

# FIRE BEHAVIOR PREDICTION USING MACHINE LEARNING ALGORITHMS

Vinicius Barros RODRIGUES<sup>1</sup>  
Fillipe Tamiozzo Pereira TORRES<sup>1</sup>

- **ABSTRACT:** Wildfires can affect ecosystem structure and threaten human lives. Understanding fire behavior and predicting fire activities is a crucial issue to mitigate fire impacts. Machine Learning is currently an important tool for the modeling, analysis, and visualization of environmental data and wildfire events. In this study, we assessed the performance of two machine learning algorithms for modeling and predicting fire intensity, the height of flames, and fire rate of spreading in *Eucalyptus urophylla* (Myrtaceae, Myrtales) and *Eucalyptus grandis* (Myrtaceae, Myrtales) plantations spatially located in Viçosa - MG, Brazil. The Random Forest showed to be the best algorithm for fire modeling, with climatic conditions, and moisture of the combustible material being the variables that significantly affect the prediction of fire behavior.
- **KEYWORDS:** Modeling; wildfire; random forest; fire prediction.

## 1 Introduction

Wildfires frequency and their impacts are increasing (SAN-MIGUEL-AYANZ *et al.*, 2012). They can change ecosystem structures and threaten human lives, which makes fire behavior and wildfire prediction crucial issues to mitigate fire impacts (BOWMAN *et al.*, 2009). In this context, persons in public management must decide the most effective distribution in scenarios with a limited amount of resources (MAVSAR *et al.*, 2013).

The Machine Learning (ML) is a modern and important tool for the modelling, analysis, and visualization, including used in environmental data and wildfire events (KANEVSKI *et al.*, 2009; O'CONNOR *et al.*, 2017). According to Rodrigues and de la Riva (2014), the ML models also have shown good predictive accuracy applied for other disciplines. Those models usually show good generalization abilities, even when modelling high dimensional and complex nonlinear phenomena (HASTIE *et al.*, 2009).

The most ML analysis and statistical models applied to wildfires analyze available in the current literature included attempts to predict fire occurrences and generate fire risk maps (LEUENBERGER *et al.*, 2018). There are no studies on modelling fire behavior variables using the ML model. In this study, we assessed the performance of two machine learning algorithms (GLMNET and Random Forest) applied for modelling and predicting fire behavior (fire intensity, the height of flames, and rate of spread) in an *Eucalyptus* plantation located in Viçosa city, Minas Gerais state, Brazil. Forest plantations show an

<sup>1</sup> Universidade Federal de Viçosa – UFV, Departamento de Engenharia Florestal, CEP 36570-900, Viçosa, Minas Gerais, Brazil. E-mail: [viniciusbrbio@gmail.com](mailto:viniciusbrbio@gmail.com); [tamiozzo@ufv.br](mailto:tamiozzo@ufv.br)

isotropic variation of topography and biomass fuel. This characteristic allowed us to assess the effects of weather on fuel moisture and fire behavior. Forest fires in eucalyptus plantations is an increasing problem and are responsible for 30% of fire occurrences in Brazil (SANTOS *et al.*, 2007).

## 2 Methods

### 2.1 Fire modeling database

In this study, we used a fire behavior database provided by Torres *et al.* (2019 unpublished). These authors assessed the fire behavior in 80 burning plots within a 10-year old hybrid *Eucalyptus urophylla* (Myrtaceae, Myrtales) and *Eucalyptus grandis* (Myrtaceae, Myrtales) plantations in Viçosa, Minas Gerais state (Figure 1), Brazil, between July 1, 2016, and November 30, 2016.

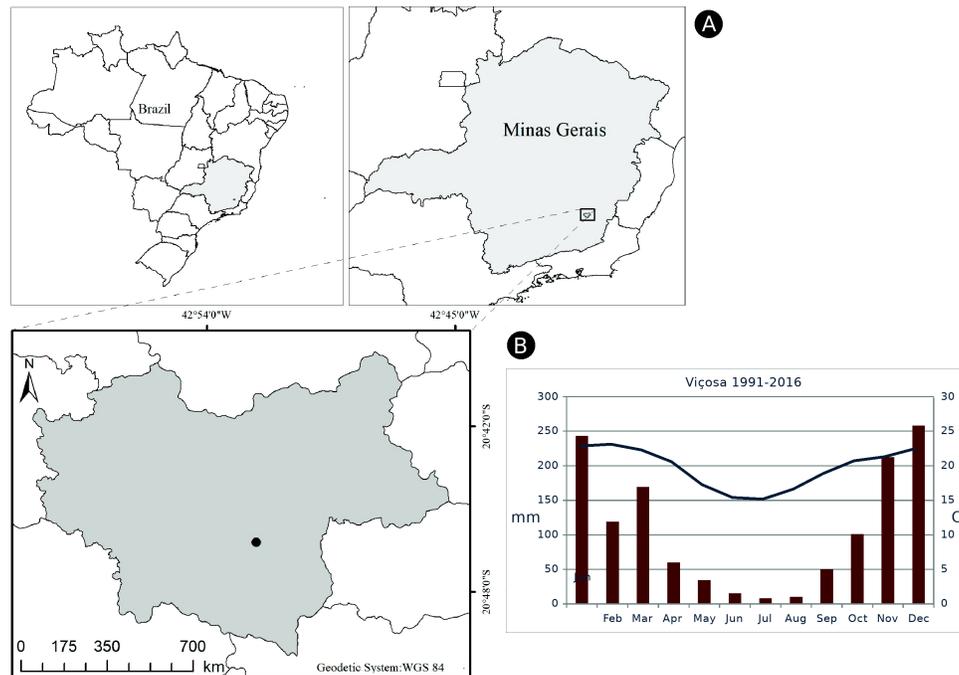


Figure 1 – A) The study area (black dot) is located in Viçosa, Minas Gerais state, Brazil; B) The mean rainfall (mm) and temperature (C°) variation through the years in the months.

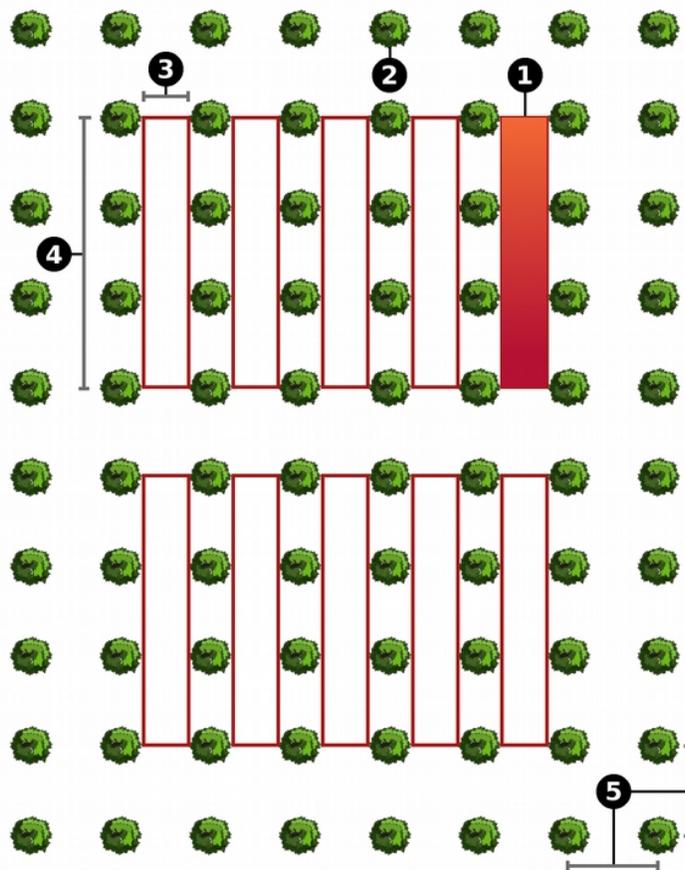


Figure 2 – Experimental design, where 1) plots for burning; 2) *Eucalyptus* spp.; 3) 2m width, 4) 10m length; and 5) 3m between the *Eucalyptus* spp.

The mean altitude in the city is 650 m and the climatic classification according to Köppen-Geiger is Cwa (humid subtropical climates with dry winters, between April and September, and wet summers, between October and March) (PEEL *et al.* 2007).

The area of each plot was 20 m<sup>2</sup> (2 m × 10 m) centralized between the planting lines where observation and measurement of combustible material (height, amount, type and humidity) and meteorological variables (precipitation and evaporation) were evaluated as well as fire behavior during burning (“Fire intensity”, “Height of flames” and “Rate of spread”).

We used in our models three directly measurable fire behavior variables (“Fire intensity”, “Height of flames” and “Rate of spread”) and selected the explanatory variables based in Torres *et al.* (2019 unpublished) models (Table 1).

Table 1 - Built models based on Torres *et al.* (TORRES *et al.* 2019 unpublished) fire behavior experiments

<b>Model 1 – Response variable: Fire intensity</b>
Explanatory variables: Days without rain + Relative humidity (%) + Wind speed (m/s) + Class 1 fuel moisture (%) + Load of Class 1 fuel (kg.m <sup>-2</sup> ) + Bed depth (cm)
<b>Model 2 – Response variable: Height of flames</b>
Explanatory variables: Days without rain + Air temperature (°C) + Class 1 fuel moisture (%) + Bed depth (cm)
<b>Model 3 – Response variable: Rate of spread</b>
Explanatory variables: T°C + Relative humidity (%) + Wind speed (m/s) + Class 1 fuel moisture (%) + Live fuel moisture (%) + Load of Class 2 fuel (kg.m <sup>-2</sup> ) + Load of miscellaneous (kg.m <sup>-2</sup> ) + Total load (kg.m <sup>-2</sup> )

## 2.2 Statistical procedures

In this study, we aimed to determine the effects of the continuous explanatory variables on the continuous response variables (i) “Fire intensity (kw.m<sup>-1</sup>)”, (ii) “Height of flames (cm)”, (iii) “Rate of spread (m/s)”.

The model-fitting was conducted using two different Machine Learning models: GLMNET and Random Forest. The GLMNET (Lasso and Elastic-Net Regularized Generalized Linear Models) fits the GLMs (Generalized Linear Models) by using penalized maximum likelihood (KUHN, 2008) and Random Forest is an algorithm based on decision trees (BREIMAN, 2001). These algorithms were chosen in our analysis based on their application for wildfire prediction reported by Tracy *et al.* (2018) and Leuenberger *et al.* (2018).

The collected and tabulated dataset was randomly split for training and testing the fire behavior models: 60% of the data was used for training and 40% for validating the models. We conducted a comparative analysis between the models based on R-squared, Mean Absolute Error (MAE) and Root Mean Squared Error (RMSE) to estimate the model performances. Subsequently, we assessed the variable importance to estimate the best model performance (BREIMAN, 2001).

In our study, the statistical analysis was conducted using the R software (R CORE TEAM, 2019) and the package "caret" (KUHN, 2008).

### 3 Results

The descriptive data of the experiment can be observed in the Table 2.

Table 2 - Descriptive data of the experimental fire plots with the measured variables. "Min" and "Max" are the minimum and maximum values sampled.

Variable	Min	Median	Mean	Max
Fire intensity (kw.m <sup>-1</sup> )	15.68	155.05	188.74	779.60
Height of flames (cm)	16.00	98.00	94.28	164.00
Rate of spread (m/s)	0.001795	0.022014	0.023905	0.068279
Days without rain	2	11	13.23	31
Air temperature (°C)	14.95	26.95	26.43	34.70
Wind speed (m/s)	0.300	2.500	2.487	6.000
Relative humidity (%)	30.00	55.00	55.15	90.00
Class 1 fuel moisture (%)	9.813	15.772	18.402	41.218
Live fuel moisture (%)	10.04	16.43	23.93	71.48
Load of Class 1 fuel (kg.m <sup>-2</sup> )	0.1952	0.6004	0.6349	1.2948
Load of Class 2 fuel (kg.m <sup>-2</sup> )	0.0	0.5520	0.5707	1.8996
Bed depth (cm)	3.950	7.050	7.324	11.500

#### 3.1 Model performance

The estimated model performances indicated a difference between the GLMNET and Random Forest models (Figure 3a, 3b, and 3c) results in predicting fire behavior. The Random Forest model showed the lowest MAE and RMSE for all predicted variables. Both models presented high R-squared values. The scatter plots for the predicted models and test data are presented in Figure 4.

The Random Forest model showed better performance for predicting wildfire behavior. The variables "Relative humidity", "Live fuel moisture", "Class 1 fuel moisture", and "Load of class 1" fuel affected the Fire intensity model prediction (Figure 3d). The variables "Days without rain" and "Miscellaneous moisture" were not significant in the Fire intensity model prediction; (b) The variables "Class 1 fuel moisture", "Days without rain" and "Air temperature" affected the "Rate of spread" model. The variable "Bed depth" was not significant in the "Rate of spread" model (Figure 3e); and (c) The "Height of flames" model was mostly affected by the "Class 1 fuel moisture". The

variables "Days without rain", "Air temperature" and "Bed depth" was not significant in the "Rate of spread model" (Figure 3f).

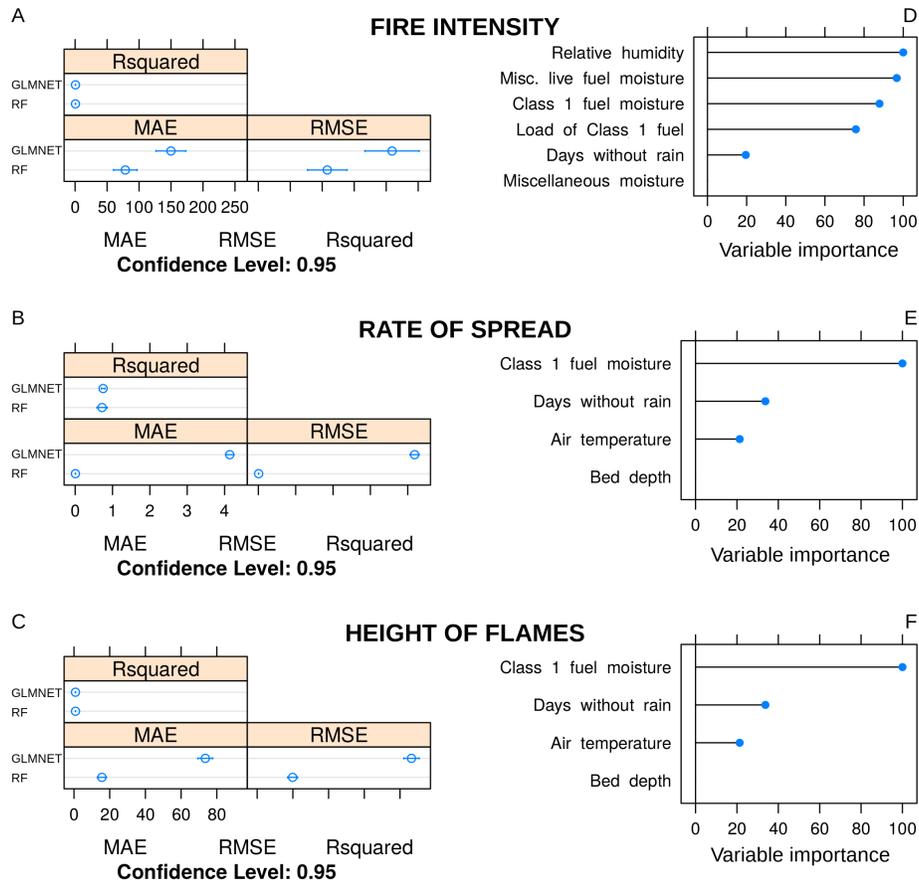


Figure 2 – Estimate of Model performances and importance of predictor variables. The left plots are the GLMNET and Random Forest (RF) estimated performance. Response variables: (a) "Fire intensity", (b) "Rate of spread", and (c) "Height of flames" (Rsquared, MAE, RMSE). The right plots are the importance of predictor variables, where (d) "Fire intensity", (e) "Rate of spread", and (f) "Height of flames".

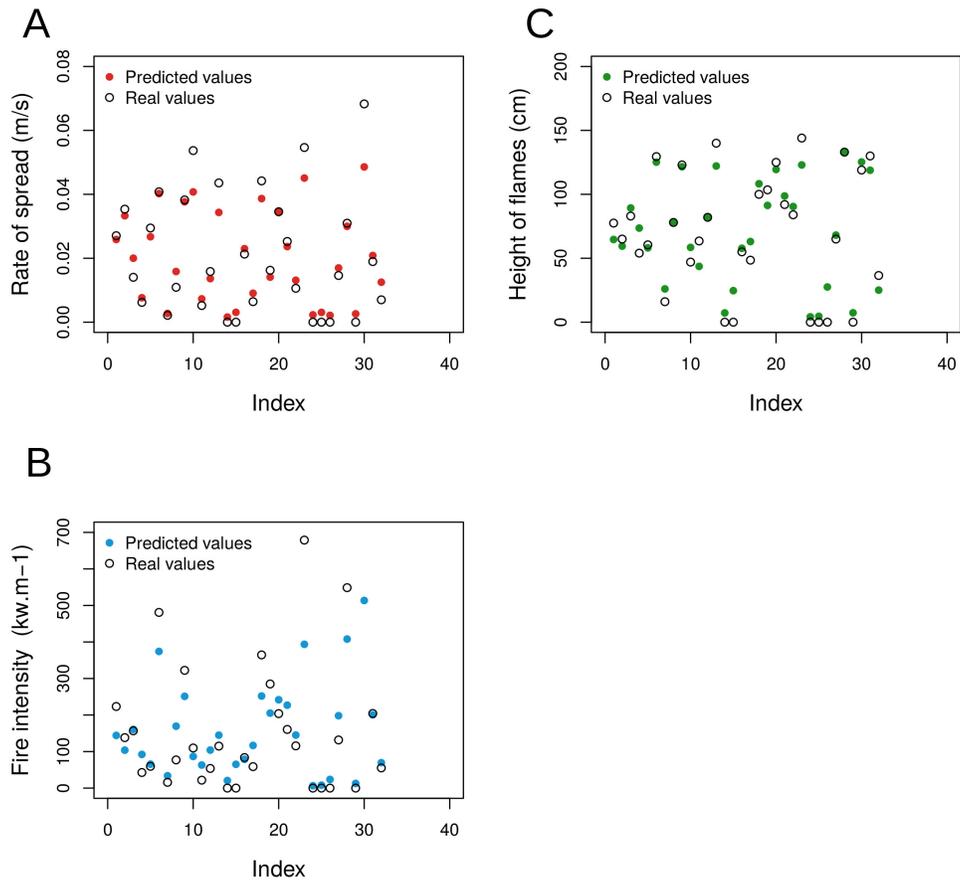


Figure 3 - Plotted values of the predicted values using Random Forest algorithm (filled points) and the real values (white points). Index (x-axis) is the sample order. Response variables: (a) “Rate of spread”, (b) “Fire intensity”, and (c) “Height of flames”.

#### 4 Discussion

This is the first approach based on Machine Learning to modeling fire behavior. The Random Forest algorithm showed better performance compared to GLMNET, similar to Rodrigues and de la Riva (2014). These authors also observed that the Random Forest algorithm showed higher accuracy using fewer predictive variables. Using fewer variables can improve and facilitate the interpretation of the statistical results (JAAFARI *et al.*,

2018), once that wildfire behavior is a complex phenomenon that can be affected by geological, topographical and environment factors (JAAFARI *et al.*, 2018).

Our results showed that the climate conditions and fuel moisture were the most important variables in the models. The fuel moisture content was the most important property that controls the flammability of fuels. The humidity of the fuel material is affected by the weather conditions, which can rapidly change. Both living and dead biomass have different water retention mechanisms and different responses to climate variations (TORRES *et al.*, 2018).

The fire-climate interaction is an intense relationship. Studies conducted by Littell and Gwozdz (2011) and Morton *et al.*, (2013) have incorporated the relationship between the meteorological elements and the fuel moisture as an independent biophysical variable, which showed higher relationship with fire behavior variables than any meteorological variable alone.

Fuel moisture determines the likelihood of fire ignition and propagation, so the estimation of fuel moisture content is a key factor for modeling fire behavior (PÉREZ-SÁNCHEZ *et al.*, 2017).

Machine Learning algorithms showed great predictive accuracy compared to traditional statistical methods (ELITH *et al.*, 2008; RODRIGUES and DE LA RIVA, 2014) and can be applied to develop new modeling approaches for fire management.

## Conclusions

The Random Forest algorithm achieved better results than GLMNET to predict fire behavior (“Fire intensity”, “Rate of spread”, and “Height of flames) for *Eucalyptus urophylla* (Myrtaceae, Myrtales) and *Eucalyptus grandis* (Myrtaceae, Myrtales) plantations. The fuel moisture was the variable that most influenced the fire variables. Our results showed that Random Forest models should be considered for further applications, such as in the elaboration of susceptibility maps.

RODRIGUES, V. B., TORRES, F. T. P. Predição do comportamento do fogo utilizando algoritmos de aprendizado de máquinas. *Rev. Bras. Biom.* Lavras, v.38, n.3, p.343-352, 2020.

- *RESUMO: Incêndios florestais podem modificar a estrutura de um ecossistema e ameaçar vidas humanas. Portanto, é crucial compreensão mais profunda da atividade e a precisão das previsões do comportamento do fogo. O Aprendizado de Máquinas (Machine Learning) é uma ferramenta moderna, com destaque em modelagens, análises e visualização de dados ambientais e eventos de incêndios florestais. No presente estudo, buscamos avaliar a performance de dois algoritmos de Machine Learning na modelagem e predição da intensidade do fogo, altura das chamas e velocidade de propagação, em uma área de *Eucalyptus urophylla* (Myrtaceae, Myrtales) e *Eucalyptus grandis* (Myrtaceae, Myrtales) em Viçosa – MG. O algoritmo de Random Forest apresentou a melhor performance na modelagem do fogo, sendo que as condições climáticas e a umidade do material combustível foram as variáveis que influenciam significativamente a predição do comportamento do fogo.*
- *PALAVRAS-CHAVE: Modelagem; incêndio florestal; random forest; predição de fogo.*

## References

- BOWMAN, D. M.; BALCH, J. K.; ARTAXO, P.; BOND W. J.; CARLSON, J. M.; COCHRANE, M. A.; D'ANTONIO, C. M.; DEFRIES, R. S.; DOYLE, J. C.; Harrison, S. P. Fire in the Earth system, *Science*, v.324, p.481–484, 2009.
- BREIMAN, L. Random forests, *Machine learning*, v.45, p.5–32, 2001.
- ELITH, J.; LEATHWICK, J. R.; HASTIE, T. A working guide to boosted regression trees. *Journal of Animal Ecology*, v.77, p.802–813, 2008.
- HASTIE, T.; TIBSHIRANI, R. F. J. Unsupervised learning. In: HASTIE, T.; TIBSHIRANI, R. F. J. (Ed.). *The elements of statistical learning*. London: Springer, 2009.
- JAAFARI, A.; ZENNER, E. K.; PHAM, B. T. Wildfire spatial pattern analysis in the Zagros Mountains, Iran: A comparative study of decision tree based classifiers. *Ecological Informatics*, v.43, p.200–211, 2018.
- KANEVSKI, M.; TIMONIN, V. P. A. *Machine Learning for Spatial Environmental Data: Theory, Applications and Software*. EPFL Press, 2009.
- KUHN, M. Building predictive models in R using the caret package. *Journal of Statistical Software*, v.5, p.1–26, 2008.
- LEUENBERGER, M.; PARENTE, J.; TONINI, M.; PEREIRA, M. G.; KANEVSKI, M. Wildfire susceptibility mapping: Deterministic vs. stochastic approaches. *Environmental Modelling & Software*, v.101, p.194–203, 2018.
- LITTELL, J. S.; GWOZDZ, R. Climatic Water Balance and Regional Fire Years in the Pacific Northwest, USA: Linking Regional Climate and Fire at Landscape Scales. In: McKenzie, D.; Miller, C. F. D. (Ed.). *The Landscape Ecology of Fire*. London: Springer, 2011.
- MAVSAR, R.; CABÁN, A. G.; VARELA, E. The state of development of fire management decision support systems in America and Europe. *Forest Policy and Economics*, v.29, p.45–55, 2013.
- MORTON, D.; COLLATZ, G.; WANG, D.; RANDERSON, J.; GIGLIO, L.; CHEN, Y. Satellite-based assessment of climate controls on US burned area. *Biogeosciences*, v.10, p.247–260, 2013.
- O'CONNOR, C. D.; CALKIN, D. E.; THOMPSON, M. P. An empirical machine learning method for predicting potential fire control locations for pre-fire planning and operational fire management. *International Journal of Wildland Fire*, v.26, p. 587-597, 2017
- PEEL, M. C.; FINLAYSON, B. L.; MCMAHON, T. A. Updated world map of the Köppen-Geiger climate classification. *Hydrology and Earth System Sciences Discussions*, v.4, p.439-473, 2007
- PÉREZ-SÁNCHEZ, J.; SENENT-APARICIO, J.; DÍAZ-PALMERO, J.; CABEZAS-CEREZO, J. D. D. A comparative study of fire weather indices in a semiarid south-

eastern Europe region. Case of study: Murcia (Spain). *Science of Total Environment*, v.590-591, p.761–774, 2017.

R CORE TEAM. *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria, 2019

RODRIGUES, M.; DE LA RIVA, J. An insight into machine-learning algorithms to model human-caused wildfire occurrence. *Environmental Modelling & Software*, v.57, p.192–201, 2014.

SAN-MIGUEL-AYANZ, J.; RODRIGUES, M.; DE OLIVEIRA, S.; PACHECO, C.; MOREIRA, F.; DUGUY, B.; CAMIA, A. Land cover change and fire regime in the European Mediterranean region. In: MOREIRA, F.; ARIANOUTSOU, M.; CORONA, P.; DE LAS HERAS, J. (Ed.). *Post-fire management and restoration of southern European forests*. Springer, 2012.

SANTOS, J. F.; SOARES, R. B. A. Perfil dos incêndios florestais no Brasil em áreas protegidas no período de 1998 a 2002. *Floresta*, v.36, p.93–100, 2007.

TRACY, J.; TRABUCCO, A.; LAWING, A.; GIERMAKOWSKI, J.; TCHAKERIAN, M.; DRUS, G.; COULSON, R. Random subset feature selection for ecological niche models of wildfire activity in Western North America. *Ecological Modelling*, v.383, p.52–68, 2018.

TORRES, F. T. P.; ROMEIRO, J. M. N.; SANTOS, A. C. D. A.; OLIVEIRA NETO, R. R.; LIMA, G. S.; ZANUNCIO, J. Fire danger index efficiency as a function of fuel moisture and fire behavior. *Science of Total Environment*, v.631-632, p.1304–1310, 2018.

Received on 02.09.2019

Approved after revised on 12.02.2020