Complementing daily fire-danger assessment using

a novel metric based on burnt area ranking

Gianni Boris Pezzatti\textsuperscript{a*}, Antonella De Angelis\textsuperscript{a}, Ismail Bekar\textsuperscript{b}, Carlo Ricotta\textsuperscript{c}, Sofia Bajocco\textsuperscript{d}, Marco Conedera\textsuperscript{a}.

\textsuperscript{a} WSL Swiss Federal Research Institute, Insubric Ecosystems Research Group, A Ramél 18, 6593 Cadenazzo, Switzerland

\textsuperscript{b} ETH Zurich, Forest Ecology, Universitätstrasse 16, 8092 Zürich, Switzerland

\textsuperscript{c} Department of Environmental Biology, University of Rome “La Sapienza”, Piazzale Aldo Moro 5, 00185 Rome, Italy

\textsuperscript{d} Consiglio per la ricerca in agricoltura e l’analisi dell’economia agraria - Council for Agricultural Research and Economics - Research Unit for Agricultural Climatology and Meteorology (CREA-CMA), Via del Caravita 7a, 00186 Rome, Italy

*Corresponding author: Gianni Boris Pezzatti. E-mail: boris.pezzatti@wsl.ch

This document is the accepted manuscript version of the following article:

This manuscript version is made available under the CC-BY-NC-ND 4.0 license http://creativecommons.org/licenses/by-nc-nd/4.0/
Abstract

Understanding the relationship between fire and weather has important implications for fire danger evaluations, firefighting and fire management. Fire weather indices are mathematical representations of a suite of meteorological variables that are often used as decision-support tools for evaluating the likelihood of fire on a given day. However, these indices are seldom evaluated for their ability to express a probability in terms of final burnt area, which is a very important fire danger component. We propose a new approach for selecting fire weather indices that represent good proxies for both the probability of fire occurrence on a particular day and of the related final burnt area. The novel performance metric was applied to historical data (11 to 31 years) from four European regions with different fire regimes in Switzerland (Canton Ticino) and Italy (Cilento in Campania, Chilivani and Campidano in Sardinia). The results confirm the suitability of the approach for selecting an appropriate fire weather index for a particular region and for providing insight into the regional meteorological control patterns of fire ignition and spread. For three of the six fire regime types analyzed (Campidano, Cilento and winter fire season in Ticino), the prediction power for burnt area of most indices was generally in accordance with estimates of fire ignition risk. For the remaining case studies, fire ignition and burnt area appear to be controlled by different meteorological or non-climatic drivers. In these cases, the selection of the most suitable index should be based on a cutoff that optimizes the two selection criteria according to local needs. From an operational point of view, this novel approach that includes the burnt area aspect can strongly support decisions on firefighting alert and preparedness as well as requested firefighting strategies.

Keywords: forest fire, burnt area, fire weather index, fire danger rating system, performance metric.
1. Introduction

Regional assessments of fire danger consider the probability of ignition and fire spread, as well as the difficulty of fire control (Vasconcelos et al., 2001; Chuvieco et al., 2003; Finney, 2005, Vasilakos et al., 2007; Moreira et al., 2011). Regardless of the ecosystem, fire ignition and spread depend on climatic factors, fuel type and distribution, and human activities (Johnson, 1992; Loepfe et al., 2011). However, at the regional scale, meteorology determines most short-term fire danger variability, and many rating systems are consequently based on fire weather indices. These indices are mathematical representations that estimate aspects of fuel flammability and related ignition probabilities or fire behavior based on factors such as air temperature, humidity and precipitation (Taylor and Alexander, 2006). Such indices also differ considerably in their complexity, and may include a memory/cumulative component or the effect of wind on (initial) fire spread (Stocks et al., 1989; Tanskanen et al., 2008). Examples include the Initial Spread Index of the Canadian Fire Weather Index System (Van Wagner, 1987), the Spread Component of the US National Fire Danger Rating System NFDRS (Deeming et al., 1977), and the McArthur Forest Fire Danger Index, which is used together with fuel load to predict the rate of spread in Australia (McArthur, 1967). This is understandable given the critical role fire spread rates play for on-the-ground firefighting tactics (Gill and Allan, 2008; Williams, 2013), and its causal relationship to the final burnt area and the related ecological and economic impacts (Conedera et al., 2003; Gill and Allan, 2008; Yates et al., 2008; Radeloff et al., 2005; Daniel et al., 2007; Reinhardt et al., 2008).

Given that the dynamics of fuel types and human activities at the regional level may affect and partially mask the role of meteorological factors, even the most complex and physically-based fire weather indices may not always be the best performing in relation to specific and local fire
activity. For instance, the type, amount, and connectivity of burnable vegetation are only
indirectly determined by climate (Vazquez and Moreno, 1993; Pyne et al., 1996; Bedía et al.,
2012; Pausas and Paula, 2012) and can be strongly altered by human land use (Guyette et al.,
2002). Humans also directly impact fire activity both by igniting or suppressing fires (Guyette et
al., 2002), and even though the role of humans is often not explicitly considered in fire danger
rating, it can be as prominent as that of weather in many fire regimes. Consequently, empirical
modeling approaches that relate fire history to a combination of fire indices, or even raw
meteorological variables, can result in increased predictive power when modeling fire
occurrence probabilities (De Angelis et al., 2015).

Since indices with greater forecast skill may allow for more efficient preparedness and allocation
of firefighting resources and more appropriate public warning systems, much effort has been
recently put into finding the most appropriate fire danger indices for selected regions (e.g.,
Matthews, 2009; Eastaugh et al., 2012; Arpaci et al., 2013; Sirca et al., 2018). Different methods
have been used to compare the fire occurrence forecasting power of single existing fire danger
indices (Viegas et al., 1999; Andrews et al., 2003), either by analyzing the indices raw values
with parametric (e.g., Mahalanobis distance) or non-parametric measures, by evaluating their
scores according to thresholds that define a predicted fire day or by means of threshold
independent rank metrics(e.g. area under the receiver operating characteristic – ROC- curve). In
contrast, less attention has been put into determining indices’ performance in relation to burnt
area on a daily basis, which may not be strongly related to raw index values. Several studies have
analyzed the relationship between fire weather indices or drought indices and burnt area on
monthly or longer time scales (e.g., Wotton 2009; Camia and Amatulli, 2009), but few have
attempted a daily or weekly evaluation, mostly by comparing empirical cumulative distribution
functions (ECDF) or by analyzing index percentiles (e.g., Riley et al., 2013; Freeborn et al.,
This is understandable also because of the lack of information on the precise daily burnt area in case of long lasting fire or on the precise ignition day in case of prolonged hold-over times.

A better understanding of the relationship between fire weather indices (or drought indices) and burnt area could help fire management services to identify large-fire days in addition to the probability of ignition. This would be an efficient way of obtaining information on large-fire days, since many such daily indices are already implemented at national and continental scales (Van Wagner, 1987; Camia et al., 2006).

In this study, we present a simple performance metric that allows to take the burnt area into consideration when evaluating the suitability of fire-danger indices to estimate daily fire danger. In particular, the proposed metric uses the potential final burnt area criterion when evaluating the performance of daily danger rating systems for a specific region. Relating the final burnt area of a fire to the fire weather at the recorded start date may not be the best suited approach for long lasting fires. However, this is the only approach applicable to most forest fire databases (Duff et al., 2016) and in most cases the initial fire weather conditions and the related fire spread is the most relevant aspect in terms of suppression difficulty and related final burnt area (Pyne et al., 1996). We thus used this procedure to complement the assessment of selected fire weather indices. We in particular tested the relationship of each single index calculated using a reference meteorological station to the final burnt area. We tested the proposed approach using a wide range of case studies representing different European fire regimes.
2. Material and methods

2.1 Study cases

In order to test the suitability of the proposed approach, we applied it in four study areas with different fire regimes: the low to intermediate fire-prone Canton Ticino (Switzerland) representing the Alpine region, and three highly fire-prone areas in Italy with different land use and human behavior patterns: (1) the coastal area of Cilento (Cilento, Vallo di Diano and Alburni National Park) in Campania; (2) the Chilivani and (3) Campidano regions of Sardinia (Fig. 1).

The low to intermediate fire-prone Canton Ticino covers an area of ~2,812 km², with forests covering ~51% and agricultural lands ~3%. The mountainous region is characterized by an ‘Insubrian’ climate, with warm, wet summers and mild, dry winters. Most of the rainfall is concentrated in June to September, but periods without rain alternate with short intervals of heavy convective precipitation. The winter season is influenced by episodes of ‘Foehn’ wind from the north, during which relative air humidity drops to < 20%. Due to its particular climate and steep topography, Ticino features three fire regimes in terms of ignition source and resulting fire characteristics (Conedera et al., 2011; Pezzatti et al., 2013). Rapidly spreading, human-caused (surface) fires are typical for December to April (non-vegetative season, hereafter called winter; Ticino-w), whereas slowly spreading fires dominate during the summer (vegetative season, May to November). Summer fire regimes are divided into two groups according to the ignition source: lightning-induced fires that are concentrated in the period from June to September mostly in coniferous forests on steep slopes and at high elevations (Conedera et al., 2006; hereafter called summer natural; Ticino-sn), and human-caused fires, mostly occurring at
low elevations and less steep slopes where most human activities take place (hereafter called summer anthropogenic; Ticino-sa).

**Fig. 1.** Study areas: Canton Ticino (Switzerland), Cilento, Chilivani, and Campidano (Italy) with their respective meteorological stations: (•) Locarno/Monti, Capo Palinuro, Chilivani, and Decimomannu.
### Table 1. Regional fire statistics for the four study areas used in the analysis.

<table>
<thead>
<tr>
<th>Region</th>
<th>Time interval *</th>
<th>Number of fires</th>
<th>Burnt Area [ha]</th>
<th>( \bar{\lambda}_F ) mean fire size [ha]</th>
<th>Number of fires</th>
<th>Burnt Area [ha]</th>
<th>( \bar{\lambda}_{LF} ) mean fire size [ha]</th>
<th>( \bar{\lambda}_F )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ticino-w</td>
<td>1992-2013 (22 y)</td>
<td>546 (24.8 y(^{-1}))</td>
<td>3617.1 (164.4 y(^{-1}))</td>
<td>6.6</td>
<td>13 (2.4%)</td>
<td>2469.5 (68.3%)</td>
<td>190.0</td>
<td>28.7</td>
</tr>
<tr>
<td>Ticino-sa</td>
<td>1992-2013 (22 y)</td>
<td>168 (7.6 y(^{-1}))</td>
<td>164.8 (7.5 y(^{-1}))</td>
<td>1.0</td>
<td>1 (0.6%)</td>
<td>62.0 (37.6%)</td>
<td>62.0</td>
<td>63.2</td>
</tr>
<tr>
<td>Ticino-sn</td>
<td>1982-2013 (32 y)</td>
<td>192 (6.0 y(^{-1}))</td>
<td>428.4 (13.4 y(^{-1}))</td>
<td>2.2</td>
<td>2 (1%)</td>
<td>230.0 (53.7%)</td>
<td>115.0</td>
<td>51.5</td>
</tr>
<tr>
<td>Cilento</td>
<td>2001-2013 (12 y)</td>
<td>1297 (108.1 y(^{-1}))</td>
<td>8429.1 (702.4 y(^{-1}))</td>
<td>6.5</td>
<td>30 (2.3%)</td>
<td>3790.0 (45%)</td>
<td>126.3</td>
<td>19.4</td>
</tr>
<tr>
<td>Chilivani</td>
<td>1996-2013 (16 y)</td>
<td>1871 (116.9 y(^{-1}))</td>
<td>12529.4 (801.8 y(^{-1}))</td>
<td>6.7</td>
<td>33 (1.8%)</td>
<td>9021.8 (72%)</td>
<td>273.4</td>
<td>40.8</td>
</tr>
<tr>
<td>Campidano</td>
<td>1996-2013 (18 y)</td>
<td>6852 (380.7 y(^{-1}))</td>
<td>37376.0 (2076.4 y(^{-1}))</td>
<td>5.5</td>
<td>136 (2%)</td>
<td>17631.5 (47.2%)</td>
<td>129.6</td>
<td>23.8</td>
</tr>
</tbody>
</table>

*some years were excluded from the dataset due to the lack of meteorological data (i.e., 2002 for Cilento; 2002 and 2008 for Chilivani)

These three regimes are characterized by different fire statistics, with most fires and burnt area occurring in the non-vegetative season for Ticino-w (Table 1). Fires larger than 50 ha account for 2.4% of all events and 68.3% of the total burnt area in Ticino-w, whereas the proportion of large summer fires is generally lower (0.6% of all events and 37.6% of total burnt area for Ticino-sa, and 1% of all events and 53.7% of the burnt area for Ticino-sn, Table 1).

The Cilento area covers about 1,810 km\(^2\) and consists of two main sub-areas. The Tyrrenian coastal portion has a Mediterranean climate, while the interior mountain area is characterized by a more temperate climate. In this study, we consider only the coastal sub-area (509 km\(^2\)) where
forests cover ~62% of the total area and heterogeneous agricultural lands 16%. The fire regime is typically Mediterranean, with a high number of anthropogenic fires from July to September. August is the peak month with respect to the number of fires, the number of days with fires, and burnt area (Fig. 2). In Cilento, fires larger than 50 ha represent only 45% of the burnt area (Table 1). To ensure consistency with the other Mediterranean regions, we considered the period from May to October as a single summer-fire season (hereafter referred to as Cilento).

**Fig. 2.** Monthly distribution of the number of days with fires, the number of fires, and burnt area per year for Canton Ticino (Switzerland), Cilento, Chilivani, and Campidano (Italy). The periods considered are 1992-2013 for Ticino-w (December to April) and Ticino-sa (May to November, human-induced fires); 1982-2013 for Ticino-sn (May to November, lightning-induced fires);
2001-2013 for Cilento; 1996-2013 for Campidano and Chilivani. Some years were excluded from the dataset due to the lack of meteorological data (2002 for Cilento; 2002 and 2008 for Chilivani).

The island of Sardinia has an area of about 24,000 km² and is characterized by a complex hilly topography and a Mediterranean climate. Similarly to Cilento, the fire regime in Sardinia is typically Mediterranean with mainly summer anthropogenic fires. We selected two sub-areas with strong differences in terms of fire regimes: the Chilivani region in the north and the Campidano region in the south. The Chilivani region (hereafter Chilivani) has a higher mean elevation (405 m asl), and the vegetation consists of shrublands (32% of the area), heterogeneous agricultural lands (27%), and arable lands (24%). The main fire season lasts from May to October, with few events in May, and, unlike Cilento, the peak fire month is July (Fig. 2). Larger fires are few in number (only 1.8%) but account for much of the burnt area (72%, Table 1). The Campidano region (hereafter Campidano) is a homogeneous plain at a mean elevation of 134 m asl in southern Sardinia, covered by arable lands (53%) and forests (16%). The main fire season extends from May to October and has a higher number of fires, days with fires, and greater total burnt area than the other regions. As in Chilivani, although the peak fire season in terms of days with fires and burnt area is July, there is a slightly higher number of fires in June (Fig. 2). Fires larger than 50 ha represent 2% of the total number of events in the summer season and cover only 47.2% of the burnt area (Table 1).
2.2. Forest fires and meteorological data

We extracted fire data from existing databases, and selected a period of analysis for each study area that experienced consistent fire management and policy (i.e., periods where firefighting decisions were guided by the same legal framework and tactical approaches, and land use and related vegetation cover were largely unchanged). For Ticino, we extracted data from the Swissfire database (Pezzatti et al., 2010). In order to improve uniformity in the database, we selected data from 1982 to 2013 for Ticino-sn (summer natural regime) beginning with the last significant change in firefighting organization and strategies (Conedera et al., 2011). For the two anthropogenic fire regimes (Ticino-sa and Ticino-w) we set the period of analysis to 1992-2013 due to the legal fire prevention acts issued between 1987 and 1991 that caused major changes in the handling of human fire-related activities (see Conedera et al., 2011 and Pezzatti et al., 2013 for more details). The periods of available data for the Italian sites were shorter, and were thus considered to be quite homogeneous in terms of fire regime. For Cilento, we referred to the fire database of the National Forest Service for the period 2001-2013, and for both study areas in Sardinia, we used the Regional Forest Service fire database for the period 1996-2013, extracting available data for May to October.

Meteorological data were collected from a representative meteorological station in each study area: Locarno/Monti (MeteoSwiss) for Ticino, and Capo Palinuro (UGM - Ufficio Generale per la Meteorologia dell’Aeronautica Militare) for Cilento. For Chilivani and Campidano, we used data from the Chilivani (RAN - Rete Agrometeorologica Nazionale) and Decimomannu (UGM) meteorological stations, respectively. Due to the lack of meteorological data in 2002 (for Cilento and Chilivani) and 2008 (for Chilivani), these years were removed from the analysis.
We considered widely used fire weather indices that combine weather variables to simulate fuel properties and related fire potential (Andrews et al., 2003). Selected indices have already been in use for the Mediterranean region for both scientific (Reineking et al., 2010; Padilla and Garcia, 2011) and operational purposes (Camia et al., 2006). The indices in question range from simple to complex, differing as a function of the calculation algorithms and the meteorological parameters taken into consideration. For example, the Angstroem Index (Chandler et al., 1983) and the Fuel Moisture Index (FMI; Sharples et al., 2009a) use daily meteorological characteristics related to air temperature and relative humidity only. The Nesterov Index (Nesterov, 1949) calculates fuel moisture based on a cumulative component that keeps track of the conditions of previous days. The Baumgartner Index (Baumgartner et al., 1967) evaluates fuel dryness based on meteorological parameters from previous days. The Keetch-Byram Drought Index (KBDSI; Keetch and Byram, 1968) has a strong memory component, making it suitable for detecting long-term drought effects on fuel. The McArthur Forest Fire Danger Index (FFDI; McArthur, 1967) includes the influence of weather conditions on fire behavior, incorporating both cumulative fuel dryness and wind speed in its calculation. The Sharples et al. (2009b) forest fire danger rating index combines the Fuel Moisture Index with wind speed (i.e., an indicator of the rate of spread). The principle that fire danger may mainly be the result of fuel moisture and wind speed also underpins the Fosberg Fire Weather Index (FFWI; Fosberg, 1978), which is based on equilibrium moisture content (i.e., fuel moisture in equilibrium with the environment relative to air temperature and humidity) and wind speed. We also used the Canadian Fire Weather Index (FWI) and its sub-components (Van Wagner, 1987) that keep track of conditions on previous days. Of the six components in the FWI system, the first three describe
different fuel moisture levels, with wind speed included (e.g., Fine Fuel Moisture Code – FFMC), or excluded (Drought Code – DC; and the Duff Moisture Code – DMC). The remaining fire-behavior components result from the combination of the first three and estimate fuel availability for combustion (Buildup Index – BUI, as the result of combining DMC and DC), and the rate of spread (Initial Spread Index – ISI, resulting from the combination of FFMC and wind speed).

These indices can be grouped into four types based on the inclusion/exclusion of the wind parameter in the index calculation (wind-included vs. wind-excluded indices), and the presence of a memory component (cumulative vs. non-cumulative indices; see Table 2).

**Table 2.** Classification of the fire weather indices used in this study. Non-cumulative indices are based only on daily observations, while cumulative indices keep track of previous-day conditions. Wind inclusion/exclusion refers to whether wind speed is included in the calculation of the index.

<table>
<thead>
<tr>
<th>Wind included</th>
<th>Wind excluded</th>
</tr>
</thead>
<tbody>
<tr>
<td>Angstroem</td>
<td>Angstroem</td>
</tr>
<tr>
<td>FMI</td>
<td>FMI</td>
</tr>
<tr>
<td>Sharples</td>
<td>Sharples</td>
</tr>
<tr>
<td>FFWI</td>
<td>FFWI</td>
</tr>
<tr>
<td>Nesterov</td>
<td>Nesterov</td>
</tr>
<tr>
<td>KBDISI</td>
<td>KBDISI</td>
</tr>
<tr>
<td>DMC, DC, BUI</td>
<td>DMC, DC, BUI</td>
</tr>
<tr>
<td>FWI, ISI, FFMC</td>
<td>FWI, ISI, FFMC</td>
</tr>
<tr>
<td>FFDI</td>
<td>FFDI</td>
</tr>
<tr>
<td>Baumgartner</td>
<td>Baumgartner</td>
</tr>
</tbody>
</table>
All fire weather indices used in this study were calculated on a daily basis (taking measurements at 12:00 for variables that required noon or early afternoon values) by means of the *Fire Weather Danger Indices Calculator* software ([https://github.com/Insubric/fire-calculator](https://github.com/Insubric/fire-calculator)) developed by WSL (Swiss Federal Institute for Forest, Snow and Landscape Research). The indices were calculated for a regional reference meteorological station that we selected as representative for the corresponding study area.

2.4 *Proposed novel performance metric*

To complement existing assessments metrics of fire weather indices on fire occurrence, we here propose a performance metric to additionally evaluate the ability of these indices to assess the probability of a fire becoming large. The proposed metric is threshold-independent and summarizes the ability of a given fire weather index to correctly rank burnt areas along the full set of fire days. Furthermore, it has been conceived in such a way that it can be applied to most fire datasets, which commonly lack details on day-to-day fire size growth over time. Despite the fact that a “snapshot” of the fire weather at time of the fire official record is often the only applicable compromise, burnt area is usually highly dependent form a possible failure of the initial attack by the fire brigade (Payne et al., 1993). Although the day-to-day fire spread conditions would be a required information when assessing the performance of fire spread models (Duff et al., 2016), we can assume that for certain contexts (e.g. anthropogenic fires in densely populated area like central Europe) the starting conditions are crucial in determining the success of the fire-fighting operations and the related final burnt area.

In detail, the calculation of the proposed performance metric requires following steps:
i) Define the fire days as the days in which at least an ignition has been officially recorded in the study region and for the period/regime considered.

ii) Calculate the daily total burnt area as the sum of the burnt areas (final sizes) of the fires first recorded on that particular day (hereafter called [final] burnt area);

iii) If needed, harmonize the index metrics by changing the sign of fire weather indices for which higher fire danger correspond to lower index values (e.g., the Angstroem Index).

iv) Order the fire days along decreasing values of the concerned fire weather index;

v) Build the burnt area curve by drawing the cumulative curve of the final burnt area (y-axis) of the ordered fire days against the proportion of fire days under consideration (x-axis). The area under this curve is therefore larger for fire weather indices that associate days with large final burnt area with high index values. This curve allows also to easily visualize the burnt area on either side of a fire danger threshold. In the example in Figure 3, we see that approximately 3% of the fire days account for approximately 30% of the burnt area, and the remaining 97% of the fire days account for 70% of the burnt area.

vi) Draw the worst and best score curves: the best possible score is determined by calculating the cumulative curve of the final burnt area after ordering the fire days along the x-axis with decreasing values of final burnt area, while the worst possible score is calculated by ordering the fire days along the x-axis with increasing values of final burnt area. In other words, the best score is reached if the fire days are ranked in the same order for both index values and final burnt area, while they are ranked in reverse order for the worst case;

vii) Calculate the two areas between the cumulative curve of the final burnt area and the worst- (A) and best-score cases (B) (Fig. 3):

\[ A = (\text{area under the cumulative curve of the final burnt area}) - (\text{area under the curve of} \]
the worst score case)

\[ B = (\text{area under the curve of the best score case}) - (\text{area under the cumulative curve of the final burnt area}) \]

viii) Rescale the metric in order to obtain a performance score between 0 and 1 that can be compared across different fire regimes. The area under the cumulative curve is rescaled according to the theoretically worst and best possible score cases of final burnt area. As shown in Fig. 3, \( A \) indicates the gain in the cumulative curve relative to the worst possible final burnt area ranking, whereas \( B \) indicates the loss relative to the theoretically best possible final burnt area ranking. We propose here calling the rescaled score value Cumulative Area Ranking Efficiency (CARE), which is obtained as

\[ \text{CARE} = \frac{A}{A+B}. \]

This new performance metric returns higher values for fire weather indices that associate days with large final burnt area with high index values: a score of 0.5 represents a random prediction, 0 the worst, and 1 the best.
Fig. 3. Example of the proposed novel performance metric based on burnt area.

The y-axis represents the proportion of cumulative final burnt area. The x-axis represents the proportion of fire days (days with at least one ignition) ranked according to decreasing fire ignition probability (danger index values). The solid line represents the cumulative burnt area. The dark grey area A represents the area delimited by the cumulative burnt area curve and the worst possible area rankings (lower dashed lines), that is, the cumulative burnt area when burnt area increases along the x-axis but fire probability decreases. The light grey area B is the area between the cumulative burnt area curves and the best possible area rankings (upper dashed lines), that is, the cumulative burnt area when both burnt area and fire probability decrease along the x-axis. The performance metric is defined as A/(A+B), thus ranging between 0 and 1 with a value of 0.5 corresponding to a random prediction (dotted line).
In order to handle the usually skewed nature of burnt area distributions, two transformations are additionally proposed. In the first option, a log-transformation is applied to the final burnt area data. The second option entails substituting final burnt areas with their ranks before constructing the cumulative curves. Based on these transformations, we obtain the Cumulative Logarithmic Area Ranking Efficiency (CLARE) and the Cumulative Rank-sum Area Ranking Efficiency (CRARE), respectively (Fig. 4). These alternative model performance metrics smooth the effect of fires that become large for non-weather reasons, such as obvious tactical errors in firefighting. Such uncommon or accidentally large fires potentially also occur on non-extreme fire days (medium to low index values), which may cause an undesired abrupt alteration in the shape of the raw cumulative curve.

Fig. 4. Example of applying different transformations to the final burnt areas when constructing the final performance metric. The y-axis represents in turn the untransformed proportion of cumulative final burnt area by CARE, its log-transformation by CLARE, and its rank-transformation by CRARE, respectively. For a detailed description of the lines and letters see caption of Figure 3.
2.5 Index performance evaluation

The overall performance of the fire weather indices was evaluated by analyzing both aspects, the ability to discriminate fire days from days without fires (binary response), and the ability to assess the probability of a fire becoming large. Several metrics exist to assess the goodness of binary classifiers, either depending on precise or optimized thresholds (measures based on the confusion matrix, like Accuracy or Cohen’s Kappa) or integrating across a wide range of specificities (e.g. area under the receiver operating characteristic – ROC- curve or the area under the precision-recall curve). Among those metrics we applied the widely used AUC for analyzing our study cases, that is, the Area Under the ROC (Receiver Operating Characteristic) Curve (Provost and Fawcett, 1998). It quantifies the goodness-of-fit of a classifier by plotting the true positive rates (in our case the fraction of correctly classified fire days) against the false positive rates (the fraction of misclassified no-fire days) for all threshold values. Results range between 0 and 1, where 1 represents perfect discrimination (i.e. correctly identifying a fire-start day in our case), while the 0.5 value corresponds to a random prediction. For specific cases like highly imbalanced dataset other metrics than AUC could have been preferred (e.g. Saito et al., 2015).

We then calculated all newly proposed cumulative area ranking efficiency metrics (CARE, CLARE, CRARE), and reported the results for CLARE in the manuscript and the others in the appendices.

In order to obtain a robust estimate of the index performance metrics, we implemented a k-fold cross-validation by partitioning the original data into k yearly sub-samples. A test fold
considered a combination of two years, and the remaining k-2 years were used for training. We repeated this process for all combinations of two years, such that each fold was in turn used for validation. Finally, all obtained results were averaged.

The significance of the performance differences among the fire weather indices was evaluated using the non-parametric paired Wilcoxon signed-rank test with and without the Holm correction of the p value for multiple comparisons, so as to provide an independent comparison with respect to the number of indices analyzed.

All analyses were run using R statistical software version 3.1.2 (R Core Team, 2014).

3. Results

The AUC average values of the k-fold test cases of single indices plotted against CLARE (Fig. 5) showed different patterns. In some cases (e.g., Ticino-w), all indices showed highly clustered AUC (from 0.73 to 0.81) and CLARE (from 0.57 to 0.65) values, whereas in other cases (e.g., Campidano), AUC (from 0.63 to 0.85) and CLARE (from 0.45 to 0.65) values are scattered. In Cilento, the values for CLARE ranged from 0.55 to 0.63, while most index values ranged from 0.67 to 0.78 for AUC. However, the two non-cumulative fire weather indices that included wind (Sharples and FFWI) showed lower values for AUC (from 0.57 to 0.58). AUC model performance for Ticino-sa (from 0.67 to 0.76) was similar to Ticino-w, but had a much broader range for CLARE (from 0.40 to 0.58). In contrast, Ticino-sn and Chilivani had similar performances when judged by AUC (0.62 to 0.79 and 0.63 to 0.78, respectively), but very
different values for CLARE (0.56 to 0.66, and 0.50 to 0.57, respectively). Ticino-sa and to some degree Campidano (DMC and DC) had CLARE values below the random threshold.

**Fig. 5.** Plots of the AUC vs. CLARE performance values for single fire weather indices used in this study (mean values obtained from the k-fold cross-validation procedure). The dashed lines represent the convex hulls of the points.

For Ticino-w, Cilento and Campidano, the performance values of AUC and CLARE indicate a coherent pattern with indices jointly showing high values for AUC and CLARE and vice versa. In Ticino-sa, Ticino-sn and Chilivani, indices were ranked differently according to the
performance metric considered. In Ticino-sa, indices with similar AUC performance had a wide range of values for CLARE, whereas in Ticino-sn and Chilivani, indices with similar CLARE values returned a wide range of AUC values.

The differences in AUC vs. CLARE mean values among indices were highly significant (see Appendix 1 and 2 for details). Significant differences were less evident in Ticino-sa based on AUC results and in Ticino-sn based on CLARE, whereas Campidano had the highest number of cases of significant differences among indices based on both performance metrics.

As a rule, cumulative indices featured the highest values of AUC and CLARE with the exception of Ticino-sa and Chilivani, where the indices scoring the highest values for CLARE were non-cumulative, although they included wind. In Ticino-w and Chilivani, cumulative and non-cumulative indices performed similarly (see Appendix 3 for more details). In Campidano, the clear differences among the groups of indices were evident, with higher performance values for AUC and CLARE when using indices that include cumulative wind effects. The non-cumulative indices that include wind (i.e., Sharples and FFWI) had low values in terms of AUC (fire occurrence criterion), even if they often reached medium to high values for the burnt area criterion. Results for CARE and CRARE (Appendix 4) were similar to those obtained by CLARE, although with a different range of values. The Chilivani region was the exception, with a visually different pattern with CARE.
In most cases, the index with the highest AUC value was the FWI, with the exception of Ticino-sn and Cilento, where the best index was DC. In contrast, the highest CLARE values did not appear to be related to any particular index (cf. Table 3 for more details).

**Table 3.** Best single indices according to the AUC and CLARE criteria. The values (mean of the test cases from the k-fold cross-validation procedure) for both performance metrics are given.

Cumulative wind-included (C-W) and wind-excluded indices (C-nW), non-cumulative wind-included (nC-W) and non-cumulative wind-excluded indices (nC-nW) selected are represented by ●.
4. Discussion

We propose a new approach for analyzing the suitability of fire weather indices to predict both fire occurrence and burnt area, and tested it in European study areas that are characterized by fire regimes differing in terms of fire seasonality, number of fires and burnt area. The performance patterns varied strongly according to the fire regime and the fire indices considered.

4.1 Regional meteorological controls on fire ignition and spread

In general, the majority of the best-performing indices for both criteria were cumulative, i.e., they accumulated past weather patterns rather than just using weather conditions on the day of the prediction. This highlights the importance of the wetting and drying dynamics of fuel over a time span of several days or weeks. The wind component appears to play a major role in Ticino-w and Campidano, with the Angstroem and the FMI index performing better than the cumulative indices that do not include wind. In spite of the clearly different ecological characteristics of the two areas, in both regimes, fire danger increases as a consequence of episodes of strong, warm-dry winds. In Ticino, winter fires burn mainly in the chestnut forest belt, typically occurring during or after periods of ‘Foehn’ winds that reduce the humidity of both air and fuel (Pezzatti et al., 2013). In the area of Campidano, with its typical Mediterranean climate where an extended summer drought is common, events of Sirocco winds can dramatically worsen the fire danger situation (Delogu, personal communication). During these strong wind events, the fuel quickly dries out and fire spreads rapidly, often overwhelming fire brigades in their attempt to suppress such events quickly.
The wind aspect is less important in the Mediterranean area of Cilento, especially regarding fire occurrence. Here, DC was the best-performing index for AUC and BUI for CLARE (Table 3), whereas the highly specific Initial Spread Index (ISI) performed poorly. The high population density of the region combined with its complex topographic and land-use patterns may reduce the role of short-term (daily) climatic conditions for the fire regime, with the persistent effect of drought on fuel being likely more important than wind for fire ignition and fire spread.

Fire indices in the other study areas showed generally lower performance in either of the two evaluation criteria (AUC or CLARE), or both (Fig. 5). In particular, the best-performing index for AUC was not among the top indices for CLARE. In Chilivani and Ticino-sa, the cumulative aspect of the index was important for fire ignition, and wind had a great influence on burnt area (Fig. 54), although no particular index performed well according to this second criterion (performances consistently < 0.6). In these regimes, ignitions may result in fire starts during prolonged dry periods, and non-climatic factors such as complex topography (e.g., Ticino) or fragmented land cover (e.g., Chilivani) may govern fire spread. Cumulative indices without a wind component performed poorly for CLARE, with several values falling below the random guess threshold in Ticino-sa. This is likely due to the strong human influence on fire ignition and spread, or a biased distribution of fire sizes because of the very small number of large fires (only one fire ≥ 50 ha in Ticino-sa).

In the lightning regime of Ticino-sn, cumulative indices without wind performed best in terms of ignition probability, but no clear pattern was identified for burnt area. Summer conditions are characterized by extended rainless periods and brief interspersed heavy precipitation events (Spinedi and Isotta, 2004), with thunderstorms that are often localized and not necessarily registered at the reference meteorological station. Since local weather conditions on the day of
ignition may not reflect dryness at the regional scale (cf. thunderstorms), the best predictions for
AUC resulted from indices with a strong memory component like the DC. As highlighted by
Conedera et al. (2006), lightning-induced fires are usually concentrated in coniferous forests at
high elevations on steep slopes, and they start as smoldering ground fires in thick mats of conifer
litter. When fuel is dry, the fire can emerge and spread rapidly under windy conditions, making
suitable conditions for fire occurrence very specific. Although wind can play a major role in the
days after ignition, this is not highlighted in our results, where final burnt area was clearly
related to the conditions at the start of the fire.

Our results highlight patterns where indices that best predicted daily fire starts also performed
well in estimating final burnt area, as in Ticino-w, Campidano and Cilento, and areas where
these aspects of the fire regime were mostly unrelated, as in Chilivani, Ticino-sa and Ticino-sn
(Fig. 5). Interestingly, regions where indices respond differently with respect to the two types of
performance metric also feature consistent differences between the average size of fires greater
than 50 ha and overall mean fire size (ratios > 30, see Table 1). Such differences imply that fire
size may not always be driven by standard climatic factors, but rather by exceptional climatic
conditions that are unrelated to fire ignition, or by non-climatic aspects such as difficulties in fire
suppression due to steep topography, poor accessibility, or multiple ignitions. In this respect,
although the investigated performance metrics (CARE, CLARE and CRARE) generally showed
a similar pattern (Appendix 4), it became evident that CLARE represents a good compromise by
mitigating the effect of a few fires that may have become very large for reasons unrelated to
climate (cf. the different pattern between CARE and CLARE in Chilivani, Appendix 4).
4.2 Applicability of the approach

Given that burnt area results from multiple drivers operating across time and space, modelling the final burnt area from a single fire event or day can be challenging (Jiang et al., 2012). Here, we used daily fire start data and the related final burnt area to evaluate the performance of fire danger indices in forecasting both aspects. Although meteorological conditions that drive burnt area are also likely to trigger fire ignitions, the contrary may not be true. As a result, a poor relationship between raw index values and burnt area is often encountered at daily resolution. While using aggregated burnt area (weekly or monthly instead of daily) could be a workaround in certain situations, daily resolution is needed, for instance, for operational fire warning and firefighting purposes. In any case, details on the daily progress of fire size (spread of fire through time) are rarely available in forest fire databases and associating the final fire burnt area with the weather data at the reported start day of the fire, irrespective of the fire duration, is often the only applicable approach (Duff et al., 2016). However, this may be a sound approximation under many circumstances, since fires spread exponentially, and in most cases the final burnt area strongly depends on the success of the initial spread rate that may cause the failure of first fire attack and subsequent fire-control measures. Furthermore, the exponential nature of the fire spreading process and the resulting highly skewed distribution of fire sizes render a linear relationship between number of fires and fire sizes improbable. Consolidating final fire sizes into a single value with reference to the start-day is not only a plausible simplification, but also a fitting procedure to develop a suitable proxy for the initial conditions that lead to large fires, which fire-fighting organizations find so challenging.

In the case of the study areas selected, the prevalence of anthropogenic fire regimes in a densely populated area makes fire detection very efficient and the coincidence of the fire ignition with
the fire record day very likely. Furthermore, in the Alpine context also the hold-over time
existing for lighting-induced fire is only exceptionally longer than a single day (Conedera et al.
(2006). This may be problematic on contrary for more remote locations such as the boreal
forests, where long hold-over periods may cause significant discrepancies between the ignition
and the fire discover date.

In cases where fire weather indices accurately predict daily fire start probabilities but are unable
to provide information regarding corresponding final burnt areas, one can assume that either the
climatic conditions that trigger ignition are substantially different from those that govern fire
spread, or that non-climatic factors (e.g., fuel type, topography, or human behavior) play a major
yet distinct role for fire ignition, burnt area, or both. In such cases, the method we have
developed is useful to evaluate whether the combination of climatic factors that drives fire
occurrence influences fire spread potential as well. Based on this information, the two
performance metrics AUC and CLARE may be weighted in order to select the fire weather index
that is most suitable and informative regarding fire regimes in specific regions. In cases where a
similar combination of climatic variables controls the probability of both fire occurrence and fire
spread, the two performance metrics can be considered to be equal (see Appendix 5 for a
graphical representation of a synthetic metric that considers both aspects). In other cases, more
importance should be given to index performance regarding fire occurrence (e.g., by using a
weighted mean), which is the effective response variable the models are built on. Finally, indices
with high performance values for both criteria can determine both when ignition probability is
high and when a large burnt area is likely.
5. Conclusion

Large fires are responsible for the majority of burnt area and suppression costs (San-Miguel-Ayanz and Camia, 2010; Liang et al., 2008). Especially in the populated mountain ranges of Europe fire size is a critical factor for post-fire natural hazards like rock fall or debris flow (Mayer et al., 2020, Conedera et al., 2003), and keeping the burnt area low is crucial for mitigating the potential consequences. To date, no easily applicable metrics exist to assess the ability of fire danger rating systems to provide information on possible final fire sizes. In response to this need, the proposed approach was conceived to assess whether a fire weather index with good fire-occurrence prediction performance is also capable of shedding light on the likelihood of large burnt areas. When both conditions are met, fire brigades can be made aware of days with a high risk of fire starts as well as the likelihood of large fire occurrence. This may result in a more danger-related level of preparedness and available means on short terms and in the choice of more suited firefighting tactics for the initial attack.

The proposed evaluation metrics may also be used for model selection criteria for more sophisticated fire weather rating systems based on widely used binomial logistic regressions (e.g., Martell et al., 1987; Vega-Garcia et al., 1995; Vasconcelos et al., 2001; Andrews et al., 2003; Dimitrakopoulos et al., 2011; Padilla and Vega-Garcia, 2011) or niche modeling approaches (Maxent) that combine different fire weather indices (e.g., De Angelis et al. 2015). However, caution is needed when applying the novel metric to fire regimes characterized by long lasting events without any information of the precise ignition date.
Acknowledgments

This research was supported by the European Project ALP FFIRS funded by the European Regional Development fund of the Alpine Space Program, the Cantons of Ticino and Valais, BAFU, and the BAFU/WSL Swiss program Forest and Climate change. We thank Curtis Gautschi and Cathy Whitlock for revising the English text.

References


interactions in Southern Europe: implications for landscape management. J. Environ. Manage. 92, 2389-2402. https://doi.org/10.1023/A:1013130528470


https://doi.org/10.1071/WF07150