büyük etkisi vardır. Geçmiş yüzyıl boyunca Türkiye'nin Akdeniz Bölgesi'nin büyük bir kısmı orman yangınları sonucunda tahrip olmuştur. Bu çalışma, orman yangınlarını etkileyen bağıl nem, maksimum sıcaklık ve rüzgâr hızı gibi meteorolojik değişkenleri belirlemeyi amaçlamaktadır. Orman yangınlarını iki gruba ayırdık. Birinci grup (F1), 10 hektardan küçük yanmış orman alanlarını ifade eder. Nadir olayları temsil eden ikinci grup (F2), 10 hektardan fazla yanmış alanları temsil eder. Veriler, Türkiye'nin Akdeniz Bölgesi'ndeki farklı lokasyonlardan 2015-2019 yılları arasında alınan ikili değerlerden oluşmaktadır (F1=0 ve F2=1). İkili veri modellemesi için, genellikle sıradan lojistik regresyon (LR) kullanılmaktadır. Ancak, nadir olay verileri kullanıldığında bu yöntem yanlı sonuçlar verebilmektedir. Bu nedenle, nadir olay verileri için kullanılabilen üç farklı modelleme tekniği kullanıldı. En iyi model olan Firth Lojistik Regresyonu (FLR) sonuçlarına göre, rüzgâr hızı ve maksimum sıcaklık, 10 hektardan büyük orman yangınlarının oluşumunda istatistiksel olarak önemli değişkenler olarak bulundu.

Anahtar Kelimeler: İklim Değişikliği, Nem, Doğal Tehlike, Cezalandırılmış Regresyon.

Cite

Türkan, S., Özel, G., Güney, C. O., Ünal, C., Şentürk, Ö., Özkan, K., (2023). "A Novel Approach to Determine the Influence of Weather Conditions on Forest Fire Risk in the Mediterranean Region of Türkiye", Mugla Journal of Science and Technology, 9(2), 1-10.

1. Introduction

Forests are highly affected by short or long-term atmospheric events. In general, weather refers to changes in atmospheric conditions, including temperature, wind, and relative humidity, among others. It is widely recognized that weather plays a significant role in determining fire exposure and fire spread [1]. Atmospheric conditions have a vital impact on the occurrence, severity, and duration of forest fires. Both human-induced and natural forest fires can only occur when atmospheric conditions are favourable. Firefighting models and systems rely on precise meteorological data. Meteorological data provides the necessary information for monitoring and evaluating every stage of fire events.

Fire experts emphasize that the primary factor influencing forest fire ignition and behaviour is the moisture content of combustible materials. Secondary factors include relative humidity, air temperature, wind direction, wind speed, and air stability. Temperature directly affects the flammability of combustible materials [2]. Forest fire risk increases when the air temperature reaches or exceeds 30°C.

The heat required for the ignition of combustible materials affects their ease of ignition and depends on the initial temperature of the material [3]. Temperature is crucial for forest fires due to its impact on the temperature and moisture content of combustible materials. Under hot and dry weather conditions, the ignition temperature of flammable materials decreases, leading to a higher risk of forest fires. In other words, high temperatures reduce the moisture content of combustible materials, making them more susceptible to ignition and contributing to the severity, speed, and intensity of ongoing forest fires.

Wind refers to the horizontal movement of air. Uneven heating of the Earth's surface by solar radiation creates different pressure areas, resulting in wind movement from high-pressure to low-pressure areas. Winds can exhibit significant changes in speed and direction over short periods. The wind direction is crucial in terms of relative humidity. An increase in wind speed directly affects the spread of fire [4]. Wind plays a significant role in fire spread, and it is often considered the main explanatory variable in models developed to estimate fire spread rate, area, and environmental conditions.

Humidity refers to the percentage of water vapour in the air, known as relative humidity. For every 10°C increase in temperature, relative humidity decreases by 50% [5]. There is a close relationship between relative humidity and the moisture content of combustible materials, which is a critical factor in fire behaviour. Relative humidity reaches its lowest levels at noon due to temperature increase during the day, leading to a decrease in the moisture content of combustible materials and making them dry. Therefore, critical times for forest fires correspond to periods of low relative humidity. Southern aspects, which receive more direct sunlight, tend to have lower relative humidity levels.

The relationship between weather conditions and the occurrence of forest fires has been the subject of numerous studies [6, 7]. However, these studies have not specifically focused on Türkiye [8, 9]. Recent studies on this topic include the examination of weather conditions and forest fires in Greece from 1894 to 2010 by Koutsias et al. [10]. Tošić et al. [11] investigated the potential influence of meteorological variables on forest fire risk in Serbia from 2000 to 2017 using stepwise regression. Mueller et al. [12] analysed the climate-fire relationships in woodland ecosystems and forests in New Mexico and Arizona from 1984 to 2015 using correlation analysis. Wang et al. [13] utilized integral regression to identify the influence of climate factors on forest fires.

In Türkiye, hundreds of hectares of forest are destroyed every year due to forest fires. Previous studies have indicated that Türkiye is among the Mediterranean countries most affected by climate change, leading to an increased fire risk [10, 14]. Detecting and extinguishing forest fires promptly is crucial in minimizing fire damage. Regions in Türkiye with a Mediterranean climate are particularly susceptible to forest fires, as they experience high temperatures and severe droughts during the summer. Meteorological conditions have a significant impact on the potential for extreme fire events [14, 15]. Large forest fires usually occur under windy, hot, and dry conditions [16, 17]. Therefore, it is of great importance to assess projected changes in climate conditions and predict future forest fires in order to implement effective forest fire management measures and mitigate the harmful impacts of such fires [18].

Support vector machines, time series, multilayer perceptron, and fuzzy logic approaches are just a few of the techniques that have been used to predict the extent of forest fires with various factors over large spatial or regional scales [19]. Beckage and Platt [20] applied a statistical model to forecast the potential for forest fires in Oregon, USA, focusing on the region's prior seasonal fire history without considering the weather. Iliadis [21] utilized a decision-support system to identify forest areas at risk in Greece. Cortez and Morais [22] employed multivariate regression, random forest, support vector machines, artificial neural networks, and decision trees to forecast forest fire losses. Cheng and Wang [23] conducted a study on future forest fires using data mining techniques. Sakr et al. [24] predicted the occurrence and size of forest fires using humidity, precipitation factors, support vector machines, and neural networks.

Logistic regression models have also been used in many studies, both to determine the causes of fires and to predict fire risk at a global level [25, 26, 27, 28, 29, 30]. However, logistic regression analysis can significantly underestimate the probability of rare events. Since forest fires are considered rare events, a new approach is needed to predict the effects of meteorological factors on these fires. To overcome the limitations of existing methodologies and improve relative risk estimations, Firth's logistic regression (FLR) was introduced by Firth [31]. The primary idea behind FLR is to introduce a term that counteracts the first-order term from the asymptotic expansion of the bias of the maximum likelihood estimation, providing a more effective score function as the sample size increases and the term approaches zero [31, 32].

This study contributes to the literature in two ways. First, it utilizes a rare event logistic regression model, which yields unbiased results for forest fire occurrence data. Second, it focuses on Türkiye's Mediterranean region, particularly the Antalya region, which has not been extensively studied in terms of forest fire activity using new statistical models and meteorological covariates. The Antalya region experiences a significant number of forest fires, including the largest forest fire in Türkiye's history in 2021, which burned nearly 60,000 hectares in Manavgat. The data used in this study specifically focuses on forest fires in the Antalya region between 2015 and 2019. To the best of our knowledge, these models have not been previously employed for predicting meteorological factors of forest fires in Antalya, Türkiye, or elsewhere.

In this study, rare event logistic regression models, including FLR, FLR with intercept correction (FLIC), and FLR with added covariate (FLAC), are explained in Section 2. Section 3 provides information on the study area and data collection, specifically classifying forest fires as less than 10 hectares (F1) and greater than 10 hectares (F2), which are considered rare events in the Antalya region between 2015 and 2019. In Section 4, the results of the study regarding the meteorological determinants of F2 are presented using rare event logistic regression models. Finally, Section 5 concludes the study.

2. Material and Methods

Many of the fundamental assumptions of linear and modified linear models are not necessary in logistic regression (LR). The LR model does not require the normality of the error distribution or a linear relationship between the independent and dependent variables. Additionally, it does not assume that the errors are homoscedastic. However, an insufficient sample size may have a negative impact on this model. Even when the sample size is sufficient, in some cases, the event rate can be quite small. When the binary-dependent variables in the logistic regression model contain considerably more zeros (non-events) than ones (events), a rare event occurs. In studies with an insufficient sample size, the parameters that need to be estimated may be biased.

The definition of a rare event in the logistic regression model is when the dependent variable has many more zeros (non-events) than ones (events). However, even when the sample size is sufficient, the event rate may still be relatively low. Imbalanced rare events have been studied in diverse fields such as political science, biology, engineering, uncommon diseases, and more [31, 33]. Rare event cases can be observed in finance, such as the bankruptcy of a debtor or fraudulent business activities of certain companies, in politics, such as vetoes or wars between countries, in medicine, such as epidemiological infections, and in engineering, such as system failures. If binary logistic regression is applied to rare event data, the probability of rare events could be underestimated. To obtain unbiased predicted probabilities in logistic regression models like FLR, FLIC, and FLAC, it is necessary to account for rare events.

2.1. Logistic Regression (LR)

The binary logistic regression is used to model the dichotomous type dependent variable and this variable follows a Bernoulli distribution. The probability that equals to one is obtained as follows:

$$P(Y_i=1)=\pi_i=\frac{\exp\left(x_i^{'}\beta\right)}{1+\exp\left(x_i^{'}\beta\right)} \tag{1}$$

Here the independent variable vector is $x_i^{'}=\left[x_{i1}\;\; x_{i2}\;\; x_{i3}\;\; \ldots.\;\; x_{ik}\right]$ and the regression coefficient vector is $\beta^{'}=\left[\beta_1\;\;\beta_2\;\;\beta_3\;\;\ldots\;\;\beta_k\right]$. The probability that $\mathrm{Y_i}$ equals to zero is obtained as follows:

$$P(Y_i=0)=1-\pi_i=\frac{1}{1+\exp\left(x_i^{'}\beta\right)} \tag{2}$$

The logistic regression's logit transformation is given by Eq. (3) as

$$\eta_i=\log\left(\frac{\pi_i}{1-\pi_i}\right)=\beta_0+x_{i1}\beta_1+x_{i2}\beta_2+\ldots.++x_{ik}\beta_k \tag{3}$$

The likelihood function is the joint Bernoulli probability distribution of $\mathrm{Y_i}$ as below:

$$L\left(\beta_0,\ldots,\beta_k;x_i\right)=\prod_{i=1}^{n}\left[\pi_i^{\mathrm{Y}_i}\left(1-\pi_i\right)^{1-\mathrm{Y}_i}\right] \tag{4}$$

Taking logarithms of Eq. (4), the log-likelihood function is obtained as follows:

$$L\left(\beta_0,\ldots,\beta_k;x_i\right)=\prod_{i=1}^{n}\left[\mathrm{Y}_i\ln\left(\pi_i\right)-\left(1-\mathrm{Y}_i\right)\ln\left(1-\pi_i\right)\right] \tag{5}$$

The score function is written as

$$\frac{\partial L\left(\beta_0,\ldots,\beta_k;x_i\right)}{\partial\beta_j}=\sum_{i=1}^{n}\left(\mathrm{Y}_i-\left(1-\pi_i\right)\right)x_{ij} \tag{6}$$

Logistic regression estimates are typically obtained using the maximum likelihood method. Due to the lack of an analytical solution for Eq. (5), iterative algorithms are employed to find the maximum likelihood estimates [34]. The odds ratio (OR) is used to measure the influence of independent variables on the risk of event occurrence. It represents the ratio of the probabilities of the event occurring to the probability of the event not occurring [35].

2.2. Firth's Logistic Regression (FLR)

The FLR is employed to obtain more accurate maximum likelihood estimates, particularly when dealing with imbalanced or rare event data. This is achieved by addressing the first-order bias of the parameter estimates. The parameter estimates in FLR are derived by solving the modified score equations, which can be formulated as follows:

$$\sum_{i=1}^{n}\left(Y_i-\pi_i+h_i\left(\frac{1}{2}-\pi_i\right)\right)x_{ij}=0 \qquad (7)$$

where h_i is the i-th diagonal element of $W^{1/2}X\left(X^{'}WX\right)^{-1}X^{'}W^{1/2}$. Here, W is also diagonal matrix $diag\left(\pi_i\left(1-\pi_i\right)\right)$ [36].

2.3. Firth's Logistic Regression with Intercept Correction (FLIC)

The FLIC is a simple modification of the FLR which provides average predicted probabilities equal to the observed proportion of events, while preserving the ability to deal with separation [37]. The intercept coefficient β_0 for FLR was estimated by different way by Puhr et al. [36] known as the FLIC and the steps for FLIC are as follows:

i. The coefficient estimates $\widehat{\beta}_{FLR}$ are obtained.
ii. The linear predictors $\eta_i = x_{i1}\widehat{\beta}_{FLR1} + x_{i2}\widehat{\beta}_{FLR2} + \ldots + x_{ik}\widehat{\beta}_{FLRk}$ are found by omitting intercept coefficient.
iii. The maximum likelihood estimate of intercept coefficient in logistic regression is given by
$$P\left(Y_i=1\right)=\left(1+\exp\left(-\beta_0-\widehat{\eta}_i\right)\right)^{-1}$$
where $\widehat{\eta}_i$ is a predictor with regression coefficient equal to one.
iv. The FLIC estimates $\widehat{\beta}_{FLIC}$ are derived as $\widehat{\beta}_{FLIC}=\left(\widehat{\beta}_0,\widehat{\beta}_{FLR,1},\ldots,\widehat{\beta}_{FLK,k}\right)$.

2.4. Firth's Logistic Regression with Added Covariate (FLAC)

The basic idea of the FLAC analysis is to discriminate between pseudo-observations and original in the alternative formulation of Firth-type estimation as iterative data augmentation procedure. The steps of the FLAC analysis as follows:

i. The diagonal elements h_i using the FLR model are obtained.
ii. The augmented data set are constructed with: The original observations are weighted by 1, $h_i/2$, and $h_i/2$ with Y_i replaced by $1-Y_i$.
iii. An indicator variable g is defined for the sub steps a, b, and c of the second step, g=0, g=0 and g=1, respectively.
iv. The FLAC estimates are derived via the maximum likelihood estimation on augmented data set adding g as the covariate.

3. Study Area and Data Collection

3.1. Study Area

Türkiye, an ancient crossroads, holds a unique geographical position as a bridge connecting southern Europe and Asia. Surrounded by water on three sides and boasting a coastline spanning 4,474 miles (7,200 km), Türkiye's Anatolian Plateau defines its landscape. The country lies between 36-42 degrees north latitude and displays diverse vegetation due to its varying temperatures. Unlike neighboring countries with north-south oriented mountains, Türkiye's mountain ranges extend east to west, a result of tectonic processes that shaped the land bridge. Approximately 80% of the country's terrain is rugged and characterized by rough topography, with a median elevation of 3700 feet (128 meters) and an average elevation of 4,400 feet (1,332 meters). The flatlands are mainly found in the river deltas of the Asiatic region [45].

In Türkiye, forest fires are particularly prevalent in areas with a Mediterranean climate. The frequency and size of these fires are expected to increase in the future due to climate change and the growing global population [38, 39, 40]. Türkiye currently has a forest cover of 28%, and forest fires are a common occurrence. Over the past eight years, an average of nearly 2,500 forest fires has been recorded annually, with each fire consuming approximately 7 acres (2.8 hectares), resulting in a total area of around 17,000 acres (7,000 hectares). Lightning accounts for 11% of these fires, while human activities are responsible for the remaining 89%. In 2018 alone, $131 million was spent on fire suppression efforts (USD). Figure 1 provides a visual representation of the areas in Türkiye with a predominantly Mediterranean climate, highlighting the regions with the highest fire hazards. Additionally, Figure 2 displays a fire risk map illustrating the regional variations in Türkiye's fire-prone areas, the extent of fire damage, and the risks of ignition. The risk levels in Figure 1 range from "1 Degree" (highest risk) to "5 Degree" (lowest risk).

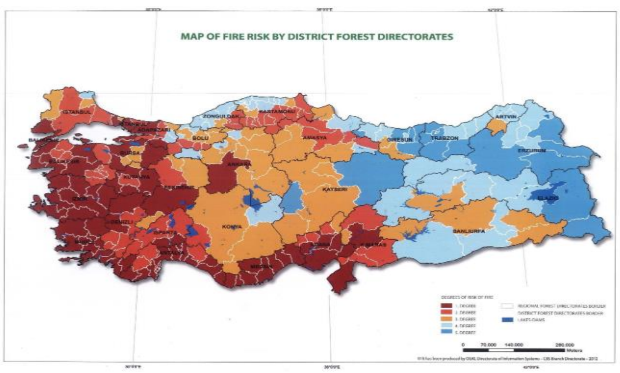

Figure 1. Map of fire risk by district forest directorates in Türkiye. (Source: The Republic of Turkish Ministry of Agriculture and Forestry of Türkiye)

Figure 2 shows that the number of forest fires and burned area size in Türkiye between years 1990-2019.

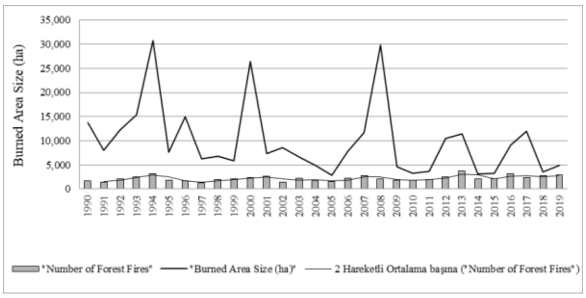

Figure 2. Number of forest fires and burned area size in Türkiye between years 1990-2019.

Antalya, located in the southern region of Türkiye, boasts the second-largest forest area in the country after Amasya. Covering a vast expanse of 1,146,062 hectares, the forests in Antalya make up a significant portion, accounting for approximately 60.43% of the region's land [41]. The dominant tree species in this area is Pinus brutia Ten, constituting a substantial 65% of the forest cover. Following closely behind are Cedrus libani at 16%, Pinus nigra at 8%, Abies cilicica at 5%, Juniperus sp. at 4%, and other deciduous species at 2%. Antalya is home to the Antalya Regional Directorate of Forestry (RDF) [42], which falls under the administration of the General Directorate of Forestry, operating within the Republic of Turkish Ministry of Agriculture and Forestry. Situated between 29° 16′ 15″ east longitudes and 36° 00′ 45″ north latitudes, the Antalya RDF plays a crucial role in managing and protecting the region's forests. To conduct this study, comprehensive fire statistics were obtained from the Antalya RDF, providing valuable insights into the historical fire occurrences in the area. Additionally, meteorological data were collected from the General Directorate of Meteorology (GDM) of Türkiye, enabling a comprehensive analysis of the relationship between weather conditions and forest fires.

The study area focused specifically on the Antalya RDF, which is one of the 28 Regional Directorates operating under the General Directorate of Forestry [42]. By highlighting the Antalya RDF on the map in red (see Figure 3), researchers were able to delineate the specific region under investigation. This study contributes to the broader understanding of forest fire dynamics in Türkiye and aids in developing effective strategies for fire prevention and management.

Figure 3. Distribution of forestry area and study area. (Source: Antalya Regional Directorate of Forestry)

The Aegean and Mediterranean regions of Türkiye are known to be highly susceptible to wildfires, particularly during the summer months. However, in May 2021, the country experienced an unprecedented event as it recorded the warmest May in over 50 years, exacerbating the risk of wildfires. This occurrence is likely attributed to the effects of climate change. Throughout the year, Türkiye witnessed a devastating wildfire season, with a staggering 2,793 fires engulfing approximately 140,000 hectares of forested land [43]. This marked the most severe wildfire season ever recorded in the country's history.

The wildfires that ravaged Türkiye in 2021 were particularly intense in the Manavgat region of Antalya Province. The fires originated on 28 July 2021, when temperatures soared to around 37 °C (99 °F) [43]. In total, more than 100 fires were reported along the Turkish Riviera, with two major fires in Muğla and Antalya still burning as of 9 August 2021. The situation was further aggravated by a heatwave and adverse weather conditions, contributing to the rapid spread of the fires.

The impact of these wildfires extended beyond the destruction of forests. The region is home to a diverse range of animal species, with approximately 11,870 species identified, of which 1,421 are unique to the Antalya region. These fires posed a severe threat to the survival of numerous animal species, further emphasizing the ecological consequences of such widespread and devastating wildfires.

As shown in Figure 4, the map provides a visual representation of the summary statistics, displaying the extent and severity of the wildfires across the affected regions. It is crucial to address the underlying factors contributing to these wildfires, including climate change, and to implement effective strategies for prevention, mitigation, and conservation to safeguard both the environment and the wildlife inhabiting these vulnerable areas.

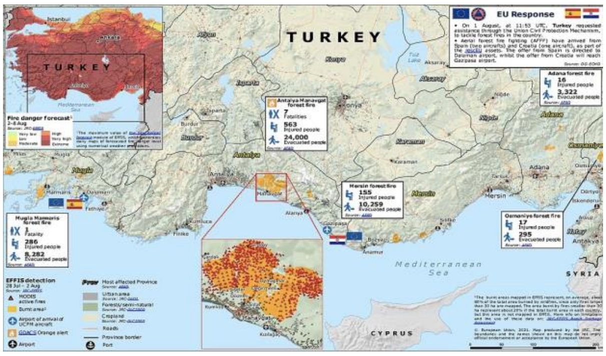

Figure 4. Summary of the forest fire statistics in the Türkiye's Mediterranean Region on August 2021. (Source: https://en.wikipedia.org/wiki/2021_Türkiye_wildfires)

3.2. Data Collection

The research area is located in the south of Türkiye. The forest fire data set is Antalya for the years between years 2005-2019. This study assesses the influence of weather conditions on the risk of forest fire occurrence in the Antalya Region using rare event logistic regression models. The response variable is composed of binary or dichotomous data defined by considering the amount of the burned forest area (i.e., 1: burned area greater than 10 hectare, 0: otherwise).

Forest fire occurrence (FRO) is defined considering the amount of the burned forest area as whether there was a major forest fire or minor forest fire. Hence, the FRO is a binary variable that takes two values: 1: The major forest fire defined as greater than 10 hectare and 0: The minor forest fires defined as less than 10 hectare. The rare class rate is 77/(2369+77)=0.03.

The Mediterranean climate in Türkiye's Mediterranean Region is characterized by the presence of short-term drying winds and heat waves, which replace the typically humid and refreshing sea breezes during the late spring to early autumn period [44]. These weather conditions play a significant role in influencing forest fires in this region, particularly in Antalya.

Three important meteorological variables that are closely monitored in relation to forest fires are relative humidity (RH), wind speed (WS), and maximum temperature (MT). These variables are recorded and analysed to understand their impact on forest fire occurrences. Meteorological data for this study was obtained from the Turkish State Meteorological Service (2022) [43]. Relative humidity (RH) is defined as the ratio of the amount of moisture present in the air to the amount of moisture required to saturate the air at the same pressure and temperature. It is a critical factor because the exchange of moisture between the air and dead forest fuels is ongoing. When relative humidity is low, fuels tend to lose moisture, while they gain moisture when relative humidity is high. Fluctuations in relative humidity have a significant effect on light fuels such as grass and pine needles, which quickly absorb and release moisture. As relative humidity decreases, these fine fuels become drier, leading to increased fire behaviour. On the other hand, heavy fuels respond to changes in humidity more slowly. Significant changes in heavy fuel moisture typically require the presence of substantial moisture, often from multiple rainfall events. Relative humidity is highest in the early morning and reaches its minimum value around midday. When relative humidity drops below 10%, it becomes extremely dangerous in terms of fire risk.

Understanding the relationship between weather factors and forest fires provides valuable insights into the conditions that contribute to fire ignition and spread. This knowledge plays a crucial role in the development of effective strategies for fire prevention, management, and mitigation. One of the significant weather factors influencing forest fires is WS, which can have destructive effects on flammable materials. High wind speeds facilitate the rapid spread of fires and hinder firefighting efforts. Another important factor is temperature, specifically the MT. High temperatures contribute to increased evaporation, which leads to reduced humidity levels. Low humidity, in turn, increases the risk of fire ignition and intensifies fire behaviour. The combination of high temperature, low humidity, and strong winds creates dangerous conditions for the occurrence and spread of forest fires.

In this study, the weather factors of RH, WS, and MT are considered as independent variables. These variables are analysed to determine their impact on forest fire occurrence in the Antalya region. Table 1 provides the descriptive statistics related to these independent variables. The mean maximum temperature in Antalya is approximately 27°C, while the average relative humidity is around 48%.

By examining these weather variables and their statistical characteristics, researchers can gain a better understanding of the prevailing climatic conditions and their influence on forest fire dynamics in the study area. This information is crucial for developing targeted and effective strategies to prevent and manage forest fires.

Table 1. Descriptive statistics of variables.

	RH (%)	WS (km/hr)	MT (°C)
Minimum	1.00	0.00	1.00
Mean	48.54	13.67	27.48
Median	50.00	12.00	28.00
Std. Dev.	17.36	9.18	7.04
Maximum	96.00	90.00	45.00
Skewness	-0.18	1.34	-0.66
Kurtosis	2.44	7.35	3.41

4. Results

In this section, ordinary logistic regression and rare-event logistic regression models have been applied to figure out which of the weather conditions are related to the risk of occurrence of the major forest fire or minor forest fire in Antalya city. The form of the models is given by

$$logit(\pi_i) = \beta_0 + \beta_1 RH_i + \beta_2 WS_i + \beta_3 MT_i$$

The results of the models were given in Table 2.

Table 2. The results of LR, FLR, FLAC and FLIC models for the forest fire data of Antalya.

	Coefficient	Standard Error	OR	χ^2 value	p-value
		LR Model			
Constant	-6.158	0.833	-	-	<0.001*
RH	-0.012	0.007	0.989	2.719	0.099
WS	0.026	0.012	1.026	4.783	0.028*
MT	0.097	0.021	1.101	20.421	<0.001*
		FLR Model			
Constant	-6.083	0.819	-	Inf	<0.001*
RH	-0.012	0.007	0.988	2.794	0.094
WS	0.026	0.012	1.026	4.670	0.031*
MT	0.094	0.021	1.099	24.094	<0.001*
		FLAC Model			
Constant	-6.092	0.821		Inf	<0.001*
RH	-0.011	0.007	0.988	2.712	0.099
WS	0.025	0.012	1.025	4.326	0.037*
MT	0.094	0.021	1.099	24.508	<0.001*
		FLIC Model			
Constant	-6.107	0.828	-	Inf	<0.001*
RH	-0.011	0.007	0.988	2.794	0.094
WS	0.026	0.011	1.026	4.670	0.031*
MT	0.095	0.021	1.099	24.094	<0.001*

*: Significant at confidence level 5%.

The results presented in Table 2 indicate that, based on the p-values obtained from the LR, FLR, FLAC, and FLIC models, two variables, namely WS and MT, have a significant impact on the risk of forest fire occurrence at a 5% significance level. This means that increases in WS and MT are associated with a higher likelihood of forest fires. On the other hand, the variable RH was found to be statistically insignificant, indicating that it does not have a significant influence on forest fire occurrence in the study area. These findings suggest that weather factors such as WS and MT play a crucial role in the occurrence and spread of forest fires, while RH may not be a significant predictor in this particular context. These results have important implications for understanding the factors contributing to forest fires and can assist in the development of targeted prevention and management strategies. By focusing on controlling and mitigating the effects of high wind speeds and maximum temperatures, authorities can work towards reducing the risk and impact of forest fires.

To evaluate and compare the performance of the different models, several classification performance criteria were utilized, including the Akaike Information Criterion (AIC), corrected AIC (AICc), class-based Brier score, and balanced accuracy. The AIC and AICc are commonly used measures for model selection based on the goodness of fit and complexity of the model. The AIC is calculated as

$$AIC = -2\log L + 2k$$

where L is value of likelihood function, k is number of parameters in the model and n is sample size. The model with the smallest AIC and AICc is the best model.

The AICc is a corrected version of the AIC that takes into account the sample size and number of parameters and defined as

$$AICc = AIC + \frac{2k(k+1)}{n-k-1}$$

By comparing the AIC and AICc values obtained for each model, we can assess their relative performance in terms of balancing goodness of fit and model complexity. Lower values of AIC and AICc indicate better model fit and parsimony.

In addition to the AIC and AICc, other performance criteria such as the class-based Brier score and balanced accuracy were also employed. The class-based Brier score measures the accuracy of the predicted probabilities of each class, while the balanced accuracy considers the overall accuracy by accounting for class imbalance. The class-based Brier score was obtained as below:

$$\text{Brier score} = \frac{1}{n}\sum_{i=1}^{n}(f_i - o_i)^2$$

where f_i is the predicted probability and o_i is the outcome. The Brier score of 0 implies that perfect accuracy otherwise 1 implies that perfect inaccuracy. The balanced accuracy is obtained based on the confusion matrix. The confusion matrix is as below:

	0 (Predicted)	1 (Predicted)
0 (Actual)	True Negative (TN)	False Positive (FP)
1 (Actual)	False Negative (FN)	True Positive (TP)

The balanced accuracy is calculated as

Balanced accuracy = [(TP/(TP+FN)) +(TN/(FP+TN))]/2

The balanced accuracy is a performance criterion that is used to assess the accuracy of a model in classifying forest fire occurrences. It takes into account the balance between sensitivity (the ability to correctly identify positive cases) and specificity (the ability to correctly identify negative cases). The balanced accuracy value ranges from 0 to 1, where a value of 1 indicates perfect classification performance, and a value of 0 indicates random guessing or poor classification. By considering this criterion, we can evaluate the models and determine which one provides the best balance between fit and complexity. A higher balanced accuracy score indicates a better model performance in accurately classifying forest fire occurrences.

This assessment allows us to compare the different models and select the one that demonstrates the highest accuracy and predictive power. Ultimately, this helps in developing effective strategies for forest fire prevention, management, and mitigation. The value of balanced accuracy ranges from 0 to 1, where 1 is the best and 0 is the worst.

By considering AIC, AICc, class-based Brier score and balanced accuracy performance criteria, we determine which model provides the best balance between fit and complexity, as well as assess the accuracy and predictive power of the models in classifying forest fire occurrences. The results are presented in Table 3.

Table 3. Performance values of the models for the forest fire data.

Model	AIC	AICc	Brier Score	Balanced Accuracy
LR	657.71	657.74	0.03	1.00
FLR	626.75	626.78	0.03	1.00
FLAC	664.05	664.08	0.03	1.00
FLIC	626.78	626.82	0.03	1.00

According to the results presented in Table 3, the FLR model emerges as the best model for predicting the impact of weather conditions on forest fires in Türkiye. This model provides valuable insights into the relationship between weather factors and the occurrence of forest fires.

Based on the FLR model, it is found that the risk of a major forest fire occurrence increases by a factor of 1.03 compared to a minor forest fire occurrence for every unit increase in WS (wind speed). This suggests that higher wind speeds contribute to a higher risk of major forest fires. Similarly, the odds ratio (OR) of MT (maximum temperature) indicates that the occurrence risk of a major forest fire increases by a factor of 1.09 compared to a minor forest fire for every unit increase in MT. This implies that higher temperatures are associated with an increased risk of major forest fires.

These findings highlight the significance of weather conditions, specifically wind speed and maximum temperature, in influencing the severity and occurrence of forest fires in Türkiye. By understanding these relationships, stakeholders and authorities can develop more effective strategies and measures for forest fire prevention and management.

5. Conclusion

Climate change has resulted in drier and hotter summers, changes in land-cover and land-use, and an increasing interface between wildland and urban areas, all of which have had a negative impact on forest fires. To effectively manage these fires, an ecosystem-based approach is crucial. This approach considers the ecological significance of fires and aims to reduce suppression costs while ensuring a sustainable response.

Accurate prediction of forest fires plays a vital role in implementing this management strategy. By accurately forecasting fire occurrences, forest managers can proactively and sustainably respond to potential threats. These forecasts also aid in better planning and resource management for fire-fighting efforts. They provide valuable data for predicting fire hazards, improving fuel management, educating the public, scheduling seasonal firefighters, and developing more effective fire management plans.

Forest fires pose a significant risk to ecosystems and the climate system in Türkiye's Mediterranean region, which has experienced considerable damage from fires in the past century. In this study, the relationship between meteorological conditions and forest fire risk in the Antalya Region was examined. The occurrence of forest fires larger than 10 hectares was treated as a rare event, and rare event logistic regression models were used to explore the relationship between fire occurrence and weather conditions.

The analysis revealed that wind speed and high air temperatures are the most favourable conditions for the occurrence of wildfires. The rare event logistic regression models demonstrated that these factors significantly influence the likelihood of forest fires larger than 10 hectares.

Comparing the rare event logistic regression models, the FLR model outperformed the FLIC and FLAC methods based on AIC values. According to the FLR model, the odds ratio of WS indicated that the risk of major forest fire occurrence increases by 1.03 times compared to minor forest fire occurrence for each unit increase in WS. Similarly, the odds ratio of MT suggested that the risk of major forest fire occurrence increases by 1.09 times for each unit increase in MT compared to minor forest fire occurrence.

In conclusion, the FLR model, with its lower AIC values, provides a better fit for predicting forest fire occurrences compared to the FLIC and FLAC methods. The model highlights the influence of wind speed and maximum temperature on the risk of major forest fires. By understanding these relationships, stakeholders can make informed decisions and take proactive measures to mitigate the impact of forest fires.

6. Acknowledgements

We would like to thank the coordinators of the project organized by the Scientific and Technological Research Council of Türkiye ("TÜBİTAK 2237-A, Doğa Bilimlerinde İstatistiksel Modelleme Teknikleri ve Uygulamaları"; Project Number: 1129B371900490) for their supports.

7. References

[1] Chandler, C., Cheney, P., Thomas, P., Trabaud, L. and Williams, D., *Fire in Forestry: Vol. II. Forest Fire Management and Organization*, Krieger Publishing Company: Malabar, FL, 1991.

[2] Schroeder, M.J. and Buck, C.C., *Fire Weather: A Guide for Application of Meteorological Information to Forest Fire Control Operations*, No. 360, US Forest Service, 1970.

[3] Burgan, R.E. and Rothermel, R.C., *BEHAVE: fire behavior prediction and fuel modeling system-FUEL subsystem*. General Technical Report INT-167. Ogden, UT: U. S. Department of Agriculture, Forest Service, Intermountain Forest and Range Experiment Station, 1984.

[4] Çanakçıoğlu, H., *Orman Koruma, İ. Ü. Orman Fakültesi* Yayın No: 411/3524, ISBN 975-404-199-7, Istanbul, 1993.

[5] Masinda, M.M., Li, F. and Qi, L., "Forest fire risk estimation in a typical temperate forest in Northeastern China using the Canadian forest fire weather index: case study in autumn 2019 and 2020", *Natural Hazards*, 111(1), 1085-1101, 2002.

[6] Holsten, A., Dominic, A.R., Costa, L. and Kropp, J.P., "Evaluation of the performance of meteorological forest fire indices for German federal states", *Forest Ecology and Management*, 287, 123–131, 2013.

[7] Pourtaghi, Z.S., Pourghasemi, H.R., Aretano, R. and Semeraro, T., "Investigation of general indicators influencing on forest fire and its susceptibility modeling using different data mining techniques", *Ecological Indicators*, 64, 72–84, 2016.

[8] Shirazi, Z., Huadong, G., Fang, C., Bo, Y. and Bin, L., "Assessing the impact of climatic parameters and their inter-annual seasonal variability on fire activity using time series satellite products in South China (2001–2014)", *Natural Hazards*, 85(3), 1393-1416, 2017.

[9] Chowdhury, E.H. and Hassan, Q.K., "Operational perspective of remote sensing-based forest fire danger forecasting systems", *Remote Sensing*, 104, 224-236, 2015.

[10] Koutsias, N., Xanthopoulos, G., Founda, D., Xystrakis, F., Nioti, F., Pleniou, M. and Mallinis, G., "On the relationships between forest fires and weather conditions in Greece from long-term national observations (1894-2010)", *International Journal of Wildland Fire,* 22(4), 493-507, 2013.

[11] Tosic, I., Mladjan, D., Gavrilov, M.B., Zivanović, S., Radaković, M.G., Putniković, S., Petrović, P., Mistridželović, I.K. and Marković, S.B., "Potential influence of meteorological variables on forest fire risk in Serbia during the period 2000-2017", *Open Geosciences*, 11(1), 414-425, 2019.

[12] Stephanie, E.M., Andrea, E.T., Ellis, Q.M., Larissa, L.Y., Jesse, D.Y. and Jose, M.I., "Climate relationships with increasing wildfire in the southwestern US from 1984 to 2015", *Forest Ecology and Management*, 460, 117861, 2020.

[13] Wang, S., Li, H. and Niu, S., "Empirical Research on Climate Warming Risks for Forest Fires: A Case Study of Grade I Forest Fire Danger Zone, Sichuan Province, China", *Sustainability*, 13, 7773, 2021.

[14] Oncel Cekim, H., Güney, C.O., Şentürk, Ö. and Özkan, K., "A novel approach for predicting burned forest area", *Natural Hazards*, 105, 2187–2201, 2021.

[15] Westerling, A.L., Hidalgo, D.R., Cayan, T.W. and Swetnam, T.W., "Warming and earlier spring increase western US forest wildfire activity", *Science*, 313(5789), 940– 943., 2006.

[16] Shikhov, A.N., Perminova, E.S. and Perminov, S.I., "Satellite-based analysis of the spatial patterns of fire-and storm-related forest disturbances in the Ural region, Russia", *Natural Hazards*, 97, 283–308, 2019.

[17] Nolan, R.H., Boer, M.M., Collins, L., Resco De Dios, V., Clarke, H., Jenkins, M., Kenny, B. and Bradstock, R.A., "Causes and consequences of eastern Australia's 2019–20 season of mega-fires", *Global Change Biology*, 26,1039–1041, 2020.

[18] Tedim, F., Leone, V., Amraoui, M., Bouillon, C., Coughlan, M., Delogu, G., Fernandes, P., Ferreira, C., McCaffrey, S. and McGee, T., "Defining extreme wildfire events: difficulties, challenges, and impacts", *Fire*, 1(1), 9, 2018.

[19] Ozbayoğlu, A.M. and Bozer, R., "Estimation of the burned area in forest fires using computational intelligence techniques", *Procedia Computer Science,* 12, 282-287, 2012.

[20] Beckage, B., Platt, W. and Panko, B., "A climate-based approach to the restoration of fire-dependent ecosystems", *Restoration Ecology*, 13(3), 429-431, 2005.

[21] Iliadis, L.S., "A decision support system applying an integrated fuzzy model for long-term forest fire risk estimation", *Environmental Modelling and Software*, 20(5), 613-621, 2005.

[22] Cortez, P. and Morais, A.D.J.R., "A data mining approach to predict forest fires using meteorological data. In New Trends in Artificial Intelligence", *Proceedings of the 13th EPIA 2007 - Portuguese Conference on Artificial Intelligence*, December, Guimarães, Portugal 25, 512-523, 2007.

[23] Cheng, T. and Wang, J., "Integrated spatio-temporal data mining for forest fire prediction", *Trans GIS*, 12(5), 591-611, 2008.

[24] Sakr, G.E., Elhajj, I.H. and Mitri, G., "Efficient forest fire occurrence prediction for developing countries using two weather parameters", *Engineering*

Applications of Artificial Intelligence, 24(5), 888-894, 2011.

[25] Vasconcelos, M.J.P., Silva, S., Tome´, M., Alvim, M. and Pereira, J.M.C., "Spatial prediction of fire ignition probabilities: comparing logistic regression and neural networks", *Photogrammetric Engineering and Remote Sensing*, 67, 73–87, 2001.

[26] Syphard, A.D., Radeloff, V.C., Keuler, N.S., Taylor, R.S., Hawbaker, T.J., Stewart, S.I. and Clayton, M.K., "Predicting spatial patterns of fire on a southern California landscape", *International Journal of Wildland Fire*, 17, 602–613, 2008.

[27] Lozano, F.J., Sua´rez-Seoane, S., Kelly, M. and Luis, E., "A multiscale approach for modeling fire occurrence probability using satellite data and classification trees: a case study in a mountainous Mediterranean region", *Remote Sensing of Environment*, 112, 708–719, 2008.

[28] Padilla, M. and Vega-Garcia, C., "On the comparative importance of fire danger rating indices and their integration with spatial and temporal variables for predicting daily human-caused fire occurrences in Spain", *International Journal of Wildland Fire*, 20, 46–58, 2011.

[29] Del Hoyo, V.L., Martın, I.M. and Vega, M.F., "Logistic regression models for human-caused wildfire risk estimation: Analysing the effect of the spatial accuracy in fire occurrence data", *The European Research Journal*, 130, 983–996, 2011.

[30] Magnussen, S. and Taylor, S.W., "Prediction of daily lightning-and human-caused fires in British Columbia", International *Journal of Wildland Fire*, 21(4), 342-356, 2012.

[31] Firth, D., "Bias reduction of maximum likelihood estimates", *Biometrika,* 80, 27–38, 1993.

[32] Heinze, G. and Schemper, M., "A solution to the problem of separation in logistic regression", *Statistics in Medicine*, 21, 2409–2419, 2002.

[33] King, G. and Zeng, L., "Logistic regression in rare events data", Political Analysis, 9(2), 137-163, 2001.

[34] Pesantez-Narvaez, J., Guillen, M. and Alcañiz, M., "RiskLogitboost Regression for Rare Events in Binary Response: An Econometric Approach", *Mathematics*, 9(5), 579, 2021.

[35] Chen, F., Peng, H., Chan, P.W., Ma, X. and Zeng, X., "Assessing the risk of windshear occurrence at HKIA using rare-event logistic regression", *Meteorological Applications*, 27 (6), e1962 doi.org/10.1002/met.1962., 2020.

[36] Puhr, R., Heinze, G., Nold, M., Lusa, L. and Geroldinger, A., "Firth's logistic regression with rare events: Accurate effect estimates and predictions?", *Statistics in Medicine*, 36, 2302–2317, 2017.

[37] Olmuş, H., Nazman, E. and Erbaş, S., "Comparison of penalized logistic regression models for rare event case", *Communications in Statistics - Simulation and Computation*, 51(4), 1578-1590, 2022.

[38] Amatulli G., Camia, G.A. and San-Miguel-Ayanz, J., "Estimating future burned areas under changing climate in the EU-mediterranean countries", *Science of The Total Environment*, 450, 209–222, 2013.

[39] Ertuğrul, M., "Orman Yangınlarının Dünyadaki ve Türkiye'deki Durumu", *Bartın Orman Fakültesi Dergisi,* 7, 43–50, 2005.

[40] Turco, M., Llasat, M.C., Von Hardenberg, J. and Provenzale, A., "Climate change impacts on wildfires in a Mediterranean environment", *Climatic Change*, 125(3), 369-380, 2014.

[41] Güney, C.O., Özkan, K. and Şentürk, Ö., "Modelling of spatial prediction of fire ignition risk in the Antalya-Manavgat district", *Journal of the Faculty of Forestry Istanbul University*, 66, 459–470, 2016.

[42] Antalya Regional Directorate of Forestry Retrieved February 03, 2022 from https://antalyaobm.ogm.gov.tr/SitePages/OGM/OGMDefault.aspx,

[43] Turkish State Meteorological Service 1930–2019 meteorological statistics. Turkish State Meteorological Service. Ankara, Türkiye Retrieved February 03, 2022, from https://www.mgm.gov.tr/veridegerlendirme/il-ve-ilceleristatistik.aspx?k=A&m=ANTALYA

[44] McCuttchan, M.H., Climatic Features as a fire determinant. USDA Forest Service General Technical Report WO, 1977.

[45] Güney, C.O., Ryan, K.C., Guney, A. and Hood, S.M., "Wildfire in Türkiye: Fire management challenges at an ancient crossroads of nature and culture", *Wildfire*, 28(3), 21-28, 2019.

[46] 2021 Turkey wildfires, Retrieved December 21, 2022 from https://en.wikipedia.org/wiki/2021_Türkiye_wildfires,