fire

MDPI

Article

Leveraging a Wildfire Risk Prediction Metric with Spatial Clustering

Ujjwal KC [1,*] and Jagannath Aryal [2,*]

[1] CSIRO | Agriculture and Food, St Lucia, Brisbane, QLD 4067, Australia
[2] Department of Infrastructure Engineering, Faculty of Engineering and Information Technology, University of Melbourne, Melbourne, VIC 3053, Australia
* Correspondence: ujjwal.kc@csiro.au (U.K.); jagannath.aryal@unimelb.edu.au (J.A.)

Abstract: Fire authorities have started widely using operational fire simulations for effective wildfire management. The aggregation of the simulation outputs on a massive scale creates an opportunity to apply the evolving data-driven approach to closely estimate wildfire risks even without running computationally expensive simulations. In one of our previous works, we validated the application with a probability-based risk metric that gives a series of probability values for a fire starting at a start location under a given weather condition. The probability values indicate how likely it is that a fire will fall into different risk categories. The metric considered each fire start location as a unique entity. Such a provision in the metric could expose the metric to scalability issues when the metric is used for a larger geographic area and consequently make the metric hugely intensive to compute. In this work, in an investigative effort, we investigate whether the spatial clustering of fire start locations based on historical fire areas can address the issue without significantly compromising the accuracy of the metric. Our results show that spatially clustering all fire start locations in Tasmania into three risk clusters could leverage the probability-based risk metric by reducing the computational requirements of the metric by a theoretical factor in thousands with a mere compromise of approximately 5% in accuracy for two risk categories of high and low, thereby validating the possibility of the leverage of the metric with spatial clustering.

Keywords: wildfire; risk metric; risk characterization; clustering; data-driven approach

Citation: KC, U.; Aryal, J. Leveraging a Wildfire Risk Prediction Metric with Spatial Clustering. *Fire* **2022**, *5*, 213. https://doi.org/10.3390/fire5060213

Academic Editor: Natasha Ribeiro

Received: 14 October 2022
Accepted: 6 December 2022
Published: 9 December 2022

1. Introduction

With an increased understanding of phenomena and advancements in computing technologies and observational sciences, natural disasters can be modeled and studied with more detail than ever before [1–4]. The precise information on the location of the disaster and the associated risks are in-demand elements in natural disaster modeling systems and the subsequent application of such systems in an operational environment. Thus, any spatial information on the disaster start locations and their propagation extent within the framework of risk quantification plays an inherent role in effective disaster management.

Consequently, in the fire disaster space, fire authorities have started widely using operational fire simulations for making better-informed decisions for wildfire management. Under current state-of-the-art of wildfire management, fire practitioners and authorities quantify the risk associated with wildfires by running several fire simulations in a geographical area by using operational fire spread models, collectively referred to as an ensemble and conducting statistical analyses [5–7]. These ensembles are computationally expensive but their recent extensive use for wildfire management has become possible due to unprecedented advancements in computing technologies such as cloud computing. The fire simulations, when aggregated on a massive scale, have created a unique opportunity to apply the evolving data-driven approach to closely estimate wildfire risks even without running a single computationally expensive simulation, although the process is highly computationally intensive otherwise.

Significant works have been done in the wildfire domain worldwide to quantify the risks associated with wildfires by using several fire danger indices. The list of some notable physical wildfire models developed in the literature includes Weber [8], FIRETEC [9], Forbes [10], Grishin [11], FIRESTAR [12], UoS [13], and WFDs [14]. Similarly, some empirical wildfire models discussed in the literature are CFS-accel [15], CALM Spinifex [16], CSIRO Grass [17], UdTM Shrub [18], Helsinki [19], and CSIRO Forest [20]. On the operational side, in Australia, fire authorities use the McArthur Forest Fire Danger Index (FFDI) [21], Grassland Fire Danger Index (GFDI) [22], and the National Fire Danger Rating System (NFDRS) [23] to provide fire danger ratings across Australia. The NFDRS was recently developed and categorizes wildfires into six different categories—low–moderate, high, very high, severe, extreme, and catastrophic. Earlier fire danger ratings were based on the simple thresholds established by Douglas [24] and McArthur [25]. Other fire danger rating systems used in Australia include "hazard sticks" [26], the difficulty of suppression tables [24], comprehensive fire danger ratings in grassland [27], and dry sclerophyll forest fuels [21,25], fire danger ratings developed specific to jarrah and karri forests in Western Australia [28] and fire danger ratings by the Australian Emergency Management Committee (AEMC) [29]. In the United States, fire authorities use the U.S National Fire Danger Rating System (U.S NFDRS) [30], which is largely based on Rothermel's fire spread model [31] and which quantifies wildfire risks as a function of different inputs such as topography, fuel type, weather conditions, and moisture content of dead and living vegetation. Canadian fire managers use Canadian Forest Fire Danger Rating System (CFFDRS) [32] to estimate wildfire risks daily. The CFFDRS has two subsystems—the Canadian Forest Fire Weather Index (FWI) system and the Canadian Forest Fire Behavior Prediction (FBP) system along with the Fire Occurrence Potential (FOP) system and the Accessory Fuel Moisture System [33]. These components take the likely number of fire ignition sources and fire behavior estimation in different fuel types to estimate fire danger ratings. Similarly, in South Africa, the Lowveld Fire Danger System [34] is used to rate wildfire risks. The system is an adaptation of the Fire Hazard Index that was developed by Laing [35] and which uses the same inputs as the McArthur FFDI. Several other fire danger rating systems are used throughout Europe. Sweden uses the Angstrom Index [36], Russia the Russian Ignition Index [37,38], France the fire danger rated using numerical risk [39,40], Portugal the Canadian Fire Weather Index for daily fire danger values on an approximately 1 km grid, Spain the Spanish Forest Fire Index (SFFI), which is an adaptation of the U.S NFDRS Ignition component. Finally, there were some efforts to introduce a wildfire danger rating system on a global scale. Such efforts include the work by Alexander [41] for a Global Wildland Fire Danger Rating System Di Giuseppe et al. [42]. This gave daily fire danger predictions together with other systems, Pettinari and Chuvieco [43] that were based on the simulation of surface fire behavior and the global fuel dataset using the Fuel Characteristic Classification Scheme [44].

In one of our previous works [45], we proposed a probability-based risk metric to characterize the risk of wildfires based on the weather conditions under which they start and the start location with a series of probability values of the fires falling into different risk categories. The work exploited the unique opportunity of applying a data-driven approach in wildfire management and validated the assumption that wildfire risks can be closely estimated even without running a single computationally expensive simulation, thanks to the evolving data-driven approach. The proposed metric was based on an inference model (Bayesian probabilistic model) that was built by using historical fire simulation data considering each fire start location as a separate entity. Despite the application of the model being computationally inexpensive, building the inference model was tedious and comparatively computationally intensive, as each possible fire start location was considered a unique entity and the probability values conditioned on fire start location had to be calculated for each location. Such a consideration can face serious scalability issues when the geographic area undertaken for wildfire risk estimation is large and has a significantly large number of possible fire start locations. Consequently, the effectiveness and usability of the proposed metric in wildfire management scenarios could be hindered.

As such, in this work, we investigate if the spatial clustering based on historical data (the data used to build the inference model) could address this scalability issue of the risk metric without significantly compromising the accuracy of the metric. The concept of spatial clusters was envisioned to leverage the risk metric as the number of spatial clusters is significantly less than the number of possible fire start locations in a geographic area, which could further reduce the overall computational requirements of the metric.

The rest of the paper is organized as follows. Section 2 briefly discusses our previously proposed risk metric and Section 3 details the methods and procedures for the study. Section 4 presents the results and discussions, and Section 5 concludes the paper with future possible works.

2. Risk Metric

Our previously proposed risk metric RM_{l_j,W_k}, for a region R with N set of possible fire start locations, $L = \{l_1, l_2, \ldots, l_N\}$, given a weather condition W at a particular time instant with k different weather parameters, $\{w_1, w_2, \ldots, w_k\}$, based on the Naive Bayes Theorem [46], is mathematically expressed as

$$RM_{L_j,W_k} = [P\{FR_1|(L_j, W_k)\}, P\{FR_2|(L_j, W_k)\}, \ldots, P\{FR_c|(L_j, W_k)\}], \quad (1)$$

where $P(FR_i|L_j, W_k)$ is the likelihood of a fire starting at location L_j under a discretized weather condition W_k falling into the risk category FR_i. By using the Naive Bayes Theorem [46], Equation (1) can further be expressed as

$$RM_{l_j,W_k} = \left[\frac{P(L_j|FR_1) \times P(W_j|FR_1) \times P(FR_i)}{P(L_j, W_k)}, \ldots, \frac{P(L_j|FR_c) \times P(W_j|FR_c) \times P(FR_c)}{P(L_j, W_k)}\right], \quad (2)$$

where $P(FR_i|W_k)$ is the likelihood of a fire starting under a discretized weather condition W_k falling into the risk category FR_i. It should be noted that the weather inputs in our study are uniformly sampled within their allowed ranges without any relevance to historical weather streams. $P(FR_i|L_j)$ is the likelihood of a fire starting at a location L_j falling into the risk category FR_i.

The risk metric gives a series of probability values for the fire to be labeled under different predefined categories, $\{FR_1, FR_2, \ldots, FR_c\}$, based on a model output O (for example, the possible area that can be burned by the fire within a particular period of time).

3. Methods and Procedures

3.1. Study Area

We chose Tasmania as the study area to validate the proposed idea of characterizing geographical regions into different risk zones by using spatial clustering. The choice was made for several reasons that include the frequent occurrence of wildfires in the region, the prevalence of readily available high-quality land datasets for the region that are usable in operational wildfire simulation tools, and a well-studied and systematic grid configuration of fire start locations in the region [7,47]. For the 2018–2019 wildfire season, Tasmania recorded 841 wildfires that burnt 310,311 hectares of forest areas [47]. Tasmania boasts high-quality land datasets available from TFS and State Emergency Service (SES) [48]. The Tasmania Fire Service (TFS) has maintained a grid configuration of 68,048 possible fire start locations at an interval of 1 km irrespective of land. Any start locations falling on the water bodies are shifted to the nearest land location. The Interim Biogeographic Regionalization for Australia Version 7 (IBRA 7) [49] divides Tasmania into nine different bioregions based on vegetation communities and ecosystems, which are shown in Figure 1.

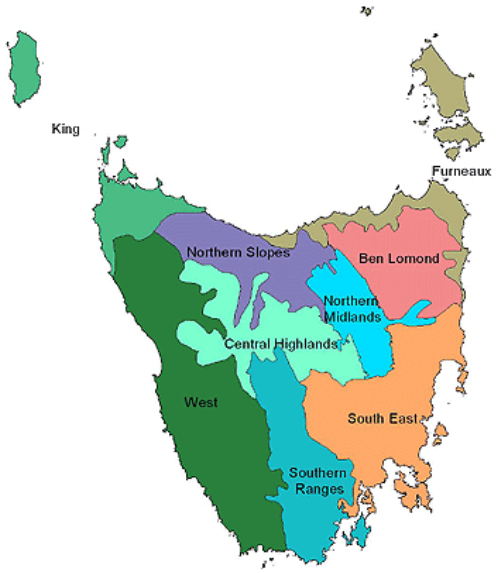


Figure 1. Nine bioregions of Tasmania. Source [49].

3.2. Wildfire Simulation Tool—Spark

We used the Spark framework [50] to run wildfire simulations and predict the spread of wildfires in different fire conditions and fuels, resulting in different rates of spread. Spark offers a flexible platform to simulate the progression of wildfires and their behaviors in different vegetation types by easily plugging in several packages and models. The packages and models can include the generation of wind fields and their topographic correction, fire ignition models, fire-line interactions, fireband transport, and fire transmission models [50]. Each wildfire simulation in Spark requires input datasets for fire behavior models, maps of land classification, fuel load, topographical datasets, and weather data to produce output metrics such as the total area burned by fire, the intensity of the fire, and the number of urban cells burned. Additionally, apart from fire progression, Spark can also predict fireband dynamics, and risk metrics for fire severity and impact. The fire simulations can run in Spark for any number of distinct fire perimeters and model firebreaks, spot fires, and coalescence between different parts of the fire over time. More information on Spark can be found at [51]. All the calculations in Spark are parallelized by using the OpenCL framework. Figure 2 shows a sample fire simulation in the Spark tool along with the regions burned by fire starting at a location at different time steps.

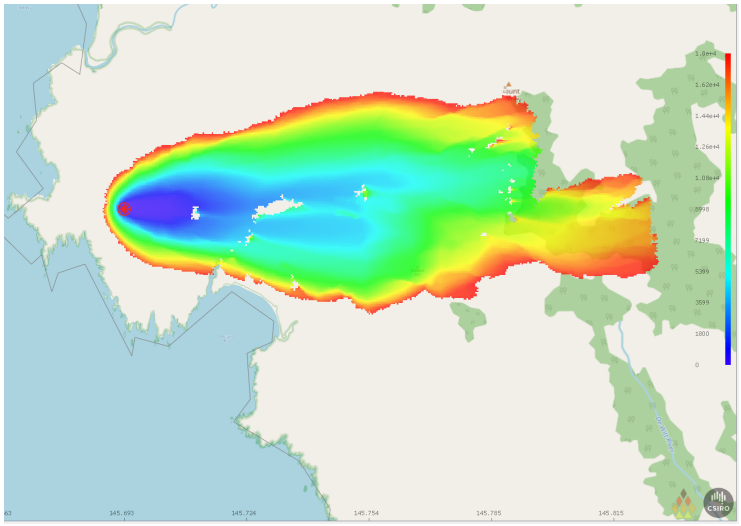


Figure 2. Visualization of an example fire simulation using Spark for a location in Tasmania, Australia.

3.3. Fire Simulations Inputs

We chose four weather inputs—temperature, relative humidity, wind speed, and wind direction—for this study following the experimental setup of one of our previous works [7,52]. The ranges for these inputs, considered for spatial clustering for risk zones, are listed in Table 1. Five equally spaced discrete values of each weather input (except wind direction) were considered along with four distinct directions (east, west, north, and south) for wind directions. Wildfires grow aggressively under weather conditions characterized by high values of wind speed, temperature, and low values of relative humidity when the wind is pushing the fire away from the water bodies. All other static inputs to fire simulations were used as per the configurations and records maintained by TFS and the Tasmanian government [53]. All fire simulations were run for five hours and the cumulative areas burned by fire in the period were reported as a simulation output.

Table 1. Range and discretization of the factors for fire weather.

Parameters (Unit)	Range	Labels with Interval
Air Temperature (°C)	[10, 40]	Low (L) [10,18] Medium (M) (18, 33) High (H) (33, 40]
Relative Humidity (%)	[10, 90]	Low (L) [70,90] Medium (M) (30, 70) High (H) [10, 30]
Wind Speed (kmh^{-1})	[10,60]	Low (L) [10,23] Medium (M) (23,48) High (H) [48, 60]

3.4. Fire Risk Categories

We considered three different fire risk categories—high, medium, and low—based on the total cumulative area burned by wildfires after five hours as listed in Table 2. Two fire risk categories, the medium and low fire risk categories, are merged to form one "low" risk category to be compared against the high-risk category. For risk estimation using the risk metric leveraged with spatial clustering for two risk categories, all the predictions for medium-risk categories were labeled as low-risk categories, unlike in our previous risk metric where the probability values for the low- and medium-risk categories were added before the comparison with the probability value for the high-risk category.

Table 2. Fire risk categories.

Fire Risk Category	Fire Area Size (ha)
High	$\geq$1057.42
Medium	[267.79, 1057.42)
Low	[0, 267.79)

3.5. Wildfire Risk Zones Assignment Using Spatial Clustering

We employed spatial clustering enabled by k-means clustering [54] based on the values of area burned by fires starting at the location to assign a risk zone to the fire start location. For a set of mathematically manipulated fire simulation outputs, i.e., fire burned areas $(O_1, \ldots, O_n)$ corresponding to fire start locations $(L_1, \ldots, L_n)$, the k-means clustering aims to categorize the n observations into k different sets $C = C_1, \ldots, C_k$ $(k \leq n)$ by minimizing the within-cluster sum of squares. The objective can be mathematically expressed as

$$\underset{C}{argmin} \sum_{i=1}^{k} \sum_{x \epsilon C_i} ||x - \mu_i||^2 = \underset{C}{argmin} \sum_{i=1}^{k} |C_i| Var C_i. \quad (3)$$

In the k-means clustering algorithm, each observation is assigned to a cluster with the nearest mean (mathematically the least squared Euclidean distance). The means for all the observations assigned to each cluster are recalculated. The algorithm converges when the assignments do not change. The cluster labels assigned for the manipulated output values are then transferred to the fire start locations to enable the spatial clustering of the start locations. The choice of k-means clustering follows the benefits the algorithm offers against its counterparts in clustering despite its computational complexities. The advantages of using k-means clustering include relatively simple implementation, scalability to large datasets, guaranteed convergence, easy adaptation to new datasets, and generalization to clusters of different shapes and sizes [55].

It should be noted that for any fire start location closer to any water bodies, the information on the fire-burned areas should be interpreted carefully. For example, under a given weather condition, a fire starting at a location with a water body to its east with the wind driving the fire toward the east ceases immediately, giving a totally unburned landscape, whereas any fire starting at the same location with the wind driving the fire in any direction but the east, may burn a significant area of land. To overcome such circumstances, we adapted the clustering mechanisms based on the mean and median values of fire-burned areas for all locations. The characteristics of the clusters obtained from clustering mechanisms were interpreted to label all fire start locations in Tasmania under three risk zones—low, medium, and high. This risk characterization of start locations is an intermediate result that is fed into the inference model to estimate the risk of any wildfires starting at these locations for a given weather condition. It should be noted that this risk characterization is solely to reduce the computational requirements of the risk metric.

3.6. Computing Environment for Wildfire Simulations

All the wildfire simulations used for this study were run by using the cloud-based frameworks as designed in [7] over the cloud infrastructure of Nectar Cloud [56] and Google Cloud [57]. Several types of cloud instances were used, as this study does not include any time-related evaluation metrics.

3.7. Fire Simulations Dataset

For 68,048 different fire start locations in Tasmania and 500 different weather input combinations, a total of 34,024,000 simulations were run with the cumulative area burned by the fire taken as the output metric. The entire can be accessed at [58].

3.8. Evaluation Metrics

Figure 3 shows the workflow of the risk metric leveraged with spatial risk clustering. The risk cluster assigned to each fire start location is the additional feature in the fire simulation dataset that was used in model building and assessment. We followed our previous evaluation setup where a conventional 75–25 train/test rule [59] with fivefold cross-validation was used and the average values were reported to assess the accuracy of the metric leveraged by the spatial clusters. The inference model estimates the risk for any fires that the model is not trained for, based on the likelihood of risk categories conditioned on the spatial clusters and the weather conditions. We compared the accuracy of the proposed cluster-based risk zone characterization against that of the baseline McArthur Forest Fire Danger Index (FFDI) [21] and our previously proposed risk metric [45]. Further details on the McArthur FFDI are included in Appendix A.1.

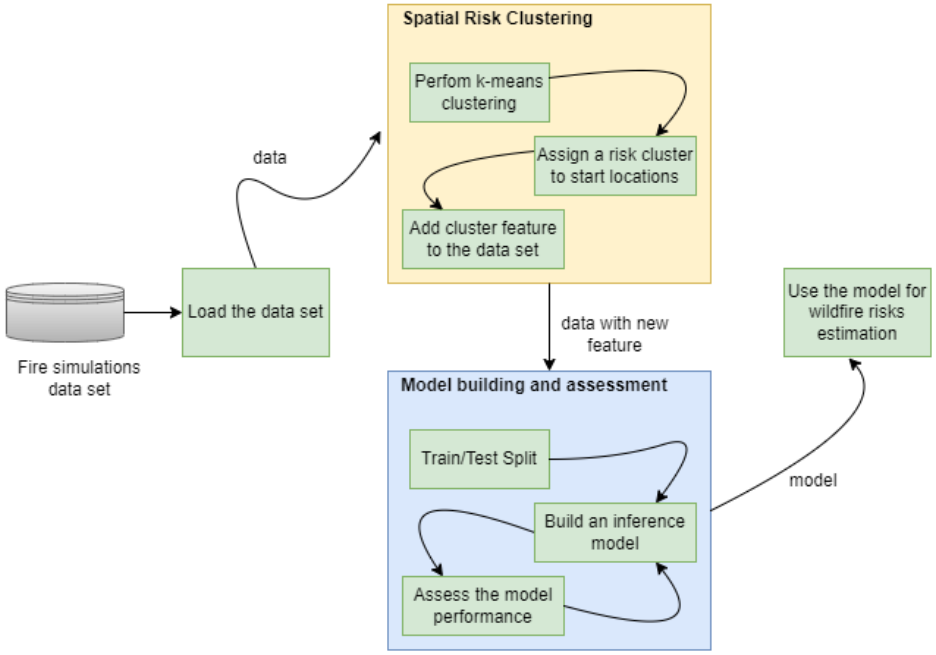


Figure 3. Workflow for the risk metric leveraged with spatial risk clustering.

We chose the McArthur FFDI as the baseline metric for evaluation as it is the closest metric to the proposed metric with the key meteorological inputs (air temperature, relative humidity, and wind speed) common to both. The long-term drought factor was fixed at the value of nine. We adapted and classified the FFDI into different fire danger rating (FDR) categories as described by Wain and Kepert [60] in Table 3 for comparison.

Table 3. Adapted fire danger rating categories of FFDI. For two risk categories during evaluation, low and medium categories are combined to form the low category.

FDR Category	FFDI Range	Adapted Risk Category
Low-Moderate	0–11	Low
High	12–23	Medium
Very High	24–49	
Severe	50–74	High
Extreme	75–99	
Catastrophic	>100	

Additionally, we also compared the ratios of underfits (predicted area less than the true values) and overfits (predicted area more than the true values). The comparison was done for three and two risk zones as initially explained in our previous work [45].

4. Results and Discussion

4.1. Risk Zone Characterization by Using Spatial Clustering

Figure 4 shows all possible fire start locations in Tasmania characterized as low-, medium-, and high-risk areas as given by the spatial clustering. Out of 68,048 fire start locations, approximately 66% of the locations were characterized as low-risk zones whereas the compositions for the medium- and the high-risk zone stood at approximately 24% and 10%, respectively. Most of the fire start locations closer to water bodies were labeled as low wildfire-risk zones whereas the fire start locations inward were labeled as high-risk zones. The high-risk zone had a range of an average fire area between 3900 and 10,400 hectares whereas the low-risk zone had the same range between 2 and 1700 hectares as seen in Figure 5. There is an overlap between the medium-risk zone with the low- and

high-risk zones in relation to the fire-burned areas. This is because the fire start locations are spatially clustered into different risk clusters and a single fire start location can start fires that can give cumulative burned area values falling into different risk categories as defined in Table 2 under various weather conditions and wind directions.

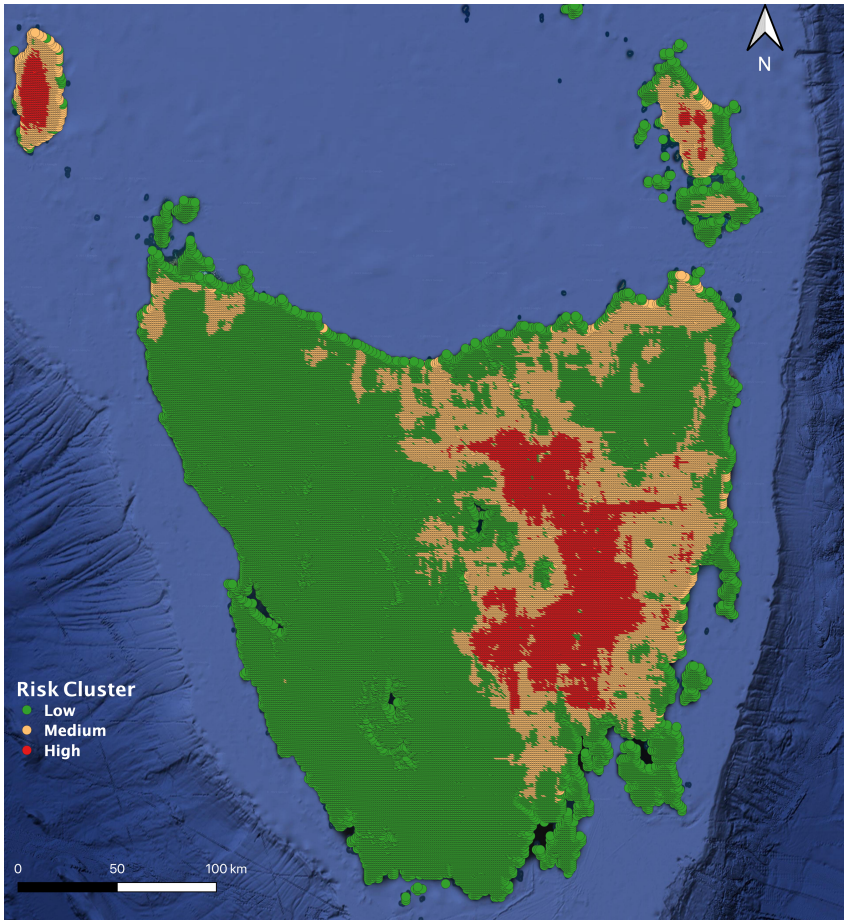


Figure 4. Risk zone characterization of all possible fire start locations in Tasmania.

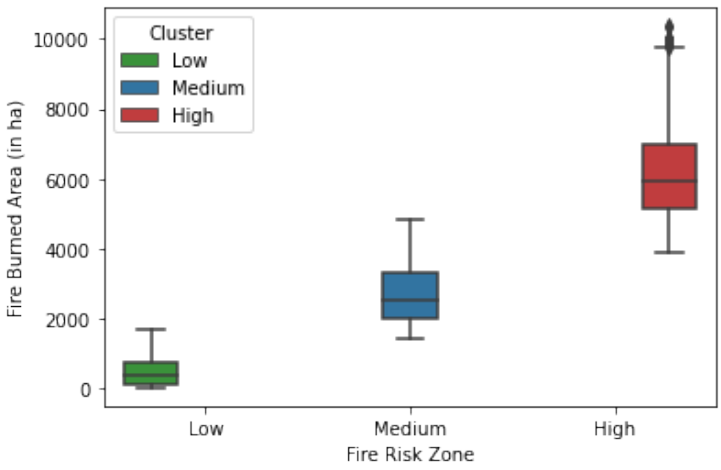


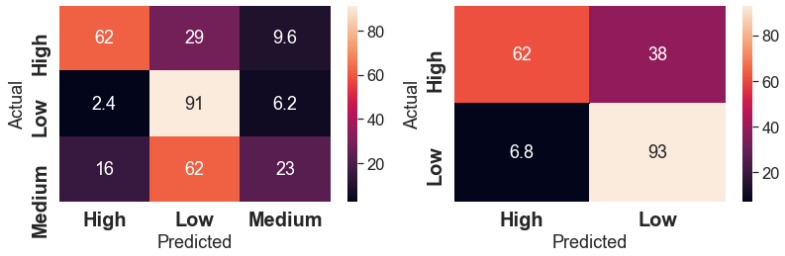
Figure 5. The statistical summary of fire areas categorized under three risk zones—low, medium, and high.

4.2. Comparison against the FFDI and the Risk Metric

Figure 6 shows the confusion matrices of the performance of the risk metric leveraged with spatial clustering for three and two distinct risk categories. The overall accuracy of the

risk metric was approximately 67% and 86% for three and two risk categories, respectively. For the three distinct risk categories, more than 90% of the low-risk fires were correctly labeled by the risk metric whereas the same for the medium- and high-risk categories stood at 23% and 62%, respectively. A large fraction of medium-risk fires were identified as low-risk fires, which could be due to the fact that the fire start locations closer to water bodies produced small fires burning a few hectares unless the wind was driving the fire away from the water bodies. The accumulation of a large number of such immediately ceasing wildfires at a start location may have contributed to the underestimation of the fire risks as the location would be spatially labeled a low-risk start location but could produce medium- or high-risk fires when the fire is growing away from the water bodies under favorable weather conditions. The underestimation of wildfire risk with the leveraged risk metric was improved when the fire risk categories were minimized to two as only approximately 10% of the total estimation was comprised of underfits, whereas the improvement for the overestimation was approximately 3%.

Table 4 shows the comparison of the cluster-based risk metric against the baseline FFDI and the previous wildfire risk metric for three and two risk categories. The accuracy of the cluster-based metric for three distinct risk categories stood at approximately 70% whereas the same for the FFDI and the previous metric stood at approximately 55% and 75% respectively. Similarly, for two distinct risk categories, the numbers stood at approximately 86%, 88%, and 77% for cluster-based, FFDI, and the previous metric, respectively. Clearly, as expected, the accuracy of the cluster-based risk metric is less than that of the previous risk metric. However, the accuracy is still better than that of the McArthur FFDI and is not considerably far from the accuracy of the previous metric. The cluster-based risk metric had a greater proportion of underfits than the previous metric for three risk categories, which could be due to a wider range of fire areas for the low-risk zone compared to the original range of low-risk zone in our previous metric. However, the proportion of underfits is less for the two risk categories. Conversely, the overfit proportion with the cluster-based metric is marginally less than in the previous metric for three risk categories, which could minimize the overestimation of resources during wildfire management.



(**a**) (**b**)

Figure 6. Confusion matrix for the performance analysis of the risk metric. (**a**) Three different fire risk categories. (**b**) Two different fire risk categories.

Table 4. Performance comparison against the McArthur FFDI.

Evaluation		**Three Categories**			**Two Categories**	
Metric	**Cluster-Based**	**Previous Metric**	**McArthur FFDI**	**Cluster-Based**	**Previous Metric**	**McArthur FFDI**
Accuracy	66.81%	74.55%	51.99%	85.32%	87.43%	76.03%
Underfit	24.98%	15.88%	38.87%	9.56%	10.27%	10.48%
Overfit	8.2%	9.66%	9.14%	5.12%	2.3%	13.49%

In this investigative study, we were able to cut down the location-specific probability calculations from 68,048 to a clustering mechanism and a significantly fewer number of 3 clusters. Such a provision in the risk metric could theoretically reduce the computational

requirements of the calculations of location-specific probability values by a factor of thousands. The most important fact in the findings is that the change in the accuracy of the metric is a mere 7% and 2% for three and two distinct risk categories, respectively. Such a significant reduction in the computational requirements proves the spatial clustering, as discussed in this paper, to be a good alternative for fire start location risk characterization in the data-driven probability-based wildfire risk metric.

This study has multiple implications for spatial information and location intelligence research in disaster risk reduction. The risk characterization at a spatial level can help in effective resource allocation for disaster management as the high-risk start locations can be prioritized during the planning phase. Additionally, such spatial risk characterization when coupled with weather estimation can quickly estimate the dangers of imminent wildfires and help in resource management and mobilization during the risk neutralization and disaster response. Furthermore, location-specific risk as obtained in this study is a key input in decision making in an operational fire management environment that can save lives and properties.

5. Conclusions and Future Research

The ever-evolving data-driven approach can be a computationally cheaper alternative for quick estimation of wildfire risks using several inference models. One such model was detailed in one of our previous works to build a probability-based risk metric, which was notably accurate in risk estimation. However, the metric was subjected to scalability issues, as each fire start location was considered a unique entity for location-related probability calculations. In this brief study, we investigated if spatial clustering could address the scalability issue of the metric without significantly compromising the accuracy of the metric. We found that the spatial clusters to characterize the risk of each start location could help solve the scalability issues by drastically reducing the number of calculations required by a theoretical factor in thousands with a mere compromise of approximately 5% in accuracy. Such an inexpensive estimation of wildfire risk with a data-driven metric can help fire authorities to prioritize resource allocation and make better-informed decisions at various stages of fire emergencies to minimize the possible losses.

We expect future works to verify the numbers around the factor by which the computational requirements of the probability-based risk metric are poised to get reduced with spatial clustering. Similarly, the influence of the number of spatial clusters on the accuracy of the metric could also be studied. This investigative study also opens up multiple directions of applied and theoretical research in spatial information applications for disaster risk reduction that can be complemented by a massive deployment of Internet of things (IoT) networks and sensors to feed real-time data into any risk-estimation metric or system to get the closest and most updated estimation for effective disaster management.

Author Contributions: Conceptualization, U.K. and J.A.; formal analysis, U.K.; investigation, U.K.; methodology, U.K. and J.A.; supervision, J.A.; validation, U.K. and J.A.; visualization, U.K.; writing—original draft, U.K.; writing—review & editing, U.K. and J.A. All authors have read and agreed to the published version of the manuscript.

Funding: The article processing charge for this article was covered by an internal grant from the University of Melbourne to Jagannath Aryal.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The fire simulation data set used for this work is available at https://doi.org/10.25919/XGAM-BS33 [58].

Acknowledgments: The authors would like to thank our academic colleagues who helped in improving the quality of the manuscript at various stages of the work.

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

Appendix A

Appendix A.1. The McArthur FFDI

The appendix is an optional section that can contain details and data supplemental to the main text—for example, explanations of experimental details that would disrupt the flow of the main text but nonetheless remain crucial to understanding and reproducing the research shown; figures of replicates for experiments of which representative data are shown in the main text can be added here if brief, or as Supplementary Data. Mathematical proofs of results not central to the paper can be added as an appendix.

The McArtur FFDI indicates the chances of fire starting and its risks in relation to its intensity and the difficulty of suppression based on the empirical relationship between short-term weather variables and drought factors. We used the relationship as defined by Noble et al. [61] and expressed FFDI as follows.

$$FFDI = 2.0 \times e^{(-0.450+0.987lnDF-0.0345RH+0.0338T+0.0234V)} \quad \text{(A1)}$$

where, DF is the drought factor, RH (%) is the relative humidity, T (°C) is the air temperature and V (kmh^{-1}) is the average wind velocity in an open flat location at a height of 10 m.

References

1. Yang, J. Convergence and uncertainty analyses in Monte-Carlo based sensitivity analysis. *Environ. Model. Softw.* **2011**, *26*, 444–457. [CrossRef]
2. Razavi, S.; Tolson, B.A.; Burn, D.H. Review of surrogate modeling in water resources. *Water Resour. Res.* **2012**, *48*, 1527. [CrossRef]
3. Yang, C.; Xu, Y.; Nebert, D. Redefining the possibility of digital Earth and geosciences with spatial cloud computing. *Int. J. Digit. Earth* **2013**, *6*, 297–312. [CrossRef]
4. Kaizer, J.S.; Heller, A.K.; Oberkampf, W.L. Scientific computer simulation review. *Reliab. Eng. Syst. Saf.* **2015**, *138*, 210–218. [CrossRef]
5. Cai, L.; He, H.S.; Liang, Y.; Wu, Z.; Huang, C. Analysis of the uncertainty of fuel model parameters in wildland fire modelling of a boreal forest in north-east China. *Int. J. Wildland Fire* **2019**, *28*, 205–215. [CrossRef]
6. KC, U.; Garg, S.; Hilton, J.; Aryal, J.; Forbes-Smith, N. Cloud Computing in natural hazard modeling systems: Current research trends and future directions. *Int. J. Disaster Risk Reduct.* **2019**, *38*, 101188. [CrossRef]
7. KC, U.; Garg, S.; Hilton, J. An efficient framework for ensemble of natural disaster simulations as a service. *Geosci. Front.* **2020**, *11*, 1859–1873. [CrossRef]
8. Weber, R. Toward a comprehensive wildlife spread model. *Int. J. Wildland Fire* **1991**, *1*, 245–248. [CrossRef]
9. Linn, R.; Reisner, J.; Colman, J.J.; Winterkamp, J. Studying wildfire behavior using FIRETEC. *Int. J. Wildland Fire* **2002**, *11*, 233–246. [CrossRef]
10. Forbes, L.K. A two-dimensional model for large-scale bushfire spread. *ANZIAM J.* **1997**, *39*, 171–194. [CrossRef]
11. Grishin, A. *Mathematical Modeling of Forest Fires and New Methods of Fighting Them*; Publishing House of Tomsk University: Tomsk, Russia, 1997.
12. Morvan, D.; Dupuy, J.L. Modeling of fire spread through a forest fuel bed using a multiphase formulation. *Combust. Flame* **2001**, *127*, 1981–1994. [CrossRef]
13. Asensio, M.; Ferragut, L. On a wildland fire model with radiation. *Int. J. Numer. Methods Eng.* **2002**, *54*, 137–157. [CrossRef]
14. Mell, W.; Jenkins, M.A.; Gould, J.; Cheney, P. A physics-based approach to modelling grassland fires. *Int. J. Wildland Fire* **2007**, *16*, 1–22. [CrossRef]
15. McAlpine, R.; Wakimoto, R. The acceleration of fire from point source to equilibrium spread. *For. Sci.* **1991**, *37*, 1314–1337.
16. Burrows, N. Experimental Development of a Fire Management Model for Jarrah (Eucalyptus Marginata) Forest. Ph.D. Thesis, Department of Forestry, Australian National University, Canberra, Australia, 1994.
17. Cheney, N.; Gould, J.; Catchpole, W.R. Prediction of fire spread in grasslands. *Int. J. Wildland Fire* **1998**, *8*, 1–13. [CrossRef]
18. Fernandes, P.A.M. Fire spread prediction in shrub fuels in Portugal. *For. Ecol. Manag.* **2001**, *144*, 67–74. [CrossRef]
19. Tanskanen, H.; Granström, A.; Larjavaara, M.; Puttonen, P. Experimental fire behaviour in managed Pinus sylvestris and Picea abies stands of Finland. *Int. J. Wildland Fire* **2007**, *16*, 414–425. [CrossRef]
20. Gould, J.S.; McCaw, W.; Cheney, N.; Ellis, P.; Knight, I.; Sullivan, A. *Project Vesta: Fire in Dry Eucalypt Forest: Fuel Structure, Fuel Dynamics and Fire Behaviour*; Csiro Publishing: Canberra, Australia, 2008.
21. McARTHUR, A. *Fire Behavior in Eucalypt Forest*; Department of Development, Forestry and Timber Bureau: Canberra, Australia, 1967.
22. McArthur, A. Weather and grassland fire behaviour. Commonwealth Department of National Development. *For. Timber Bur. Leafl.* **1966**, *100*.

23. Mathews, S.; Fox-Hughes, P.; Grootemaat, S.; Hollis, J.J.; Kenny, B.J.; Sauvage, S. *Australian Fire Danger Rating System: Research Prototype*; NSW Rural Fire Service: Lidcombe, NSW, Australia, 2019; 384p.
24. Douglas, D. *Forest Fire Weather Studies in South Australia*; Woods and Forests Department: Meadows, Australia, 1957.
25. McArthur, A. The preparation and use of fire danger tables. In Proceedings of the Fire Weather Conference, Australia Bureau of Meteorology, Melbourne, VIC, Australia, July 1958.
26. Gisborne, H. The wood cylinder method of measuring forest inflammability. *J. For.* **1933**, *31*, 673–679.
27. McArthur, A.G.; McArthur, M. *Fire Danger Rating Tables for Annual Grasslands*; Forestry and Timber Bureau: Canberra, Australia, 1960.
28. Peet, G. *Fire danger rating and controlled burning guide for the Northern Jarrah (Euc. marginata SM) Forest, of Western Australia*; Forests Department: Perth, Australia, 1965.
29. AEMC-National Bushfire Warnings Taskforce. *Australia's Revised Arrangements for Bushfire Advice and Alerts*; AEMC-National Bushfire Warnings Taskforce: Melbourne, VIC, Australia, 2009.
30. Cohen, J.D. *The National Fire-Danger Rating System: Basic Equations*; US Department of Agriculture, Forest Service, Pacific Southwest Forest and Range Rating System: Berkeley, CA, USA, 1985; Volume 82.
31. Rothermel, R.C. *A Mathematical Model for Predicting Fire Spread in Wildland Fuels*; Intermountain Forest and Range Experiment Station, Forest Service, US Department of Agriculture: Ogden, UT, USA, 1972; Volume 115.
32. Van Wagner, C.E.; Canadian Forestry Service. *Development and Structure of the Canadian Forest Fireweather Index System*; Forestry Technical Report 35; Canadian Forestry Service: Ottawa, ON, Canada, 1987.
33. Stocks, B.; Lawson, B.; Alexander, M.; Van Wagner, C.; McAlpine, R.; Lynham, T.; Dubé, D. The Canadian system of forest fire danger rating. In Proceedings of the Conference on Bushfire Modelling and Fire Danger Rating Systems, Canberra, Australia, 11–12 July 1988; pp. 11–12.
34. Meikle, S.; Heine, J. A fire danger index system for the Transvaal Lowveld and adjoining escarpment areas. *South Afr. For. J.* **1987**, *143*, 55–56. [CrossRef]
35. Laing, M. *Forecasting Bush and Forest Fire Weather in Rhodesia*; Meteorological Notes, Series B; Department Meteorological Services: Rhodesia, 1978.
36. Chandler, C.C. Risk rating for fire prevention planning. *J. For.* **1961**, *59*, 93–96.
37. Nesterov, V. *Flammability of the Forest and Methods for Its Determination, (Gorimost lesa i metodi eio opredelenia)*; Goslesbumizdat, USSR State Ind.: Moscow, Russia, 1949.
38. Nesterov, V. *Forest Fires and Methods of Fire Risk Determination*; Goslesbumizdat: Moscow, Russia, 1949.
39. Sol, B. Numerical meteorological hazard of forest fires in the Méditerran Rêgionéenne: Examination of the summer 1988 test and proposals for improvements. Work note SMIR/SE, N°1, France, 1989.
40. Drouet, J.C.; Sol, B. Development of a numerical index of meteorological risk of forest fires. *Dossier* **1993**, *14*, 155–162. Forets and fires 1991–1992.
41. Alexander, M. *Feasibility Study for the Setting up of a Global Wildland Fire Danger Rating System*; European Commission, Joint Research Centre, Institute of Environment and Sustainability: Ispra, Italy, 2010.
42. Di Giuseppe, F.; Pappenberger, F.; Wetterhall, F.; Krzeminski, B.; Camia, A.; Libertá, G.; San Miguel, J. The potential predictability of fire danger provided by numerical weather prediction. *J. Appl. Meteorol. Climatol.* **2016**, *55*, 2469–2491. [CrossRef]
43. Pettinari, M.L.; Chuvieco, E. Fire behavior simulation from global fuel and climatic information. *Forests* **2017**, *8*, 179. [CrossRef]
44. Ottmar, R.D.; Sandberg, D.V.; Riccardi, C.L.; Prichard, S.J. An overview of the fuel characteristic classification system—quantifying, classifying, and creating fuelbeds for resource planning. *Can. J. For. Res.* **2007**, *37*, 2383–2393. [CrossRef]
45. Ujjwal, K.; Hilton, J.; Garg, S.; Aryal, J. A probability-based risk metric for operational wildfire risk management. *Environ. Model. Softw.* **2022**, *148*, 105286.
46. Berry, D.A.; Berry, D.A. *Statistics: A Bayesian Perspective*; Number 04; QA279. 5, B4; Duxbury Press: Belmont, CA, USA, 1996.
47. Tasmania Fire Service. *State Fire Commission Annual Report*; Tasmania Fire Service: Tasmania, Australia, 2019.
48. *Monitoring, and Program*, TasVeg 3.0; Tasmanian Department of Primary Industries, Parks, Water & Environment: Hobart, Australia, 2013.
49. IBRA7. Available online: http://www.environment.gov.au/system/files/pages/5b3d2d31-2355-4b60-820c-e370572b2520/files/bioregions-new.pdf (accessed on 12 May 2021).
50. Miller, C.; Hilton, J.; Sullivan, A.; Prakash, M. SPARK–A bushfire spread prediction tool. In Proceedings of the International Symposium on Environmental Software Systems, Melbourne, VIC, Australia, 25–27 March 2015; pp. 262–271.
51. Spark: Predicting Bushfire Spread. Available online: https://data61.csiro.au/en/Our-Research/Our-Work/Safety-and-Security/Disaster-Management/Spark (accessed on 12 May 2021).
52. Ujjwal, K.; Garg, S.; Hilton, J.; Aryal, J. A cloud-based framework for sensitivity analysis of natural hazard models. *Environ. Model. Softw.* **2020**, *134*, 104800.
53. Tasmania. List Data. Available online: https://listdata.thelist.tas.gov.au/opendata/ (accessed on 12 March 2021).
54. Likas, A.; Vlassis, N.; Verbeek, J.J. The global k-means clustering algorithm. *Pattern Recognit.* **2003**, *36*, 451–461. [CrossRef]
55. Steinley, D. K-means clustering: A half-century synthesis. *Br. J. Math. Stat. Psychol.* **2006**, *59*, 1–34. [CrossRef]
56. Nectar Cloud. Available online: https://nectar.org.au/research-cloud/ (accessed on 12 May 2018).
57. Google Cloud. Available online: https://cloud.google.com/ (accessed on 12 May 2020).

58. KC, U.; Garg, S.; Hilton, J.; Aryal, J. *Fire Simulation Data Set for Tasmania*; CSIRO: Clayton South, VIC, Australia, 2021. [CrossRef]
59. Ayyadevara, V.K. Basics of machine learning. In *Pro Machine Learning Algorithms*; Springer: New York, NY, USA, 2018; pp. 1–15.
60. Wain, A.; Kepert, J.; Tory, K. *A Comprehensive, Nationally Consistent Climatology of Fire Weather Parameters for Australia*; Technical report, CAWCR Technical Report; Centre for Australian Weather and Climate Research: Melbourne, Australia, 2013.
61. Noble, I.; Gill, A.; Bary, G. McArthur's fire-danger meters expressed as equations. *Aust. J. Ecol.* **1980**, *5*, 201–203. [CrossRef]