Modeling the Forest Fire Occurrence in Some Regions of Portugal. A First Approach

M. Filomena Teodoro[1,2(✉)]

[1] CINAV, Center of Naval Research, Naval Academy, Portuguese Navy, 2810-001 Almada, Portugal
mteodoro64@gmail.com
[2] CEMAT, Center for Computational and Stochastic Mathematics, Instituto Superior Técnico, Lisbon University, 1048-001 Lisbon, Portugal
maria.teodoro@tecnico.ulisboa.pt, maria.alves.teodoro@marinha.pt

Abstract. The forest surveillance carried in Portugal presents some difficulties. To improve and solve some of these problems the use of new technologies such as unmanned aerial vehicle systems (UAVS) can be implemented in an efficient way. To do so we need to determine the risk of occurrence of a forest fires at a certain time in a certain region. With such goal, we have used several statistical techniques, such as ARIMA or GLM approach. We built adequate models in certain regions but for others it was impossible to determine a good quality model. The work is still going on, we expect to enlarge the area where we can get a good quality prevision of risk fire.

Keywords: Forest fire · Risk · Logit · ARIMA

1 Introduction

The Portuguese forest spaces, in addition to harboring, an unquestionable biogenic wealth, it is estimated that they are responsible for generating wealth in the order a thousand million euros/year, adding that they support direct jobs. Fires, according to several studies, are today the phenomenon susceptible to greater destruction that forest areas face. According to data of the National Forestry Authority, statistical data and reports alluding to the burned areas in forest areas, since 1980 some millions of hectares had burned. A more detailed analysis reveals that the greatest number of occurrences occur in the more densely populated districts (Porto and Lisbon). In contrast, the districts with the lowest population density have a lower number of occurrences, although they have substantially larger areas covered by fires. The districts located within the country are a "good" example of this phenomenon. For the same period of time, the districts that have the largest burned area are Guarda, Vise and Caste lo Branco. In Portugal, the Castelo Branco district stands out in terms of forest

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2022
O. Gervasi et al. (Eds.): ICCSA 2022 Workshops, LNCS 13377, pp. 653–666, 2022.
https://doi.org/10.1007/978-3-031-10536-4_43

stands, as they present in the same period a greater burned area. Naturally, the availability of a fire probability map, in the short/medium term, it will be a valuable instrument to support the decision in the election of zones for treatment of fuels, location of certain infrastructures, or the disposition of fire suppression means available in a given territory. The spatial estimate of occurrence of fire derives from a cartographic zoning, in a given territory, with different levels of probability that the phenomenon will occur [1,2].

The thematic relevance of "forest fire risk" cartography has deserved the greater attention both by the scientific community and the technical community. The definition of the concept of "risk" does not, however, result from unanimous and universally recognized agreement. The concepts of risk and hazardousness have been the subject of attempts at differentiation.

The term hazard, translated in some situations by dangerousness, is usually used to describe the characteristics of forest fuels at a given location and in a certain time period. The term is intended to express above all the potential of fire behavior. The term risk is commonly accepted and derives from the probability occurrence of the phenomenon, in a certain place and at a certain time, considering the nature and incidence of the causative agents. In other words, it is assumed that area has a certain fire potential taking into account the difficulty of determining its ignition probability [3–5].

In addition to the terminological context, the term "risk" has been applied both in the "structural risk" as well as "dynamic risk". By structural risk, or local risk, according with to some authors, it is understood as the interaction of a set of factors with the history of forest fires in a given area and over a period of significant time, such as land occupation and use, fuel load, orography, infrastructures for fire prevention and detection, population density, among others. On the other hand, dynamic risk essentially presupposes the detection of alteration the flammability of forest fuels during the fire season, such as the wind speed and direction or temperature [1,6,7]. In the same context, the authors of [8] describe a study about occurrence of wildfires in the Iberian Peninsula based on harmonized data from national forest inventories.

From a conceptual perspective, the implementation of probability methodologies occurrence of forest fires, necessarily presupposes the analysis of factors that explain the appearance, development and impact of fire on populations, material goods and natural resources.

Any fire probability methodology implies the combination of two or more variables, which normally derive from forest fuels, orography and various infrastructures that assume human interaction.

In [9] the author intended the construction of forest fire probability models, with regional specificity, as well as the elaboration of scenarios of spatial probability of occurrence of fires. The considered areas of study were completely different in terms of orography and land occupation and use. The models estimated by logistics regression analysis used as explanatory variables physical parameters, land occupation and location space of certain infrastructures susceptible to human interaction with fire. The two obtained models evidenced regional specific details allowing the construction of probability scenarios of occurrence of

fire, which showed spatial agreement levels close both regions, although in different validation periods. In [10], the authors performed a cross regional study in Southern Europe, estimating fire ignition probability using Hurdle models. The authors of [11] used SARIMA models to forecasting the number of hotspots.

The present work focuses its objectives on the development of probabilistic models to estimate the risk of forest fire, using the technique of logistic regression [9,12,13] and ARIMA approach [11,14–16], employing the specificity and regional suitability of the models developed taking into account the importance of the various variables considered.

The manuscript presents a first approach where is done a descriptive statistical analysis considering the data from some regions of Portugal, followed by the estimation of Logit models, taking into consideration the work presented in [9,12].

This article is divided in five sections. An introduction starts the manuscript; in Sect. 2 are presented some details about wildfires occurrence. Follows some theoretical issues about Logit models and ARIMA techniques with a brief description of the used methodology; in Sect. 4 is performed the empirical application; in Sects. 5 are detailed and discussed some remarks.

2 Preliminaries

Vegetation fires occur and spread with the presence of unwanted fires. It swoops over the surface and devours plants no matter the cause artificial (accidental, intentional or negligent) or due to lightning, plant fires are considered natural disasters. Economic and human damage that can destroy property as well as environmental damage endangering material and people's lives. In this sense, it is determinant studying the dynamics of vegetation fire propagation as a study area experimental in nature. It stems from the human need to understand and consequently develop mechanisms to anticipate and control this phenomenon [17]. Fire propagation models are tools that can simulate fire scenarios. The propagation of fire as a function of factors regulating the behavior of fire that much is proposed in [18]. The integration of these models with data from geographic information systems is a qualitative leap forward in the development of computational tools to help create decisions on actions including vegetation fire monitoring, prevention and suppression [18]. Accordingly with [17], the fire behavior is determined by the interaction of three sets of factors: vegetation, atmospheric conditions and topography.

3 Methodology

3.1 ARIMA

The identification of an ARIMA model to model the data can be considered one of the most critical phases when using ARIMA approach. For a stationary time

series[1] the selection of the model to be used is mainly based on the estimated auto-correlations and partial auto-correlations, which we will use to compare with the theoretical quantities and identify a possible model for the data.

The auto-regressive models with order p AR(p), the moving average with order q MA(q) and their combination, ARMA(p, q) models have their auto-correlation functions (ACF) with a certain specific feature, similarly to a finger print:

- The ACF of an autoregressive process with order p is infinite in extent that decays according to a damped exponential/sinusoidal;
- The ACF of a moving average with order q process is finite, i.e. presents a cut after the lag q;
- The ACP of ARMA process (p, q) is like a mixture of the processes described in previous items, the ACF has infinite that decays according to exponents/damped sinusoidals after the lag $q - p$.

The idea is to identify a pattern that behaves with the same profile that some theoretical model. In particular, the it is useful to identify MA models but it is not so simple to identify other kind of models. As a possible solution, we can compute the partial auto-correlation function (PACF). This function corresponds to the correlation of X_t, X_{t-k+1} removing the effect of the observations $X_{t-1}, X_{t-2}, \ldots, X_{t-k-1}$ and is denoted by ϕ_{kk}. In the case a stationary time series we can use the Yule-Walker equations to compute the PACF. Again, the PACF have a specific profile for each process like a proper finger print:

- The PACF of a MA(d) is infinite in extent that decays according to a damped exponential/sinusoidal (similarly to the behaviour of an ACP from a AR(d) process;
- The PACF of AR(p) the process is finite, i.e. presents a cut after the lag q, like the behavior of an ACP from a MA(p) process;
- The PACF of ARMA process (p, q) is similar to an ACF from a MA(q) process.

A general method for finding a.c. for a stationary process with f.a. is using the Yule-Walker equations. This method seems to fail in the case of non-stationary time series (the irregular component is significant). To solve this issue we differentiate the non stationary series so many times (d times) as necessary to get a stationary series. After these differences of successive terms of the chain, applied d times, we can apply the same technique: identify which model(s) are identified from ACF and PACF. A model that represents a non-stationary series, differenciated d times, with an auto-regressive component with order p and a moving average component with order q is represented as an $ARIMA(p, d, q)$.

ARIMA models are also capable of modeling series that have a seasonal component, being described as $SARIMA(p, d, q)(P, D, Q)_m$, where the first parentheses refers to the non-seasonal part of the model ($ARIMA(p, d, q)$) and the second to the seasonal part, with m the number of seasonal periods.

[1] A time series that develops in time around a constant mean.

To estimate the best models between several proposals, we usually apply the information criteria AIC, BIC: the best models have the lowest values of AIC and BIC. Also the log of likelihood function is a good statistic to evaluate the quality of the estimated models: the lowest value means a better model.

After selection, the models need to be validated. One of the rules is to analyze the residuals (the ith residual is the difference between the ith obervation and its estimate). Residuals are supposed to be Gaussian and non-correlated. To verify this can be used several testes and other techniques (Llung-Box test, box-Pierce teste, Kolmogorov-Smirnov test, Bera and Jarcke test, some graphics, e.g. boxplots, qq plots.

The estimates precision evaluation is another step to include in all process. For that we can compute the usual measures: MAPE, MADE, etc.

In general the procedure of all process is composed by the following iterates:

1. Models formulation: use of ACF and PACF;
2. Models adjustment: estimation of model parameters, application of suitability measures of estimates;
3. Validation of models: selection of variables, diagnostics, residual analysis and interpretation.
4. Analysis of precision and update the models.

3.2 LOGIT Model

Follows some details about Logistic regression as a particular case of general linear models. In the classical linear model, a vector X with p explanatory variables $X = (X_1, X_2, \ldots, X_p)$ can explain the variability of the variable of interest Y (response variable), where $Y = Z\beta + \epsilon$. Z is a specification matrix with size $n \times p$ (sometimes $Z = X$, considering an unitary vector in first column), β a parameter vector and ϵ a vector of random errors ϵ_i, independent and identical distributed to a reduced Gaussian.

The data are in the form (y_i, x_i), $i = 1, \ldots, n$, as result of observation of (Y, X) n times. The response variable Y has expected value $E[Y|Z] = \mu$.

In GLM, where the model is an extension of classical model, the response variable, following an exponential family distribution [19], does not need to be Gaussian, and a transformation of expected value of response variable is related with explanatory variables. A detailed description of GLM ca be found in [19,20].

Another extension from the classical model is the fact that the function which relates the expected value and the explanatory variables can be any differentiable function. Each Y_i has expected value $E[Y_i|x_i] = \mu_i = b'(\theta_i)$, $i = 1, \ldots, n$.

Also, it is defined a differentiable and monotone link function g that relates the random component with the systematic component of the response variable.

The expected value μ_i is related with the linear predictor $\eta_i = z_i^T \beta_i$ through the relation

$$\mu_i = h(\eta_i) = h(z_i^T \beta_i), \qquad \eta_i = g(\mu_i) \tag{1}$$

where h is a differentiable function, $g = h^{-1}$ is the link function, β is a vector of parameters with size p (the same size of the number of explanatory variables) and Z is a specification vector with size p.

We can found different link functions in GLM. When the random component of response variable has a Poisson distribution, the link function is logarithmic and the model is log-linear. In particular, when the linear predictor $\eta_i = z_i^T \beta_i$ coincides with the canonical parameter θ_i, $\theta_i = \eta_i$, which implies $\theta_i = z_i^T \beta_i$, the link function is denominated canonical link function.

Usually, in presence of dichotomous data, discrete models are usually estimated by logistic or probit regression [19].

The logistic regression is widely applied in models with dichotomous data and/or proportions. It is supposed that the response variable Y verifies $Y_t \cap Bin(1, \pi_t)$, at time $t = i$ i.e.

$$f(y_i|x_i) = (\pi_i)^{y_i} (1 - \pi_i)^{1-y_i}, \qquad y_i = 0, 1, \tag{2}$$

where each time i as associated a specification vector z_i which results from p explanatory variables $x = (x_1, x_2, \ldots, x_p)$. Knowing that

$$E[Y_i] = \pi_i, \qquad \theta_i = ln\left(\frac{\pi_i}{1 - \pi_i}\right)$$
$$and \qquad \theta_i = \eta_i = z_i^T \beta_i \ , \tag{3}$$

we conclude that the logistic function is the link function g. The probability of success π_i, the probability of fire occurrence ($P(Y = 1)$) is given by

$$\pi_i = \frac{e^{z_i^T \beta_i}}{1 + e^{z_i^T \beta_i}}. \tag{4}$$

Notice that $E[Y_i] = \pi_i = \mu_i \in [0, 1]$.

Another distribution function can be a good candidate to the link function inverse. One can suppose that the success probability π_i and the covariate vector are related by

$$\pi_i = \Phi(\eta_i) = \Phi(z_i^T \beta_i), \tag{5}$$

where $\phi(.)$ is the Gaussian distribution $N(0, 1)$. In this case, we obtain the link function $g(\mu_i) = \Phi^{-1}(\mu_i)$, the probit function, and the estimated model is named the probit model.

3.3 Extended LOGIT Model

In [21] the authors estimated the forest Fire Probability Mapping in Eastern Serbia using a Logistic regression approach versus a random forest method. The authors of [12] used an extended version of Logit model with spatial and time dependence. The use of a spatio-temporal model is done knowing that if there exist a fire in a certain localization (a,b) at $t = i$ the existent fuel material is burned, the probability of occurrence a new fire at the same point (a, b) in next weeks or months will be low [22]. Also similar topography, weather conditions and vegetation conduce to identical fire risk in distinct positions.

Defining $Y_{iab} = 1$ when there exists a fire at time $t = i$ in position (a, b) and $Y_{iab} = 0$ when there not exists a fire at time $t = i$ in position (a, b) the extended logistic model is based on

$$f(y_{iab}|x_{iab}) = (\pi_{iab})^{y_{iab}} (1 - \pi_{iab})^{1-y_{iab}}, \qquad y_{iab} = 0, 1, \tag{6}$$

where

$$E[Y_{iab}] = \pi_{iab}, \qquad \theta_{iab} = ln\left(\frac{\pi_{iab}}{1 - \pi_{iab}}\right)$$
$$and \qquad \theta_{iab} = \eta_{iab} = z_{iab}^T \beta_{iab}. \tag{7}$$

The probability of fire occurrence π_{iab} at time $t = i$ in position (a, b) is given by

$$\pi_{iab} = \frac{e^{\theta_{iab}}}{1 + e^{\theta_{iab}}}. \tag{8}$$

4 Empirical Application

4.1 Data

In order to achieve the proposed objectives, it was used the available data during the last 10 years, from 9/2/2011 to 31/12/2021. In Table 1 we can find the variables used in the determined models: year (ano), date and initial hour of the fire occurrence (DHIni), duration of the fire (Dur), burned area (AreaTotal), administrative zone (Concelho), local (Local) and respective latitude (Lat) and longitude (Lon). Some information, considered important in literature, was not available, for example, kind of forest, altitude, declive, humidity level, etc. We expect to complete our data base in a short time.

Summarizing some information, in Fig. 1 is displayed the number of fire occurrences between 2011 and 2020 (in blue) and during 2021 (in orange). We can observe that the districts with bigger density of population have more occurrences (Porto, Lisboa, Aveiro, Braga). We can notice that, in some districts, during 2021 the number of occurrences are almost half of the occurrences in previous 10 years.

Table 1. Available data. Period: from 9/2/2011 to 31/12/2021.

	A	D	E	F	K	L	M	N	O
1	Ano	Tipo	DHIni	Dur	AreaTotal	Concelho	Local	lat	lon
2	2011	Florestal	9/2/2011 16:37	02h18m	0,13	Montemor-o-Novo	Vale Cerejo	38,816	-8,423
3	2011	Florestal	2/4/2011 16:53	01h37m	0,2	Vila Viçosa	MONTE DA ALMAGF	38,726	-7,405
4	2011	Agrícola	14/4/2011 11:38	01h7m	0,174	Montemor-o-Novo	Estrada da Amoreira	38,658	-8,147
5	2011	Agrícola	28/4/2011 15:21	45m	0,46	Portel	Tapada do Serrado	38,312	-7,703
6	2011	Agrícola	5/2/2011 14:25	01h35m	0,055	Crato	Quinta da Nave	39,321	-7,723
7	2011	Florestal	3/3/2011 14:20	02h20m	0,2	Celorico da Beira	Quinta do Gil - Galis	40,594	-7,421
8	2011	Agrícola	4/3/2011 18:00	45m	0,06	Celorico da Beira	Ramisga	40,635	-7,309
9	2011	Florestal	7/4/2011 15:37	02h48m	0,2	Celorico da Beira	Quinta da Carriça-C	40,594	-7,41
10	2011	Florestal	13/4/2011 21:00	35m	1,3	Figueira de Castelo Rodr	Cabreiriças	40,875	-6,866
11	2011	Florestal	7/2/2011 14:55	02h50m	3	Celorico da Beira	PRADOS - PNSE (PRA	40,562	-7,378
12	2011	Florestal	8/2/2011 13:57	01h28m	6	Figueira de Castelo Rodr	Rodo de Cebolas	40,869	-6,827
13	2011	Florestal	8/2/2011 19:34	02h56m	0,5	Celorico da Beira	Quinta de São Bento	40,674	-7,3
14	2011	Florestal	28/2/2011 17:26	54m	0,01	Celorico da Beira	IC7 (EN17) Km 121	40,561	-7,479
15	2011	Florestal	6/4/2011 12:48	32h12m	46	Águeda	Lourizela	40,641	-8,318
16	2011	Florestal	5/2/2011 16:45	01h42m	0,73	Felgueiras	CHUQUEIRO (Choqu	41,335	-8,262
17	2011	Florestal	6/2/2011 16:15	55m	0,44	Paredes	Alto da Vila	41,226	-8,381
18	2011	Florestal	7/2/2011 11:20	40m	0,2	Paredes	Penhas Altas	41,233	-8,416
19	2011	Florestal	3/3/2011 17:16	58m	0,395	Paredes	QUINTA DA IGREJA	41,189	-8,378
20	2011	Florestal	3/3/2011 20:04	36m	0,11	Penafiel	LG. RANHA (BUSTÊL	41,23	-8,271
21	2011	Florestal	5/3/2011 9:32	02h8m	0,68	Penafiel	MATO (Matos)	41,228	-8,277
22	2011	Florestal	5/3/2011 13:49	01h11m	0,178	Penafiel	Campo Grande	41,143	-8,351
23	2011	Florestal	5/3/2011 15:00	01h20m	1,087	Paredes	Tourilhe	41,251	-8,362
24	2011	Florestal	7/3/2011 16:51	01h58m	0,8	Felgueiras	L.SARDOAL (Sardoal	41,374	-8,159
25	2011	Florestal	7/3/2011 19:17	01h3m	0,08	Penafiel	L.PEDRANTIL (Pedra	41,22	-8,233
26	2011	Florestal	8/3/2011 12:38	01h42m	0,07	Matosinhos	R. PASSADOURO (Ru	41,24	-8,703
27	2011	Florestal	10/3/2011 13:46	02h14m	1,351	Penafiel	Penoucos	41,226	-8,171
28	2011	Florestal	10/3/2011 15:00	01h25m	0,39	Vila Do Conde	LG. CASAL (Ponte de	41,396	-8,687
29	2011	Florestal	10/3/2011 15:17	02h28m	1,15	Póvoa de Varzim	RUA CRESCENCIOS (	41,379	-8,763
30	2011	Florestal	10/3/2011 16:25	04h10m	3,15	Vila Do Conde	RUA NOVA (RIO MA	41,403	-8,679

When we consider the burned area per district (see Fig. 2) between 2011 and 2020 (in blue) and during 2021 (in orange), the major burned area is found in the districts with lower population density (for example, Castelo Branco).

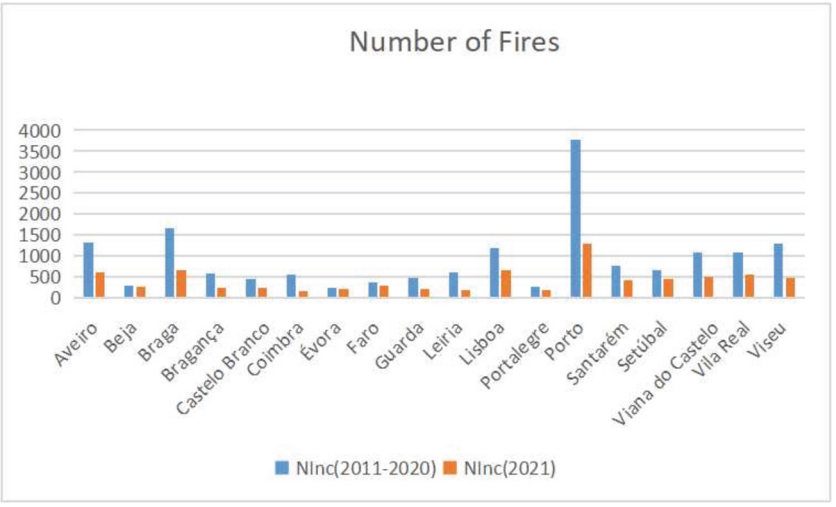

Fig. 1. Number of fire occurrences between 2011 and 2020 (in blue) and during 2021 (in orange). (Color figure online)

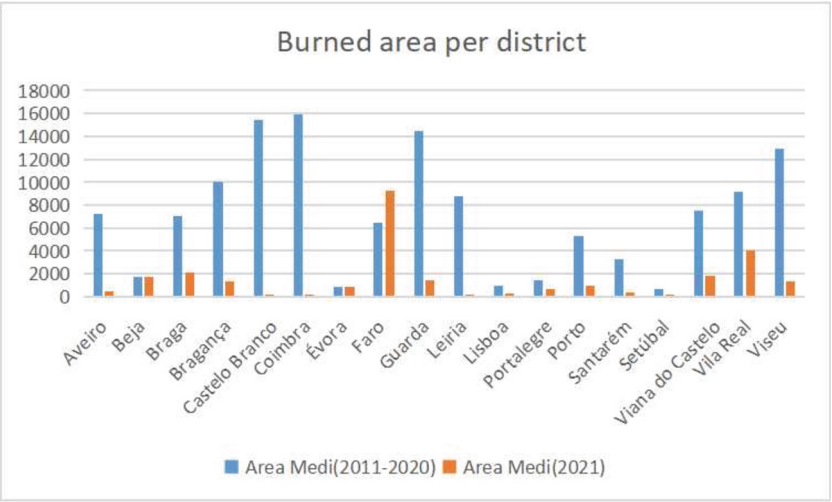

Fig. 2. Burned area per district between 2011 and 2020 (in blue) and during 2021 (in orange). (Color figure online)

We find in Fig. 3 the time at the beginning of fire from 22h to 06h. It suggests a human act due the amount of fires starting during the night period. This fact is also evidenced when we consult Fig. 4, where is displayed a detail from Fig. 3, the occurrences between midnight and 2AM.

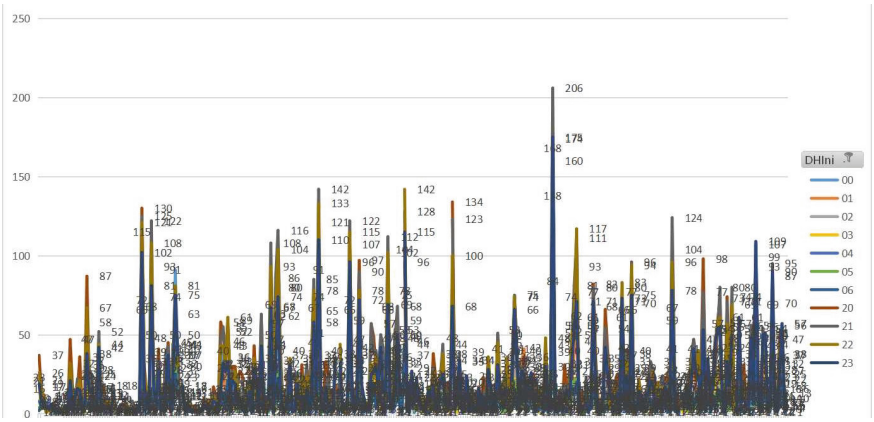

Fig. 3. Fire occurrences when start fire time is during the night period.

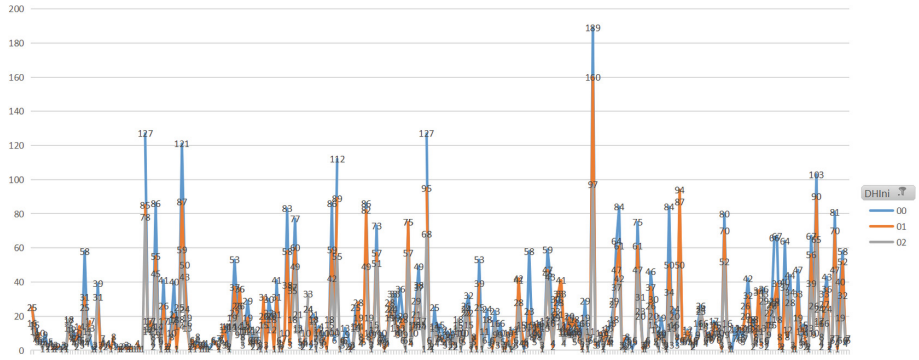

Fig. 4. Fire occurrences when start fire time is between midnight and 2 AM.

We also plot (see Fig. 5) the count of fire occurrences from 07 h to 19 h.

At Figs. 7 and 6 are presented the number of fire occurrences during the day and between midnight and 02 AM in 2021 respectively.

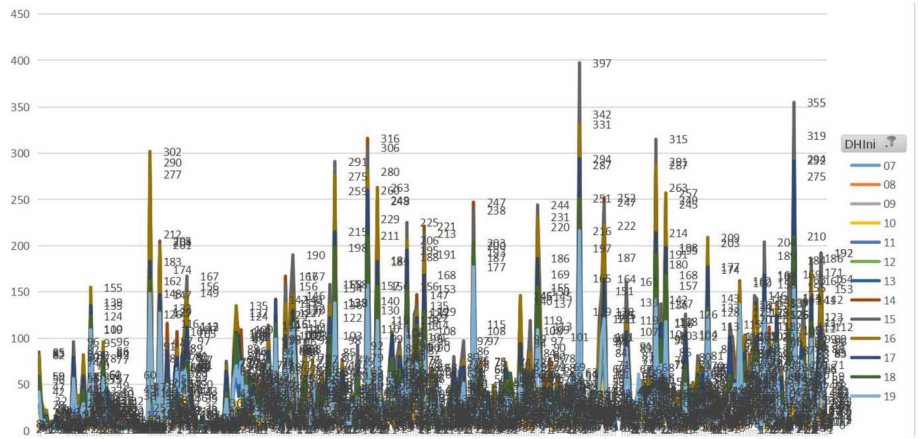

Fig. 5. Fire occurrences when start fire time is during the day.

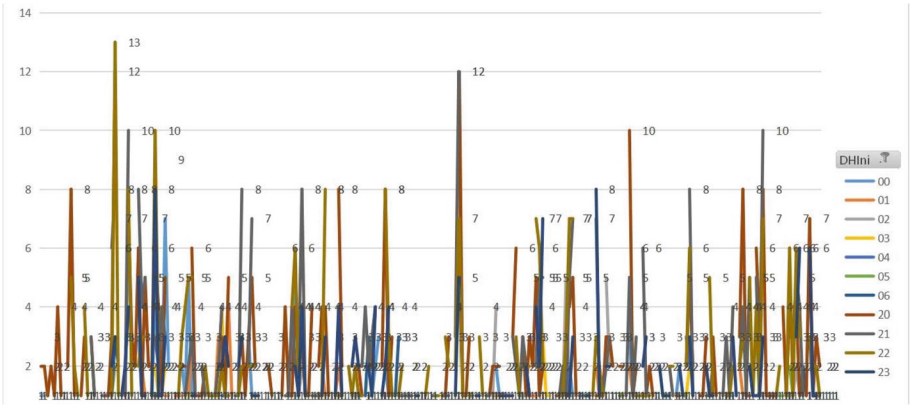

Fig. 6. Fire occurrences when start fire time is during the night period, during 2021.

4.2 Models

Under the aim of building the models, we have codified some variables, namely the duration of fire, and the burned area was classified in intervals, the other variables, latitude, longitude and occurrence time were considered without codification in a first approach.

When one builds a model to estimate the fire risk probability it is necessary to select the region of study. When a certain administrative region is selected to perform the study, several methods are tested to get adequate quality in the pretended forecast. As an initial approach we selected two districts: Castelo Branco and Vila Real. Both districts have a rural profile and a low population density. We have applied a generalized least squares (LOGIT model) and a Sarima app-

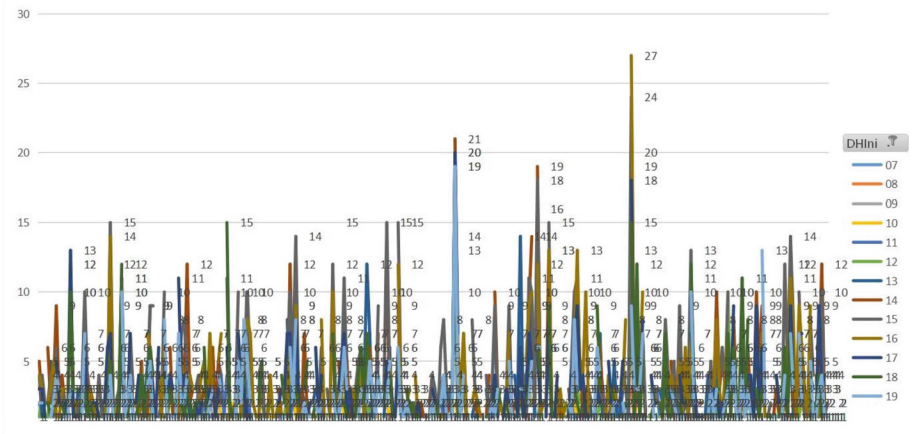

Fig. 7. Fire occurrences when start fire time is during the day in 2021.

Table 2. Codification of variables.

Variables	Description	Codification
Y	Fire occurrence	1 = Yes
		0 = No
x_1	Burned area	$'1' if \ \leq 0.1$
		'2'= (]0.1, 0.5])
		'3'= (]0.5, 1])
		'4'= (]1, 2])
		'5'= (]2, 4])
		'6'= (]4, 6])
		'7'= (]6, 8])
		'8'= (]8, 10])
		'9'= (]10, 15])
		'10'= (]15, 20])
		'11'= (]20, 30])
		'13'= (]40, 50])
		'14'(> 15)
x_2	Duration	$'1' = (]0, 0.5])$
		$'2' = (]0.5, 1])$
		$'3' = (]1, 2])$
		$'4' = (]2, 4])$
		$'5' = (]4, 6]')$
		$'6' = (]6, 8])$
		$'7' = (]8, 10])$
		$'8' = (]10, 15])$
x_3	Start fire occurrence time	Positive real
x_4	Latitude	
x_5	Longitude	

roach. Here, are reproduced some details about the logit approach. The first step was to get some association measures between the dependent variable and the remaining variables. It was obtained a significant association between the occurrences and the burned area (with delay), also between the occurrences and the duration time. It was considered the date of the occurrence as explanatory variable (in polynomial form or square root form). All usual steps of model validation accordingly with GLM approach were performed. About one of the used measures, the adjusted determination coefficient (R^2_{adj}), we have followed the author of [23], where in this kind of logistic approach, it is considered adequate a model with a $R^2_{adj} > 0.2$. When we consider the Castelo Branco district, we got an adequate logistic model. The best approach to Vila Real district data was

obtained by a Sarima approach. The statistical analysis details of the obtained models will be presented in an extended version of this article.

5 Final Remarks

The forest surveillance carried in Portugal presents some difficulties. To improve and solve some of these problems the use of new technologies such as unmanned aerial vehicle systems (UAVS) can be implemented in an efficient way. To do so we need to determine the risk of occurrence of a forest fires at a certain time in a certain region. With such goal, we have used several statistical techniques, such as ARIMA or GLM approach. We were able to get adequate models in some districts, namely Vila Real and Castelo Branco. For some other districts we were enable to get adequate forecasts of the probability of occurrence of a fire. We built adequate models in certain regions but for others it was impossible to determine a good quality model. The work is still at the begining, we expect to enlarge the data base with the details about the kind of tree, accessibility, altimetry, sun exposure, slope, etc., so we can complete our study.

Acknowledgements. This work was supported by Portuguese funds through the *Center of Naval Research* (CINAV), Portuguese Naval Academy, Portugal and *The Portuguese Foundation for Science and Technology* (FCT), through the *Center for Computational and Stochastic Mathematics* (CEMAT), University of Lisbon, Portugal, project UID/Multi/04621/2019. The author also aknowledges her colleagues P. Silva and C. Pina by their availability and ideas discussion.

References

1. Colin, P., Jappiot, M., Mariel, A.: Protection des forêts contre l'incendie : fiches techniques pour les pays du bassin méditerranéen (2001)
2. Vesseron, P., Delarue, F., Rosenberg, P., Sappin, M.: lans de prévention des risques naturels (ppr) - risques d'incendies de forêt, methodogic Guide
3. Chuvieco, E. In: Wildland Fire Danger Estimation and Mapping. The Role of Remote Sensing Data. Series in Remote Sensing, vol. 4. World Scientific, São Carlos, SP (2003)
4. Salinero, E., Isabel, M.: Nuevas tecnologías para la estimación del riesgo de incendios forestales. CSIC, Madrid (2004)
5. Keane, R., et. al.: A method for mapping fire hazard and risk across multiple scales and its application in fire management. Ecological Modelling **221**(1), 2–18 (2010)
6. Freire, S., Carrão, H., Caetano, M.: Produção de cartografia de risco de incêndio florestal com recurso a imagens de satélite e dados auxiliares. IGP, Lisboa (2002)
7. Bergonse, R., Bidarra, J.: Probabilidade bayesiana e regressão logística na avaliação da susceptibilidade de ocorrência de incêndios de grande magnitude. Finisterra **45**(89), 79–104 (2010)
8. Nunes, L., et al.: Analysis of the occurrence of wildfires in the iberian peninsula based on harmonised data from national forest inventories. Ann. For. Sci. **76**(27), 1–17 (2019)

9. Silva, A.S.: Modelação e mapeamento da probabilidade de incêndio florestal. Master thesis, Instituto Politécnico de Castelo Branco (2011)
10. D'Este, M., et. al.: Modeling fire ignition probability and frequency using hurdle models: a cross-regional study in Southern Europe. Ecological Processes **9**(1), 1–14 (2020)
11. Espinosa, M.M., Prado, S.M., Ghellere, M.: Uso do modelo sarima na previsão do número de focos de calor para os meses de junho a outubro no estado de mato grosso. Ciência e Natura, UFSM **32**(2), 7–21 (2010)
12. Preisler, H.K., et. al.: Probability based models for estimation of wildfire risk. Int. J. Wildland Fire **13**(2), 133–142 (2003)
13. Hamilton, J.D.: Time Series Analysis. Princeton University Press, New Jersey (1994)
14. Box, G., Jenkins, G., Reinsel, G.C.: Time Series: Forecasting and Control. Prentice Hall, New Jersey (1994), 3rd edn
15. Chatfield, C.: The Analysis of Time Series: An Introduction. Chapman and Hall/CRC, Boca Raton FL (2004), 6th Edition
16. Cleveland, R.B., Cleveland, W.S., McRae, J.E., Terpenning, I.: Stl: a seasonal-trend decomposition procedure based on loess. J. Official Stat. **6**(1), 3–73 (1990)
17. Pyne, J.S., Andrews, P.L., Laven, R.D.: Introduction to wildland fire. Wiley, New York (1996), 2nd Edition
18. Pastor, E., Zarate, L., Planas, E., Arnaldos, J.: Mathematical models and calculation systems for the study of wildland fire behaviour. Prog. Energy Combust. Sci. **29**(2), 139–153 (2003)
19. Turkman, M.A., Silva, G.: Modelos Lineares Generalizados da teoria a prática. Sociedade Portuguesa de Estatística, Lisboa (2000), 2nd Edition
20. Nelder, J.A., Wedderburn, R.W.M.: Linear models. J. Roy. Stat. Soc. **35**, 370–384 (1972)
21. Milanović, S., et al.: Forest fire probability mapping in eastern serbia: Logistic regression versus random forest method. Forests **2**(5), 1–17 (2021)
22. Ricardo, A.C.: Modelação da Probabilidade de Ocorrência de Incêndio em Povoamentos Florestais de Portugal Continental. Master thesis, Instituto Superior de Agronomia, UTL (2010)
23. Hosmer, D., Lemeshow, S., Sturdivant, R.X.: Applied Logistic Regression. John Wiley, New York (2013), 3rd Edition