Article

Verification of Red Flag Warnings across the Northwestern U.S. as Forecasts of Large Fire Occurrence

Joshua Clark [1],*, John T. Abatzoglou [2], Nicholas J. Nauslar [3] and Alistair M.S. Smith [4]

[1] Wildfire Division, Washington State Department of Natural Resources, 1111 Washington Street SE, Olympia, WA 98502, USA

[2] Management of Complex Systems, College of Engineering, University of California, Merced, 5200 North Lake Road, Merced, CA 95343, USA; jabatzoglou@ucmerced.edu

[3] Predictive Services, National Interagency Fire Center, 3833 Development Avenue, Boise, ID 83705, USA; nnaulsar@blm.gov

[4] Department of Forest, Rangeland, and Fire Sciences, College of Natural Resources, University of Idaho, 875 Perimeter Drive, Moscow, ID 83844, USA; alistair@uidaho.edu

* Correspondence: josh.clark@dnr.wa.gov

Abstract: Red Flag Warnings (RFWs) issued by the National Weather Service in the United States (U.S.) are an important early warning system for fire potential based on forecasts of critical fire weather that promote increased fire activity, including the occurrence of large fires. However, verification of RFWs as they relate to fire activity is lacking, thereby limiting means to improve forecasts as well as increase value for end users. We evaluated the efficacy of RFWs as forecasts of large fire occurrence for the Northwestern U.S.—RFWs were shown to have widespread significant skill and yielded an overall 124% relative improvement in forecasting large fire occurrences than a reference forecast. We further demonstrate that the skill of RFWs is significantly higher for lightning-ignited large fires than for human-ignited fires and for forecasts issued during periods of high fuel dryness than those issued in the absence of high fuel dryness. The results of this first verification study of RFWs related to actualized fire activity lay the groundwork for future efforts towards improving the relevance and usefulness of RFWs and other fire early warning systems to better serve the fire community and public.

Keywords: fire weather; forecast verification; fire danger forecasting

1. Introduction

Wildland fire plays an important role as a natural and increasingly anthropogenic disturbance found in most vegetated ecosystems globally [1,2] and generally serves to promote healthy, resilient landscapes [3]. However, when fire threatens to directly impact human life and property, it may be considered hazardous [4], prompting land management action to mitigate impacts on ecosystems and communities [5,6]. In recent decades, longer periods of critical fire weather [7] coupled with human-caused fire activity [8] and increased fuel loads as a result of fire suppression activities [9] have expanded the threat of hazardous fires in the United States (U.S.), making land management objectives more difficult and costlier to achieve, while also placing the safety of fire suppression personnel and the public at greater risk. These effects have been demonstrated by recent fire events in the U.S. [10–12] and are expected to continue due to anthropogenic climate change [4,13] and the growth of wildland–urban interfaces [14].

Many early warning systems have been developed for fire hazards, including the development of numerous fire danger indices [15]. Comprehensive fire danger systems integrate weather, fuels,

and climate information to generate daily predictions regarding the fire environment and fire behavior characteristics [16–18]. These outputs are generally considered in fire management plans as actionable criteria that prompt some prevention, preparedness, or resource allocation decision [15] and have seen widespread adoption in wildland fire management communities in North American and Australasia [16,17,19,20], with other countries like the United Kingdom in the process of developing similar tools [21]. In the U.S., systems such as the Severe Fire Danger Index and Santa Ana Wind Threat Index, leverage advancements in gridded meteorological forecast data along with fuels and fire behavior information to provide local-to-regional predictions of fire danger [22,23], while other systems predict fire danger directly from fire weather conditions, independent of fuels [24,25]. Similarly, utility providers incorporate a matrix of meteorological and fuel moisture criteria to de-energize electrical lines as a fire-prevention strategy [26]. While these systems serve as important tools for understanding hazardous fire risk, many are generated independent of real-time human input that may be important for deciphering rapidly changing fire environment factors that drive the most extreme fire activity.

Red Flag Warnings (RFWs), which are issued by meteorologists at the U.S. National Oceanic and Atmospheric Administration (NOAA) National Weather Service (NWS), are forecasts of critical fire weather conditions that could lead to abundant new ignitions and/or rapid spread and growth on existing fires (i.e., extreme fire behavior), but are not actually forecasts of fire activity (Figure 1) [27]. Forecasters at local weather forecast offices (WFOs) issue RFWs on short timescales (imminent to 48-hr lead times) and consider numerous quantitative and qualitative meteorological parameters (e.g., relative humidity, wind speed, atmospheric stability, dry frontal passage, lightning) along with some measure of fuel dryness. These parameters are considered as criteria for RFW issuance (Table 1) based on the expert opinion and analyses at WFOs and vary geographically by sub-regional fire weather zones (FWZs). Although this arrangement allows for greater flexibility to accommodate local conditions conducive to area fire initiation and growth, it has resulted in many different criteria across FWZs, complicating the messaging intent of these forecasts. Furthermore, limited studies have been conducted to verify the performance of RFWs and we are not aware of any peer-reviewed studies that discuss possible refinements to these forecasts.

```
URGENT - FIRE WEATHER MESSAGE
NATIONAL WEATHER SERVICE SPOKANE WA
440 AM PDT FRI AUG 28 2015

...RED FLAG WARNING FOR VERY WINDY CONDITIONS AND LOW RELATIVE
HUMIDITIES SATURDAY...


IDZ101-WAZ684-686-687-282300-
/O.NEW.KOTX.FW.W.0020.150829T1800Z-150830T0300Z/
NORTHERN AND CENTRAL IDAHO PANHANDLE (ZONE 101)-
EAST WASHINGTON OKANOGAN/METHOW VALLEYS (ZONE 684)-
EAST WASHINGTON NORTHEAST (ZONE 686)-
EAST WASHINGTON OKANOGAN HIGHLANDS (ZONE 687)-
440 AM PDT FRI AUG 28 2015

...RED FLAG WARNING IN EFFECT FROM 11 AM TO 8 PM PDT SATURDAY FOR
WIND AND LOW RELATIVE HUMIDITY FOR NORTHERN WASHINGTON AND NORTH
IDAHO...

THE NATIONAL WEATHER SERVICE IN SPOKANE HAS ISSUED A RED FLAG
WARNING FOR WIND AND LOW RELATIVE HUMIDITY...WHICH IS IN EFFECT
FROM 11 AM TO 8 PM PDT SATURDAY.

* AFFECTED AREA: FIRE WEATHER ZONE 101 NORTHERN AND CENTRAL
  IDAHO PANHANDLE (ZONE 101)...FIRE WEATHER ZONE 684 EAST
  WASHINGTON OKANOGAN/METHOW VALLEYS (ZONE 684)...FIRE WEATHER
  ZONE 686 EAST WASHINGTON NORTHEAST (ZONE 686) AND FIRE WEATHER
  ZONE 687 EAST WASHINGTON OKANOGAN HIGHLANDS (ZONE 687).

* WINDS: SOUTHWEST 25 TO 35 MPH WITH GUSTS UP TO 60 MPH. GUSTS AS
  HIGH AS 70 MPH IN THE MOUNTAINS.

* RELATIVE HUMIDITIES: 20 TO 30 PERCENT IN THE VALLEYS. 30 TO 40
  PERCENT IN THE MOUNTAINS.

* IMPACTS: VERY WINDY CONDITIONS AND LOW RELATIVE HUMIDITIES WILL
  RESULT IN RAPID GROWTH ON EXISTING FIRES OR ANY NEW IGNITIONS.

PRECAUTIONARY/PREPAREDNESS ACTIONS...

A RED FLAG WARNING MEANS THAT CRITICAL FIRE WEATHER CONDITIONS
ARE EITHER OCCURRING NOW...OR WILL SHORTLY. A COMBINATION OF
STRONG WINDS AND LOW RELATIVE HUMIDITY WILL CREATE EXTREME FIRE
GROWTH POTENTIAL.
```

Figure 1. An example of a Red Flag Warning issued by the National Weather Service (NWS) Spokane forecast office for windy and dry conditions on 28 August 2015.

Table 1. Select criteria for issuing Red Flag Warnings (RFWs) at different weather forecast offices (WFOs) taken from the 2019 Northwest Area Fire Weather Annual Operating Plan available at https://gacc.nifc.gov/nwcc/content/pdfs/2020_Region_6_AOP_public.pdf.

WFO Pendleton All FWZs	"LIGHTNING: Abundant lightning in conjunction with sufficiently dry fuels (fuels remain dry or critical during and after a lightning event). Warnings are not typically issued for isolated coverage events. Warnings not typically issued for events that will be accompanied by significant rain (greater than 0.25 inches). However, if a lightning event will occur with significant rain, but is then followed by very hot and dry conditions, a warning may be issued if holdover/sleeper fires are a concern."
WFO Portland FWZs 605, 607, and 660	"One station (RAWS) must report 35% humidity or less AND 10-minute wind speed of 10 mph AND/OR gusts to 20 mph or more for four hours in an 8- hour block, AND at least TWO other stations reporting 35% humidity or less AND 10- minute wind of 10 mph AND/OR gusts to 20 mph for at least TWO hours. Key RAWS: Horse Creek, Log Creek, Wanderer's Peak, Kosmos, Canyon Creek, Orr Creek, Elk Rock, and 3-Corner Rock. NOTE: Includes stations from zone 659." (Only valid during nighttime hours)
WFO Spokane All FWZs	"An unusually unstable atmosphere. This would be associated with a strong thermal trough which typically forms along the east slopes of the Washington Cascades."

RFWs are similar to other NWS early warning products for natural hazards (e.g., tornadoes, severe thunderstorms, flash floods) as they are issued for high-impact events that have the potential to threaten life and property. Many verification studies have been conducted on the performance of these other hazard forecasts to increase the quality of the system and value to end users [28–31]. The NWS maintains internal verification statistics for RFWs, but these are computed for weather criteria—not the occurrence of hazardous fires (unlike warnings for tornadoes or flash floods)—and therefore may have reduced value for end users. RFWs are unique warning forecasts not only in this regard, but also because the issuance of RFWs may actually prevent the hazard from occurring in the case of human-caused fires or through increased situational awareness of fire suppression during RFWs. We hypothesize that fire prevention efforts tied to the issuance of RFWs reduce fire hazards and result in reduced quantitative performance for RFWs, as measured through the lens of fire activity.

Here, we consider occurrences of new large fires (LFs) as hazardous fires and treat RFWs as dichotomous forecasts of these LF occurrences across subregions of the Northwestern U.S. Since LF occurrence is complicated by many non-meteorological factors, we examine how RFW performance varies across different LF size thresholds, fire causes, and land cover. Finally, we assess how forecast performance varies as a function of relative fuel dryness, which, while being a criterion considered in many RFWs, is used heterogeneously across zones. While some studies have shown the importance of meteorological conditions tied with RFW criteria on fire activity [32], this study provides a first known effort to evaluate the added value of RFW forecasts for actualized LF activity that is of key importance for fire suppression. The results of this study may help to both refine RFWs and identify reasons why forecast performance and skill varies across different regions and LF characteristics.

2. Materials and Methods

2.1. Study Location

The study area is composed of the county warning areas of eight NWS WFOs of the northwest contiguous U.S.: Seattle (SEW), Spokane (OTX), Missoula (MSO), Pocatello (PIH), Boise (BOI), Medford (MFR), Pendleton (PDT), and Portland (PQR) (Figure 2). This covers a diverse set of regions we herein refer to as the Northwest, including the coastal areas of Washington and Oregon, the Cascade Mountains and Blue Mountains, much of the Northern Rockies, the Columbia Basin, and portions of the Great Basin. Fire season in the Northwest is generally shorter and more well-defined compared to other fire-prone regions across the U.S. [33], with ~86.1% of all wildfires and ~99.7% of total burned area occurring in the months of June–October from 2006 to 2015. Across the region, the number of lightning-ignited fires and human-ignited fires is nearly the same, although lightning fires comprise

~83.6% of the total burned area while human-ignited fires account for only ~16.42% of the total burned area [34]. Smaller and primarily human-ignited fires are more common west of the Cascade Mountains due to less lightning frequency [35,36] and a larger wildland–urban interface component [8].

2.2. Datasets

Spatial data of RFWs were obtained from the Iowa State University Iowa Environmental Mesonet (IEM) archive of NWS watches and warnings [37]. This dataset includes the WFO, FWZ, issuance date/time, and expiration date/time with each warning. We simplified RFWs that span multiple days into records for each day so that if an RFW was active for any portion of a calendar day it was considered an RFW day for that FWZ. The resulting dataset contains 8940 RFW forecast days between 2006 and 2015.

Point location fire records for the same 10-year period were obtained from the Fire Program Analysis (FPA) fire occurrence database [34]. These data include the fire discovery date, final fire size, and fire cause information, aggregated from numerous federal, state, and local fire reporting systems. Records were reduced to the study area, resulting in 64,122 fires that were assigned to the corresponding FWZ that existed at the time of the fire discovery date. This latter point is important as FWZ boundaries have changed throughout the entire period due to anticipated improvements in the quality of forecasts and the ability to adequately warn affected populations [38]. We designated LFs for each FWZ as the largest 10% of fires within each zone [39,40], but also assess the sensitivity of the results in relation to fire size by considering the largest 5% and 20% of fires. Rather than quantifying skill in terms of the number of individual LFs, we identified LF days (LFDs) on the basis of having at least one LF occurrence within a FWZ on a given day. This means that multiple LFs occurring within a FWZ for a given day were not considered separate events, which would otherwise result in misleading performance statistics since RFWs are issued for the FWZ. The resulting number of LFDs across the entire study region were 10,402, 5341, and 2775 when fire sizes were ≥ 80th, 90th, 95th percentile sizes, respectively. For the purpose of tabulating LFs, we assumed that fires reached size thresholds on the date of discovery as detailed fire progression information was unavailable.

We additionally explore how RFW performance varies by fire cause, land cover, and a measure of relative fuel dryness. These subsequent tests are constrained to LFs above the 90th percentile size threshold. To more effectively determine performance for just human- and lightning-caused fires, we eliminated those which had an 'unknown' cause, leaving 2309 lightning-caused and 2890 human-caused LFDs. However, we retained these 'unknown' caused fires in the other tests to maximize the frequency of events available for computing performance statistics. We then classified LFs as burning primarily in forested lands and those primarily burning in non-forested lands using the 0.5-km moderate resolution imaging spectroradiometer (MODIS)-based global land cover climatology dataset [41]. Due to the absence of fire perimeters, we approximated perimeters by assuming a circular fire with an area equivalent to the final fire size. Fires which had >50% pixels classified as forested were assigned as such and those <50% were assigned as non-forested vegetation. A total of 3073 LFDs were classified as forested while the remaining 2466 were considered non-forested.

Lastly, energy release component (ERC) percentile values were assigned to each fire using co-located ~4-km gridded data from gridded meteorological data (gridMET) [42] to represent fuel dryness. These percentiles were calculated by pooling data for the entire calendar year during the 2006–2015 period. ERC is defined as the total available energy within the flaming front of a fire calculated from the U.S. National Fire Danger Rating System (NFDRS), here using Model G (dense conifer) as the fuel model input. ERC is often used in fire business decision making as it is a good measure of cumulative fire danger because it gives higher weighting to heavier fuel types that tend to reflect seasonal drying trends [43], and is not as sensitive to daily fluctuations in fire weather variables as shown in other NFDRS indices (e.g., burning index). Several studies have shown a strong relationship between interannual fire activity and ERC [44,45]. Most commonly, the locally defined 90th percentile ERC threshold is used to represent high fire danger across particular geographic regions [22,46,47], and we adopt this threshold in our analysis. A total of 2160 LFDs were classified as ERC ≥ 90th percentile while 3307 were classified as ERC < 90th percentile.
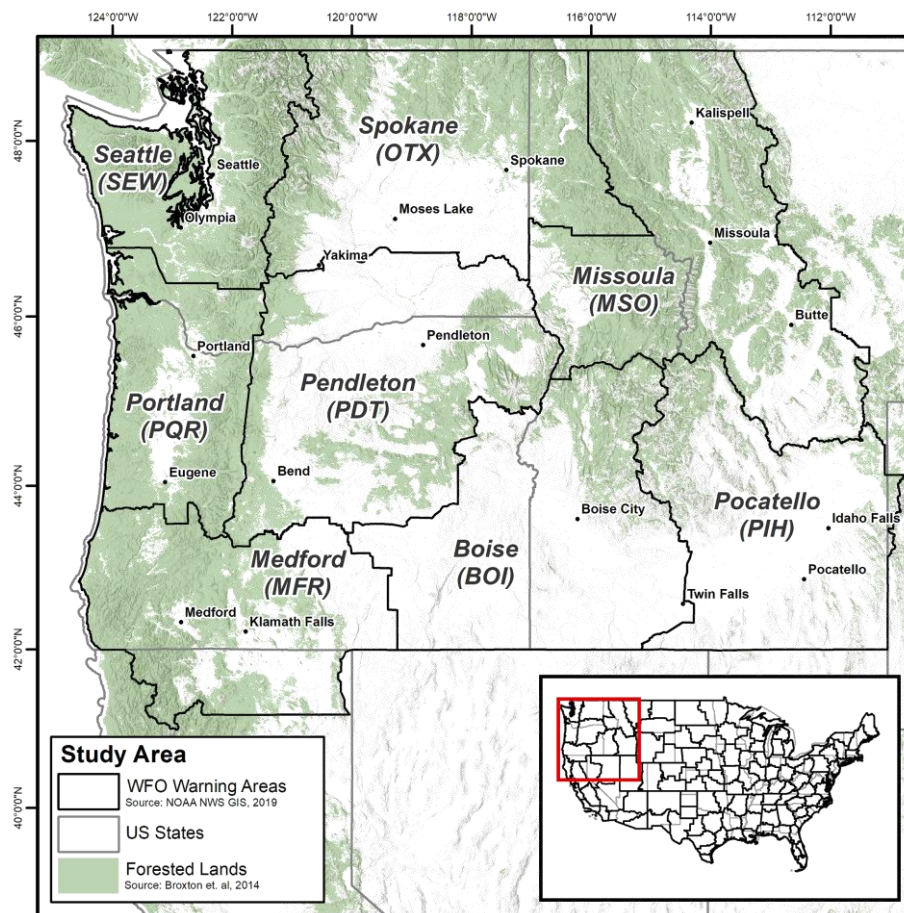
Figure 2. Map of the study area with forestlands defined using the 0.5 km moderate resolution imaging spectroradiometer (MODIS)-based global land cover climatology dataset [41] and NWS weather forecast office (WFO) warning areas based on fire weather zones.

2.3. Analysis Methods

To quantify forecast performance, we treated RFWs as dichotomous forecasts for LFDs and constructed a contingency table that shows the frequency of forecasts and LFDs. The contingency table is a standard forecast verification tool used to compute performance statistics of nonprobabilistic forecasts of discrete predictands [48–50] and has the general form:

		Observed Event?	
		Yes	No
Forecast Issued?	Yes	Correct Positive (a)	False Alarm (b)
	No	Miss (c)	Correct Negative (d) *

* Not calculated.

Here, correct positive forecasts (i.e., hits) are RFWs where an LFD occurred on or within one day following the forecast in the same FWZ. This 2-day period was used to accommodate typical delays in fire reporting and is consistent with other studies that consider meteorological conditions immediate to fire occurrence using LF databases [51]. If an RFW was issued but no LFs were observed coincident with or one day after the forecast date, the forecast is classified as a false alarm. A forecast is classified as a miss when there was no forecast issued for a FWZ during the 2-day period but an LFD occurred. We acknowledge that considering a 2-day period in the compilation of our performance statistics differs from traditional methods of measuring forecast skill of meteorological elements. Moreover, our treatment of RFW forecast skill is specific to the occurrence of new LFDs; RFWs issued coincident with days of rapid growth on existing fires in the absence of new LF are treated as false alarms.

Since RFWs are forecasts of high-risk, statistically rare events (similar to tornadoes, flash floods) calculating correct negative forecasts are of little value due to the overwhelming amount of days where no event was observed or forecast [29,30,52]. Thus, we omit correct negatives in our analysis [53]. An example of this classification scheme for a particular day is given in Figure 3. On 10 August 2015, 32 RFWs (shown as hits or false alarms) were issued and LFs which occurred on 10–11 August were used to classify forecasts. The resulting number of hits, misses, and false alarms for this RFW day were 11, 11, and 21, respectively. Table S1 summarizes the number of RFW days as well as LFDs over the 10-year dataset for each WFO.
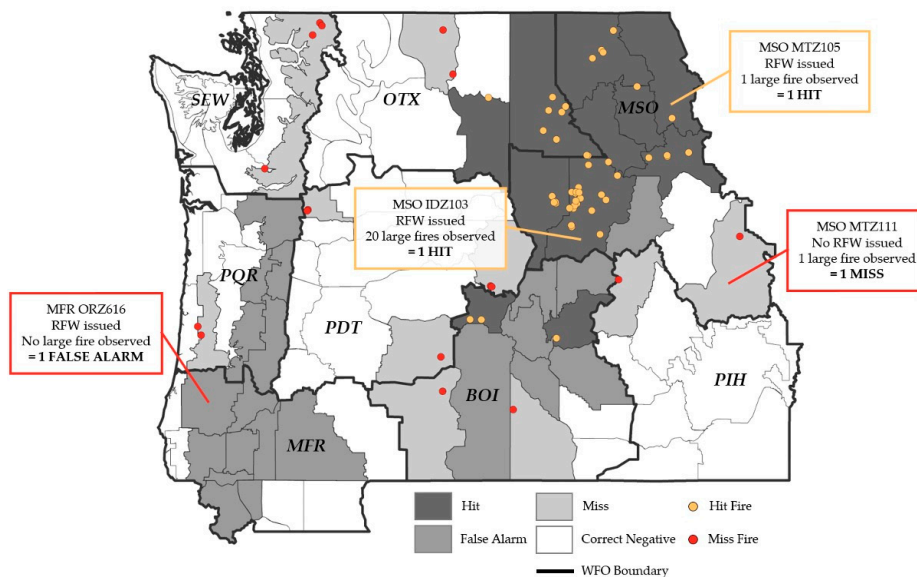

Figure 3. An example of the RFW and large fire day classification scheme pairing RFWs issued for 10 August 2015 and new large fires from 10–11 August 2015. Thirty-two RFWs were issued, resulting in 11 hits, 11 misses, and 21 false alarms across the study area.

From the contingency table, we computed the following performance measures: probability of detection (POD), false alarm ratio (FAR), and critical success index (CSI). The relevant equations and definitions of these metrics are listed in Table 2. These measures are frequently reported together in the case of rare event forecasts because of a shared lack of consideration given to correct negatives [54]. Furthermore, if we consider the formulae in Table 2, we can rearrange terms to show that CSI is a nonlinear function of POD and FAR given by

$$CSI = \left[(POD)^{-1} + (1 - FAR)^{-1} - 1 \right]^{-1},$$

For rare events such as LFs, it is important to note that values will seldom approach CSI = 1 (and may instead be much closer to CSI = 0) due to the decreased frequency of events and increased frequency of times where no event was forecast or observed.

Most commonly, forecast systems are evaluated against persistence forecasts, random forecasts, or climatological forecasts by determining the 'skill score' of the compared forecasts for a particular performance metric [50]. These skill scores provide an analytical guide to measure the added value of a forecast relative to a reference forecast. In its most generic form, the skill score with respect to a performance metric (*M*) is defined as [55]

$$M_{ss} = \frac{M_{FORECAST} - M_{REFERENCE}}{M_{PERFECT} - M_{REFERENCE}},$$

where $M_{FORECAST}$ is the performance metric (such as POD or CSI) of the forecast system being evaluated, $M_{REFERENCE}$ is the same metric but for the reference forecast, and $M_{PERFECT}$ is the scalar value of a

perfect forecast for that particular metric (for example, $POD_{PERFECT}$ = 1 if POD was being evaluated). If M_{SS} > 0, improvement over the reference forecast can be inferred.

Table 2. Forecast metrics calculated in this study.

Performance Metric	Measures	Equation	Possible Range and Perfect Score
Probability of detection (POD)	Fraction of observed events that were correctly forecast	$POD = \frac{a}{a+c}$	0–1 1
False alarm ratio (FAR)	Fraction of predicted events that did not actually occur	$FAR = \frac{b}{a+b}$	0–1 0
Critical success index (CSI)	Fraction of correctly forecasted events without consideration to correct negatives	$CSI = \frac{a}{a+b+c}$	0–1 1

Since RFWs and LFDs within the study area are generally confined to the summer months, reference forecasts need to preserve this seasonality for fair comparisons. Climatological and persistence forecasts are commonly used as reference forecasts for skill score calculations, particularly for continuous variables [56–58]. For our case, a persistence-based reference would not be suitable given the serial correlation of these forecasts from one day to the next that would make it difficult to separate the skill of the actual forecasts from the reference forecast [59]. A randomly generated forecast provides a reference completely independent of the observations (a truly no-skill forecast) but may sacrifice the seasonal relationship we are trying to preserve if the forecasts are allowed to be drawn from any time in the year. Thus, we define a 'random climatology' forecast to achieve an independent set of reference forecasts that retain a similar seasonal distribution to the actual RFWs. To achieve this, a resampling procedure was implemented where actual RFW days were resampled for any date within ± 15 days among all years of the study period, resulting in a reference set of the same size. These reference forecasts were then assessed against the observations using the same procedures and performance metrics as before and skill scores were computed against the actual forecasts. This process was repeated one hundred times to obtain a robust sample of statistics and skill scores and median values were selected for presentation in the results. To test the statistical significance of RFW forecast skill versus the reference, we reject the null hypothesis of RFW skill being the same as the reference forecast if 95% of the sample skill scores (M_{SS}) are greater than zero.

We also report results in terms of relative improvement between the RFWs and reference as another measure of skill since skill scores may be unintuitive to fire managers and other readers external to the weather forecasting community. Relative improvement is expressed as the percentage difference from the reference forecast as:

$$Rel.\ Improvement\ (\%) = \frac{M_{FORECAST} - M_{REFERENCE}}{M_{REFERENCE}} \times 100,$$

In addition to reporting performance metrics and skill for each WFO, we compute area-wide statistics by aggregating the number of hits, misses, and false alarms from each WFO into one contingency table representing all eight offices.

3. Results

3.1. Performance as Forecasts of LF Occurrence

RFWs exhibited skill as forecasts of LFDs across WFOs in the Northwestern U.S. (Figure 4). Area-wide POD was 0.29, while POD$_{SS}$ was 0.18 and POD relative improvement was 124.1% over the reference forecast. Six of the eight WFOs demonstrated >100% POD relative improvements. POD was the lowest for two offices covering the populated and mesic portions of the study area (SEW and PQR), although relative improvements in capturing LFDs in these areas were high. Herein, all forecasts were found to have significant skill unless otherwise noted.
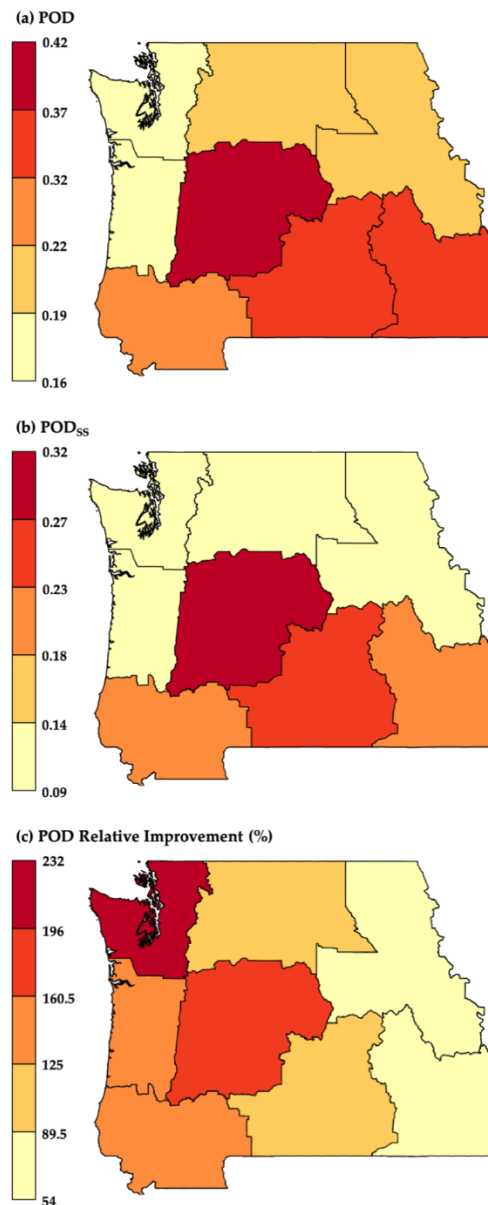
Figure 4. Red Flag Warning skill for large fire days (≥ 90th percentile) for each WFO showing (**a**) POD, (**b**) POD_{SS}, and (**c**) relative improvement over the random climatology reference forecast.

3.2. Fire Size

Area-wide POD increased with larger fire size thresholds, although this was countered by increased FAR values due to decreasing event frequency (Table 3; Table S2a–c). Similarly, area-wide skill scores and relative improvement increased with larger fire size thresholds; POD_{SS} increased from 0.13 for 80th percentile LFDs to 0.23 for 95th percentile LFDs, while the area-wide relative improvement in POD increased from 99.4% for 80th percentile LFDs to 138.2% for 95th percentile LFDs. Differences in POD as a function of fire size varied among the WFOs; across all fire size thresholds, the highest performance was shown for regions with the largest fire sizes (BOI, PIH, and PDT) while a notably lower performance was observed for regions west of the Cascade Mountains (PQR and SEW). Five WFOs demonstrated relative improvements over the reference forecast > 100% for LFDs regardless of fire size threshold.

Table 3. Area-wide Probability of Detection (POD), POD skill scores, and POD relative improvement in Red Flag Warnings when large fire day (LFD) was defined at 80th, 90th, and 95th percentile fire size thresholds.

	LF Size ≥ 80th Perc.	LF Size ≥ 90th Perc.	LF Size ≥ 95th Perc.
POD	0.24	0.29	0.34
POD$_{SS}$	0.13	0.18	0.23
Relative Improvement	99.4%	124.1%	138.2%

3.3. Fire Cause

Area-wide POD, POD$_{SS}$, and the relative improvement in POD were notably higher for lightning-caused LFDs than for human-caused LFDs (Figure 5a). For example, the area-wide POD was 0.46 for lightning-caused LFDs and 0.17 for human-caused LFDs. Furthermore, POD$_{SS}$ was 0.34 for lightning-caused LFDs and only 0.08 for human-caused LFDs. While skill scores for human-caused LFDs were low, they showed substantial relative improvement (~78.4%) in POD over the reference forecast.
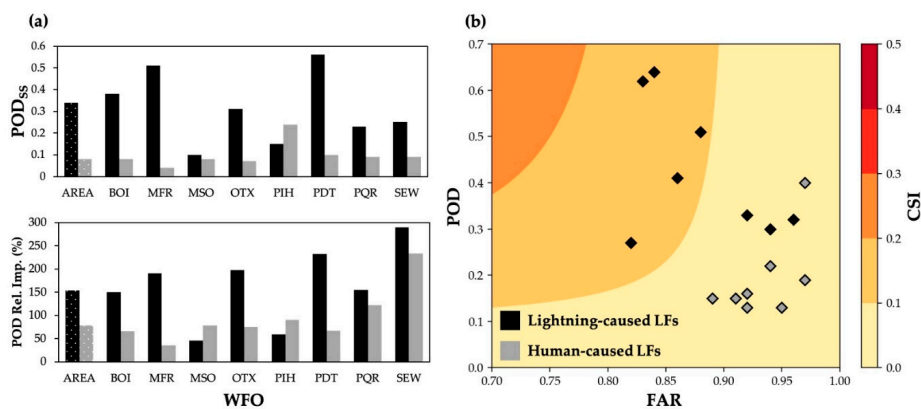

Figure 5. POD skill scores and relative improvement (**a**) and performance metrics (**b**) of RFWs for lightning- and human-caused large fire days across the WFOs. For (**b**), critical success index (CSI) is found to generally increase with higher POD and decreasing false alarm ratio (FAR).

Similar to the area-wide results, we generally found higher skill for lightning-caused LFDs than human-caused LFDs across WFOs (Table S3a,b). Figure 5b shows this as points for each WFO where higher POD and lower FAR result in increased CSI. WFOs with the fewest lightning-caused LFDs (PIH, PQR, and SEW) showed the lowest POD and highest FAR. Although more human-caused LFDs occurred across the region, FAR values for these fires were higher than lightning-caused LFDs for five of the eight WFOs. Skill scores demonstrated improvements for both fire-cause types across regions, although the improvement among human-caused LFDs was low.

3.4. Land Cover Type

We found little discernible difference in performance and skill scores calculated for LFDs stratified by forest and non-forest land cover (Table 4). Area-wide POD was slightly higher for non-forested LFDs but FAR and CSI were nearly the same. Both land cover types showed area-wide relative improvements > 125% above the reference. We found statistically significant skill for all RFWs by land cover except for non-forested LFDs in SEW. The results for individual WFOs are provided in Table S4a,b.

3.5. Fire Danger

RFW forecasts conditioned on being coincident with high fire danger (ERC ≥ 90th percentile, area-wide POD of 0.42) exhibited superior skill to RFW forecasts coincident with lesser fire danger (ERC < 90th percentile, area-wide POD of 0.23) (Figure 6a). Furthermore, area-wide POD$_{SS}$ coincident with high fire danger was nearly double that of RFW, coincident with lesser fire danger. Area-wide

POD relative improvement over the reference was 142.1% for LFDs with high fire danger and 116.2% for those LFDs with lesser fire danger.

Table 4. As in 3a but for forested and non-forested large fire land cover types using a 90th percentile fire size threshold.

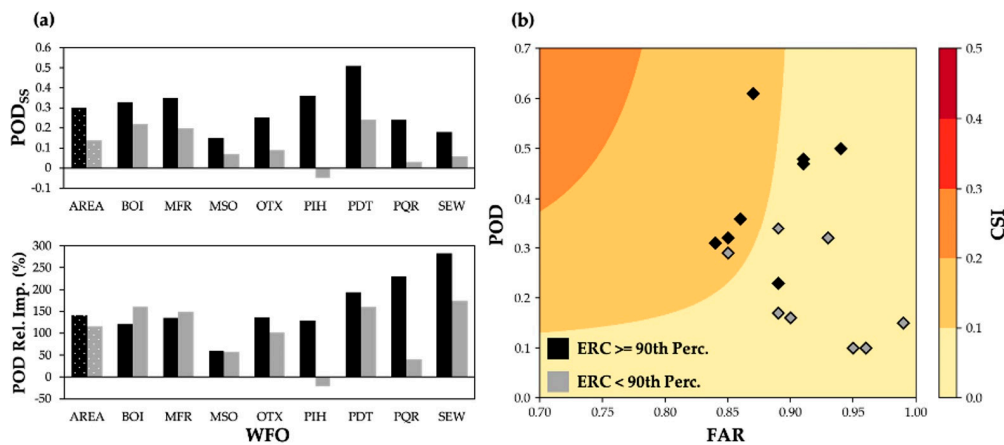|	Forested LFD	Non-Forested LFD
POD	0.28	0.34
POD$_{SS}$	0.18	0.23
Relative Improvement	129.0%	131.3%

Figure 6. As in Figure 5, but for RFWs when large fire days occurred with ERC ≥ 90th percentile and ERC < 90th percentile values.

POD and POD$_{SS}$ were ubiquitously higher for RFWs coincident with high fire danger than lesser fire danger across WFOs (Figure 6a, Table S5a,b). Similar to findings for differences in skill metrics between human- and lightning-caused LFDs, we showed improvements in POD, FAR, and CSI between RFWs issued during lesser fire danger and during high fire danger across WFOs (Figure 6b). For example, we found a POD of 0.61 at for the Pendleton WFO during high fire danger, which is much higher than the POD during lesser fire danger (0.34) and showed a 195% improvement over the reference forecast. We found statistically significant skill for all RFWs conditioned by fire danger except for RFWs issued by PIH, coincident with the ERC < 90th percentile.

4. Discussion

4.1. Performance in the Context of Rare Event Forecasting

We found that RFWs have skill as forecasts for the occurrence of new LFDs across the Northwestern U.S. and further demonstrate substantial improvement from reference forecasts (Figure 7). This is an important finding that indicates the added value RFWs provide to fire managers and the public. As was expected, overall performance metrics were low due to the rare nature of LFDs, and the fact that we constrained our definition of fire activity to new LFs rather than also accounting for growth on existing fires or the number of new ignitions. In addition, forecasts of high-risk, rare events may be prone to hedging, where the cost of a missed forecast exceeds the forecaster's risk tolerance, leading to the issuance of more forecasts and a greater number of false alarms [60,61]. Other forecasts of high-risk, rare events such as flash floods [30] and earthquakes [62,63] similarly demonstrate the consequences of hedging and lower forecast skill. We mitigate biases introduced through hedging by demonstrating a substantial relative improvement in RFWs compared with reference forecasts. However, complementary analyses show that the majority of days with new LF starts do not coincide with the issuance of RFW across WFOs, highlighting the challenges inherent in the stochastic nature of fire occurrence (Table S1).
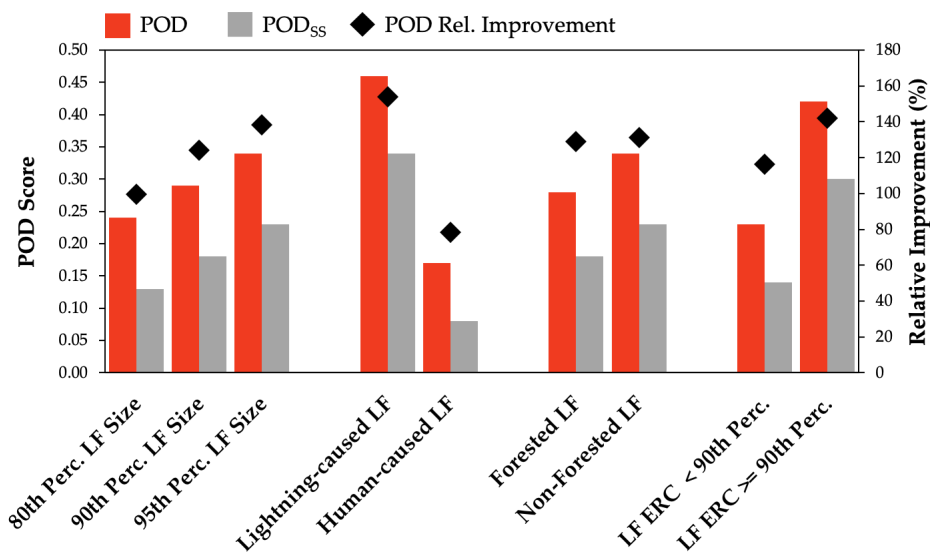
Figure 7. Summary of area-wide POD, POD$_{SS}$, and POD relative improvement for the four classifications of large fire day types.

4.2. Lightning- and Human-Caused Fires

The majority (55.6%) of LFs across the study area were human-caused. However, the performance of RFWs for the occurrence of human-caused LFDs was quite low with only one WFO having POD > 0.25. Conversely, performance was notably higher for days with lightning-ignited large fires with all WFOs having POD > 0.25 and three WFOs having POD > 0.5. The interaction between lightning and fire occurrence is well understood [36], although there is some debate about which factors are most critical for determining fire ignition potential [64]. Previous research has shown that the presence of dry thunderstorms, low fuel moisture, and fuel type impact the ignition efficiency of lightning [65]. The probability of dry thunderstorms as agents of fire ignition is often included in the issuance of RFWs. Improvements in dry thunderstorm forecasting and continued research on lightning characteristics (e.g., polarity, residence time, and amplitude) for ignition potential (see [66]) are likely to increase RFW performance for lightning-caused fires.

By contrast, RFWs do not explicitly include predictors of human ignitions. The lesser skill of RFW for capturing human-caused LFDs is consistent with prior research that shows that human-caused fires are more difficult to predict [67,68] and occur over a more diverse set of fuel moistures and broader period of the year than lightning-caused fires [8]. Other factors beyond fuels, weather, and topography tend to influence human-caused ignitions such as road, population, and railroad density, days of the week, holidays, and socioeconomic status [69].

The issuance of RFWs may alter human activity resulting in degraded skill. RFWs for non-lightning events (e.g., hot and dry conditions, high winds) may act as preventative measures for fire managers and the public that reduce the number of new large human-caused fires. For example, the issuance of RFWs can promote action by local land agencies to restrict campfire usage, limit silvicultural and agricultural burning, and bolster suppression capability in the affected areas. In this case, are false alarms actually indicative of a successful forecast? Contrarily, there are claims that illegal burners and arsonists may view RFWs as a window of opportunity to maximize their efforts, although research suggests this may not be well-founded [70]. Collectively, these factors highlight the reasons for the reduced performance and skill of RFWs for human-caused LFDs. More research is needed here to discern the efficacy of RFWs in fire prevention efforts.

4.3. Fuel Dryness as a Prerequisite for RFW Issuance

We found improved RFW skill for new LFDs conditioned on fuel dryness. LFs generally occur when fuels are more receptive due to weather and climate drivers [45,51] and RFWs are intended

to consider some measure of fuel dryness. Our results reinforce the added value of RFW forecasts that explicitly integrate objective measures of fuel dryness. For this study, we chose ERC because of its representativeness of seasonal drying trends and widespread usage by regional fire management, although we acknowledge that other fire danger indices may be more appropriate for different geographic areas or times of the year depending on the dominant drivers of fire activity for the expected event. For example, in western Washington, downslope wind events that bring hot and dry winds from the Columbia Basin across the Cascades can occur throughout the year and are a known critical fire weather pattern when they co-occur with dry fuels [33,71]. These events typically occur on short timescales and thus the fuels response and fire risk may be better resolved by 10-/100-hour fuel moisture values than ERC. Additional studies that examine a variety of fuel aridity metrics throughout the year may aid in the efficacy of RFWs.

4.4. Limitations of Assessing and Interpreting Skill of RFWs

RFWs are issued for areas of relatively homogenous climate and fuels. We caution that comparisons of performance between WFOs need to consider the climate, fuels, and frequency of events that differ markedly across FWZs. Furthermore, WFOs with larger FWZ sizes could show artificially better skill because of a larger pool of LFDs. Previous studies have shown similar results where performance increases with the scale of the geographic area considered, but tends to result in decreased value to the end user [63]. Additional analyses that examine the specific criteria for RFW by zones as well as the host of biophysical and human factors may help shed light on differences in forecast skill.

Lastly, we considered RFWs as forecasts of new LFDs, although, in reality, RFWs may be issued for weather conditions (e.g., atmospheric instability, high winds, low relative humidity) that promote growth on existing fires and heighten the potential for rapid rates of fire spread for new ignitions. Our explicit treatment of RFWs as forecasts for new LFDs results in a low estimate of skill as we classify RFW days that may have rapid growth on existing fires but no new LFs as false alarms. Recent geospatial datasets of daily fire incident status reports (SIT-209s, [72]) and burned area information from satellite imagery [73] aim to provide fire growth information for larger fires and thus may be useful for the future evaluation of RFWs and other fire hazard warning systems.

5. Conclusions

We found the skills of RFWs for meaningful measurements of fire activity broadly across the Northwestern U.S. We additionally demonstrated that RFW skill was significantly better for lightning-caused LFDs and when issued coincident to high fuel dryness (area-wide POD relative improvements of 154% and 142.1% over our reference forecast, respectively). While this is the first known study on RFW skill married to actual fire activity, our measures of skill are specific only to the occurrence of new, large fires. However, these forecasts may have value for fire early warning systems beyond those reported here.

Our results provide a means for discussing the quality of RFWs and highlight recommendations for improving RFWs while preserving value to end users. The definition of RFWs and criteria used to issue these forecasts should be explicit, centrally documented, and easily verifiable. We discovered many different WFO interpretations of the RFW definition and found numerous qualitative RFW criteria that are challenging to verify directly. RFW criteria should include a measure of fuel dryness (e.g., ERC, 1000-hr fuel moisture) and be flexible enough to accommodate different weather regimes that drive fire activity throughout the year. Furthermore, improved empirical analyses of the weather and fuels conditions that lead to significant ignitions and rapid rates of spread need to be performed to determine appropriate local RFW criteria, and these criteria should be quantifiable and concise so that performance may be easily assessed and improved upon. Such analyses may draw upon studies that have identified meteorological and fuel moisture thresholds important for ignitions and rapid spread rates [51,65,74]. We emphasize that RFWs are directly linked to fire occurrence and spread within the NWS Fire Weather Services Product Specification [27], and argue that RFW issuance criteria and verification should

include actual fire activity. We have demonstrated the performance of RFW as forecasts of large fire occurrence and hypothesize that further improvements could be made by incorporating daily fire spread data and revising RFW criteria to include our aforementioned recommendations.

RFWs and other early warning systems for fire may benefit from incorporating a probabilistic framework that is better suited for high-risk, rare event forecasting [61,75] and quantitative risk assessments [76]. A probabilistic forecast would be especially useful for the fire community, where decisions are commonly made with high economic cost and human risk factors [77,78]. The NOAA Forecasting a Continuum of Environment Threats (FACETs) framework seeks to supplement or replace existing NWS deterministic products with high-resolution probabilistic information [79] and may be applicable for moving RFWs in this direction. Beyond RFWs, other fire early warning systems should consider providing probabilistic information that identifies the range of potential scenarios that would ultimately lead to better, more actionable decisions by end users.

Supplementary Materials: The following are available online at http://www.mdpi.com/2571-6255/3/4/60/s1, Table S1: Summary statistics for the eight WFOs showing the number of days where an RFW was issued, days when an LF occurred, the mean LF size for the WFO, the number of human- and lightning-caused LF days, and the ratio of LFs occurring on or within one day following an RFW day. Table S2: Complete RFW performance statistics, skill scores, and relative improvement values for large fire days ≥ 80th percentile, ≥ 90th percentile, and ≥ 95th percentile sizes. Table S3: Complete RFW performance statistics, skill scores, and relative improvement values for lightning-caused and human-caused large fire days ≥ 90th percentile sizes. Table S4: Complete RFW performance statistics, skill scores, and relative improvement values for forested and non-forested large fire days ≥ 90th percentile sizes. Table S5: Complete RFW performance statistics, skill scores, and relative improvement values for large fire days ≥ 90th percentile sizes where ERC≥90th percentile values and ERC < 90th percentile values.

Conflicts of Interest: The authors declare no conflict of interest.

List of Acronyms

BOI	Boise (weather forecast office)
CSI	Critical success index
ERC	Energy release component
FACET	Forecasting a Continuum of Environment Threats
FAR	False alarm ratio
FPA	Fire Program Analysis
FWZs	Fire weather zones
gridMET	Gridded meteorological data
IEM	Iowa Environmental Mesonet
LF	Large fire
LFD	Large fire day
MFR	Medford (weather forecast office)
MODIS	Moderate resolution imaging spectroradiometer
MSO	Missoula (weather forecast office)
NOAA	National Oceanic and Atmospheric Administration
NWS	National Weather Service
OTX	Spokane (weather forecast office)
PDT	Pendleton (weather forecast office)
PIH	Pocatello (weather forecast office)
POD	probability of detection
PQR	Portland (weather forecast office)
RAWS	Remote Automatic Weather Stations

RFWs	Red Flag Warnings
SEW	Seattle (weather forecast office)
U.S.	United States
WFOs	weather forecast offices

References

1. Bowman, D.M.; Balch, J.K.; Artaxo, P.; Bond, W.J.; Carlson, J.M.; Cochrane, M.A.; D'Antonio, C.M.; DeFries, R.S.; Doyle, J.C.; Harrison, S.P.; et al. Fire in the Earth system. *Science* **2009**, *324*, 481–484. [CrossRef] [PubMed]

2. Bowman, D.M.; Williamson, G.J.; Abatzoglou, J.T.; Kolden, C.A.; Cochrane, M.A.; Smith, A.M.S. Human exposure and sensitivity to globally extreme wildfire events. *Nat. Ecol. Evol.* **2017**, *1*, 0058. [CrossRef] [PubMed]

3. Vaillant, N.M.; Kolden, C.A.; Smith, A.M.S. Assessing landscape vulnerability to wildfire in the USA. *Curr. For. Rep.* **2016**, *2*, 201–213. [CrossRef]

4. Moritz, M.A.; Batllori, E.; Bradstock, R.A.; Gill, A.M.; Handmer, J.; Hessburg, P.F.; Leonard, J.; McCaffrey, S.; Odion, D.C.; Schoennagel, T.; et al. Learning to coexist with wildfire. *Nature* **2014**, *515*, 58–66. [CrossRef]

5. Smith, A.M.S.; Kolden, C.A.; Bowman, D.M. Biomimicry can help humans to sustainably coexist with fire. *Nat. Ecol. Evol.* **2018**, *2*, 1827–1829. [CrossRef] [PubMed]

6. Smith, A.M.S.; Kolden, C.A.; Paveglio, T.; Cochrane, M.A.; Mortitz, M.A.; Bowman, D.M.; Hoffman, C.M.; Lutz, J.; Queen, L.P.; Hudak, A.T.; et al. The science of firescapes: Achieving fire resilient communities. *BioScience* **2016**, *66*, 130–146. [CrossRef]

7. Jolly, W.M.; Cochrane, M.A.; Freeborn, P.H.; Holden, Z.A.; Brown, T.J.; Williamson, G.J.; Bowman, D.M. Climate-induced variations in global wildfire danger from 1979 to 2013. *Nat. Commun.* **2015**, *6*, 7537. [CrossRef]

8. Balch, J.K.; Bradley, B.A.; Abatzoglou, J.T.; Nagy, R.C.; Fusco, E.J.; Mahood, A.L. Human expansion of the fire niche across the United States. *Proc. Natl. Acad. Sci. USA* **2017**, *114*, 2946–2951. [CrossRef]

9. Parks, S.A.; Miller, C.; Parisien, M.A.; Holsinger, L.M.; Dobrowski, S.Z.; Abatzoglou, J.T. Wildland fire deficit and surplus in the western United States, 1984–2012. *Ecosphere* **2015**, *6*, 275. [CrossRef]

10. Brewer, M.J.; Clements, C.B. The 2018 Camp Fire: Meteorological analysis using in situ observations and numerical simulations. *Atmosphere* **2020**, *11*, 47. [CrossRef]

11. Nauslar, N.J.; Abatzoglou, J.T.; Marsh, P.T. The 2017 North Bay and Southern California Fires: A case study. *Fire* **2018**, *1*, 18. [CrossRef]

12. Balch, J.K.; Schoennagel, T.; Williams, A.P.; Abatzoglou, J.T.; Cattau, M.E.; Mietkiewicz, N.P.; St. Denis, L.A. Switching on the Big Burn of 2017. *Fire* **2018**, *1*, 17. [CrossRef]

13. Barbero, R.; Abatzoglou, J.T.; Larkin, N.K.; Kolden, C.A.; Stocks, B. Climate change presents increased potential for very large fires in the contiguous United States. *Int. J. Wildl. Fire* **2015**, *24*, 892–899. [CrossRef]

14. Radeloff, V.C.; Helmers, D.P.; Kramer, H.A.; Mockrin, M.H.; Alexandre, P.M.; Bar-Massada, A.; Butsic, V.; Hawbaker, T.J.; Martinuzzi, S.; Syphard, A.D.; et al. Rapid growth of the US wildland-urban interface raises wildfire risk. *Proc. Natl. Acad. Sci. USA* **2018**, *115*, 3314–3319. [CrossRef] [PubMed]

15. de Groot, W.J.; Wotton, B.M.; Flannigan, M.D. Wildland fire danger rating and early warning systems. In *Wildfire Risks, Hazards, and Disasters*; Paton, D., Buergelt, P.T., McCaffrey, S., Tedim, F., Eds.; Elsevier, Inc.: Amsterdam, The Netherlands, 2014; p. 268. ISBN 978-012-410-434-1.

16. Deeming, J.E.; Burgan, R.E.; Cohen, J.D. *The 1978 National Fire-Danger Rating System: Technical Documentation*; General Technical Report INT-169 1983; U.S. Department of Agriculture Forest Service Intermountain Forest and Range Experiment Station: Ogden, UT, USA, 1983; p. 44.

17. Wagner, C.E. *Development and Structure of the Canadian Forest Fire Index System*; Ottawa Forestry Tech. Rep. 35; Canadian Forestry Service, Headquarters: Ottawa, ON, Canada, 1987; p. 35.

18. Boer, M.M.; Nolan, R.H.; De Dios, V.R.; Clarke, H.; Price, O.F.; Bradstock, R.A. Changing weather extremes call for early warning of potential for catastrophic fire. *Rev. Geophys.* **2017**, 1196–1202.

19. Sharples, J.J.; McRae, R.H.D.; Weber, R.O.; Gill, A.M. A simple index for assessing fire danger rating. *Environ. Model. Softw.* **2009**, *24*, 764–774. [CrossRef]

20. Anderson, S.A.J. Future Options for Fire Behaviour Modelling and Fire Danger Rating in New Zealand. *Proc. R. Soc. Qld.* **2009**, *115*, 119–128.

21. Gazzard, R.; McMorrow, J.; Aylen, J. Wildfire policy and management in England: An evolving response from Fire and Rescue Services, forestry, and cross-sector groups. *Phil. Trans. R. Soc. B* **2016**, *371*, 20150341. [CrossRef] [PubMed]

22. Jolly, W.M.; Freeborn, P.H.; Page, W.G.; Butler, B.W. Severe Fire Danger Index: A forecastable metric to inform firefighter and community wildfire risk management. *Fire* **2019**, *2*, 47. [CrossRef]

23. Rolinski, T.; Capps, S.B.; Fovell, R.G.; Cao, Y.; D'Agostino, B.J.; Vanderburg, S. The Santa Ana Wildfire Threat Index: Methodology and operational implementation. *Weather Forecast.* **2016**, *31*, 1881–1897. [CrossRef]

24. Srock, A.F.; Charney, J.J.; Potter, B.E.; Goodrick, S.L. The Hot-Dry-Windy Index: A new fire weather index. *Atmosphere* **2019**, *9*, 279. [CrossRef]

25. Erickson, M.J.; Charney, J.J.; Colle, B.A. Development of a fire weather index using meteorological observations within the Northeast United States. *J. Appl. Meteor. Climatol.* **2016**, *55*, 389–402. [CrossRef]

26. Abatzoglou, J.T.; Smith, C.M.; Swain, D.L.; Ptak, T.; Kolden, C.A. Population exposure to pre-emptive de-energization aimed at averting wildfires in Northern California. *Environ. Res. Lett.* **2020**, *15*, 094046. [CrossRef]

27. NWS. *National Weather Service Instruction 10-401. Fire Weather Services Product Specification*; National Weather Service: Silver Spring, MD, USA, 2017; p. 24. Available online: https://www.nws.noaa.gov/directives/sym/pd01004001curr.pdf (accessed on 27 March 2020).

28. Brooks, H.E.; Correia, J. Long-term performance metrics for National Weather Service tornado warnings. *Weather Forecast.* **2018**, *33*, 1501–1511. [CrossRef]

29. Obermeier, H.M.; Anderson, M.R. Verification and analysis of impact-based tornado warnings in the Central Region of the National Weather Service. *Electron. J. Sev. Storms Meteor.* **2014**, *10*, 1–20.

30. Clark, R.A.; Gourley, J.J.; Flamig, Z.L.; Hong, Y.; Clark, E. CONUS-wide evaluation of National Weather Service flash flood guidance products. *Weather Forecast.* **2014**, *29*, 377–392. [CrossRef]

31. McGovern, A.; Gagne, D.J., 2nd; Williams, J.K.; Brown, R.A.; Basara, J.B. Enhancing understanding and improving prediction of severe weather through spatiotemporal relational learning. *Mach. Learn.* **2014**, *95*, 27–50. [CrossRef]

32. Lindley, T.T.; Vitale, J.D.; Burgett, W.S.; Beierle, M.-J. Proximity meteorological observations for wind-driven grassland wildfire starts on the southern High Plains. *Electron. J. Sev. Storms Meteor.* **2011**, *6*, 1–27.

33. Werth, P.A.; Potter, B.E.; Alexander, M.E.; Clements, C.B.; Cruz, M.G.; Finney, M.A.; Forthofer, J.M.; Goodrick, S.L.; Hoffman, C.; Jolly, W.M.; et al. *Synthesis of Knowledge of Extreme Fire Behavior: Volume 2 for Fire Behavior Specialists, Researchers, and Meteorologists*; U.S Department of Agriculture Forest Service Pacific Northwest Research Station General Tech. Rep. PNW-GTR-891; U.S. Department of Agriculture: Portland, OR, USA, 2016; p. 258.

34. Short, K.C. *Spatial Wildfire Occurrence Data for the United States, 1992–2015 [FPA_FOD_20170508]*, 4th ed.; Forest Service Research Data Archive; U.S. Department of Agriculture Forest Service Rocky Mountain Research Station: Fort Collins, CO, USA, 2017. Available online: https://www.fs.usda.gov/rds/archive/catalog/RDS-2013-0009.4 (accessed on 7 August 2019).

35. Agee, J.K. *Fire Ecology of Pacific Northwest Forests*, 2nd ed.; Island Press: Washington, DC, USA, 1996; p. 505. ISBN 978-155-963-230-0.

36. Abatzoglou, J.T.; Kolden, C.A.; Balch, J.K.; Bradley, B.A. Controls on interannual variability in lightning-caused fire activity in the western US. *Environ. Res. Lett.* **2016**, *11*, 4. [CrossRef]

37. IEM. Iowa Environmental Mesonet. Archived NWS Watch/Warnings. Available online: https://mesonet.agron.iastate.edu/request/gis/watchwarn.phtml (accessed on 30 May 2019).

38. NWS. *National Weather Service Instruction 10-407, Fire Weather Services Zone Change Process*; NWS: Silver Spring, MD, USA, 2018; p. 6. Available online: https://www.nws.noaa.gov/directives/sym/pd01004007curr.pdf (accessed on 27 March 2020).

39. Nagy, R.C.; Fusco, E.; Bradley, B.A.; Abatzoglou, J.T.; Balch, J.K. Human-related ignitions increase the number of large wildfires across U.S. ecoregions. *Fire* **2018**, *1*, 4. [CrossRef]

40. Dennison, P.E.; Brewer, S.C.; Arnold, J.D.; Moritz, M.A. Large wildfire trends in the western United States, 1984–2011. *Geophys. Res. Lett.* **2014**, *41*, 2928–2933. [CrossRef]

41. Broxton, P.D.; Zeng, X.; Sulla-Menashe, D.; Troch, P.A. A global land cover climatology using MODIS data. *J. Appl. Meteor. Climatol.* **2014**, *53*, 1593–1605. [CrossRef]

42. Abatzoglou, J.T. Development of gridded surface meteorological data for ecological applications and modelling. *Int. J. Climatol.* **2013**, *33*, 121–131. [CrossRef]

43. Freeborn, P.H.; Cochrane, M.A.; Jolly, W.M. Relationships between fire danger and the daily number and daily growth of active incidents burning in the northern Rocky Mountains, USA. *Int. J. Wildl. Fire* **2015**, *24*, 900–910. [CrossRef]

44. Abatzoglou, J.T.; Kolden, C.A. Relationships between climate and macroscale area burned in the western United States. *Int. J. Wildl. Fire* **2013**, *22*, 1003–1020. [CrossRef]

45. Barbero, R.; Abatzoglou, J.T.; Steel, E.A. Modeling very large-fire occurrences over the continental United States from weather and climate forcing. *Environ. Res. Lett.* **2014**, *9*, 12. [CrossRef]

46. Heinsch, F.A.; Andrews, P.L.; Kurth, L.L. Implications of using percentiles to define fire danger levels. Extended Abstract P1.5. In Proceedings of the Eighth Symposium on Fire and Forest Meteorology, Kalispell, MT, USA, 12–15 October 2009.

47. Dalton, M.M.; Abatzoglou, J.T.; Evers, L.; Hegewisch, K. *Projected Changes in the Energy Release Component Under Climate Change in Northwest Predictive Services Areas*; A report to the Bureau of Land Management Oregon-Washington State Office; The Oregon Climate Change Research Institute (OCCRI): Corvallis, OR, USA, 2015; p. 88.

48. Murphy, A.H.; Winkler, R.L. A general framework for forecast verification. *Mon. Weather. Rev.* **1987**, *115*, 1330–1338. [CrossRef]

49. Doswell, C.A.; Davies-Jones, R.; Keller, D.L. On summary measures of skill in rare event forecasting based on contingency tables. *Weather Forecast.* **1990**, *5*, 576–585. [CrossRef]

50. Wilks, D.S. *Statistical Methods in the Atmospheric Sciences*, 3rd ed.; Academic Press: Cambridge, MA, USA, 2011; p. 676. ISBN 978-012-385-022-5.

51. Abatzoglou, J.T.; Balch, J.K.; Bradley, B.A.; Kolden, C.A. Human-related ignitions concurrent with high winds promote large wildfires across the USA. *Int. J. Wildl. Fire* **2018**, *27*, 377–386. [CrossRef]

52. Vaughan, M.T.; Tang, B.H.; Bosart, L.F. Climatology and analysis of high-impact, low predictive skill severe weather events in the northeast United States. *Weather Forecast.* **2017**, *32*, 1903–1919. [CrossRef]

53. Mason, I.B. Binary events. In *Forecast Verification: A Practitioner's Guide in Atmospheric Science*, 1st ed.; Jolliffe, I.T., Stephenson, D.B., Eds.; John Wiley & Sons Ltd.: West Sussex, UK, 2003; pp. 37–76. ISBN 978-047-086-441-8.

54. Schaefer, J.T. The critical success index as an indicator of warning skill. *Weather Forecast.* **1990**, *5*, 570–575. [CrossRef]

55. Murphy, A.H. Skill Scores Based on the Mean Square Error and Their Relationships to the Correlation Coefficient. *Mon. Weather Rev.* **1988**, *116*, 2417–2424. [CrossRef]

56. Potts, J.M. Basic concepts. In *Forecast Verification: A Practitioner's Guide in Atmospheric Science*, 1st ed.; Jolliffe, I.T., Stephenson, D.B., Eds.; John Wiley & Sons Ltd.: West Sussex, UK, 2003; pp. 13–36. ISBN 978-047-086-441-8.

57. Mason, S.J. On using "climatology" as a reference strategy in the Brier and ranked probability skill scores. *Mon. Weather Rev.* **2004**, *132*, 1891–1895. [CrossRef]

58. Warner, T.T. *Numerical Weather and Climate Prediction*, 1st ed.; Cambridge University Press: New York, NY, USA, 2011; p. 550. ISBN 978-051-176-324-3.

59. Mittermaier, M.P. The potential impact of using persistence as a reference forecast on perceived forecast skill. *Weather Forecast.* **2008**, *23*, 1022–1031. [CrossRef]

60. Murphy, A.H.; Winkler, R.L. Forecasters and probability forecasts: Some current problems. *Bull. Am. Meteorol. Soc.* **1971**, *52*, 239–247. [CrossRef]

61. Murphy, A.H. Probabilities, odds, and forecasts of rare events. *Weather Forecast.* **1991**, *6*, 302–307. [CrossRef]

62. Holliday, J.R.; Nanjo, K.Z.; Tiampo, K.F.; Rundle, J.B.; Turcotte, D.L. Earthquake forecasting and its verification. *Nonlinear Process. Geophys.* **2005**, *12*, 965–977. [CrossRef]

63. Shcherbakov, R.; Turcotte, D.L.; Rundle, J.B.; Tiampo, K.F.; Holliday, J.R. Forecasting the locations of future large earthquakes: An analysis and verification. *Pure Appl. Geophys.* **2010**, *167*, 743–749. [CrossRef]

64. Nauslar, N.J. Examining the lightning polarity of lightning caused wildfires. In Proceedings of the 23rd International Lightning Detection Conference, 5th International Lightning Detection Conference, Tucson, AZ, USA, 18–21 March 2014; p. 15.

65. Rorig, M.L.; Ferguson, S.A. Characteristics of lightning and wildland fire ignition in the Pacific Northwest. *J. Appl. Meteorol.* **1999**, *38*, 1565–1575. [CrossRef]

66. Schultz, C.J.; Nauslar, N.J.; Wachter, J.B.; Hain, C.R.; Bell, J.R. Spatial, temporal and electrical characteristics of lightning in reported lightning-initiated wildfire events. *Fire* **2019**, *2*, 18. [CrossRef]

67. Martínez, J.; Vega-Garcia, C.; Chuvieco, E. Human-caused wildfire risk rating for prevention planning in Spain. *J. Environ. Manag.* **2009**, *90*, 1241–1252. [CrossRef] [PubMed]

68. Magnussen, S.; Taylor, S.W. Prediction of daily lightning-and human-caused fires in British Columbia. *Int. J. Wildl. Fire* **2012**, *21*, 342–356. [CrossRef]

69. Costafreda-Aumedes, S.; Comas, C.; Vega-Garcia, C. Human-caused fire occurrence modelling in perspective: A review. *Int. J. Wildl. Fire* **2017**, *26*, 983–998. [CrossRef]

70. Mees, R. Is arson associated with severe fire weather in Southern California? *Int. J. Wildl. Fire* **1991**, *1*, 97–100. [CrossRef]

71. Brewer, M.C.; Mass, C.F.; Potter, B.E. The West Coast thermal trough: Climatology and synoptic evolution. *Mon. Weather Rev.* **2012**, *140*, 3820–3843. [CrossRef]

72. St. Denis, L.A.; Mietkiewicz, N.P.; Short, K.C.; Buckland, M.; Balch, J.K. All-hazards dataset mined from the US National Incident Management System 1999–2014. *Sci. Data* **2020**, *7*, 64. [CrossRef]

73. Andela, N.; Morton, D.C.; Giglio, L.; Paugam, R.; Chen, Y.; Hantson, S.; van der Werf, G.R.; Randerson, J.T. The Global Fire Atlas of individual fire size, duration, speed and direction. *Earth Syst. Sci. Data* **2019**, *11*, 529–552. [CrossRef]

74. Podur, J.; Wotton, B.M. Defining fire spread event days for fire-growth modelling. *Int. J. Wildl. Fire* **2011**, *20*, 497–507. [CrossRef]

75. Gneiting, T.; Katzfuss, M. Probabilistic forecasting. *Annu. Rev. Stat. Appl.* **2014**, *1*, 125–151. [CrossRef]

76. Casati, B.; Wilson, L.J.; Stephenson, D.B.; Nurmi, P.; Ghelli, A.; Pocernich, M.; Damrath, U.; Ebert, E.E.; Brown, B.G.; Mason, S. Forecast verification: Current status and future directions. *Met. Apps.* **2008**, *15*, 3–18. [CrossRef]

77. Worsnop, R.P.; Scheuerer, M.; Hamill, T.M. Extended-range probabilistic fire-weather forecasting based on ensemble model output statistics and ensemble copula coupling. *Mon. Weather Rev.* **2020**, *148*, 499–521. [CrossRef]

78. Noonan-Wright, E.K.; Opperman, T.S.; Finney, M.A.; Zimmerman, G.T.; Seli, R.C.; Elenz, L.M.; Calkin, D.E.; Fiedler, J.R. Developing the US Wildland Fire Decision Support System. *J. Combust.* **2011**, 168473. [CrossRef]

79. Rothfusz, L.P.; Schneider, R.; Novak, D.; Klockow-McClain, K.; Gerard, A.E.; Karstens, C.; Stumpf, G.J.; Smith, T.M. FACETs: A proposed next-generation paradigm for high-impact weather forecasting. *Bull. Am. Meteorol. Soc.* **2018**, *99*, 2025–2043. [CrossRef]