Technical Note

Mean Composite Fire Severity Metrics Computed with Google Earth Engine Offer Improved Accuracy and Expanded Mapping Potential

Sean A. Parks [1],* , Lisa M. Holsinger [1], Morgan A. Voss [2], Rachel A. Loehman [3] and Nathaniel P. Robinson [4]

[1] Aldo Leopold Wilderness Research Institute, Rocky Mountain Research Station, US Forest Service,
    790 E. Beckwith Ave., Missoula, MT 59801, USA; lisamholsinger@fs.fed.us
[2] Department of Geography, University of Montana, Missoula, MT 59812, USA; morgan.voss@umontana.edu
[3] Alaska Science Center, US Geological Survey, 4210 University Drive, Anchorage, AK 99508, USA;
    rloehman@usgs.gov
[4] W.A. Franke College of Forestry and Conservation & Numerical Terradynamic Simulation Group,
    University of Montana, Missoula, MT 59812, USA; Nathaniel.Robinson@umontana.edu
*   Correspondence: sean_parks@fs.fed.us; Tel.: +1-406-542-4182

check for
updates

Abstract: Landsat-based fire severity datasets are an invaluable resource for monitoring and research purposes. These gridded fire severity datasets are generally produced with pre- and post-fire imagery to estimate the degree of fire-induced ecological change. Here, we introduce methods to produce three Landsat-based fire severity metrics using the Google Earth Engine (GEE) platform: The delta normalized burn ratio (dNBR), the relativized delta normalized burn ratio (RdNBR), and the relativized burn ratio (RBR). Our methods do not rely on time-consuming a priori scene selection but instead use a mean compositing approach in which all valid pixels (e.g., cloud-free) over a pre-specified date range (pre- and post-fire) are stacked and the mean value for each pixel over each stack is used to produce the resulting fire severity datasets. This approach demonstrates that fire severity datasets can be produced with relative ease and speed compared to the standard approach in which one pre-fire and one post-fire scene are judiciously identified and used to produce fire severity datasets. We also validate the GEE-derived fire severity metrics using field-based fire severity plots for 18 fires in the western United States. These validations are compared to Landsat-based fire severity datasets produced using only one pre- and post-fire scene, which has been the standard approach in producing such datasets since their inception. Results indicate that the GEE-derived fire severity datasets generally show improved validation statistics compared to parallel versions in which only one pre-fire and one post-fire scene are used, though some of the improvements in some validations are more or less negligible. We provide code and a sample geospatial fire history layer to produce dNBR, RdNBR, and RBR for the 18 fires we evaluated. Although our approach requires that a geospatial fire history layer (i.e., fire perimeters) be produced independently and prior to applying our methods, we suggest that our GEE methodology can reasonably be implemented on hundreds to thousands of fires, thereby increasing opportunities for fire severity monitoring and research across the globe.

Keywords: burn severity; change detection; Landsat; dNBR; RdNBR; RBR; composite burn index (CBI); MTBS

1. Introduction

The degree of fire-induced ecological change, or fire severity, has been the focus of countless studies across the globe [1–5]. These studies often rely on gridded metrics that use pre- and post-fire

imagery to estimate the amount of fire-induced change; the most common metrics are the delta normalized burn ratio (dNBR) [6], the relativized delta normalized burn ratio (RdNBR) [7], and the relativized burn ratio (RBR) [8]. These metrics generally have a high correspondence ($r^2 \geq 0.65$) to field-based measures of fire severity [9–12], making them an attractive alternative to expensive and time-consuming collection of post-fire field data. These satellite-inferred fire severity metrics are often produced using Landsat Thematic Mapper (TM), Enhanced Thematic mapper Plus (ETM+), and Operational Land Imager (OLI) imagery due to their combined temporal depth (1984-present) and global coverage, although they can be produced from other sensors such as the Moderate Resolution Imaging Spectroradiometer (MODIS) [13] and Sentinal2A [14].

However, producing satellite-inferred fire severity datasets can be challenging, particularly if severity data are needed for a large number of fires (>~20) or over broad spatial extents. For example, expertise in remote sensing technologies and software is necessary, indicating the need for a remote-sensing specialist or a substantial investment of time to learn such technologies and software. Furthermore, fire severity datasets have traditionally been produced using one pre-fire and one post-fire Landsat image [15,16], which requires careful attention to scene selection. Image selection can be time consuming in terms of identifying scenes with no clouds covering the fire of interest and avoiding scenes affected by a low sun angle and those with mismatched phenology between pre- and post-fire conditions [6,17]. Even when careful attention to image selection has been achieved, some images (those from Landsat ETM+ acquired after 2003) and the resulting gridded severity datasets will have missing data due to the failure of the Scan Line Corrector [18].

Challenges in producing satellite-inferred severity datasets have likely hampered development of regional to national fire severity products in many countries. The exception is in the United States (US), where Landsat-derived severity metrics have been produced for all 'large' fires (those ≥400 ha in the western US and ≥250 ha in the eastern US) that have occurred since 1984 [19]. This effort, undertaken by the US government, is called the Monitoring Trends in Burn Severity (MTBS) program and has mapped the perimeter and severity of over 20,000 fires. The MTBS program has provided data for numerous scientific studies ranging from those involving <10 fires [20–22] to those involving >1000 fires [2,23,24] and for topics such as fuel treatment effectiveness, climate change impacts, and time series analyses [25–28]. The fire severity datasets produced by the MTBS program have clearly advanced wildland fire research in the US. Although some studies involving the trends, drivers, and distribution of satellite-inferred fire severity are evident outside of the US [4,5,15,29,30], the number and breadth of such studies are relatively scarce and restricted compared to those conducted in the US. We suggest that, if spatially and temporally comprehensive satellite-inferred severity metrics were more widely available in other countries or regions, opportunities for fire severity monitoring and research would increase substantially.

In this paper, we present methods to quickly and easily produce Landsat-derived fire severity metrics (dNBR, RdNBR, and RBR). These methods are implemented within the Google Earth Engine (GEE) platform. As opposed to the standard approach in which one pre-fire and one post-fire Landsat scene are identified and used to produce these fire severity datasets, we use a mean compositing approach in which all valid pixels (e.g., cloud-free) over a pre-specified date range are stacked and the mean value for each pixel over each stack is calculated. Consequently, there is no need for a priori scene selection, which substantially speeds up the time necessary to produce fire severity datasets. The main caveat, however, is that a fire history GIS dataset (i.e., polygons of fire perimeters) must be available and produced independent of this process. Where fire history datasets are currently available or can easily be generated, our methods provide a means to produce satellite-inferred fire severity products similar to those distributed by the MTBS program. We also validate the severity metrics produced with our GEE methodology by evaluating the correspondence of dNBR, RdNBR, and RBR to a field-based measure of severity and measure the classification accuracy when categorized as low, moderate, and high severity. These validations were conducted on 18 fires in the western US [8] and were compared to parallel validations of fire severity datasets using one pre-fire and post-fire scene.

Code and a sample fire history GIS dataset are provided to aid users in replicating and implementing our methods.

2. Materials and Methods

2.1. Processing in Google Earth Engine

We produced the following Landsat-based fire severity metrics for each of the 18 fires that are described in Section 2.2; the perimeter of each fire was obtained from the MTBS program [19]. All fire severity metrics are based on the normalized burn ratio (NBR; Equation (1)) and include the: (i) Delta normalized burn ratio (dNBR; Equation (2)) [6]; (ii) relativized delta normalized burn ratio (RdNBR; Equation (3)) [7]; and (iii) relativized burn ratio (RBR; Equation (4)) [8]. These are produced using Landsat TM, ETM+, and OLI imagery.

$$NBR = \left(\frac{NIR - SWIR}{NIR + SWIR} \right) \tag{1}$$

$$dNBR = \left(NBR_{prefire} - NBR_{postfire} \right) \times 1000 \tag{2}$$

$$RdNBR = \begin{cases} \frac{dNBR}{|NBR_{prefire}|^{0.5}}, & \left| NBR_{prefire} \right| \geq 0.001 \\ \frac{dNBR}{|0.001|^{0.5}}, & \left| NBR_{prefire} \right| < 0.001 \end{cases} \tag{3}$$

$$RBR = \frac{dNBR}{NBR_{prefire} + 1.001} \tag{4}$$

where *NIR* (Equation (1)) is the near infrared band and *SWIR* (Equation (1)) is the shortwave infrared band. The $NBR_{prefire}$ qualifier in *RdNBR* (Equation (3)) is necessary because the equation fails when $NBR_{prefire}$ equals zero and produces very large values when it approaches zero.

Within GEE, mean pre- and post-fire NBR values (Equation (1)) across a pre-specified date range (termed a 'mean composite') were calculated per pixel across the stack of valid pixels (e.g., cloud- and snow-free pixels). For fires that occurred in Arizona, New Mexico, and Utah, the date range is April through June; for all other fires, the date range is June through September (Figure 1). These date ranges are based on various factors including the fire season, expected snow cover, expected cloud cover and latitude. We used the Landsat Surface Reflectance Tier 1 datasets, which among the bands, includes a quality assessment mask to identify those pixels with clouds, shadow, water, and snow. This mask is produced by implementing a multi-pass algorithm (called 'CFMask') based on decision trees and is described in detail by Foga et al. [31]. As such, pixels identified as cloud, shadow, water, and snow were excluded when producing the mean composite pre- and post-fire NBR. The resulting pre- and post-fire NBR mean composite images are then used to calculate dNBR, RdNBR, and RBR (Equations (2)–(4)). Our mean compositing approach renders the need for a priori scene selection unnecessary.

We also produced alternative versions of each severity metric in which we account for potential phenological differences between pre- and post-fire imagery, also known as the 'dNBR$_{offset}$' [6]. The dNBR$_{offset}$ is the average dNBR of pixels outside the burn perimeter (i.e., unburned) and is intended to account for differences between pre- and post-fire imagery that arise due to varying conditions in phenology or precipitation between respective time periods. Incorporating the dNBR$_{offset}$ is advisable when making comparisons among fires [7,8]. For each fire, we determined the dNBR$_{offset}$ by calculating the mean dNBR value across all pixels located 180 m outside of the fire perimeter; informal testing indicated that a 180 m distance threshold adequately quantifies dNBR differences among unburned pixels. A simple subtraction of the fire-specific dNBR$_{offset}$ from each dNBR raster incorporates the dNBR$_{offset}$ [17]. The dNBR (with the offset) is then used to produce RdNBR and RBR (Equations (3) and (4)).

2.2. Validation

We aimed to determine whether our GEE methodology (specifically the mean compositing method) produced Landsat-based fire severity datasets with equivalent or higher validation statistics than severity datasets produced using one pre-fire and one post-fire scene (i.e., the standard approach since these metrics were introduced). This validation has three components (described below), all of which rely on 1681 field-based severity plots covering 18 fires in the western US that burned between 2001 and 2011; these are the same plots and fires that were originally evaluated by Parks et al. [8] (Figure 1) (Table 1). The field data represent the composite burn index (CBI) [6], which rates factors such as surface fuel consumption, soil char, vegetation mortality, and scorching of trees. CBI is rated on a continuous scale from zero to three, with CBI = 0 reflecting no change due to fire and CBI = 3 reflecting the highest degree of fire-induced ecological change. The fires selected by Parks et al. [8] and used in this study (Table 1) met the following criteria: (i) They had at least 40 field-based CBI plots; and (ii) at least 15% of the plots fell into each class representing low, moderate, and high severity. Of the 1681 field-based CBI plots, 30% are considered low severity (CBI < 1.25), 41% are moderate severity (CBI $\geq$ 1.25 and < 2.25), and 29% are high severity (CBI $\geq$ 2.25).

Figure 1. Location and names of the 18 fires included in the validation of the delta normalized burn ratio (dNBR), relativized delta normalized burn ratio (RdNBR), and relativized burn ratio (RBR). Forested areas in the western United States (US) are shown in gray shading. Inset shows the study area in relation to North America.

The first validation evaluates the correspondence of each severity metric to the CBI data for each fire. Exactly following Parks et al. [8], we extracted GEE-derived dNBR, RdNBR, and RBR values using bilinear interpolation and then used nonlinear regression in the R statistical environment [32] to evaluate the performance of each severity metric. Specifically, we quantified the correspondence of each severity metric (the dependent variable) to CBI (the independent variable) as the coefficient of determination, which is the R^2 of a linear regression between predicted and observed severity values. We conducted this analysis for each fire and reported the mean R^2 across the 18 fires. We then conducted a parallel analysis but used MTBS-derived severity datasets. This parallel analysis allows for a robust comparison of severity datasets produced using one pre-fire and one post-fire image (e.g., MTBS-derived metrics) with the mean compositing approach as achieved with GEE. This validation was conducted on the severity metrics without and with the dNBR$_{offset}$.

Table 1. Summary of fires analyzed in this study; this table is from originally Parks et al. 2014 [8].

Fire Name	Year	Number of plots	Overstory species (in order of prevalence)	Historical Fire Regime [33]		
				Surface	Mixed	Replace
Tripod Cx (Spur Peak) [1]	2006	328	Douglas-fir, ponderosa pine, subalpine fir, Engelmann spruce	80–90%	<5%	5–10%
Tripod Cx (Tripod) [1]	2006	160	Douglas-fir, ponderosa pine, subalpine fir, Engelmann spruce	>90%	<5%	<5%
Robert [2]	2003	92	Subalpine fir, Engelmann spruce, lodgepole pine, Douglas-fir, grand fir, western red cedar, western larch	5–10%	30–40%	40–50%
Falcon [3]	2001	42	Subalpine fir, Engelmann spruce, lodgepole pine, whitebark pine	0%	30–40%	60–70%
Green Knoll [3]	2001	54	Subalpine fir, Engelmann spruce, lodgepole pine, Douglas-fir, aspen	0%	20–30%	70–80%
Puma [4]	2008	45	Douglas-fir, white fir, ponderosa pine	20–30%	70–80%	0%
Dry Lakes Cx [3]	2003	49	Ponderosa pine, Arizona pine, Emory oak, alligator juniper	>90%	0%	0%
Miller [5]	2011	94	Ponderosa pine, Arizona pine, Emory oak, alligator juniper	80–90%	5–10%	0%
Outlet [6]	2000	54	Subalpine fir, Engelmann spruce, lodgepole pine, ponderosa pine, Douglas-fir, white fir	30–40%	5–10%	50–60%
Dragon Cx WFU [6]	2005	51	Ponderosa pine, Douglas-fir, white fir, aspen, subalpine fir, lodgepole pine	60–70%	20–30%	5–10%
Long Jim [6]	2004	49	Ponderosa pine, Gambel oak	>90%	0%	0%
Vista [6]	2001	46	Douglas-fir, white fir, ponderosa pine, aspen, subalpine fir	20–30%	70–80%	0%
Walhalla [6]	2004	47	Douglas-fir, white fir, ponderosa pine, aspen, subalpine fir, lodgepole pine	60–70%	20–30%	<5%
Poplar [6]	2003	108	Douglas-fir, white fir, ponderosa pine, aspen, subalpine fir, lodgepole pine	20–30%	20–30%	40–50%
Power [7]	2004	88	Ponderosa/Jeffrey pine, white fir, mixed conifers, black oak	>90%	0%	0%
Cone [7]	2002	59	Ponderosa/Jeffrey pine, mixed conifers	80–90%	<5%	<5%
Straylor [7]	2004	75	Ponderosa/Jeffrey pine, western juniper	>90%	0%	<5%
McNally [7]	2002	240	Ponderosa/Jeffrey pine, mixed conifers, interior live oak, scrub oak, black oak	70–80%	10–20%	0%

Composite burn index (CBI) data sources: [1] Susan Prichard, US Forest Service, Pacific Northwest Research Station; [2] Mike McClellan, Glacier National Park; [3] Zack Holden, US Forest Service, Northern Region; [4] Joel Silverman, Bryce Canyon National Park; [5] Sean Parks, US Forest Service, Rocky Mountain Research Station, Aldo Leopold Wilderness Research Institute; [6] Eric Gdula, Grand Canyon National Park; [7] Jay Miller, US Forest Service, Pacific Southwest Region.

Our second validation is nearly identical to that described in the previous paragraph but plot data from all 18 fires was combined (*n* = 1681). That is, instead of evaluating on a per-fire basis, we evaluated the plot data from all fires simultaneously. Following Parks et al. [8], this evaluation used a five-fold cross-validation. That is, five evaluations were conducted with 80% of the plot data used to train each nonlinear model and the remaining 20% used to test each model. The resulting coefficients of determination (R^2) and standard errors for the five testing datasets were averaged.

The third validation evaluates the classification accuracy when categorizing the satellite- and field-derived severity datasets into three discrete classes representing low, moderate, and high severity. To do so, we grouped the CBI plot data into severity classes using well-recognized CBI thresholds: Low severity corresponds to CBI values ranging from 0–1.24, moderate severity from 1.25–2.24, and high severity from 2.25–3.0 [7]. We then identified thresholds specific to each metric (with and without incorporating the $dNBR_{offset}$) corresponding to the low, moderate, and high CBI thresholds using nonlinear regression models as previously described. However, the nonlinear models used to produce low, moderate, and high severity thresholds for this evaluation used all 1681 plots combined and did not use the cross-validated versions. We measured the classification accuracy (i.e., the percent correctly classified) with 95% confidence intervals using the 'caret' package [34] in the R statistical environment [32]. We also produced confusion matrices for each severity metric and report the user's and producer's accuracy for each severity class (low, moderate, and high).

Finally, it is worth noting that we did not directly use the fire severity datasets distributed by the MTBS program. Our reasoning is that the MTBS program does not distribute the RBR. Furthermore, the MTBS program incorporates the $dNBR_{offset}$ into the RdNBR product but does not distribute RdNBR without the $dNBR_{offset}$. The MTBS program does, however, distribute the imagery used to produce each fire severity metric. In order to make valid comparisons to the GEE-derived datasets, we opted to use the pre- and post-fire imagery distributed by the MTBS program to produce dNBR, RdNBR, and RBR, with and without the $dNBR_{offset}$, for each of the 18 fires. All processing of MTBS-derived fires was accomplished with the 'raster' package [35] in the R statistical environment [32].

2.3. Google Earth Engine Implementation and Code

We provide a sample code and a geospatial fire history layer to produce a total of six raster datasets (dNBR, RdNBR, and RBR; with and without the $dNBR_{offset}$) for each of the 18 previously described fires. This code produces severity datasets that are clipped to a bounding box representing the outer extent of each fire. We designed the code to use imagery from one year before and one year after each fire occurs and to use a pre-specified date range for image selection for each fire, as previously described. These parameters can easily be modified to suit the needs of different users, ecosystems, and fire regimes.

3. Results

Using GEE, we were able to quickly produce dNBR, RdNBR, and RBR (with and without the $dNBR_{offset}$) for the 18 fires analyzed. The entire process was completed in approximately 1 h; fires averaged about 15,000 hectares in size and ranged from 723–60,000 hectares. This timeframe included a few minutes of active, hands-on time and about 60 min of GEE computational processing. This timeframe should be considered a very rough estimate, however, because GEE processing time varies widely among fires (larger fire sizes require more computational processing) and because production time depends on available resources shared among users within GEE's cloud-based computing platform [36]; nonetheless, processing time is very fast with fairly low investment in terms of human labor.

The mean compositing approach, in conjunction with the exclusion pixels classified as cloud, shadow, snow, and water, resulted in a variable number of valid Landsat scenes used in producing each pre- and post-fire NBR image. The average number of stacked pixels used to produce pre- and

post-fire NBR was about 11. This varied by fire and ranged from 2–20 for pre-fire NBR and from 6–20 for post-fire NBR.

Our first validation, in which correspondence between CBI and each severity metric was computed independently for each fire, shows that there is not a substantial improvement between the MTBS- and GEE-derived fire severity metrics (Table 2).

Table 2. Mean R^2 of the correspondence between CBI and each MTBS- and GEE-derived fire severity metric across the 18 fires. MTBS: Monitoring Trends in Burn Severity; GEE: Google Earth Engine. The correspondence between CBI and the severity metrics were computed for each of the 18 fires and the mean R^2 is reported here.

	Mean R^2 without $dNBR_{offset}$		Mean R^2 with $dNBR_{offset}$	
	MTBS-Derived	**GEE-Derived**	**MTBS-Derived**	**GEE-Derived**
dNBR	0.761	0.768	0.761	0.768
RdNBR	0.736	0.764	0.751	0.759
RBR	0.784	0.791	0.784	0.790

When the correspondence between CBI and each severity metric for 1681 plots covering 18 fires was evaluated simultaneously using a five-fold cross-validation (our second evaluation), the R^2 was consistently higher for the GEE-derived fire severity datasets as compared to the MTBS-derived datasets (Table 3; Figure 2). Furthermore, the inclusion of the $dNBR_{offset}$ increased the correspondence to CBI for all fire severity metrics except for GEE-derived RdNBR (Table 3). All terms in the nonlinear regressions for all severity metrics (those with and without the $dNBR_{offset}$) were statistically significant ($p < 0.05$) in all five folds of the cross-validation.

Table 3. R^2 of the five-fold cross-validation of the correspondence between CBI and each MTBS- and GEE-derived fire severity metric for 1681 plots across 18 fires; standard error shown in parentheses. The values characterize the average of five folds and represent the severity metrics excluding and including the $dNBR_{offset}$.

	R^2 without $dNBR_{offset}$ (Standard Error)		R^2 with $dNBR_{offset}$ (Standard Error)	
	MTBS-Derived	**GEE-Derived**	**MTBS-Derived**	**GEE-Derived**
dNBR	0.630 (0.026)	0.660 (0.025)	0.655 (0.026)	0.682 (0.025)
RdNBR	0.616 (0.026)	0.692 (0.025)	0.661 (0.026)	0.669 (0.026)
RBR	0.683 (0.025)	0.722 (0.024)	0.714 (0.025)	0.739 (0.024)

The GEE-derived fire severity datasets generally resulted in an improvement over the comparable MTBS-derived datasets in terms of overall classification accuracy (Table 4); Inclusion of the $dNBR_{offset}$ provided additional improvement for the most part (Table 4). The only exception is for the GEE-derived RdNBR, in which the classification accuracy was slightly lower when using the $dNBR_{offset}$ (Table 4). The confusion matrices for each fire severity metric (with and without the $dNBR_{offset}$) indicate that the user's and producer's accuracies are usually higher with the GEE-derived metrics compared to the MTBS-derived metrics (Tables 5 and 6). The thresholds we used to classify plots as low, moderate, or high severity are shown in Table 7; these may be useful for others who implement our GEE methodology and want to classify the resulting datasets.
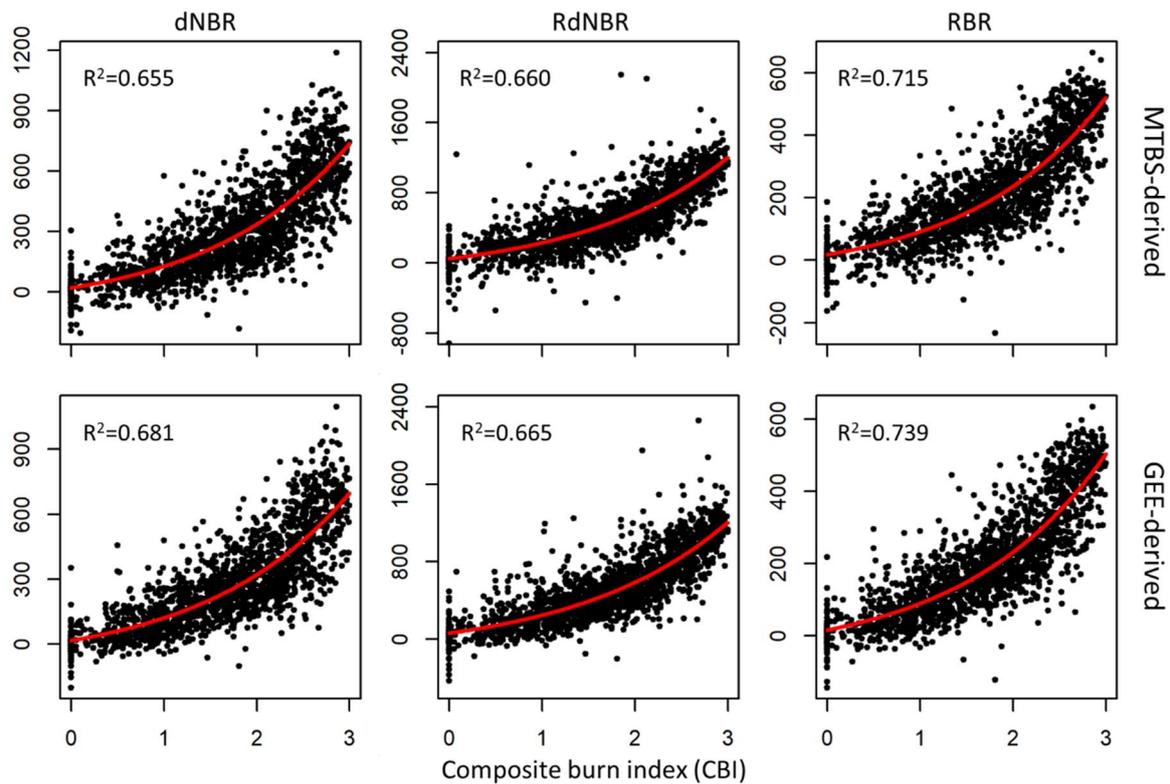
Figure 2. Plots show each MTBS- (top row) and GEE-derived (bottom row) severity metric and the corresponding field-based CBI. All severity metrics include the dNBR$_{offset}$. Red lines show the modeled fit of the nonlinear regressions for all 1681 plots. The model fits and the resulting R^2 shown here were not produced using cross-validation and therefore may differ slightly from the results shown in Table 3. Extreme RdNBR values are not shown to improve visual appearance of the RdNBR panels.

Table 4. Classification accuracy (percent correctly classified) and 95% confidence intervals (CI) for the three fire severity metrics (with and without the dNBR$_{offset}$). Each fire severity metric is classified into categories representing low, moderate, and high severity based on index-specific thresholds (see Table 7) and compared to the same classes based on composite burn index thresholds.

		Without dNBR$_{offset}$		With dNBR$_{offset}$	
		Accuracy (%)	95% CI	Accuracy (%)	95% CI
dNBR	MTBS-derived	69.6	67.3–71.8	70.2	68.0–72.4
	GEE-derived	71.3	69.0–73.4	71.7	69.5–73.9
RdNBR	MTBS-derived	71.4	69.2–73.5	73.6	71.4–75.6
	GEE-derived	73.4	71.2–75.5	73.1	71.0–75.3
RBR	MTBS-derived	72.4	71.1–74.5	73.5	71.4–75.6
	GEE-derived	73.5	71.4–75.6	74.1	72.0–76.2

Table 5. Confusion matrices for classifying as low, moderate, and high severity using the severity metrics computed without the dNBR$_{offset}$. Confusion matrices for MTBS-derived metrics are on the left and confusion matrices for GEE-derived metrics are on the right. UA: user's accuracy; PA: producer's accuracy.

		Reference CBI Class						**Reference CBI Class**			
		Low	Mod.	High	UA			Low	Mod.	High	UA
	Low	401	159	18	69.4		Low	407	139	13	72.8
Classified using MTBS-derived dNBR	Mod.	91	412	114	66.8	Classified using GEE-derived dNBR	Mod.	87	438	123	67.6
	High	5	124	357	73.5		High	3	118	353	74.5
	PA	80.7	59.3	73.0			PA	81.9	63.0	72.2	
		Reference CBI class						**Reference CBI class**			
		Low	Mod.	High	UA			Low	Mod.	High	UA
	Low	366	142	7	71.1		Low	385	136	5	73.2
Classified using MTBS-derived RdNBR	Mod.	119	451	99	67.4	Classified using GEE-derived RdNBR	Mod.	105	465	100	69.4
	High	12	102	383	77.1		High	7	94	384	79.2
	PA	73.6	64.9	78.3			PA	77.5	66.9	78.5	
		Reference CBI class						**Reference CBI class**			
		Low	Mod.	High	UA			Low	Mod.	High	UA
	Low	380	127	12	73.2		Low	403	130	9	74.4
Classified using MTBS-derived RBR	Mod.	113	462	102	68.2	Classified using GEE-derived RBR	Mod.	90	464	111	69.8
	High	4	106	375	77.3		High	4	101	369	77.8
	PA	76.5	66.5	76.7			PA	81.1	66.8	75.5	

Table 6. Confusion matrices for classifying as low, moderate, and high severity using the severity metrics computed with the dNBR$_{offset}$. Confusion matrices for MTBS-derived metrics are on the left and confusion matrices for GEE-derived metrics are on the right. UA: user's accuracy; PA: producer's accuracy.

		Reference CBI Class						**Reference CBI Class**			
		Low	Mod.	High	UA			Low	Mod.	High	UA
	Low	397	156	13	70.1		Low	402	141	10	72.7
Classified using MTBS-derived dNBR	Mod.	98	425	118	66.3	Classified using GEE-derived dNBR	Mod.	92	451	126	67.4
	High	2	114	358	75.5		High	3	103	353	76.9
	PA	79.9	61.2	73.2			PA	80.9	64.9	72.2	
		Reference CBI class						**Reference CBI class**			
		Low	Mod.	High	UA			Low	Mod.	High	UA
	Low	378	133	5	73.3		Low	390	137	5	73.3
Classified using MTBS-derived RdNBR	Mod.	112	467	92	69.6	Classified using GEE-derived RdNBR	Mod.	101	460	104	69.2
	High	7	95	392	79.4		High	6	98	380	78.5
	PA	76.1	67.2	80.2			PA	78.5	66.2	77.7	
		Reference CBI class						**Reference CBI class**			
		Low	Mod.	High	UA			Low	Mod.	High	UA
	Low	390	135	6	73.4		Low	386	123	7	74.8
Classified using MTBS-derived RBR	Mod.	105	460	97	69.5	Classified using GEE-derived RBR	Mod.	107	481	103	69.6
	High	2	100	386	79.1		High	4	91	379	80.0
	PA	78.5	66.2	78.9			PA	77.7	69.2	77.5	

Table 7. Threshold values for each fire severity metric corresponding to low (CBI = 0–1.24), moderate (CBI = 1.25–2.25), and high severity (CBI = 2.25–3).

		MTBS-Derived			GEE-Derived		
		Low	**Moderate**	**High**	**Low**	**Moderate**	**High**
Excludes dNBR$_{offset}$	dNBR	$\leq$186	187–429	$\geq$430	$\leq$185	186–417	$\geq$418
	RdNBR	$\leq$337	338–721	$\geq$722	$\leq$338	339–726	$\geq$727
	RBR	$\leq$134	135–303	$\geq$304	$\leq$135	136–300	$\geq$301
Includes dNBR$_{offset}$	dNBR	$\leq$165	166–440	$\geq$411	$\leq$159	160–392	$\geq$393
	RdNBR	$\leq$294	295–690	$\geq$691	$\leq$312	313–706	$\geq$707
	RBR	$\leq$118	119–289	$\geq$289	$\leq$115	116–282	$\geq$283

4. Discussion

The Google Earth Engine (GEE) methodology we developed to produce Landsat-based measures of fire severity is an important contribution to wildland fire research and monitoring. For example, our methodology will allow those who are not remote sensing experts, but have some familiarity with GEE, to quickly produce fire severity datasets (Figure 3). This benefit is due to the efficiency and speed of the cloud-based GEE platform [37,38] and because no a priori scene selection is necessary. Furthermore, compared to the standard approach in which only one pre- and post-fire scene are used, the GEE mean composite fire severity datasets exhibit higher validation statistics in terms of the correspondence (R^2) to CBI and higher classification accuracies for most severity classes. This suggests that mean composite severity metrics more accurately represent fire-induced ecological change, likely because the compositing method is less biased by pre- and post-fire scene mismatch and image characteristics inherent in standard processing. The computation and incorporation of the dNBR$_{offset}$ within GEE further improves, for the most part, the validation statistics of all metrics.

The improvements in the validation statistics of the GEE-derived severity metrics over the MTBS-derived severity metrics, when evaluated on a per-fire basis, are more or less negligible (see Table 2). This suggests that if practitioners and researchers are interested in only one fire [20,39], it does not matter if fire severity metrics are produced using the mean compositing approach or using one pre-fire and one post-fire image (e.g., MTBS). It is also worth noting that the improvements in the validation statistics of the GEE-derived severity metrics over the MTBS-derived severity metrics, when all plots are evaluated simultaneously, are not statistically significant in most cases. That is, the overall classification accuracy of the GEE-derived metrics overlap the 95% confidence intervals of the MTBS-derived metrics in all comparisons except that of RdNBR without the dNBR$_{offset}$ (Table 4). Although the user's and producer's accuracy is oftentimes higher for the GEE-derived severity metrics (Tables 5 and 6), this is not always the case for all severity classes. In particular, the producer's accuracy (but not the user's accuracy) is generally higher for the MTBS-derived metrics when evaluating the high severity class. Nevertheless, the modest improvement in most validation statistics of the GEE-derived metrics, together with the framework and code we distribute in this study, will likely provide the necessary rationale and tools for producing fire severity datasets in counties that do not have national programs tasked with producing such datasets (e.g., MTBS in the United States).
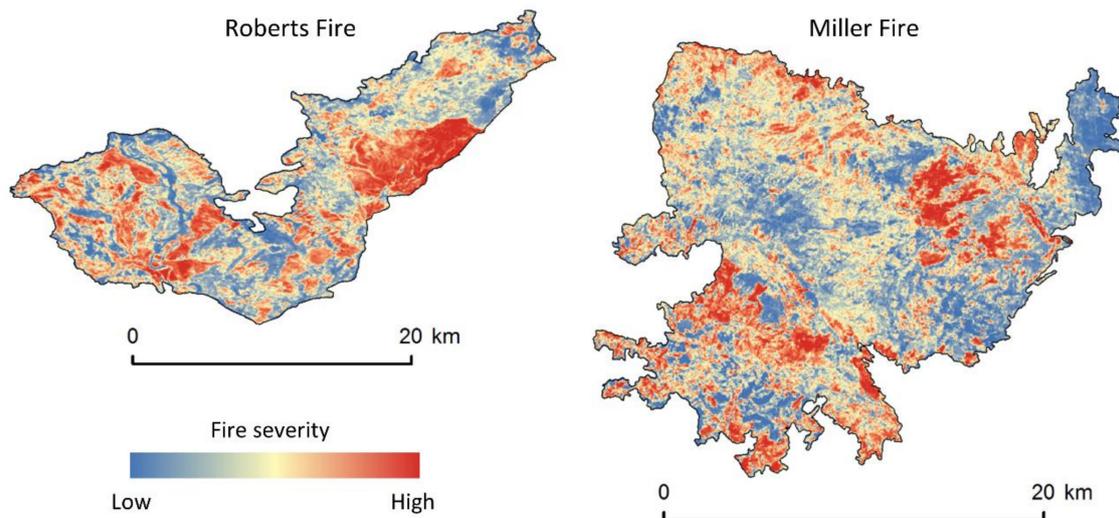
Figure 3. Example shows the RBR (includes the $dNBR_{offset}$) for two of the fires (Roberts and Miller) we evaluated. See Figure 1 to reference the locations of these fires.

The Monitoring Trends in Burn Severity (MTBS) program in the US, which produces and distributes Landsat-based fire severity datasets [19], has enabled scientists to conduct research involving hundreds to thousands of fires [2,24,40,41]. Outside of the US, where programs similar to MTBS do not exist, most fire severity research is limited to only a handful of fires, the exceptions being Fang et al. [15] (n = 72 fires in China) and Whitman et al. [42] (n = 56 fires in Canada). We suggest that the GEE methodology we developed will allow users in regions outside of the US to efficiently produce fire severity datasets for hundreds to thousands of fires in their geographic areas of interest, thereby providing enhanced opportunities for fire severity monitoring and research. Although fire history datasets (i.e., georeferenced fire perimeters) are a prerequisite for implementing our GEE methodology, such datasets have already been produced and used in scientific studies in Portugal [43], Spain [44], Canada [45], portions of Australia [46], southern France [47], the Sky Island Mountains of Mexico [48], and likely elsewhere. Therefore, the GEE methods developed here provide a common platform for assessing fire-induced ecological change and can provide more opportunities for fire severity monitoring and research across the globe.

The fires we analyzed primarily burned in conifer forests and were embedded within landscapes comprised of similar vegetation. As such, our approach to incorporating the $dNBR_{offset}$ that used pixels in a 180 m 'ring' around the fire perimeter may not be appropriate everywhere and we urge caution in landscapes in which fires burn vegetation that is not similar to that of the surrounding lands. For example, our methods for calculating and implementing the $dNBR_{offset}$ would not be appropriate if a fire burned a forested patch that was surrounded by completely different vegetation such as shrubland or agriculture. In such cases, we recommend that fire severity datasets exclude the $dNBR_{offset}$ as it may not improve burn assessments. Similarly, the low, moderate, and high severity thresholds identified in this study (Table 7) are likely only applicable to forested landscapes in the western US, and other thresholds may be more suitable to other regions of the globe and in different vegetation types. Finally, our choice of developing post-fire imagery from the period one-year after the fire may not be appropriate for all ecosystems. Arctic tundra ecosystems, for example, might be better represented by imagery derived immediately after the fire or after snowmelt but prior to green-up the year following the fire [49]. The GEE approach can be easily modified to select dates that best suit each ecosystem.

5. Conclusions

In this paper, we present practical and efficient methodology for producing three Landsat-based fire severity metrics: dNBR, RdNBR, and RBR. These methods rely on Google Earth Engine and provide expanded potential in terms of fire severity monitoring and research in regions outside of the US that do not have a dedicated program for mapping fire severity. In validating the fire severity metrics, our goal was not to compare and contrast individual metrics (e.g., dNBR vs. RBR) [11,12] nor to critique products produced by the MTBS program. Instead, we aimed to evaluate differences between the GEE-based mean compositing approach to the standard approach in which one pre-fire and post-fire Landsat scene are used to produce severity datasets. The GEE-based severity datasets generally achieved higher validation statistics in terms of correspondence to field data and overall classification accuracy. The inclusion of the $dNBR_{offset}$ generally provided additional improvements in these validation statistics for most fire severity metrics regardless of whether they were MTBS- or GEE-derived. This provides further evidence that inclusion of the $dNBR_{offset}$ should be considered when multiple fires are of interest [8,17]. Our evaluation included fires over a large spatial extent (the western US) and with varied fire regime attributes, ranging from those that are predominantly surface fire regimes to those that are stand-replacing regimes. Consequently, the higher validation statistics reported here for the GEE-derived composite-based fire severity datasets should provide researchers and practitioners with increased confidence in these products.

Author Contributions: S.A.P. conceived of the study, conducted the statistical validations, and wrote the paper. L.M.H. aided in designing the study, developed GEE code, and contributed to manuscript writing. R.A.L. aided in designing the study and contributed to manuscript writing. M.A.V. and N.P.R. developed GEE code and contributed to manuscript writing.

Code Availability: The code to implement our methods is available here: https://code.earthengine.google.com/ c76157be827be2f24570df50cca427e9. The code is set up to run on the 18 fires highlighted in this paper (Figure 1) and will produce dNBR, RdNBR, and RBR with and without the $dNBR_{offset}$.

References

1. Parks, S.A.; Parisien, M.A.; Miller, C.; Dobrowski, S.Z. Fire activity and severity in the western US vary along proxy gradients representing fuel amount and fuel moisture. *PLoS ONE* **2014**, *9*, e99699. [CrossRef] [PubMed]

2. Dillon, G.K.; Holden, Z.A.; Morgan, P.; Crimmins, M.A.; Heyerdahl, E.K.; Luce, C.H. Both topography and climate affected forest and woodland burn severity in two regions of the western US, 1984 to 2006. *Ecosphere* **2011**, *2*, 130. [CrossRef]

3. Veraverbeke, S.; Lhermitte, S.; Verstraeten, W.W.; Goossens, R. The temporal dimension of differenced Normalized Burn Ratio (dNBR) fire/burn severity studies: The case of the large 2007 Peloponnese wildfires in Greece. *Remote Sens. Environ.* **2010**, *114*, 2548–2563. [CrossRef]

4. Fernández-Garcia, V.; Santamarta, M.; Fernández-Manso, A.; Quintano, C.; Marcos, E.; Calvo, L. Burn severity metrics in fire-prone pine ecosystems along a climatic gradient using Landsat imagery. *Remote Sens. Environ.* **2018**, *206*, 205–217. [CrossRef]

5. Fang, L.; Yang, J.; White, M.; Liu, Z. Predicting Potential Fire Severity Using Vegetation, Topography and Surface Moisture Availability in a Eurasian Boreal Forest Landscape. *Forests* **2018**, *9*, 130. [CrossRef]

6. Key, C.H.; Benson, N.C. Landscape assessment (LA). In *FIREMON: Fire Effects Monitoring and Inventory System*; General Technical Report RMRS-GTR-164-CD; U.S. Department of Agriculture, Forest Service, Rocky Mountain Research Station: Fort Collins, CO, USA, 2006.

7. Miller, J.D.; Thode, A.E. Quantifying burn severity in a heterogeneous landscape with a relative version of the delta Normalized Burn Ratio (dNBR). *Remote Sens. Environ.* **2007**, *109*, 66–80. [CrossRef]

8. Parks, S.A.; Dillon, G.K.; Miller, C. A new metric for quantifying burn severity: The relativized burn ratio. *Remote Sens.* **2014**, *6*, 1827–1844. [CrossRef]

9. Holden, Z.A.; Morgan, P.; Evans, J.S. A predictive model of burn severity based on 20-year satellite-inferred burn severity data in a large southwestern US wilderness area. *For. Ecol. Manag.* **2009**, *258*, 2399–2406. [CrossRef]

10. Wimberly, M.C.; Reilly, M.J. Assessment of fire severity and species diversity in the southern Appalachians using Landsat TM and ETM+ imagery. *Remote Sens. Environ.* **2007**, *108*, 189–197. [CrossRef]

11. Veraverbeke, S.; Lhermitte, S.; Verstraeten, W.W.; Goossens, R. Evaluation of pre/post-fire differenced spectral indices for assessing burn severity in a Mediterranean environment with Landsat Thematic Mapper. *Int. J. Remote Sens.* **2011**, *32*, 3521–3537. [CrossRef]

12. Soverel, N.O.; Perrakis, D.D.B.; Coops, N.C. Estimating burn severity from Landsat dNBR and RdNBR indices across western Canada. *Remote Sens. Environ.* **2010**, *114*, 1896–1909. [CrossRef]

13. Beck, P.S.A.; Goetz, S.J.; Mack, M.C.; Alexander, H.D.; Jin, Y.; Randerson, J.T.; Loranty, M.M. The impacts and implications of an intensifying fire regime on Alaskan boreal forest composition and albedo. *Glob. Chang. Biol.* **2011**, *17*, 2853–2866. [CrossRef]

14. Mallinis, G.; Mitsopoulos, I.; Chrysafi, I. Evaluating and comparing Sentinel 2A and Landsat-8 Operational Land Imager (OLI) spectral indices for estimating fire severity in a Mediterranean pine ecosystem of Greece. *GISci. Remote Sens.* **2018**, *55*, 1–18. [CrossRef]

15. Fang, L.; Yang, J.; Zu, J.; Li, G.; Zhang, J. Quantifying influences and relative importance of fire weather, topography, and vegetation on fire size and fire severity in a Chinese boreal forest landscape. *For. Ecol. Manag.* **2015**, *356*, 2–12. [CrossRef]

16. Cansler, C.A.; McKenzie, D. How robust are burn severity indices when applied in a new region? Evaluation of alternate field-based and remote-sensing methods. *Remote Sens. Mol.* **2012**, *4*, 456–483. [CrossRef]

17. Key, C.H. Ecological and sampling constraints on defining landscape fire severity. *Fire Ecol.* **2006**, *2*, 34–59. [CrossRef]

18. Picotte, J.J.; Peterson, B.; Meier, G.; Howard, S.M. 1984–2010 trends in fire burn severity and area for the conterminous US. *Int. J. Wildl. Fire* **2016**, *25*, 413–420. [CrossRef]

19. Eidenshink, J.C.; Schwind, B.; Brewer, K.; Zhu, Z.-L.; Quayle, B.; Howard, S.M. A project for monitoring trends in burn severity. *Fire Ecol.* **2007**, *3*, 3–21. [CrossRef]

20. Kane, V.R.; Cansler, C.A.; Povak, N.A.; Kane, J.T.; McGaughey, R.J.; Lutz, J.A.; Churchill, D.J.; North, M.P. Mixed severity fire effects within the Rim fire: Relative importance of local climate, fire weather, topography, and forest structure. *For. Ecol. Manag.* **2015**, *358*, 62–79. [CrossRef]

21. Stevens-Rumann, C.; Prichard, S.; Strand, E.; Morgan, P. Prior wildfires influence burn severity of subsequent large fires. *Can. J. For. Res.* **2016**, *46*, 1375–1385. [CrossRef]

22. Prichard, S.J.; Kennedy, M.C. Fuel treatments and landform modify landscape patterns of burn severity in an extreme fire event. *Ecol. Appl.* **2014**, *24*, 571–590. [CrossRef] [PubMed]

23. Parks, S.A.; Holsinger, L.M.; Panunto, M.H.; Jolly, W.M.; Dobrowski, S.Z.; Dillon, G.K. High-severity fire: Evaluating its key drivers and mapping its probability across western US forests. *Environ. Res. Lett.* **2018**, *13*, 044037. [CrossRef]

24. Keyser, A.; Westerling, A. Climate drives inter-annual variability in probability of high severity fire occurrence in the western United States. *Environ. Res. Lett.* **2017**, *12*, 065003. [CrossRef]

25. Arkle, R.S.; Pilliod, D.S.; Welty, J.L. Pattern and process of prescribed fires influence effectiveness at reducing wildfire severity in dry coniferous forests. *For. Ecol. Manag.* **2012**, *276*, 174–184. [CrossRef]

26. Wimberly, M.C.; Cochrane, M.A.; Baer, A.D.; Pabst, K. Assessing fuel treatment effectiveness using satellite imagery and spatial statistics. *Ecol. Appl.* **2009**, *19*, 1377–1384. [CrossRef] [PubMed]

27. Parks, S.A.; Miller, C.; Abatzoglou, J.T.; Holsinger, L.M.; Parisien, M.-A.; Dobrowski, S.Z. How will climate change affect wildland fire severity in the western US? *Environ. Res. Lett.* **2016**, *11*, 035002. [CrossRef]

28. Miller, J.D.; Safford, H.D.; Crimmins, M.; Thode, A.E. Quantitative evidence for increasing forest fire severity in the Sierra Nevada and southern Cascade Mountains, California and Nevada, USA. *Ecosystems* **2009**, *12*, 16–32. [CrossRef]

29. Whitman, E.; Parisien, M.-A.; Thompson, D.K.; Hall, R.J.; Skakun, R.S.; Flannigan, M.D. Variability and drivers of burn severity in the northwestern Canadian boreal forest. *Ecosphere* **2018**, *9*. [CrossRef]

30. Ireland, G.; Petropoulos, G.P. Exploring the relationships between post-fire vegetation regeneration dynamics, topography and burn severity: A case study from the Montane Cordillera Ecozones of Western Canada. *Appl. Geogr.* **2015**, *56*, 232–248. [CrossRef]

31. Foga, S.; Scaramuzza, P.L.; Guo, S.; Zhu, Z.; Dilley, R.D.; Beckmann, T.; Schmidt, G.L.; Dwyer, J.L.; Hughes, M.J.; Laue, B. Cloud detection algorithm comparison and validation for operational Landsat data products. *Remote Sens. Environ.* **2017**, *194*, 379–390. [CrossRef]

32. R Core Team. *R: A Language and Environment for Statistical Computing*; R Foundation for Statistical Computing: Vienna, Austria, 2016; Available online: https://www.r-project.org/ (accessed on 1 July 2017).

33. Rollins, M.G. LANDFIRE: A nationally consistent vegetation, wildland fire, and fuel assessment. *Int. J. Wildl. Fire* **2009**, *18*, 235–249. [CrossRef]

34. Kuhn, M. Caret package. *J. Stat. Softw.* **2008**, *28*, 1–26.

35. Hijmans, R.J.; van Etten, J.; Cheng, J.; Mattiuzzi, M.; Sumner, M.; Greenberg, J.A.; Lamigueiro, O.P.; Bevan, A.; Racine, E.B.; Shortridge, A.; et al. *Package 'Raster'*; R. Package, 2015.

36. Gorelick, N.; Hancher, M.; Dixon, M.; Ilyushchenko, S.; Thau, D.; Moore, R. Google Earth Engine: Planetary-scale geospatial analysis for everyone. *Remote Sens. Environ.* **2017**, *202*, 18–27. [CrossRef]

37. Kennedy, R.E.; Yang, Z.; Gorelick, N.; Braaten, J.; Cavalcante, L.; Cohen, W.B.; Healey, S. Implementation of the LandTrendr Algorithm on Google Earth Engine. *Remote Sens.* **2018**, *10*, 691. [CrossRef]

38. Robinson, N.P.; Allred, B.W.; Jones, M.O.; Moreno, A.; Kimball, J.S.; Naugle, D.E.; Erickson, T.A.; Richardson, A.D. A Dynamic Landsat Derived Normalized Difference Vegetation Index (NDVI) Product for the Conterminous United States. *Remote Sens.* **2017**, *9*, 863. [CrossRef]

39. Lydersen, J.M.; Collins, B.M.; Brooks, M.L.; Matchett, J.R.; Shive, K.L.; Povak, N.A.; Kane, V.R.; Smith, D.F. Evidence of fuels management and fire weather influencing fire severity in an extreme fire event. *Ecol. Appl.* **2017**, *27*, 2013–2030. [CrossRef] [PubMed]

40. Reilly, M.J.; Dunn, C.J.; Meigs, G.W.; Spies, T.A.; Kennedy, R.E.; Bailey, J.D.; Briggs, K. Contemporary patterns of fire extent and severity in forests of the Pacific Northwest, USA (1985–2010). *Ecosphere* **2017**, *8*, e01695. [CrossRef]

41. Stevens, J.T.; Collins, B.M.; Miller, J.D.; North, M.P.; Stephens, S.L. Changing spatial patterns of stand-replacing fire in California conifer forests. *For. Ecol. Manag.* **2017**, *406*, 28–36. [CrossRef]

42. Whitman, E.; Batllori, E.; Parisien, M.-A.; Miller, C.; Coop, J.D.; Krawchuk, M.A.; Chong, G.W.; Haire, S.L. The climate space of fire regimes in north-western North America. *J. Biogeogr.* **2015**, *42*, 1736–1749. [CrossRef]

43. Fernandes, P.M.; Loureiro, C.; Magalhães, M.; Ferreira, P.; Fernandes, M. Fuel age, weather and burn probability in Portugal. *Int. J. Wildl. Fire* **2012**, *21*, 380–384. [CrossRef]

44. Trigo, R.M.; Sousa, P.M.; Pereira, M.G.; Rasilla, D.; Gouveia, C.M. Modelling wildfire activity in Iberia with different atmospheric circulation weather types. *Int. J. Climatol.* **2016**, *36*, 2761–2778. [CrossRef]

45. Parisien, M.-A.; Miller, C.; Parks, S.A.; Delancey, E.R.; Robinne, F.-N.; Flannigan, M.D. The spatially varying influence of humans on fire probability in North America. *Environ. Res. Lett.* **2016**, *11*, 075005. [CrossRef]

46. Price, O.F.; Penman, T.D.; Bradstock, R.A.; Boer, M.M.; Clarke, H. Biogeographical variation in the potential effectiveness of prescribed fire in south-eastern Australia. *J. Biogeogr.* **2015**, *42*, 2234–2245. [CrossRef]

47. Fox, D.M.; Carrega, P.; Ren, Y.; Caillouet, P.; Bouillon, C.; Robert, S. How wildfire risk is related to urban planning and Fire Weather Index in SE France (1990–2013). *Sci. Total Environ.* **2018**, *621*, 120–129. [CrossRef] [PubMed]

48. Villarreal, M.L.; Haire, S.L.; Iniguez, J.M.; Montaño, C.C.; Poitras, T.B. Distant Neighbors: Recent wildfire patterns of the Madrean Sky Islands of Southwestern United States and Northwestern México. *Fire Ecol.* **2018**, in press.

49. Kolden, C.A.; Rogan, J. Mapping wildfire burn severity in the Arctic tundra from downsampled MODIS data. *Arct. Antarct. Alp. Res.* **2013**, *45*, 64–76. [CrossRef]