Fire danger: the skill provided by the ECMWF Integrated Forecasting System

Francesca Di Giuseppe[1], Claudia Vitolo[1], Blazej Krzeminski[1], Jesus San-Miguel[2], and Florian Pappenberger[1]

[1]European Centre for Medium-range Weather Forecasts, Reading, United Kingdom
[2]European Commission, Joint Research Centre, Ispra, Italy

Correspondence: Francesca Di Giuseppe (F.DiGiuseppe@ecmwf.int)

Abstract. In the framework of the EU Copernicus program the European Centre for Medium-range Weather Forecast (ECMWF) on behalf of the Joint Research Centre (JRC) is forecasting daily fire weather indices using its medium range ensemble prediction system. The use of weather forecast in place of local observations can extend early warnings up to 1-2 weeks allowing for greater proactive coordination of resource-sharing and mobilization within and across countries. In addition, the use of an ensemble system allows to perform a probabilistic assessment of the forecast uncertainties which can boost confidence in the decision process during emergency situations. Using one year of pre-operational service in 2017 here we assess the capability of the system globally and analyze in detail three major events in Chile, Portugal and California. We also present examples on how fire forecast products could be tailored to provide uncertainties of fire weather conditions and information in a probabilistic fashion.

1 Introduction

The prediction of fire danger conditions allows forestry agencies to implement fire prevention, detection, and pre-suppression action plans before fire damages. However, in many countries fire danger rating relies on observed weather data which only allows for daily environmental monitoring of fire conditions (Taylor and Alexander, 2006). Even when this estimation is enhanced with the combined use of satellite data, such as hot spots for early fire detection, and land cover and fuel conditions it normally only provides 4- to 6-hour warnings. By using forecast conditions from advanced numerical weather models, early warning could be extended up to 1-2 weeks allowing for greater coordination of resource-sharing and mobilization within and between countries.

Due to the improved skills of weather forecasting, the use of numerical weather prediction offers a real opportunity to enhance early warning capabilities (Roads et al., 2005; Mölders, 2008, 2010). In recent years institutions such as Natural Resources Canada (NRC) and the US National Oceanic and Atmospheric Administration (NOAA) have implemented regional fire danger forecasting systems based on their operational weather forecasts (Bedia et al., 2018). The Global Fire Early Warning System is also an international initiative, promoted by the Canadian Parternship for Wildland Fire Science and the United Nation Office for Disaster Risk Reduction, to provide fire danger forecast up to 10 days ahead using the Canadian operational weather forecasting system (http://canadawildfire.ualberta.ca/gfews). Parallel initiatives are promoted by the European Commission under the umbrella of the Copernicus Emergency Management Service (CEMS), namely the European Fire Forecast Information System (EFFIS, http://effis.jrc.ec.europa.eu/) and its global counterpart the Global wildfire Information System (GWIS, http://gwis.jrc.ec.europa.eu/). Both systems rely on the Canadian Fire Weather Index (FWI) (Van Wagner et al., 1974, 1985) to rate fire danger and on numerical weather predictions to provide forecasted fire danger information at the European and global levels (San-Miguel-Ayanz et al., 2002).

Systems such as the FWI detect dangerous weather conditions conducive of uncontrollable fires rather than modelling the probability of ignition and fire behaviours. The FWI (developed in Canada) is specifically calibrated to describe the fire behaviour in a jack pine stand (*Pinus banksiana*) typical of the Canadian forests. However, its simplicity of implementation has made it a popular choice in many countries and it has shown to perform reasonably well in ecosystems very dissimilar to the boreal forest (Di Giuseppe et al., 2016; de Groot et al., 2007). The FWI calculation only relies on weather forcings and no information on the actual vegetation status is taken into account. When weather forecasts are used in place of observations, uncertainties can be introduced. Sources of uncertainty can be: (i) the limited predictability of atmospheric dynamics which is strongly affected by the initial state and (ii) the misrepresentation of physical processes. In the former case, errors are randomly distributed around the true state (Orrell et al., 2001); in the latter, errors produce systematic deviations from the true state. In both cases, errors in the weather forecast may be amplified or damped by nonlinear transformations in the fire weather model (Erickson et al., 2018). Thus, for example, a dry bias in the model in a certain region will lead to the persistent prediction of higher fire danger values compared to what would be calculated using local observations.

Handling random errors in weather forecasts is traditionally done through the use of ensemble prediction systems where several simulations are performed starting from slightly different initial conditions and model configurations (Molteni et al., 1996; Buizza et al., 1999). The forecast is then interpreted as probabilistic rather than deterministic. While it has been shown that the probabilistic information contained in an ensemble prediction system might be difficult to interpret for end-users (Pappenberger et al., 2013), ensembles can also boost confidence in the decision process during emergency situations as a cost-loss analysis can be associated to the different scenarios (Cloke et al., 2017). Moreover, ensemble predictions can have more information value than the single deterministic simulation (Richardson, 2000; Zhu et al., 2002). Systematic biases, on the other hand, can be reduced by model improvements. Still appropriate post-processing (bias correction) of the atmospheric model (Piani et al., 2010; Di Giuseppe et al., 2013a, b) or post-processing of the sectoral application outputs (Raftery et al., 2005) can correct resolved processes and improve the final forecast skill.

Given the above considerations, in this paper we assess the performance of the fire danger forecasting system developed for the Copernicus Emergency Management Service at European Centre for Medium-range Weather Forecasts (ECMWF) to predict the FWI values and the probability of detection of fire during one year of operation in 2017. As the Fire Weather Index is the main index of this system we will concentrate on this model component.

1.1 FWI calculation

1.1.1 General concept

The Fire Weather Index system is composed of six variables describing fuel moisture and fire behavior, as influenced by weather (Van Wagner et al., 1987). Three fuel moisture codes provide the moisture content of dead woody debris of different diameter classes laying on three fuel beds. The dead fine fuels is described by the Fine Fuel Moisture Code (FFMC) which has a fast response time to weather modification (hourly time scales). The surface organic matter of moderate density, such as the fermentation layer of forest soils (Hood, 2010) is calculated using the Duff Moisture Code (DMC). DMC has a slower response to the weather forcing, conventionally set to 12 days. Finally the deep, compact soil organic layers, such as the humus layer of forest soils Hood (2010) is expressed with a Drought Code (DC). This is the slower evolving component of the FWI codes and has a lag-time set to 52 days in the original FWI formulation. The FWI system also provides fire behavior indexes, in terms of rate of fire spread (Initial Spread Index, ISI), fuel available for combustion (Buildup Index, BUI), and head fire intensity (Fire Weather Index, FWI). The FWI component combines the ISI and BUI and can be considered as the most general indicator of fire danger.

A comprehensive description of the FWI system, the interaction between the various components and how these are used in fire management can be found in (Van Wagner et al., 1987; Wotton, 2009). In the interest of brevity we have here only recalled the basic concepts underlying the FWI system and how different components might be more appropriate to describe fire danger depending on the local fuel characteristics and the type of information required. Abatzoglou et al. (2018) showed that FWI exhibits strong correlative relationships to burned area across most non-arid eco-regions globally, while Bowman et al. (2017) highlighted how high FWI values are often associated to the most extreme fire activities recorded using Fire Radiative Power observations. As FWI has been shown to provide a good metric for quantifying fire danger globally, the proposed analysis of forecast skills will concentrate on this index (Di Giuseppe et al., 2016; de Groot et al., 2007).

1.1.2 FWI forecast

For each day indexes of the FWI rating system are calculated operationally at ECMWF using real-time (RT) forecasts. A full description of the modeling components can be found in Di Giuseppe et al. (2016). The high resolution (HRES) and the ensemble prediction systems (ENS) provide weather forecasts which extend up to 10 days in the future. The atmospheric forcings have a temporal resolution of 3 hours and a spatial resolution of 9km for the high resolution run and 18 km for the ensemble prediction simulations. While the HRES is a single (deterministic) model integration, the ENS provides 51 realizations from perturbed initial conditions and different model physics (Buizza et al., 1999). These ENS forecasts are used to assess uncertainties in the prediction.

A model integration at any nominal time simulates atmospheric conditions at a different local time, depending on the location. FWI calculations are usually performed at 12:00 local time because the model was calibrated using measurements at 12:00 against fire behavior the most active window (between 14:00-16:00) (Van Wagner et al., 1987). Therefore to produce a snapshot at 12:00 local time, a temporal and spatial collage of 24 hours time model simulations is performed. Atmospheric fields are cut into 3-hourly time strips using the closest 3-hour forecast outputs and then concatenated together so that the final field is representative of the conditions around the local noon within the 3 hour resolution available (Di Giuseppe et al., 2016).

1.1.3 FWI reference and benchmark

As many forestry agencies still rely on observed meteorological data to provide fire danger, a first assessment of the quality of forecasted FWI will rely on the comparison with observations. Despite several meteorological observations are available through the Global Telecommunication System (GTS) SYNOP network, only a subgroup of stations have at least 30 days of recordings at local noon during 2017 (spatial coverage is given in figure **??**). Many fire prone regions, such as Australia, would not be covered by this comparison. In order to overcome this limitation, a reference dataset of FWI modelled values is also used. This is constructed using the ERA5 reanalysis dataset. ERA5 is the latest of ECMWF reanalysis products which was released at the beginning of 2019. It substitutes the previous ERA-Interim database (Dee et al., 2011; Vitolo et al., 2019) providing a much improved spatial resolution and an extensive increment of assimilated observations. Simulations begin in 1979 and are updated in quasi real time. Fields have a spatial resolution of about 30 km and hourly time resolution. Outputs from ERA5 undergo the same temporal interpolation described in the previous section to provide the model with a composite fire reanalysis product at 12:00 local time. It has to be noted that, compared to local observations, a reanalysis provides a dynamically consistent estimate of the climate state at each time step and can, to a large extent, be considered a good proxy for observed meteorological conditions. Moreover, by combining different observations, reanalysis datasets extend well beyond the natural life of single observational networks and they can provide a wider spatial coverage than using local observations.

In addition, from ERA5 we derive a climatological benchmark simulation (called CLIM hereafter), considered as a "mean year" for the FWI using the period 1980-2018. At pixel level, every day the median FWI is calculated over the period spanning 4 days before/after a given date. While CLIM has no expected predictive skill for a specific year, it retains information on the yearly climatological variation of FWI. CLIM is what could be used in the absence of either a monitoring (based on observations) or a forecast system (based on weather forecast). It should therefore score better or equal of the forecast on time ranges beyond the limits of predictability. CLIM is used in this study as a benchmark to rank the expected improvements provided by a forecasting system

1.2 Observed fire events

While national inventories of wildfire activities exist in many countries, they can be heterogeneous and lack the temporal span desirable for the validation of a fire danger system at the global scale. Satellite observations can supply a valid alternative especially as they cover remote areas where in-situ observations are sparse (Flannigan and Haar, 1986; Giglio et al., 2003; Schroeder et al., 2008). Daily maps of fire radiative power (FRP) (Kaufman et al., 2003; Wooster et al., 2005) are available from ECMWF since 2003 through the Global Fire Assimilation System (GFAS) (Kaiser et al., 2012; Di Giuseppe et al., 2017, 2018). This dataset has been developed in the framework of the Copernicus Atmosphere Monitoring Services (CAMS) and uses observations from the MODIS sensors on board of Terra and Aqua platforms and assumptions on fire evolution to calculate a continuous record of active fires. The GFAS dataset integrates all available FRP observations available in a day over a regular

$0.1\,\mathrm{deg}$ grid. According to Wooster et al. (2005), this provides an indication of the cumulative dry mass available for burning which can be then put into a relationship with fire emissions. In this paper, the FRP products are only used as an observations of fire events. However, FRP values are ignored and only used to derive a mask of fire occurrence based on a minimum detection criteria: $FRP > 0.5Wm^{-2}$ (Kaiser et al., 2012). A "hit" is recorded if the fire forecast predicts fire danger above the 95$^{\text{th}}$ percentile of its historical values (provided by the ERA5 simulations) when a fire really occurred.

1.3 Score metrics

The performance of the the fire forecasting systems to reproduce observed FWI values is assessed using deterministic and probabilistic scores. Both the synop database and ERA5 are treated as a proxy for observations in the evaluation. To asses the quality of the computation we use traditional deterministic skill scores such as the mean bias (MB) and the mean absolute error (MAE). For a probabilistic assessment, the continuous ranked probability score is also employed (CRPS; Hersbach (2000)). These are defined as:

$$MB = \sum_{p=1}^{cases} [F_{n=HRES} - O] \tag{1}$$

$$MAE = \frac{1}{cases} \sum_{p=1}^{cases} \sqrt{[(F_{n=HRES} - O)^2]} \tag{2}$$

$$CRPS = \frac{1}{cases} \sum_{t=1}^{cases} \int_{-\inf}^{+\inf} \left[F_n - O)^2\right] dn \tag{3}$$

where F is the forecast at time step t of N number of forecasts and O is the observed value. While the MB, MAE and the ACC are applied to a single forecasts, the high resolution forecast, the CRPS takes into account the whole distribution of possible values predicted by the ensemble. It is the continuous extension of the ranked probability score, where F_n is the cumulative distribution function of the predicted ensemble values. Then, the CRPS compares the cumulative probability distribution of the FWI forecasted by the ensemble forecast system to the observation. In this sense the CRPS is sensitive to the mean forecast biases as well as the spread of the ensemble (Hersbach, 2000).

While conventional skills score can be employed to assess the quality of the FWI computation, the verification of the FWI as a fire indicator is instead extremely challenging. First, as widely explained, FWI is not a physical measure of fire activity but of its potential danger, if one were ignited. Therefore high fire danger, while being correctly forecasted, might not result in active fires if there is no ignition. From the verification point of view this means that the identification of false alarms is not meaningful and the verification should mainly rely on hits and misses. Secondly, fires are rare events and, as for any other infrequent phenomena, the verification statistics are heavily influenced by the small number of hits when compared to the total. Still, when the cost of a missed event is high, for example in terms of human lives, the deliberate over-forecasting may be justified (Richardson, 2000; Cloke et al., 2017).

In these cases a positively oriented score such as "hit rate" may be more useful. Also forecast quality does not always equals forecast value (Richardson, 2000). A forecast has high quality if it predicts the observed conditions well according to some

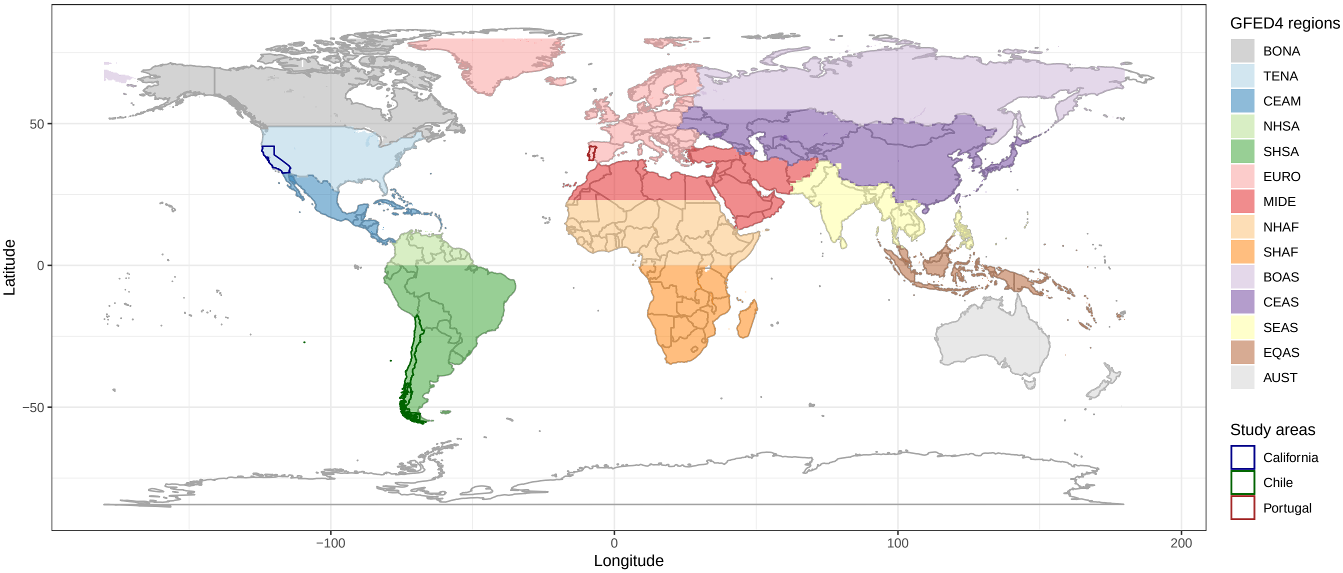

Figure 1. GFED4 regional classification and the 3 countries selected to showcase the fire forecast performances (California, Chile, and Portugal).

objective or subjective criteria. It has value if it helps the user to make a better decision in terms of protective actions (Cloke et al., 2017). For example predicting high temperature and low precipitation in desert areas might be accurate but carries low information content and therefore limited value. Following these arguments and to gain an appreciation of the potential value of the forecasting system globally we use as a metric the Probability of Detection (POD), which measures the fraction of the observed events that were correctly forecasted ($POD = hits/(hits + misses)$). Therefore, POD only takes into account observed fires and, unlike other skill scores such as the Brier score, does not suffer from the artificial vanishing due to the high number of correct negative and false alarms (see Stephenson et al. (2008); Ferro and Stephenson (2011) for a discussion on these problems).

1.4 Fire regions

The global assessment of the fire forecast skills is mostly provided as an average over selected regions even if the calculation of the various scores is performed at pixel level by interpolating the model grid over the benchmarks. For an assessment at the continental scale, we use the fire macro-regions defined by the Global Fire Emission Database, GFED4 (Giglio et al., 2013). These macro-regions are characterized by different fire regimes and are very roughly homogeneous in their burning emissions contribution (Giglio et al., 2013). Inside these regions we also select 3 areas at national/regional level - California, Portugal and Chile - which experience recurrent intense fire episodes and saw major events taking place in 2017 (Figure 1). Events in these locations are also analyzed in detail.

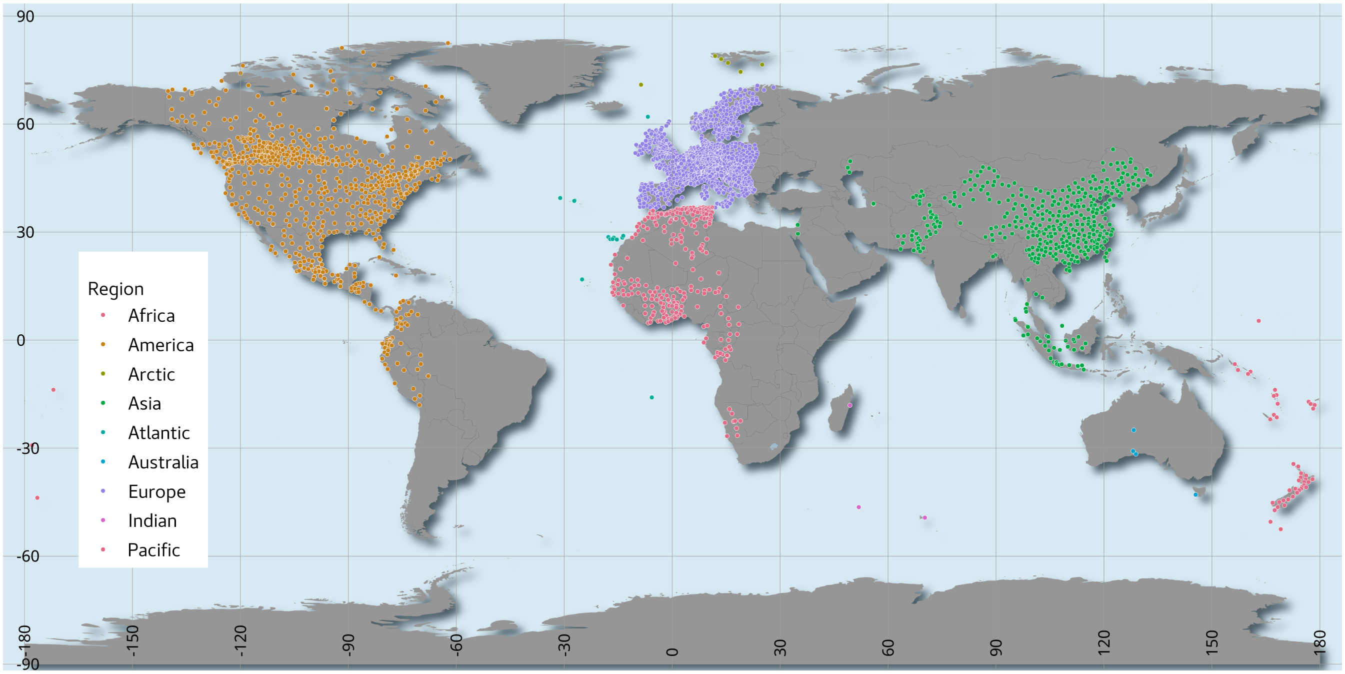

Figure 2. Spatial distribution of weather stations from the synop network which have at least 30 observations recorded at local noon in 2017.

2 Results

2.1 Skill in the FWI prediction

The first assessment looks at the capability of ECMWF fire forecast to reproduce the same FWI values as would be estimated from the network of local stations but 2, 6 and 10 days ahead. The selected stations (figure 2) have at least 30 records during 2017 at local noon and are used to perform an analysis of bias and anomaly correlation at different lead times. For comparison also FWI calculation using ERA5 is included. This provide an estimations of the limit of predictability when using forcing from model simulations in place of observed values (Di Giuseppe et al., 2016). As expected there is a performance degradation going towards longer lead times and mean biases are limited to few units even at day 10. However, depending on the calibration procedure adopted few units could mean a mismatch in danger level classification. The mean absolute error (figure **??**b) provides information on the residual amplitudes. FWI from reanalysis have the largest skills as expected and the mean absolute errors rapidly increases with lead times. However the distribution of MAE values clearly shows that in selected cases predictive skills can be achievable even at day 10.

```
\begin{figure*}
    \centering
        \includegraphics[width=\textwidth]{station_error.png}
```

\caption{Comparison between modelled FWI and observed FWI value. FWI are calculated from

\end{figure*}

Despite its importance the analysis performed using the synop network it is still pointwise and does not cover all the regions where fires are relevant. Moreover mean biases and MAE are based on the high resolution forecast. These skill metrics do not provide information about the performance of the ensemble forecasting as a whole.

Figure **??**

PUT here crps skills scores

2.2 Skill in detecting fire events

Conventionally, we assume that an active fire is correctly predicted if the FWI is greater than the 95^{th} percentile of its distribution of values here defined using the ERA5 database. Figure~**??** shows the mean POD for all events in 2017 at forecast day 2, 6 and day 10. Figure~**??** also shows the POD that could be achieved in the absence of a forecasting system when just using a mean climatological FWI estimation. Given the intrinsic limitations of the POD as skill metric, CLIM provides an useful benchmark to understand the incremental skill provided by the forecast.

At day 2, high latitudes are characterized by POD greater than or equal to 0.5. These are mostly temperate regions where vegetation is dominated by forests and fuel is abundant and where fire danger is moisture limited. In these regions the FWI is a good predictor of fire danger (Di Giuseppe et al., 2016). It has to be noted that the FWI does not take into account management measures that could introduce a relevant number of “false-alarm”. Central America, the Middle East and the northern hemisphere areas, Africa are characterized by a POD in the range 0.3-0.4 as in most of the tropics, where, fires usually occur in grass-shrub lands. Here fuel is scarce and weather plays a less relevant controlling role. Also it has to be noted that the statistics here are likely to be contaminated by many agricultural and prescribed fires that are considered “events” and which would dilute some of the skill in regions where annual cropland is high or are heavily managed. The spatial distribution of POD at day 6 is very similar to the corresponding figures at day 2, just slightly shifted towards lower values.

One important exception is the extremely good performance of the fire forecast in Equatorial Asia where the system seems to have a predictability of 0.9 even at day 6. Also de Groot et al. (2007) highlighted how FWI is not a good indicator in this area and a fire early warning system should mostly rely on the drought code. Still, there are factors that could contribute to this enhanced predictability. First fires in this region are mainly caused by humans for the purposes of cleaning the land for establishing plantations (Field et al., 2009; Benedetti et al., 2016). Therefore they occur every season as soon as fuel conditions are favorable and the ignition is less of a stochastic component. This also means that the POD statistics are built on a much higher number of events. Moreover, the strength and prevalence of these fires are strongly influenced by large-scale climate patterns like El Ni{~n}o Field et al. (2004) which have been proven being highly predictable (Zhu et al., 2015). The other relevant fact is the almost null skill globally when a climatological FWI is used highlighting the added benefit of a forecasting system

\begin{figure}[htb]

```
\centering
\includegraphics[width=0.5\textwidth]{POD2.png}\\[0.5cm]
\includegraphics[width=0.5\textwidth]{POD6.png}\\[0.5cm]
\includegraphics[width=0.5\textwidth]{POD10.png}\\[0.5cm]
\includegraphics[width=0.5\textwidth]{PODclima.png}\\[0.5cm]
\caption{Global area averaged Probability of Detection (POD) for  day 2, day 6 and day 10 fo
\end{figure}
```

2.3 2017 case studies

Figure **??** provides an averaged assessment of the global performances of the forecasted FWI as a generic indicator of fire danger. There are regional and seasonal variations to this skill. Also it is important to understand how the information provided could be used in real cases when the forecast is intended to aid emergency responses. Here we will analyses three cases of fire events that took place in 2017, which proved to be an extreme year for fire all across the globe.

The year 2017 started with an extended fire in central Chile that lasted almost all of January. Strong winds, high temperatures and long-term drought conditions led to an event that has been described as the worst wildfire in Chilean history (Bowman et al., 2018). Fires in the central regions of O'Higgins, Maule and B'io B'io south of Santiago were difficult to control. Although fire activities where recorded since July 2016 they became particularly intense in January 2017. In June, between 17 and 18, another devastating fire hit Portugal. It claimed more than 60 lives mostly recorded in the Pedr'og~ao Grande area, 50 km southeast of Coimbra. A persistent heatwave had been building in the region, with temperatures above 40C, which are highly unusual for the season. Moreover, relative humidity levels below 30% had a role to the intensification of the deflagration and the spread of the wildfire, which raged out of control for several days (Boer et al., 2017). Finally in Octoberextensive wildfires raced just north of the San Francisco Bay Area in California causing historic levels of death and destruction. These named “Wine Country” wildfires were the most destructive in California history, with 44 deaths; the loss of 9,000 buildings; damage to approximately 21,000 structures; $10 billion of insured losses; and substantially greater total economic loss (Nauslar et al., 2018; Mass and Ovens, 2019).

For each case study, the affected area is identified as the minimum area including all detected active fires (cells with $FRP > 0.5Wm^{-2}$) during the selected time window. Figure **??** shows the information that could have been provided for the study areas by the 10-day fire danger high resolution forecasts (HRES), had these been already available. Each plot shows on the x-axis the dates in which FRP was observed and, on the y-axis, the dates forecasts were issued. The cell in the bottom left corner shows the percentage of pixels in the study area that are expected to be above the 95^{th} percentile of the FWI climatology for that pixel. The forecasts for day 2 to day 10 are on the same row. The forecasts issued on the following day are one row above and so forth. The dashed lines shows the observed fire radiative power (see also secondary y-axis).

The reader is reminded that active fires are triggered by highly unpredictable events (ignition) which are not accounted for in the FWI system. The FWI is not supposed to provide the exact localization of the event but an indication of potential fire activity. Large areas can be affected by anomalous conditions in the proximity of where the event really occurred. However it

Table 1. Events summary table.

Country	Region	Start date	End date	Main event	Location
Chile	O'Higgins, Maule, Bío Bío	01-01-2017	31-01-2017	26-01-2017	36° 46'S; 73° 03'W
Portugal	Pedrogao Grande	01-06-2017	30-06-2017	18-06-2017	39° 55'N ; 8° 08' W
USA	California	21-09-2017	20-10-2017	09-10-2017	38° 34'N; 122° 34' W

is noticeable the capability of the forecast to detect the increase in fire danger associated to the three events. For the Chile case, for example from mid-January between 70 and 80% of the area exceeded the high danger threshold. The FRP spike (occurred on 26th January) highlighting that most of the region was classified at very high danger almost 10 days ahead. Similarly, the Portugal fire could have been predicted ten days before, but the California event only two to three days.

```
\begin{figure}
    \centering
    \begin{subfigure}[b]{0.45\textwidth}
        \centering
        \includegraphics[width=\textwidth]{Chile_boxy_map}
        \caption{Chile}
    \end{subfigure}
    \begin{subfigure}[b]{0.45\textwidth}
        \centering
        \includegraphics[width=\textwidth]{Portugal_boxy_map}
        \caption{Portugal}
    \end{subfigure}
    \begin{subfigure}[b]{0.45\textwidth}
        \centering
        \includegraphics[width=\textwidth]{California_boxy_map}
        \caption{California}
    \end{subfigure}
    \caption{Comparison of Fire Radiative Power (gray dashed line with axis on the right han
\end{figure}
```

One of the advantages of developing a probabilistic ensemble prediction system (ENS) for fire danger is the possibility to assess the confidence in the model forecast, at least from the point of view of the uncertainties in weather forcings. Admittedly, the use of this type of information is not straightforward as many studies have highlighted (Pappenberger et al., 2013; Palmer, 2000). This difficulty steams from the fact that in many response systems an activation is conditional to the excedance of deterministic threshold. Instead, in a probabilistic approach what is provided is the {*probability} of exceedance a certain values and this is often perceived as a degradation of the information values (Richardson, 2000). However by using the probabilistic component of the system, it is possible to understand how consistent the fire danger forecast is with respect to slightly different forecast scenarios. To highlight the information that can be added with the use of the ensemble prediction system, Figure* **??** *shows the progression of 10-day forecasts for the major fire events (corresponding to the peak of observed FRP in figure* **??***) in the three case studies. In this case we concentrate on a specific location (model grid box) whose coordinates are specified in table 1. In this plot, there is very high consistency between HRES and ENS, with the former falling generally close to the ensemble median. In addition, the bulk of ensemble members falls in the yellow region which identifies very high danger. It is this consistency and persistence of very high danger over time that could have helped users gain confidence in the forecast, had that been already available. While these events were all considered catastrophic in terms of impacts, it is noticeable that, at all forecast ranges, fire danger would have been classified as very-high but not extreme by the exclusive use of the high resolution forecast. However in all three cases by looking at the full distribution of FWI values provided by the ensemble, it is clear that some of the forecasts were hinting at the possibility for extreme conditions.*

Thus is there an enhanced accuracy in the forecast produced with the use of the ensemble forecasting system? Figure **??** *shows the probability of detection calculated over the three fire regions (see red box in Figure* **??** *for the geographical extents of these areas) using all the ensemble members. The high resolution forecast (red dots) is not in all cases the most skillful prediction despite the increased resolution. It is interesting to notice how in the California case, for example, many of the ensemble members perform consistently better than the high resolution run up to 6 days ahead.*

```
\begin{figure}
    \centering
    \begin{subfigure}[b]{0.45\textwidth}
        \centering
        \includegraphics[width=\textwidth]{Chile_ENSfwi}
        \caption{Chile}
    \end{subfigure}
    \begin{subfigure}[b]{0.45\textwidth}
        \centering
        \includegraphics[width=\textwidth]{Portugal_ENSfwi}
        \caption{Portugal}
    \end{subfigure}
    \begin{subfigure}[b]{0.45\textwidth}
        \centering
        \includegraphics[width=\textwidth]{California_ENSfwi}
        \caption{California}
    \end{subfigure}
```

```
    \caption{Fire Weather Index distribution at the location and date specified as "main eve
\end{figure}

\begin{figure}
    \centering
    \begin{subfigure}[b]{0.45\textwidth}
        \centering
        \includegraphics[width=\textwidth]{Chile_ENS}
        \caption{Chile}
    \end{subfigure}
    \begin{subfigure}[b]{0.45\textwidth}
        \centering
        \includegraphics[width=\textwidth]{Portugal_ENS}
        \caption{Portugal}
    \end{subfigure}
    \begin{subfigure}[b]{0.45\textwidth}
        \centering
        \includegraphics[width=\textwidth]{California_ENS}
        \caption{California}
    \end{subfigure}
    \caption{Probability of detection (POD) as a function of forecast horizon for the ensemb
\end{figure}
```

3 Conclusions

In the last years, ECMWF has been involved in the EFFIS development by providing weather forcing and fire danger calculations using its medium-range weather forecasts. Global fields of FWI are calculated daily using the high-resolution (9 km) forecast up to 10 days ahead. The 18 km resolution ensemble prediction system provides additional 51 realizations based on slightly different initial conditions and/or using different model configurations (Molteni et al., 1996). These datasets are freely available in line with the data and information policy of the Copernicus program which intends to provide users with free, full and open access to environmental data.

Using one year of operational service in 2017 we have showcased the potential of the use of weather forecasts to support the monitoring of fire danger conditions and planning in case of a potential emergency. By applying a model based definition of warning levels to the FWI we have shown that it provides a probability of detection (POD) for fire activity that is often above 60% at day 2. Mid and high latitude forested areas, where fuel is abundant have the highest predictability while in *savana/shrub-land regions the relationship between FWI and fire occurrence weakens. With the exception of Equatorial Asia which shows a POD in the range of 0.9 the average mean POD falls below the 0.6 threshold by day 6 on average highlighting that this might be consider a rough estimation of the limit for a skillful prediction which could justify the use of prediction at this stage. However there are regional and case-specific variations to this limit. Using the three large fire events of 2017 in Chile, Portugal and California, as examples, we have shown that accurate forecasts can be achieved up to 10 days ahead.*

Another interesting aspect attached to the use of weather forecasts is the use of probabilistic information. The quantification of forecast uncertainties through the use of ensemble predictions is something still pretty new in fire forecasting. However it opens great opportunities in terms of adding a confidence level to the the fire prediction.

Code availability. The code to generate the results of this work is available as vignette of the caliver R package: Verification of fire danger classes.

Author contributions. FDG wrote the . . .

Competing interests. The authors declare no competing interests.

Acknowledgements. This work was founded by the EU Project ANYWHERE (Contract 700099) and the Global Fire Contract 389730 between the Joint Research Centre and ECMWF.

References

Abatzoglou, J. T., Williams, A. P., Boschetti, L., Zubkova, M., and Kolden, C. A.: Global patterns of interannual climate–fire relationships, Global change biology, 24, 5164–5175, 2018.

Bedia, J., Golding, N., Casanueva, A., Iturbide, M., Buontempo, C., and Gutiérrez, J. M.: Seasonal predictions of Fire Weather Index: Paving the way for their operational applicability in Mediterranean Europe, Climate Services, 9, 101 – 110, https://doi.org/https://doi.org/10.1016/j.cliser.2017.04.001, http://www.sciencedirect.com/science/article/pii/S2405880716300826, climate services in practice: what we learnt from EUPORIAS, 2018.

Benedetti, A., Di Giuseppe, F., Flemming, J., Inness, A., Parrington, M., Rémy, S., and Ziemke, J.: Atmospheric composition changes due to the extreme 2015 Indonesian fire season triggered by El Niño [in "State of the Climate in 2015"], Bull. Amer. Meteor. Soc, 97, S56–S57, 2016.

Boer, M. M., Nolan, R. H., Resco De Dios, V., Clarke, H., Price, O. F., and Bradstock, R. A.: Changing Weather Extremes Call for Early Warning of Potential for Catastrophic Fire, Earth's Future, 5, 1196–1202, https://doi.org/10.1002/2017EF000657, https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1002/2017EF000657, 2017.

Bowman, D. M., Williamson, G. J., Abatzoglou, J. T., Kolden, C. A., Cochrane, M. A., and Smith, A. M.: Human exposure and sensitivity to globally extreme wildfire events, Nature Ecology & Evolution, 1, 0058, 2017.

Bowman, D. M., Moreira-Muñoz, A., Kolden, C. A., Chávez, R. O., Muñoz, A. A., Salinas, F., González-Reyes, Á., Rocco, R., de la Barrera, F., Williamson, G. J., et al.: Human–environmental drivers and impacts of the globally extreme 2017 Chilean fires, Ambio, pp. 1–13, 2018.

Buizza, R., Milleer, M., and Palmer, T.: Stochastic representation of model uncertainties in the ECMWF ensemble prediction system, Quarterly Journal of the Royal Meteorological Society, 125, 2887–2908, 1999.

Cloke, H. L., Pappenberger, F., Smith, P. J., and Wetterhall, F.: How do I know if I've improved my continental scale flood early warning system?, Environmental Research Letters, 12, 044 006, 2017.

de Groot, W. J., Field, R. D., Brady, M. A., Roswintiarti, O., and Mohamad, M.: Development of the Indonesian and Malaysian fire danger rating systems, Mitigation and Adaptation Strategies for Global Change, 12, 165–180, 2007.

Dee, D., Uppala, S., Simmons, A., Berrisford, P., Poli, P., Kobayashi, S., Andrae, U., Balmaseda, M., Balsamo, G., Bauer, P., et al.: The ERA-Interim reanalysis: Configuration and performance of the data assimilation system, Quarterly Journal of the Royal Meteorological Society, 137, 553–597, 2011.

Di Giuseppe, F., Molteni, F., and Dutra, E.: Real-time correction of ERA-Interim monthly rainfall, Geophysical Research Letters, 40, 3750–3755, 2013a.

Di Giuseppe, F., Molteni, F., and Tompkins, A. M.: A rainfall calibration methodology for impacts modelling based on spatial mapping, Quarterly Journal of the Royal Meteorological Society, 139, 1389–1401, 2013b.

Di Giuseppe, F., Pappenberger, F., Wetterhall, F., Krzeminski, B., Camia, A., Libertá, G., and San Miguel, J.: The potential predictability of fire danger provided by numerical weather prediction, Journal of Applied Meteorology and Climatology, 55, 2469–2491, 2016.

Di Giuseppe, F., Rémy, S., Pappenberger, F., and Wetterhall, F.: Improving Forecasts of Biomass Burning Emissions with the Fire Weather Index, Journal of Applied Meteorology and Climatology, 56, 2789–2799, 2017.

Di Giuseppe, F., Rémy, S., Pappenberger, F., and Wetterhall, F.: Using the Fire Weather Index (FWI) to improve the estimation of fire emissions from fire radiative power (FRP) observations, Atmospheric Chemistry and Physics, 18, 5359–5370, https://doi.org/10.5194/acp-18-5359-2018, https://www.atmos-chem-phys.net/18/5359/2018/, 2018.

Erickson, M. J., Colle, B. A., and Charney, J. J.: Evaluation and Postprocessing of Ensemble Fire Weather Predictions over the Northeast United States, Journal of Applied Meteorology and Climatology, 57, 1135–1153, https://doi.org/10.1175/JAMC-D-17-0180.1, https://doi.org/10.1175/JAMC-D-17-0180.1, 2018.

Ferro, C. A. and Stephenson, D. B.: Extremal dependence indices: Improved verification measures for deterministic forecasts of rare binary events, Weather and Forecasting, 26, 699–713, 2011.

Field, R. D., Wang, Y., Roswintiarti, O., et al.: A drought-based predictor of recent haze events in western Indonesia, Atmospheric Environment, 38, 1869–1878, 2004.

Field, R. D., van der Werf, G. R., and Shen, S. S.: Human amplification of drought-induced biomass burning in Indonesia since 1960, Nature Geoscience, 2, 185–188, 2009.

Flannigan, M. D. and Haar, T. V.: Forest fire monitoring using NOAA satellite AVHRR, Canadian Journal of Forest Research, 16, 975–982, 1986.

Giglio, L., Descloitres, J., Justice, C. O., and Kaufman, Y. J.: An enhanced contextual fire detection algorithm for MODIS, Remote sensing of environment, 87, 273–282, 2003.

Giglio, L., Randerson, J. T., and Werf, G. R.: Analysis of daily, monthly, and annual burned area using the fourth-generation global fire emissions database (GFED4), Journal of Geophysical Research: Biogeosciences, 118, 317–328, 2013.

Hersbach, H.: Decomposition of the continuous ranked probability score for ensemble prediction systems, Weather and Forecasting, 15, 559–570, 2000.

Hood, S. M.: Mitigating old tree mortality in long-unburned, fire-dependent forests: a synthesis, DIANE Publishing, 2010.

Kaiser, J., Heil, A., Andreae, M., Benedetti, A., Chubarova, N., Jones, L., Morcrette, J.-J., Razinger, M., Schultz, M., Suttie, M., et al.: Biomass burning emissions estimated with a global fire assimilation system based on observed fire radiative power, Biogeosciences, 9, 527–554, 2012.

Kaufman, Y., Ichoku, C., Giglio, L., Korontzi, S., Chu, D., Hao, W., Li, R.-R., and Justice, C.: Fire and smoke observed from the Earth Observing System MODIS instrument-products, validation, and operational use, International Journal of Remote Sensing, 24, 1765–1781, 2003.

Mass, C. F. and Ovens, D.: The Northern California Wildfires of 8–9 October 2017: The Role of a Major Downslope Wind Event, Bulletin of the American Meteorological Society, 100, 235–256, https://doi.org/10.1175/BAMS-D-18-0037.1, https://doi.org/10.1175/BAMS-D-18-0037.1, 2019.

Mölders, N.: Suitability of the Weather Research and Forecasting (WRF) model to predict the June 2005 fire weather for Interior Alaska, Weather and Forecasting, 23, 953–973, 2008.

Mölders, N.: Comparison of Canadian Forest Fire Danger Rating System and National Fire Danger Rating System fire indices derived from Weather Research and Forecasting (WRF) model data for the June 2005 Interior Alaska wildfires, Atmospheric Research, 95, 290–306, 2010.

Molteni, F., Buizza, R., Palmer, T. N., and Petroliagis, T.: The ECMWF ensemble prediction system: Methodology and validation, Quarterly journal of the royal meteorological society, 122, 73–119, 1996.

Nauslar, N., Abatzoglou, J., and Marsh, P.: The 2017 North Bay and Southern California Fires: A Case Study, Fire, 1, 18, 2018.

Orrell, D., Smith, L., Barkmeijer, J., and Palmer, T.: Model error in weather forecasting, Nonlinear processes in geophysics, 8, 357–371, 2001.

Palmer, T. N.: Predicting uncertainty in forecasts of weather and climate, Reports on Progress in Physics, 63, 71, 2000.

Pappenberger, F., Wetterhall, F., Dutra, E., Di Giuseppe, F., Bogner, K., Alfieri, L., and Cloke, H. L.: Seamless forecasting of extreme events on a global scale, pp. 3–10, missing, 2013.

Piani, C., Haerter, J., and Coppola, E.: Statistical bias correction for daily precipitation in regional climate models over Europe, Theoretical and Applied Climatology, 99, 187–192, 2010.

Raftery, A. E., Gneiting, T., Balabdaoui, F., and Polakowski, M.: Using Bayesian model averaging to calibrate forecast ensembles, Monthly weather review, 133, 1155–1174, 2005.

Richardson, D. S.: Skill and relative economic value of the ECMWF ensemble prediction system, Quarterly Journal of the Royal Meteorological Society, 126, 649–667, 2000.

Roads, J., Fujioka, F., Chen, S., and Burgan, R.: Seasonal fire danger forecasts for the USA, International Journal of Wildland Fire, 14, 1–18, 2005.

San-Miguel-Ayanz, J., Barbosa, P., Schmuck, G., Liberta, G., and Schulte, E.: Towards a coherent forest fire information system in Europe: the European Forest Fire Information System (EFFIS), Forest fire research and wildland fire safety. Millpress, Rotterdam, ISBN, pp. 90–77 017, 2002.

Schroeder, W., Prins, E., Giglio, L., Csiszar, I., Schmidt, C., Morisette, J., and Morton, D.: Validation of GOES and MODIS active fire detection products using ASTER and ETM+ data, Remote Sensing of Environment, 112, 2711–2726, 2008.

Stephenson, D., Casati, B., Ferro, C., and Wilson, C.: The extreme dependency score: a non-vanishing measure for forecasts of rare events, Meteorological Applications, 15, 41–50, 2008.

Taylor, S. W. and Alexander, M. E.: Science, technology, and human factors in fire danger rating: the Canadian experience., International Journal of Wildland Fire, 15, 121–135, 2006.

Van Wagner, C., Pickett, T., et al.: Equations and FORTRAN program for the Canadian forest fire weather index system, vol. 33, Canadian Forestry Service, Headquarters, Ottawa, 1985.

Van Wagner, C. et al.: Development and structure of the Canadian forest fire weather index system, vol. 35, Canadian Forestry Service, Headquarters, Ottawa, 1987.

Van Wagner, C. E. et al.: Structure of the Canadian forest fire weather index, Environment Canada, Forestry Service, 1974.

Vitolo, C., Di Giuseppe, F., Krzeminski, B., and San-Miguel-Ayanz, J.: A 1980–2018 global fire danger re-analysis dataset for the Canadian Fire Weather Indices, Scientific data, 6, 190 032, 2019.

Wooster, M. J., Roberts, G., Perry, G., and Kaufman, Y.: Retrieval of biomass combustion rates and totals from fire radiative power observations: FRP derivation and calibration relationships between biomass consumption and fire radiative energy release, Journal of Geophysical Research: Atmospheres (1984–2012), 110, 2005.

Wotton, B. M.: Interpreting and using outputs from the Canadian Forest Fire Danger Rating System in research applications, Environ Ecol Stat, 16, 107–131, 2009.

Zhu, J., Huang, B., Cash, B., Kinter, J. L., Manganello, J., Barimalala, R., Altshuler, E., Vitart, F., Molteni, F., and Towers, P.: ENSO Prediction in project Minerva: sensitivity to atmospheric horizontal resolution and ensemble size, Journal of Climate, 28, 2080–2095, 2015.

Zhu, Y., Toth, Z., Wobus, R., Richardson, D., and Mylne, K.: The economic value of ensemble-based weather forecasts, Bulletin of the American Meteorological Society, 83, 73–84, 2002.