FIRE BEHAVIOR PREDICTION USING MACHINE LEARNING ALGORITHMS

Vinícius Barros RODRIGUES¹
Filipe Tamiozzo Pereira TORRES¹

• 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

¹ Universidade Federal de Viçosa – UFW, Departamento de Engenharia Florestal, CEP 36570-900, Viçosa, Minas Gerais, Brazil. E-mail: viniciusbrbio@gmail.com; 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 m² (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

<table>
<thead>
<tr>
<th>Model</th>
<th>Response variable</th>
<th>Explanatory variables</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model 1</td>
<td>Fire intensity</td>
<td>Days without rain + Relative humidity (%) + Wind speed (m/s) + Class 1 fuel moisture (%) + Load of Class 1 fuel (kg.m^{-2}) + Bed depth (cm)</td>
</tr>
<tr>
<td>Model 2</td>
<td>Height of flames</td>
<td>Days without rain + Air temperature (°C) + Class 1 fuel moisture (%) + Bed depth (cm)</td>
</tr>
<tr>
<td>Model 3</td>
<td>Rate of spread</td>
<td>T°C + Relative humidity (%) + Wind speed (m/s) + Class 1 fuel moisture (%) + Live fuel moisture (%) + Load of Class 2 fuel (kg.m^{-2}) + Load of miscellaneous (kg.m^{-2}) + Total load (kg.m^{-2})</td>
</tr>
</tbody>
</table>

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^{-1})”, (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.

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

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” (R squared, MAE, RMSE). The right plots are the importance of predictor variables, where (d) “Fire intensity”, (e) “Rate of spread”, and (f) “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.


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 influenciaram significativamente a predição do comportamento do fogo.

PALAVRAS-CHAVE: Modelagem; incêndio florestal; random forest; predição de fogo.
References


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-


Received on 02.09.2019
Approved after revised on 12.02.2020