Automatic Calibration of Forest Fire Weather Index For Independent Customizable Regions Based on Historical Records

Jorge S. S. Júnior*, João Paulo*, Jérôme Mendes*, Daniela Alves† and Luís Mário Ribeiro†
* Institute of Systems and Robotics, Department of Electrical and Computer Engineering,
University of Coimbra, Pólo II, PT-3030-290 Coimbra, Portugal
Email: jorge.silveira@isr.uc.pt, jpaulo@isr.uc.pt, jermendes@isr.uc.pt
† Association for the Development of Industrial Aerodynamics, Coimbra, Portugal PT–3030–289
Email: danielaalves@adai.pt, luis.mario@adai.pt

***Abstract*—Wildfire Decision Support Systems are critical tools for civil protection authorities in the management of all wildfire stages, including prevention. To timely act and apply the necessary preventive measures to reduce the fire danger in wildfires, many proposed calibration studies of the Canadian Forest Fire Weather Index System (CFFWIS) have been performed mainly based on techniques that still depend on manual and empirical analysis, being limited to exploiting a few regions. This paper proposes a methodology for automatic calibration of the CFFWIS to obtain a fire danger measurement that best suits the specific characteristics of a given region. The proposed methodology, applied to 769 regions from Europe, is based on the *k*-means clustering technique to automatically identify patterns in the data sets composed of elements of the CFFWIS and wildfire records. The results of the automatic calibration of the CFFWIS on each of the 769 regions reinforce the versatility of the proposed methodology, which can be adapted to different regions.**

***Keywords*—Automatic calibration, CFFWIS, k-means clustering, Wildfires, Fire Danger Classes.**

I. Introduction

Wildfires are phenomena that inevitably can occur due to unknown factors, natural causes (as dry weather, lightning), accidental human actions (as by bonfires, railways), and arson [1]. According to the 19th issue of the European Commission's annual report on wildfires for the year 2018 [2], wildfires over 30 hectares (*ha*) of burnt area were observed in 38 countries from Europe, Middle East, and North Africa, obtaining a total mapping of 204.861 [ha]. Although this value was lower than the long-term average and that reported in 2017, 1.376.090 [ha], there was a presence of large wildfires in more countries than usual, where more than 250 people were injured or killed. Furthermore, this report emphasizes the damage caused to protected sites such as Natura2000, with 36% of the burnt area of the total, with Portugal being one of the countries most affected. Due to the problems caused by the wildfires phenomena, it is essential to obtain knowledge of the wildfire danger for a given day and region, contributing to a quick decision making that can prevent these occurrences and considerably reduce human and forest losses (wildfire management). In this sense, fire danger rating systems have been introduced, which have the role of evaluating the meteorological and soil moisture conditions that favor the ignition of wildfires, in addition to determining, analytically, fire danger into classes with variations between Low, Moderate, and High. Among them, the National Fire Danger Rating System (NDFRS) used in the United States [3], the McArthur Forest Fire Danger Index (FFDI) used in eastern Australia [4], and the Canadian Forest Fire Danger Rating System (CFFDRS) used in Canada [5].

The Canadian Forest Fire Weather Index System (CFFWIS), a CFFDRS subsystem, was developed for Canada with an indication of six classes: Very Low, Low, Moderate, High, Very High, and Extreme. Outside Canada, it has been extensively explored, accompanied by necessary calibration studies taking into account the distinction between meteorological and soil moisture conditions to adapt to specific regions. Several works have been focused on the calibration of the CFFWIS. In [6], it is presented a methodology to calibrate the CFFWIS in the districts of Continental Portugal, based on the statistical daily values of FWI (Fire Weather Index), and wildfire records (number of occurrences and burnt area) for each district. In [7], a procedure for operational generation of daily maps of fire danger over Mediterranean Europe and calibration of CFFWIS is proposed, using probabilities of fire duration exceeding to determine thresholds leading five classes of fire danger. In [8], the calibration of CFFWIS is made through measurement of the forest fuel moisture content in the field for two Mediterranean regions (Algarve, Portugal and Peloponnese, Greece), using mathematical models. In [9], a percentile-based calibration of the CFFWIS is proposed with optimization to the United Kingdom conditions, through exceeding analysis of CFFWIS components with seasonal variation and land cover type. In all these works, it is noted the difficulty of the authors in covering a long period of fire records, as they still depend on manual and empirical analyses about very specific regions. Also, the variation in the values of the elements of CFFWIS and wildfire records due to the use of data from different meteorological stations, in certain regions, justifies the adjustment of the danger classification to

978-1-7281-8708-2/20/$31.00 ©2020 IEEE
DOI 10.1109/AIKE48582.2020.00011

meet this variation, as shown in [6].

In this sense, the present work proposes a methodology for automatic calibration of the Canadian Forest Fire Weather Index System (CFFWIS) adapted individually for a given region. The number of fire danger classes is defined according to the seasons, divided between Hot Season (Summer and Autumn) and Cold Season (Winter and Spring). Thus, for the present study, five danger classes for Hot Season and three danger classes for Cold Season are chosen according to experts' knowledge regarding the pattern of occurrences [10]. To define the five and three classes of fire danger in Hot Season and Cold Season, respectively, the proposed methodology uses as input variables the daily FWI values, the daily number of fire occurrences, and the daily burnt area of a given region, which are used in the *k*-means clustering method. The objective is to determine danger classes that correspond to the obtained clusters. The *k*-means clustering method promotes unsupervised learning of data sets composed of CFFWIS and wildfire records for each region, considering a strong relationship between these elements. In this way, it can automatically identify distinct pattern levels by the defined clusters (fire danger classes) through the similarity between the data. By associating clusters with danger classes, thresholds are defined between these classes for each region. These thresholds take into account the FWI value, which is the measurement estimated daily by the meteorological weather systems, which is matched with a fire danger class. The proposed methodology was applied to calibrate the CFFWIS in 769 regions from Europe, where each region was divided into Hot Season and Cold Season. The results have shown that the proposed methodology has performed the automatic calibration of the CFFWIS on each of the 769 regions successfully, reinforcing the versatility of the proposed methodology to define the fire danger classes to different regions without requiring manual or empirical analysis. To the best of our knowledge, there are no studies that use techniques to automatically determine fire risk classes using CFFWIS for a given region, making the present study important to prove that clustering-based learning, specifically based on k-means clustering, can be used to achieve this goal.

This work follows the following structure. In Section II, an overview about CFFWIS is presented. Section III presents the input variables and desired target for CFFWIS calibration, the *k*-means clustering method adapted for this calibration, and the general framework containing all fundamental elements for the proposed methodology. Then, Section IV presents the results of the implementation of automatic calibration of this system using data sets of 769 regions from Europe provided by the European Commission Joint Research Centre (JRC), as well as validation by analysing large wildfires. Finally, Section V presents the final remarks, as well as proposals for improving this work.

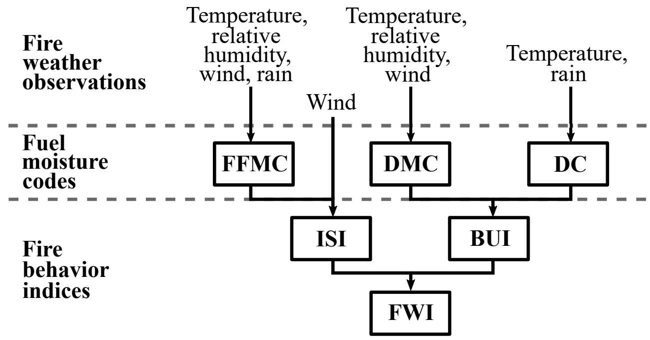


Fig. 1: Structure of the Canadian Forest Fire Weather Index System.

II. The Canadian Forest Fire Weather Index System

The Canadian Forest Fire Danger Rating System (CFFDRS) is a wildfire danger assessment system mainly used in Canada developed since 1968, presenting two subsystems: the Canadian Forest Fire Weather Index System (CFFWIS) and the Canadian Forest Fire Behavior Prediction System (CFFBPS) [11]. Years of CFFDRS research in Canada have resulted in a final CFFWIS structure presented in [5], consisting of six elements related to environmental conditions and vegetation characteristics:

- Fine Fuel Moisture Code (FFMC): represents the moisture content of litter and other cured fine fuels;
- Duff Moisture Code (DMC): consists of average moisture content of loosely compacted organic layers;
- Drought Code (DC): represents the average moisture content of deep layers of compact organic matter;
- Initial Spread Index (ISI): the combination of the FFMC and wind speed that consists of the expected rate of fire spread;
- Buildup Index (BUI): the combination of the DMC and DC that represents a numeric rating of the total amount of fuel available for combustion;
- Fire Weather Index (FWI): the combination of the DMC and DC that consists of a numeric rating of fire intensity.

These six elements are represented in Figure 1, adapted from [12].

The CFFWIS requires meteorological parameters as input variables, i.e., air temperature and relative humidity, wind speed, and accumulated precipitation in the last 24 hours, with the element FWI as the main output, being an indicator of fire behavior and danger. Knowing that the most significant daily value of the danger index occurs up to the daily maximum, which occurs at noon in local time, the values of these meteorological parameters are obtained around that time.

The first studies considered FWI values on a scale from zero to 16, called the D-scale. However, this scale did not have enough interpretation in terms of fire intensity. Thus, new studies were developed based on the determination of new logarithmic scales to approximate fire danger rating appropri-

Tab. I: Fire danger classes with respective FWI values in Canada and the district of Coimbra, Portugal.

Danger Class	Canada [5]	Coimbra, Portugal [6]
Very Low	$FWI < 2$	-
Low	$2 \leq FWI \leq 5$	$FWI < 15$
Moderate	$5 \leq FWI \leq 9$	$15 \leq FWI \leq 22$
High	$9 \leq FWI \leq 17$	$22 \leq FWI \leq 30$
Very High	$17 \leq FWI \leq 30$	$30 \leq FWI \leq 45$
Extreme	$FWI \geq 30$	$FWI \geq 45$

ately. The current scale used is the S-scale, also logarithmic, conceptually representing the frontal fire intensity, which has a range starting at zero and then grows indefinitely,according to the severity of the fire [12].

Detailed information on the influence of each element of CFFWIS, as well as mathematical formulations, can be found in [5], [13]. Fire danger classes are generally determined by analysing FWI values as thresholds between those classes. The CFFWIS was developed for Canada with an indication of six classes: Very Low, Low, Moderate, High, Very High, and Extreme. In regions around the world, these classes can be reduced/adjusted according to individual weather characteristics. For example, in [6], analysis of the daily accumulation of occurrences (number of wildfires and burnt area) is made based on percentiles, where it was found that it was only necessary to indicate five classes (disregarding ”Very Low” class) to calibrate the CFFWIS for the districts of Portugal. Table I shows some examples of FWI values chosen to determine the fire danger classes for two different situations: from the study that resulted in the final structure of CFFWIS for Canada in [5], and from the aforementioned study adapted to Portugal [6], taking as an example the district of Coimbra.

Figure 2 shows a mapping obtained on the September 2nd, 2020, containing a distribution of Conjunctural and Meteorological Risk (RCM) in Continental Portugal, provided by the Portuguese Institute for Sea and Atmosphere (IPMA), composed by FWI (updated once a day by the IPMA) and rural fire danger index (under the responsibility of the Institute for Nature Conservation and Forests - ICNF) [14], [15].

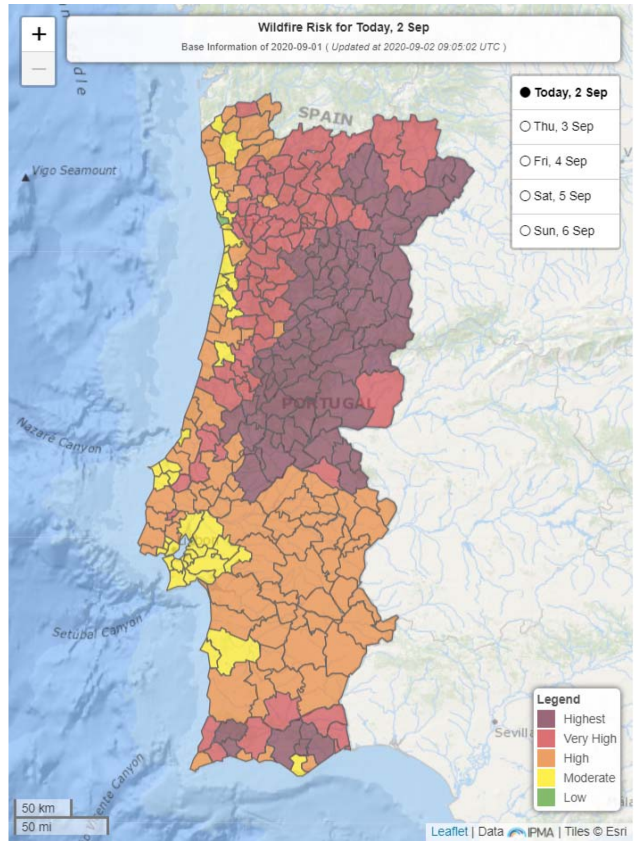


Fig. 2: Conjunctural and Meteorological Risk in Continental Portugal, mapped on September 2nd, 2020 [14].

III. Automatic Calibration of CFFWIS

In this section, the input variables and desired target are defined (Section III-A), and the clustering method, *k*-means, is adapted to the CFFWIS calibration context (Section III-B). Finally, in Section III-C, it is presented the algorithm that summarizes the step-by-step of the proposed automatic calibration of the CFFWIS, defining the fire danger classes for a given region.

A. Input Variables and Desired Target

One way to classify certain regions of countries is by using spatial units. For instance, the system used in the European Union (EU) is the Nomenclature of Territorial Units for Statistics (NUTS) classification, which is a hierarchical system for dividing up the EU's economic territory [10]. The socio-economic analyses can be performed in three levels: major regions (NUTS1), basic regions with regional policies (NUTS2), and small regions for specific diagnoses (NUTS3). The present work will be focused on the third level (NUTS3), which has the format “ABXXX”, where the first two characters represent the country code, and the rest represent a specific administrative area [16].

The choice of input variables reflects the interpretation of the FWI values of the Canadian Forest Fire Weather Index System. As FWI is a daily numerical representation of the intensity of the danger of a wildfire, and because it is composed of components that indicate favorable meteorological and soil moisture conditions for ignition, it can follow the daily amount of wildfire records (occurrences and burnt area). Also, the behavior of wildfires can vary depending on the season. Thus, the fire danger analysis can be performed considering two main seasons, the Hot Season (Summer and Autumn) and Cold Season (Winter and Spring). In this sense, the chosen input variables for the present study, for a given region and each season are:

- the daily FWI values,
- the daily number of fire occurrences,
- the daily burnt area.

The desired target of the proposed automatic calibration is to obtain thresholds from the FWI values to determine the fire danger classes for a given region. These thresholds are determined using a clustering method that is discussed below.

B. Clustering Method

Clustering techniques can be used to efficiently classify data sets based on similar characteristics, being mainly used for unsupervised learning of systems by partitioning data sets into subsets (clusters). They can be done by classifying them as belonging exclusively to these groups (hard partitioning) or through weighting belonging to these groups (soft partitioning) [17], [18]. The proposed methodology to automatically calibrate the CFFWIS is based on a hard partitioning technique, the *k*-means clustering [19].

Let a data set be composed by samples of the input variables aforementioned, given by (1). Each k sample consists of these 3 input variables grouped into a column vector $\mathbf{x}_k = [x_{1,k}, x_{2,k}, x_{3,k}]^T$. A set of $k = 1, 2, \ldots, N$ samples is given by [19]:

$$\mathbf{X} = \begin{bmatrix} x_{1,1} & x_{1,2} & \cdots & x_{1,N} \\ x_{2,1} & x_{2,2} & \cdots & x_{2,N} \\ x_{3,1} & x_{3,2} & \cdots & x_{3,N} \end{bmatrix}, \tag{1}$$

where the observed variables are the input variables presented in the previous subsection, where x_1 represents the daily values of FWI, x_2 the daily number of fire occurrences, and x_3 the daily burnt area. The sample values are counted per day, for a total of $k = 1, 2, \ldots, N$ days.

Let N_i be the number of samples of the i-th cluster ($i = 1, 2, \ldots, c$, with $c = 5$ for Hot Season and $c = 3$ for Cold Season)[1], and $\mathbf{I}^i$ be the set of sample positions that belong to the i-th cluster [20]. The distance between a data samples $\mathbf{x}_k$ and the centers $\mathbf{v}_i$ of i-th cluster is represented by a standard Euclidean norm [19]:

$$d_{i,k}^2 = (\mathbf{x}_k - \mathbf{v}_i)^T (\mathbf{x}_k - \mathbf{v}_i), \tag{2}$$

with

$$\mathbf{v}_i = \begin{cases} \dfrac{\sum_{k \in \mathbf{I}^i} \mathbf{x}_k}{N_i}, & \text{if } N_i > 0 \\ \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}, & \text{otherwise} \end{cases}, \tag{3}$$

where $\mathbf{v}_i = [v_{1,i}, v_{2,i}, v_{3,i}]^T \in \Re^{3\times1}$ has elements represented by the average FWI value, average fire occurrences and average burnt area, respectively, for each cluster, according (3).

The objective of *k*-means clustering method is to assigns each data sample to one cluster exclusively, considering the minimum distance, $d_{i,k}^2$ (2), between them. Iteratively, the distances are computed, and the centers of the clusters are updated until a stop condition be reached. A simple stop condition to be used, it is if there is no variation in the values of the centers between one iteration and another, that is:

$$\sum_{j=1}^{3} \sum_{i=1}^{c} \| v_{i,j}^{(l)} - v_{i,j}^{(l-1)} \| \approx 0, \tag{4}$$

where l and $l-1$ superscripts are, respectively, the actual and previous iteration.

After reaching the condition (4), and considering that each cluster represents a different fire danger class, the calibration of the CFFIWS for a given region is done through the indication of thresholds with the FWI values that separate these classes:

$$\boldsymbol{\theta} = \{\theta^1, \theta^2, \ldots, \theta^{c-1}\}, \tag{5}$$

being $c = 5$ for Hot Season, and $c = 3$ for Cold Season.

The proposed way to obtain a threshold is by averaging the maximum FWI value for one class and the minimum FWI value for the next class, i.e.:

$$\theta^{i|i=1,2,\ldots,c-1} = \frac{\max\left(x_{1,k|k\in\mathbf{I}^i}\right) + \min\left(x_{1,k|k\in\mathbf{I}^{i+1}}\right)}{2}. \tag{6}$$

The fire danger classes for a given region are defined based on FWI values and the thresholds calculated as follows:

- Class "1": FWI $< \theta^1$;
- Class "2": $\theta^1 <$ FWI $< \theta^2$;

  ⋮
- Class "$c-1$": $\theta^{c-2} <$ FWI $< \theta^{c-1}$;
- Class "c": FWI $> \theta^{c-1}$.

C. Proposed Calibration Methodology

Algorithm 1 presents the step-by-step automatic calibration of CFFWIS using the *k*-means clustering method presented in Section III-B. In this algorithm, there are two procedures, one for the Hot Season and one for the Cold Season.

IV. Results and Discussion

In this section, the automatic calibration of the Canadian Forest Fire Weather Index System for European regions (NUTS3) is implemented through data sets composed between the components of CFFWIS and wildfire records. Then, the proposed methodology is validated by analysing large wildfires in terms of fire danger classification, where the efficiency of the proposed calibration framework based on *k*-means clustering is presented by comparing it to another clustering method, the Fuzzy *c*-Means (FCM) [21], [22].

A. Data sets description

To determine fire danger classes, two data sets provided by the European Commission Joint Research Centre (JRC) are used: a set consisting of wildfire records (number of occurrences and burnt area), and another composed of values of the components of the Canadian Forest Fire Weather Index System, with all these data distinguished by 769 NUTS3 regions and date. The 769 NUTS3 regions presented on the data set are divided between 22 countries from Europe, between the years 2006 and 2015. The fire records set is

[1]The choice of these values comes from the experts' knowledge of the pattern of wildfire occurrences over the seasons [10].

Algorithm 1: Automatic Calibration of the Canadian Forest Fire Weather Index System for a given region.

Inputs: N data samples $\{(x_{1,k}, x_{2,k}, x_{3,k})\}_{k=1}^{N}$ being x_1 the daily FWI values, x_2 the daily number of fire occurrences, and x_3 the daily burnt area for a given region;

Outputs: Fire danger classes based on FWI values for a given region and Season;

Procedure:

1. Organize the input variables as a data set with the structure in (1);

2. Split the data set between Hot Season (Summer and Autumn seasons) and Cold Season (Winter and Spring seasons) according to the date;

for Hot Season **or** Cold Season **do**

3. Set the number of fire danger classes (clusters): if Hot Season, $c = 5$; if Cold Season, $c = 3$;

begin

4. Choose from the data set at random and distinguishable centers for all c clusters;

repeat

5. Compute the distances between data samples and all clusters using (2);

6. Obtain $\mathbf{I}^i$, which contains the data sample positions that belong to the i-th cluster by selecting minimal distances from (2);

7. Update clusters centers using (3);

until Termination condition: $\sum_{j=1}^{3}\sum_{i=1}^{c}\|v_{i,j}^{(l)} - v_{i,j}^{(l-1)}\| \approx 0$, with l representing the actual iteration and $l-1$ the previous iteration;

end

8. Calculate thresholds $\theta^{i|i=1,2,\ldots,c-1}$ between the fire danger classes using (6), order them in ascending order according (5).

end

composed of 245.509 data samples, while the CFFWIS is composed of 2.495.997 data samples. Hence, these data sets are merged using NUTS3 regions and their date as keys (common identifiers).

As the CFFWIS is related to the climatic conditions of a given region, these countries may present different characteristics, and as such the values that represent the danger classes vary. For this, it is considered to separate the days of the year between two main seasons: the "Hot Season", which comprises the period between Summer and Autumn seasons (that is, between May 15th and September 30th), and the "Cold Season", which comprises the period between Winter and Spring seasons (that is, between October 1st and May 14th).

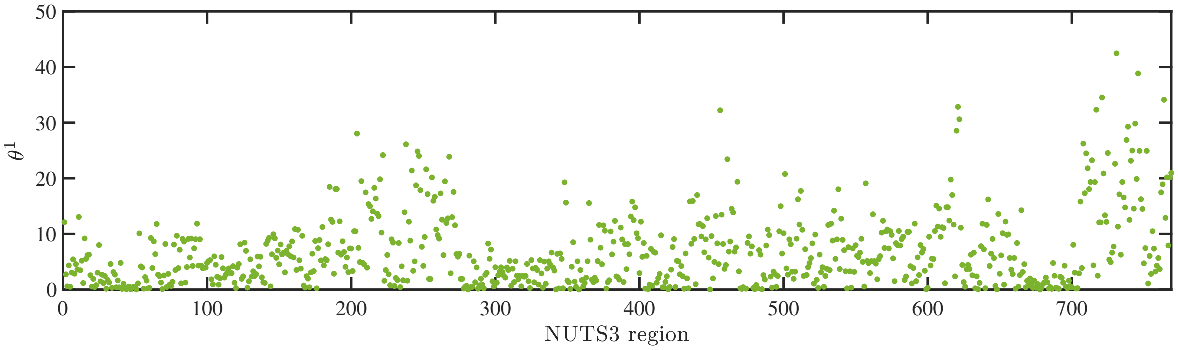
B. Results

For the present study, the danger classes defined for the Hot Season are 1) Very Low, 2) Low, 3) Moderate, 4) High, and 5) Very High, while the danger classes defined for Cold Season are 1) Low, 2) Moderate and 3) High. Also, 769 NUTS3 with data referring to the Hot Season and 769 NUTS3 with data referring to the Cold Season were selected. After implementing the proposed methodology (Algorithm 1) for each NUTS3 region, in both seasons, the thresholds were obtained as the final results of the calibration to represent the desired danger classes. Figure 3 presents the Figures 3a–3d, where each one determines the thresholds for each NUTS3 in the Hot Season between a given fire danger class and the next class. Figure 4 presents the Figures 4a–4b, where each one determines the thresholds for each NUTS3 in the Cold Season between a given fire danger class and the next class. In both figures, the 769 NUTS3 regions are distributed among the following countries using values on the x-axis:

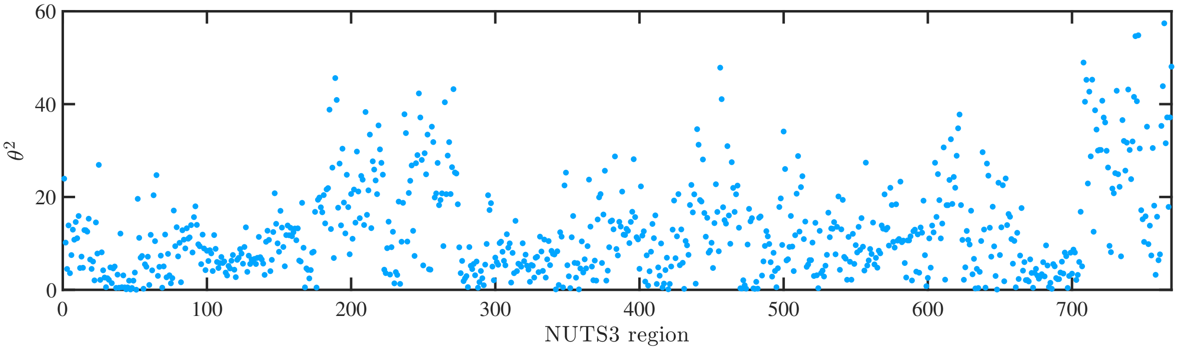
- Bulgaria (BG): from 1 to 28;
- Switzerland (CH): from 29 to 51;
- Cyprus (CY): from 52 to 52;
- Czech Republic (CZ): from 53 to 66;
- Germany (DE): from 67 to 167;
- Estonia (EE): from 168 to 172;
- Greece (EL): from 173 to 222;
- Spain (ES): from 223 to 274;
- Finland (FI): from 275 to 292;
- France (FR): from 293 to 376;
- Croatia (HR): from 377 to 383;
- Hungary (HU): from 384 to 403;
- Italy (IT): from 404 to 513;
- Lithuania (LT): from 514 to 523;
- Latvia (LV): from 524 to 529;
- Poland (PL): from 530 to 595;
- Portugal (PT): from 596 to 623;
- Romania (RO): from 624 to 665;
- Sweden (SE): from 666 to 686;
- Slovenia (SI): from 687 to 698;
- Slovakia (SK): from 699 to 706;
- Turkey (TR): from 707 to 769.

C. Validation

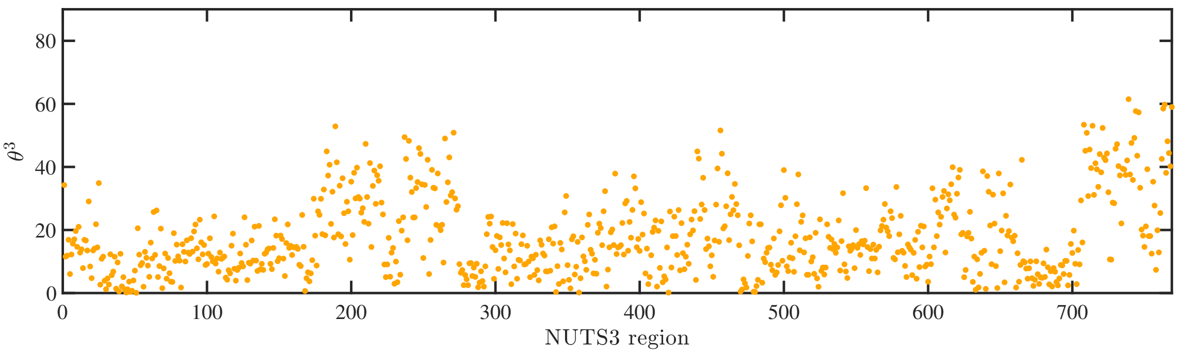
To validate the thresholds obtained in Section IV-B, two analyses are carried out: 1) the choice of samples with burnt area $B.A. \geq 1000\,[\text{ha}]$ in Hot and the choice of samples with burnt area $B.A. \geq 500\,[\text{ha}]$ in Cold Season for each NUTS3, and 2) the choice of 10% of samples that have the highest values of burnt area for each NUTS3 and each Season. These two analyses are done to observe and classify the daily FWI values corresponding to the large wildfires, where they are expected to belong to the "High" and "Very High" danger classes in the Hot Season, and to the "Moderate" and "High" danger classes, following the definition of these classes in Section IV-B. The classification of FWI values follows the distribution of danger classes based on thresholds presented at the end of Section



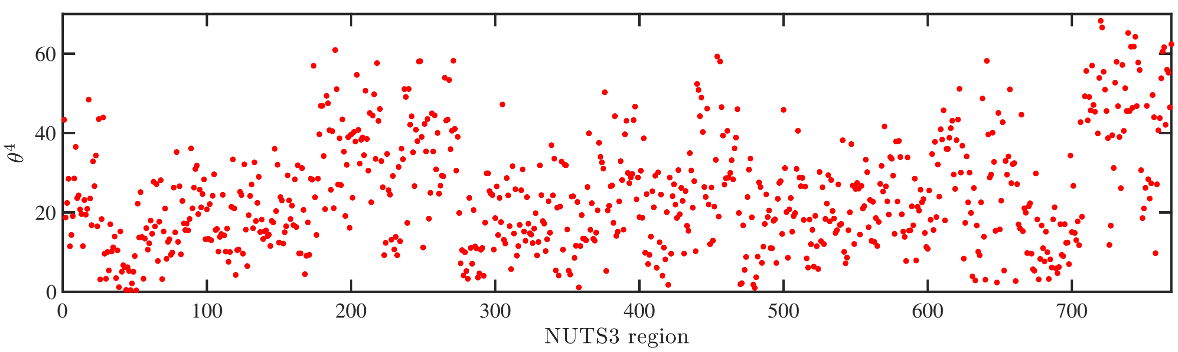
(a) FWI threshold θ^1 (6) between Very Low and Low, for all 769 NUT3 regions.



(b) FWI threshold θ^2 (6) between Low and Moderate, for all 769 NUT3 regions.

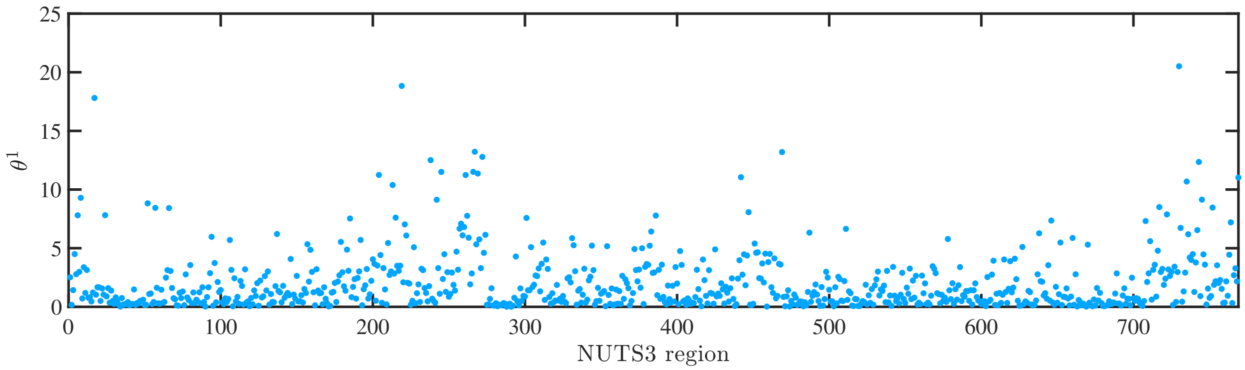


(c) FWI threshold θ^3 (6) between Moderate and High, for all 769 NUT3 regions.

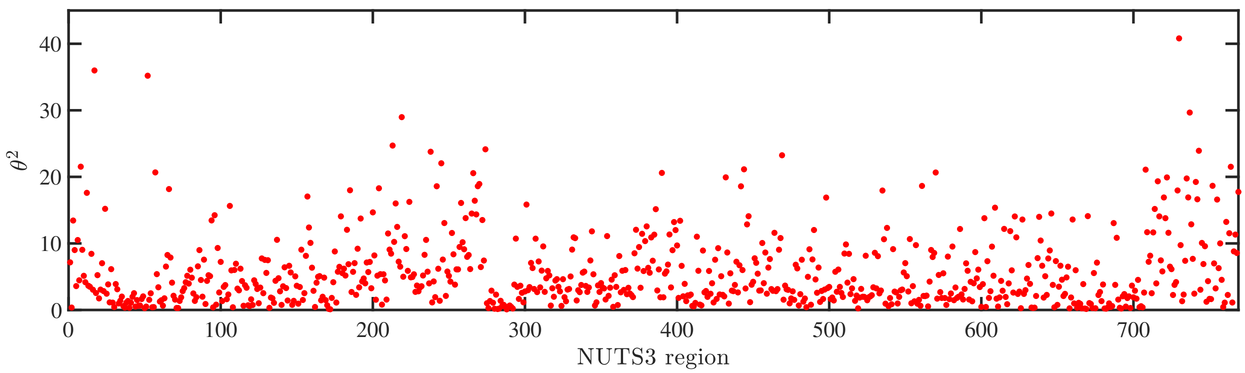


(d) FWI threshold θ^4 (6) between High and Very High, for all 769 NUT3 regions.

Fig. 3: FWI thresholds for the 769 NUT3 regions in Hot Season.



(a) FWI threshold θ^1 (6) between Low and Moderate, for all 769 NUT3 regions.



(b) FWI threshold θ^2 (6) between Moderate and High, for all 769 NUT3 regions.

Fig. 4: FWI thresholds for the 769 NUT3 regions in Cold Season.

III-B. For the first analysis, 338 samples from 119 NUTS3 in Hot Season were selected (for $B.A. \geq 1000\,[\mathrm{ha}]$), while 173 samples from 46 NUTS3 in Cold Season were selected (for $B.A. \geq 500\,[\mathrm{ha}]$). For the second analysis (10% of the highest burnt area), 14.556 samples from 752 NUTS3 in Hot Season were selected, while 9.028 samples from 734 NUTS3 in Cold Season were selected.

To check the efficiency of using k-means on the proposed automatic calibration methodology of CFFWIS, it becomes interesting to compare it with another well-known clustering method, such as fuzzy c-means (FCM), a soft partitioning technique [21], [22]. In FCM, the samples have a membership degree to all clusters instead of just one exclusively (as in k-means). However, using FCM in place of k-means for CFFWIS calibration, samples are assigned to each danger class based on the highest membership degree, and then choose thresholds between those classes.

The percentages of large wildfires in the Hot Season classified in the danger classes "High" and "Very High", as well as the percentages of large wildfires in the Cold Season classified in the danger classes "Moderate" and "High", using k-means and fuzzy c-means on the proposed automatic calibration methodology of CFFWIS, are shown in Table II.

Analysing the results in Table II for the proposed calibration using k-means, it is noticed a better performance of the analysis of burnt area with a minimum value (B.A. $\geq 1000/500$ [ha]) than choosing the 10% of samples with the highest burnt area. This fact must occur due to the possibility that within these 10% samples whose information related to them does not configure as days that large wildfires occurred, justifying a lower fire danger classification for these cases. As in the Cold Season, the number of defined danger classes is smaller, the percentages presented in the two analyses present values close to each other. Table II also shows that the proposed calibration methodology involving k-means presented better results than the same methodology using fuzzy c-means. This fact is because the case study requires a precise determination of danger classes, separated between well-defined thresholds, which matches the idea of exclusivity of a given sample belonging to a class proposed by k-means.

V. Conclusions

The methodology proposed in this work for automatic calibration of the Canadian Forest Fire Weather Index System (CFFWIS) adapted to specific regions contributed with a new perspective in defining fire danger classes. Adapting the k-means clustering method to the calibration context, it became possible to obtain important characteristics for an analysis of fire behavior using FWI values and wildfire records for various regions, distinguishing between Hot and Cold Season.

Tab. II: Percentage values of correct classification in the danger classes "High" and "Very High" in Hot Season, and "Moderate" and "High" in Cold Season, obtained through the proposed calibration methodology; and the calibration methodology with the replacement of *k*-means by fuzzy *c*-means.

	Proposed Calibration		Using fuzzy *c*-means	
Season	**B.A.** $\geq$ **1000/500 [ha]**	**Highest 10% of B.A.**	**B.A.** $\geq$ **1000/500 [ha]**	**Highest 10% of B.A.**
Hot	90.53%	74.61%	66.27%	58.72%
Cold	82.66%	85.67%	66.47%	71.37%

Also, the proposed method applied to each of the 769 NUTS3 showed positive results, which can be seen when classifying large wildfires. The validation step, through two different analyses of burnt area, showed the superiority of the proposed methodology using *k*-means in comparison with the fuzzy *c*-means applied for the same purpose. Future work may explore other clustering techniques to be applied in the CFFWIS calibration, in addition to new analyses that may contribute to promoting greater efficiency in the determination of fire danger indices.

Acknowledgements

This work was carried out under the project IMFire - Intelligent Management of Wildfires, ref. PCIF/SSI/0151/2018, and was fully funded by national funds through the Ministry of Science, Technology and Higher Education. Besides, the authors acknowledge the European Commission Joint Research Centre for granting authorization to use its data sets for the present study.

References

[1] J. San-Miguel-Ayanz, E. Schulte, G. Schmuck, A. Camia, P. Strobl, G. Liberta, C. Giovando, R. Boca, F. Sedano, P. Kempeneers, D. McInerney, C. Withmore, S. S. de Oliveira, M. Rodrigues, T. Durrant, P. Corti, F. Oehler, L. Vilar, and G. Amatulli, "Comprehensive Monitoring of Wildfires in Europe: The European Forest Fire Information System (EFFIS)," in *Approaches to managing disaster-Assessing hazards, emergencies and disaster impacts*," IntechOpen, March 2012.

[2] J. San-Miguel-Ayanz, T. Durrant, R. Boca, G. Libertà, A. Branco, D. de Rigo, D. Ferrari, P. Maianti, T. A. Vivancos, D. Oom, H. Pfeiffer, D. Nuijten, and T. Leray, "Forest Fires in Europe, Middle East and North Africa 2018," Joint Research Centre, Tech. Rep., October 2019.

[3] J. E. Deeming, R. E. Burgan, and J. D. Cohen, *The National Fire-Danger Rating System – 1978*, Intermountain Forest and Range Experiment Station, Forest Service, US, vol. 39, October 1977.

[4] A. G. McArthur, "The preparation and use of fire danger tables," in *Proceedings of tre Fire Weather Conference, Commonwealth of Australian Bureau of Meteorology*, July 1958.

[5] C. E. Van Wagner, "Development and Structure of the Canadian Forest Fire Weather Index System," *Canadian Forestry Service, Forestry Technical Report*, 1987.

[6] D. X. Viegas, R. M. Reis, M. G. Cruz, and M. T. Viegas, "Calibração do Sistema Canadiano de Perigo de Incêndio para Aplicação em Portugal," *Silva Lusitana*, vol. 12, no. 1, pp. 77–93, June 2004.

[7] C. C. DaCamara, T. J. Calado, S. L. Ermida, I. F. Trigo, M. Amraoui, and K. F. Turkman, "Calibration of the Fire Weather Index over Mediterranean Europe based on fire activity retrieved from MSG satellite imagery," *International Journal of Wildland Fire*, vol. 23, no. 7, pp. 945–958, August 2014.

[8] S. Chelli, P. Maponi, G. Campetella, P. Monteverde, M. Foglia, E. Paris, A. Lolis, and T. Panagopoulos, "Adaptation of the Canadian fire weather index to Mediterranean forests," *Natural Hazards*, vol. 75, no. 2, pp. 1795–1810, January 2015.

[9] M. C. de Jong, M. J. Wooster, K. Kitchen, C. Manley, R. Gazzard, and F. F. McCall, "Calibration and evaluation of the Canadian Forest Fire Weather Index (FWI) System for improved wildland fire danger rating in the United Kingdom," *Natural Hazards and Earth System Sciences*, vol. 16, no. 5, pp. 1217–1237, May 2016.

[10] D. Alves, L. M. Ribeiro, and D. X. Viegas, "Calibration of the Canadian FWI system for the territory of Europe," in *Advances in Forest Fire Research*," Imprensa da Universidade de Coimbra, pp. 33–43, November 2018.

[11] B. J. Stocks, B. D. Lawson, M. E. Alexander, C. E. Van Wagner, R. S. McAlpine, T. J. Lynham, and D. E. Dubé, "Canadian forest fire danger rating system: an overview," *The Forestry Chronicle*, vol. 65, no. 4, pp. 258–265, August 1989.

[12] W. J. de Groot, "Interpreting the Canadian Forest Fire Weather Index (FWI) System," in *Proc. of the Fourth Central Region Fire Weather Committee Scientific and Technical Seminar*, pp. 3–14, April 1987.

[13] C. E. Van Wagner and T. L. Pickett, "Equations and FORTRAN program for the Canadian forest fire weather index system," *Canadian Forestry Service, Forestry Technical Report*, 1985.

[14] IPMA. (2020) Fire risk Mainland. [Online]. Available: https://bit.ly/30Cxr1U. Accessed: September 2nd, 2020.

[15] ——. (2020) Fire Risk - Forest Fire Risk Index - RCM. [Online]. Available: https://bit.ly/2GqxwPg. Accessed: September 2nd, 2020.

[16] European Parliament, "Regulation (EC) No 1059/2003 of the European Parliament and of the Council of 26 May 2003 on the establishment of a common classification of territorial units for statistics (NUTS)," Official Journal of the European Union, June 2018.

[17] M. Kearns, Y. Mansour, and A. Y. Ng, "An Information-Theoretic Analysis of Hard and Soft Assignment Methods for Clustering," in *Learning in graphical models*," Springer, pp. 495–520, 1998.

[18] J. Mendes, S. Pinto, R. Araújo, and F. Souza, "Evolutionary Fuzzy Models for Nonlinear Identification," in *Proc. 17th IEEE International Conference on Emerging Technologies and Factory Automation (ETFA 2012)*, pp. 1–8, September 2012.

[19] J. MacQueen, "Some methods for classification and analysis of multivariate observations," in *Proceedings of the 5th Berkeley Symposium on Mathematical Statistics and Probability*, vol. 1, no. 14, pp. 281–297, June–July 1967.

[20] K. S. Al-Sultana and M. M. Khan, "Computational experience on four algorithms for the hard clustering problem," *Pattern recognition letters*, vol. 17, no. 3, pp. 295–308, March 1996.

[21] J. C. Bezdek, R. Ehrlich, and W. Full, "FCM: The fuzzy c-means clustering algorithm," *Computers & Geosciences*, vol. 10, no. 2–3, pp. 191–203, 1984.

[22] J. Mendes, R. Araújo, T. Matias, R. Seco, and C. Belchior, "Evolutionary Learning of a Fuzzy Controller for Industrial Processes," in *Proc. of the The 40th Annual Conference of the IEEE Industrial Electronics Society (IECON 2014)*, pp. 139–145, October-November 2014.