

NRC Publications Archive Archives des publications du CNRC

A literature review on WUI fire susceptibility using machine learning Duong, Dana; Mozaffari, M. Hamed; Gomaa, Islam; Bénichou, Noureddine

For the publisher's version, please access the DOI link below./ Pour consulter la version de l'éditeur, utilisez le lien DOI ci-dessous.

Publisher's version / Version de l'éditeur:

<https://doi.org/10.4224/40003221>

Client Report (National Research Council of Canada. Construction); no. A1-020246-01, 2024

NRC Publications Archive Record / Notice des Archives des publications du CNRC :

<https://nrc-publications.canada.ca/eng/view/object/?id=733e2543-485a-4e9e-bca6-37678e2168f1>

<https://publications-cnrc.canada.ca/fra/voir/objet/?id=733e2543-485a-4e9e-bca6-37678e2168f1>

Access and use of this website and the material on it are subject to the Terms and Conditions set forth at

<https://nrc-publications.canada.ca/eng/copyright>

READ THESE TERMS AND CONDITIONS CAREFULLY BEFORE USING THIS WEBSITE.

L'accès à ce site Web et l'utilisation de son contenu sont assujettis aux conditions présentées dans le site

<https://publications-cnrc.canada.ca/fra/droits>

LISEZ CES CONDITIONS ATTENTIVEMENT AVANT D'UTILISER CE SITE WEB.

Questions? Contact the NRC Publications Archive team at

PublicationsArchive-ArchivesPublications@nrc-cnrc.gc.ca. If you wish to email the authors directly, please see the first page of the publication for their contact information.

Vous avez des questions? Nous pouvons vous aider. Pour communiquer directement avec un auteur, consultez la première page de la revue dans laquelle son article a été publié afin de trouver ses coordonnées. Si vous n'arrivez pas à les repérer, communiquez avec nous à PublicationsArchive-ArchivesPublications@nrc-cnrc.gc.ca.

AIRC·CMRC CONSTRUCTION

A Literature Review on WUI Fire Susceptibility using Machine Learning

Author(s): Dana Duong, M. Hamed Mozaffari, Islam Gomaa,
Noureddine Bénichou

Report No.: A1-020246-01

Report Date: 11-09-2022

Contract No.: A1-020246




© (2022) Her Majesty the Queen in Right of Canada,
as represented by the National Research Council Canada.

Cat. No. xxxxxxxxxxxx-PDF
ISBN xxxxxxxxxxxxxxxx



A Literature Review on WUI Fire Susceptibility Using Machine Learning

Author


Dana Duong, Student, PhD Candidate

Approved

Armstrong, Marianne
Digitally signed by Armstrong, Marianne
DN: cn=Armstrong, Marianne, c=CA,
o=NRC, ou=NRC-CNRC, email=
marianne.armstrong@cnrc-nrc.gc.ca
Date: 2023.08.25 08:44:35 -0400

Marianne Armstrong, CRBE Initiative Leader
NRC Construction Research Centre

Report No: A1-020246-01
Report Date: 11 09 2022
Contract No: A1-020246
Program: xxxxxxxxxxxxxxxxxxxx

26 pages

Copy no. 1 of 1

This report may not be reproduced in whole or in part without the written consent of the National Research Council Canada and the Client.

(This page is intentionally left blank)

Table of Contents

Table of Contents.....	i
List of Figures	ii
Nomenclature	ii
Executive Summary	3
1 Introduction.....	4
2 Background.....	5
3 Literature Review.....	13
4 Conclusions, Recommendations and Future Work.....	19
References	20

List of Figures

Figure 1. Examples of the support vector machine task.	6
Figure 2. Example of a Bayesian network.	7
Figure 3. Example of a decision tree.	8
Figure 4. Example of <i>k</i> -nearest neighbours algorithm.	8
Figure 5. Example of <i>k</i> -means algorithm.	9
Figure 6. Examples of generalized linear models.	9
Figure 7. Example of a hidden Markov model.	10
Figure 8. Schematic of: a) a neuron b) one hidden layer perceptron.	11
Figure 9. Example case of a one-hidden layer perceptron.	11
Figure 10. Convolutional neural network architecture. [towardsdatascience.com]	12

Nomenclature

Acronym

ML	Machine learning
DL	Deep learning
SVM	Support vector machine
BN	Bayesian networks
NB	Naïve Bayes
DT	Decision tree
RF	Random forest
GLM	Generalized linear model
KNN	<i>k</i> -nearest neighbour
KM	<i>k</i> -means
ANN	Artificial neural network
HMM	Hidden Markov model
CNN	Convolutional neural network
WUI	Wild-Urban Interface
CRBE	Center for Research in Building and Environment
STEM	Science, Technology, Engineering and Mathematics

Executive Summary

This report comprises a literature review on the use of machine learning (ML) in fire susceptibility/hazard mapping, fuel mapping/characterization and experimental work on ember-driven ignition. This review is part of the CRBE WUI Fire Hazard Assessment tool project. Previously, a literature review on the use of ML in fire dynamics and detection was conducted (Report A1-017985). The scope of the review involves:

1. Providing an overview of theoretical fundamentals on the ML applied in the literature.
2. Identifying the literature involved in ML in fire susceptibility/hazard mapping and fuel mapping/characterization.
3. Identifying the literature involved in experimental ember-driven ignition in as a guide to the data required to develop the ML-based tool.

Machine learning has been recently utilized in various applications and fields of science and technologies with impressive results. With this in mind, several studies employed machine learning for WUI fire incidents. The aim of previous studies in this literature have been to use machine learning methods for prediction of fire susceptibility (hazard) maps as well as maps of fuel characteristics. Overall, the trend found that ML-based methods, and in particular deep learning methods, were outperforming traditional methods such as linear and logistic regression in both fire susceptibility/hazard mapping and fuel mapping/characterization. With regards to the proposed CRBE WUI Ember Hazard Assessment tool project, this literature review shows that no study has examined the strict use of ML in the prediction of ignition probabilities/risk based on ember-driven ignition, but a similar study using machine learning to predict firebrand areal mass density and firebrand areal number density was found. Many studies have experimentally documented solely ember-driven ignition, and this data can be used in a ML algorithm. However, this literature review also identified gaps in the literature that can be addressed by additional experimental work that would enhance the proposed ML-based tool. Additionally, the literature presented novel work coupling thermal radiation with ember-driven ignition. These studies serve as a foundation for understanding the mechanisms behind WUI fires and as a foundation to a more complex ML-based algorithm that can predict coupled thermal radiation and ember-driven ignition.

1 Introduction

Six of the last ten most damaging fires involving structures over the last 100 years have been due to wild-urban interface (WUI) fires (NIST, 2021). The wild-urban interface is defined as the location where structures and communities meet or intermingle with undeveloped wildlands (NIST, 2021). Structures ignited from firebrands, commonly known as embers, has been shown from post-fire damage assessments to be a major source of structural ignitions in WUI fires (NIST, 2021). Although this has been shown to be a major problem, codes and standards do not adequately address the issue of firebrand exposure (NIST, 2021). Most structure ignitions are caused either indirectly by embers that ignite low-intensity surface fires next to a home or directly by ember ignitions on the home, not by direct flame heating (i.e., radiation or flame contact). Thus, the creation of new codes, standards or tools that attempt to address the issue of firebrand exposure is of great importance to public safety. This work attempts to tease out the needs of the latter, and in particular examine the literature of machine learning used in WUI fire science and fire safety.

Machine learning (ML) has become an ever-increasingly-used technology applied in both academia and industry outside of computer science (Butler *et al.*, 2018; Liakos *et al.*, 2018; Carleo *et al.*, 2019; Wei *et al.*, 2019; Brunton *et al.*, 2020; Meuwly, 2021). This is attributed to not only the wide array of techniques involved, but also due to the ability of ML to make fast and accurate predictions to the applied situation, when trained and tested properly. Machine learning is a subfield of artificial intelligence and is the study of computer algorithms/models that improve through experience (Murphy, 2012; Bengio *et al.*, 2017; LeCun *et al.*, 2015; Ray, 2019). This experience is obtained when the algorithms learn estimates of their parameters through describing underlying relationships in a given set of data. The model is often validated by plugging the estimated parameters back into the algorithm and check if the algorithm can predict a portion of the data when another portion is given. The above procedures are often referred as training and testing of the algorithms, while the data used in each procedure are termed the training data and the testing data.

There are three subfields of machine learning algorithms: **supervised learning, unsupervised learning, and reinforcement learning**. Supervised learning is when the algorithms treat a set of observations as an input/predictor and learn the underlying mapping of them to a set of associated observations, termed the output/response. In contrary, unsupervised learning is when the algorithms learn the underlying patterns from a set of observations with no associated output/response provided. Lastly, reinforcement learning is when the algorithms learn to make decisions to maximize rewards and minimize penalties based on actions taken by the agent who will be completing a task. This review focuses on supervised and unsupervised learning.

Supervised and unsupervised learning algorithms are used to solve a few general tasks: classification, regression and clustering. Classification and regression are supervised learning tasks where algorithm learn the underlying mapping of an input to a discrete or continuous output, respectively. On the other hand, clustering is an unsupervised learning task where algorithms

learn to group/label datasets with similar properties. There are two specific examples of application in computer science that demonstrate the utilization of these tasks. Object detection deals with distinguishing between objects in an image or video. It is more specific in what it identifies, applying classification to distinct objects in an image/video and using bounding boxes to tell us where each object is in an image/video. On the other hand, segmentation is a type of labeling where each pixel in an image is labeled with given concepts. Whole images are divided into pixel groupings which can then be labeled and classified, with the goal of simplifying an image or changing how an image is presented to a model, making it easier to analyze. Segmentation models provide the exact outline of the object within an image, unlike classification models, where the model identifies what is in an image, and detection models, which place a bounding box around specific objects. The main focus on this background review will be on classification and clustering and their methods and algorithms.

This literature review is a component of the ongoing CRBE project, Fire Hazard Assessment Tools and Standard for Structures in WUI Areas, which is connected to the overarching WUI Standards project. This literature review explores the use of machine learning to WUI fire susceptibility and risk, and ember ignition probability prediction. WUI fire susceptibility is a broad topic, ranging from susceptibility mapping, to forecasting WUI fire occurrence, and to ignition probability of wildland fuel beds (Jain *et al.*, 2020; Tymstra *et al.*, 2020; Mell *et al.*, 2010). This review will also examine the available experimental literature data for ember ignition probability prediction, while addressing the gaps that can be filled. The examination of the available experimental work is relevant for this project because the data will be used to train a ML-based classifier for ember ignition risk prediction tool, another component in the project. Thus, a thorough review of the experimental literature will be important. This report comprises of a theory section, which provides a high-level overview of the theoretical background to some of the ML methods used in the fire susceptibility literature. The literature review section, where the literature on the use of ML in fire susceptibility and experimental work on ember ignition prediction is examined in detail; and the conclusion section, where some conclusions and recommendations will be drawn from the current literature to apply to the ignition risk classifier.

2 Machine learning Algorithms Used in Fire Susceptibility

Support vector machines (SVMs) are classifiers that determine the decision boundary in an n -dimensional space between different data classes and are generally used in supervised learning. The SVM determines the decision boundary by maximizing the Euclidian distance, known as the margin, between the closest data points of each class to the boundary. Figure 1 presents two examples of an SVM classification task. Two data sets, C_1 and C_2 , are separated by a hyperplane that is determined by the SVM.

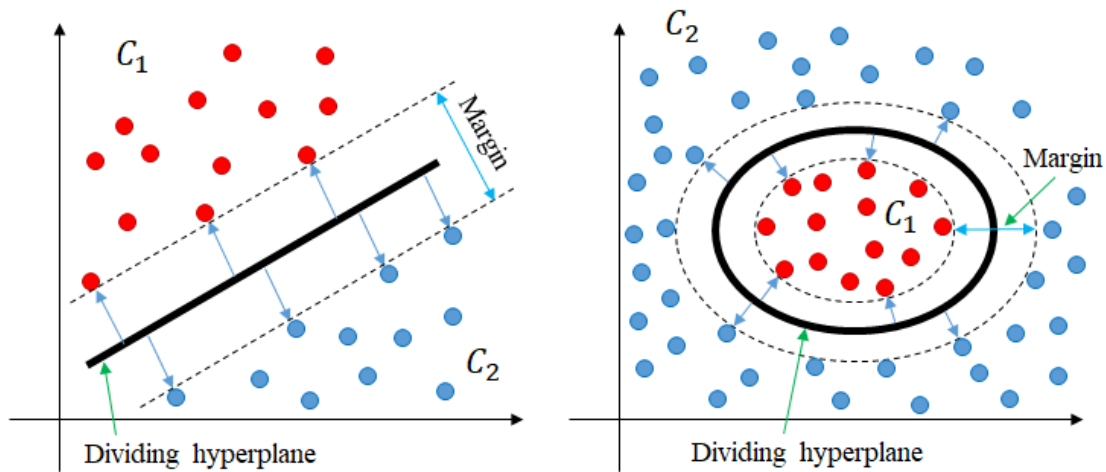


Figure 1. Examples of the support vector machine task.

Bayesian networks (BNs) are a type of probabilistic graphical model that uses Bayes' theorem, which relates the conditional and marginal probabilities of random variables, to determine relationships between variables. BNs are generally used in supervised learning classification. The goal of a Bayesian network is to model causation by representing the conditional dependencies by edges (or arrows) in a direct graph (Figure 2). Naïve Bayes is a subset of Bayesian networks, but with the assumption of the input features being independent from one another and having equal contributions to the output. The Naïve Bayes (NB) classifier is simple to implement but can lead to errors in prediction if variables are not independent. Figure 2 presents an example of a Bayesian network. Starting at the first node (Cloudy), there is a 30% probability that the day is cloudy. The first node in the second layer tells the reader the probability that it will be raining on that day. If the day is cloudy, that is, if $C = \text{True (T)}$, there's a 70% chance that it will rain. In the other case, if $C = \text{False (F)}$, there is a 0.5% chance that it will rain on that day. Similar logic applied to the sprinkler node: If the day is cloudy, there is a small chance (10%) that the sprinkler will be placed on the lawn, while if the day is clear, there will be an 80% chance that the sprinkler will be placed on the lawn. The line that connects the *Cloudy* and *Rain* or *Cloudy* and *Sprinkler* nodes represents a causal relationship. In the case of the Naïve Bayes network, arrows are generally in the reverse causal direction. In the Naïve Bayes network example in Figure 2, arrows would point out of the *Wet Grass* node, to the *Rain* and *Sprinkler* nodes, and the *Cloudy* feature node would be removed, as it has no contribution to the *Wet Grass* node.

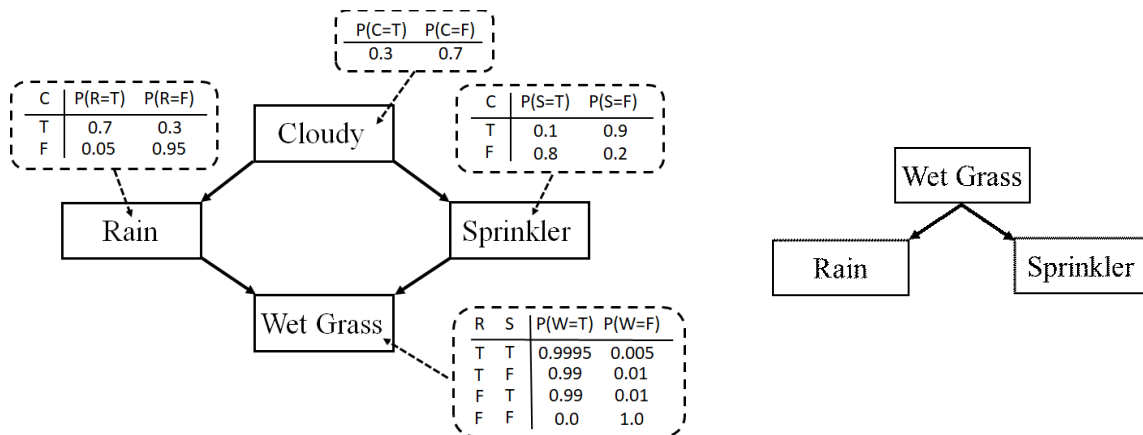


Figure 2. Example of a Bayesian network (left). Example of a Naive Bayesian network (right).

Decision trees (DTs) are a type of non-parametric, sequential, and hierarchical graphical model used in ML for both supervised classification and regression predictions. They are structured with branches, decisions nodes and leaf nodes. The decision (or internal) node splits the tree into different branches, with each branch representing the path the ML algorithm has taken to another decision node. At the end of the path, the tree is terminated with a leaf node, which represents the output of the ML algorithm. The main goal of the DT algorithm is the represent the task at hand with the most compact tree possible. The tree is built in a process called *induction*, where all of the hierarchical decision boundaries are set. Removing the unnecessary structures from the DT is performed in a process called *pruning*. Some major advantages of DTs are that they are highly interpretable, even to non-experts and can lead of insights that would be otherwise difficult to determine, and they can be constructed easily. They can also handle both categorical and numerical data. Some disadvantage of DTs are that they are unstable, as small changes in the data can lead to large changes in the structure of the tree and they are relatively inaccurate compared to other methods such as artificial neural networks. Figure 3 presents an example of a DT. In this example, the goal is to determine the day's action. The first decision node is to determine whether one has work to do. The node then has two branches, "No" or "Yes". If "No", then the algorithm moves down to the next decision node, the weather outlook. If "Yes", then the path terminates at the leaf node "Stay in", to complete the work. This process occurs until reaching a leaf node, representing the output of the algorithm. Random forests (RFs) are ensemble models that are an extension of DTs, where many DTs are trained and ensembled. Once ensembled (or bagged), the decision (or classification) with the most votes is determined to be the model output.

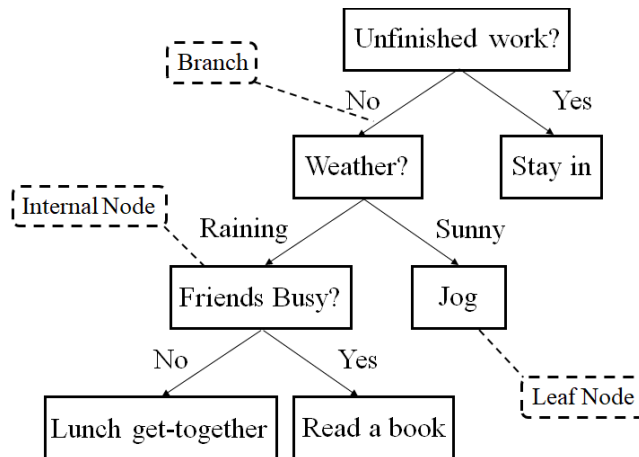


Figure 3. Example of a decision tree.

K-nearest neighbours (KNNs) algorithm is a non-parametric classification algorithm used in supervised learning, based on the idea that similar data points are close to each other based on some property or feature. KNNs determine the similarity of data points by using the Euclidian distance between the k closest data points. Figure 4 presents an example of a KNN algorithm. Suppose we have some data point, represented by the green diamond in Figure 4a that needs to be classified. In the KNN algorithm, one of the parameters is to specify the number of nearest neighbours to compare the unknown data point with. Suppose we had chosen the nearest three data points, represented by the black circle in Figure 4b. The green diamond would then be classified as a red dot, as this is most common neighbour of the three data points. Suppose we expanded it to the nearest five points of data (Figure 4c). The center data point of interest will be labeled as a blue dot, based on the closest neighbours in this case.

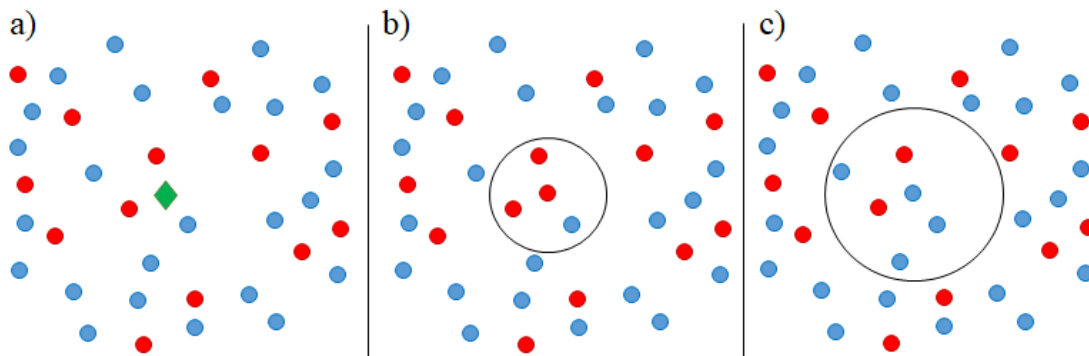


Figure 4. Example of k -nearest neighbours algorithm.

K-means (KMs) is a clustering algorithm used in unsupervised learning. It is similar to the KNN algorithm, but the difference is that k refers to the number of cluster centroids to cluster the data around. More importantly, the label of the data are known in the KNN algorithm (coloured in the figure) and hence serves as the output in a classification task. In the KMs algorithm, the labels are unknown (uncoloured in the figure) and hence it is a clustering task. The goal of the algorithm

is to determine the k number of centroids and cluster the data to the nearest centroid while keeping the centroids as small as possible. Figure 5 presents an example of a k -means algorithm, with three centroids (represented by the \otimes) classifying the data clusters.

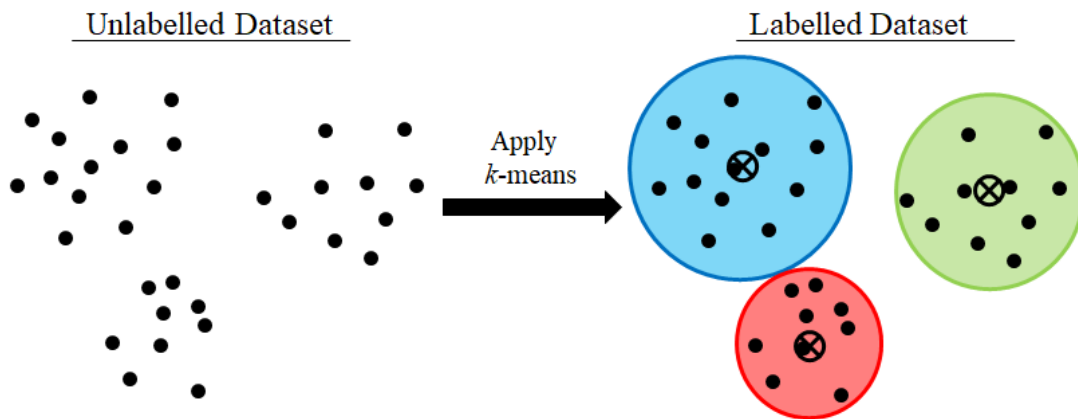


Figure 5. Example of k -means algorithm.

Generalized linear models (GLMs) can be used in both supervised classification and regression tasks. The most common GLM is linear regression or multi-variate linear regression. Another common GLM is logistic regression. The goal of the GLM is to determine a linear mapping between the inputs/features and outputs, while minimizing the error between the prediction and the output data. Figure 6 presents an illustration of a linear regression (a) and a logistic regression (b). The red lines represent the model and the blue dots represent the data, with input x and output y .

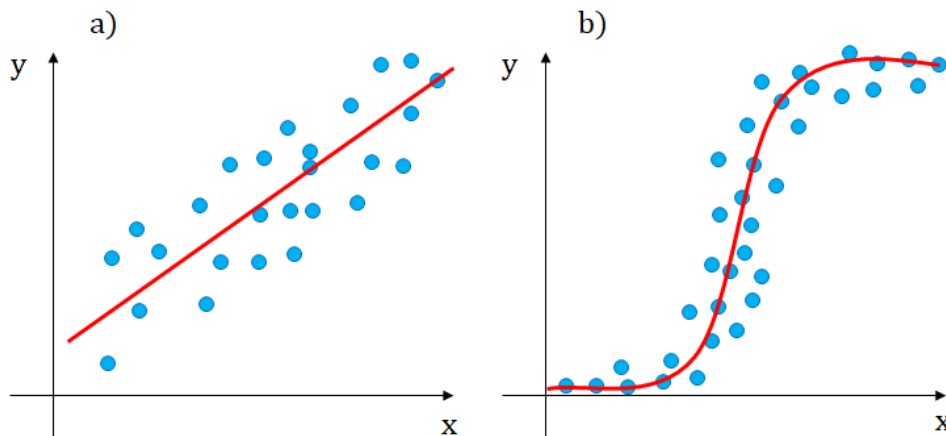


Figure 6. Examples of generalized linear models.

Hidden Markov models (HMMs) are unsupervised learning probabilistic models that cluster sequences (as opposed to feature inputs), based on Markov processes (Smyth, 1996). However, they are often trained using supervised learning methods when training data is available. Markov

processes are random processes that are indexed by time with the property that the future state is independent of the past state, but dependent on the present state. The goal is to find the best set of states and probabilities to derive the maximum likelihood estimate of parameters given the set of output sequences. Figure 7 is an example of an HMM. In this example, suppose two friends Jane and Jill are working in an office with no windows. Jane has a rough understanding of Jill's emotions (*Happy* or *Grumpy*) as Jill walks into the office based on the weather (*Sunny* or *Rainy*). Jane does not know how the weather is outside, but based on Jill's responses, the weather can be probabilistically inferred based on Jill's emotions. In the example, *Sunny* and *Rainy* nodes are the hidden states, while the state-dependent outputs (*Happy* and *Grumpy*) are visible.

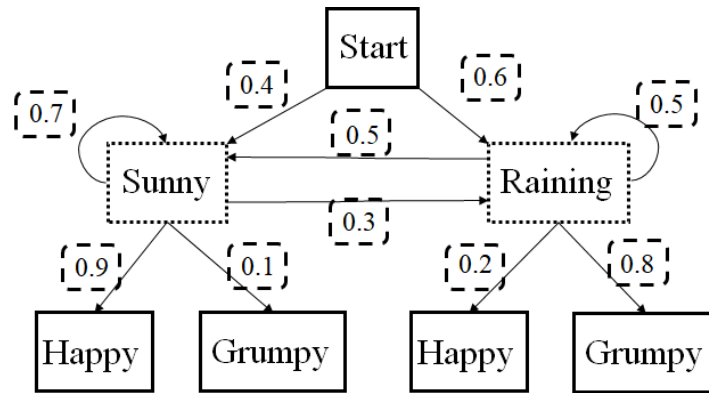


Figure 7. Example of a hidden Markov model.

Artificial neural networks (ANNs) are computing systems inspired by and are an attempt to mimic the behaviour of the human brain. They are used in supervised, unsupervised and reinforcement learning, for a wide array of tasks. The basic units of an ANN are neurons (Figure 8a). Neurons are composed of three parts: the input, the activation function and the output. For observations $i = 1, \dots, n$ and neurons $j = 1, \dots, m$, the input, x_j is where the data or the output of the previous neuron is fed into. The data is then passed through an activation function g , which performs a calculation based on applied weights, w_{ij} , multiplied to the input values for some function g , resulting in the output y_i described by the equation:

$$y_i = g \left(\sum_{j=1}^m w_{ij} x_j \right)$$

Some activation functions include the sigmoid, hyperbolic tangent and rectified linear unit. Figure 8a presented a simplified single-input-single-output neuron. Connecting many of these neurons in parallel in multiple layers results in a fully-connected network. A simplified fully-connected network with one output node at the end of the network is called a perceptron (Figure 8b). The first layer of the perceptron is the input layer, where the dataset is passed into the neural network. The following layers are called hidden layers. These layers perform a calculation based on the applied weights and input values from the previous layer for some activation function. Finally, the last layer is called the output layer. This final layer outputs the prediction h_θ , from the neural

network. When there are two or more hidden layers, the network is considered a deep neural network, and in the case of the perceptron, it would be a multi-layer perceptron. Suppose the task of the perceptron is to classify whether a material is combustible or not (Figure 9). Examples of inputs could be chemical composition, local temperature, and local surrounding oxygen content. The values of these inputs are passed into the hidden layer, a_i^2 , where the each input is multiplied by some weight and passed through an activation function. These are then summed in the output layer. In this example, the output layer would represent if the material would combust or not (a probability between 0 and 1). Many other types of artificial neural networks such as convolutional neural networks (CNNs) and recurrent neural networks exist, with the main differences being how the layers are connected, the types of layers in each network, the data-type of the input data, the data-type of the output data, and the task to be completed with each network. For the purpose of this review, a high-level introduction to convolutional neural networks will be reviewed.

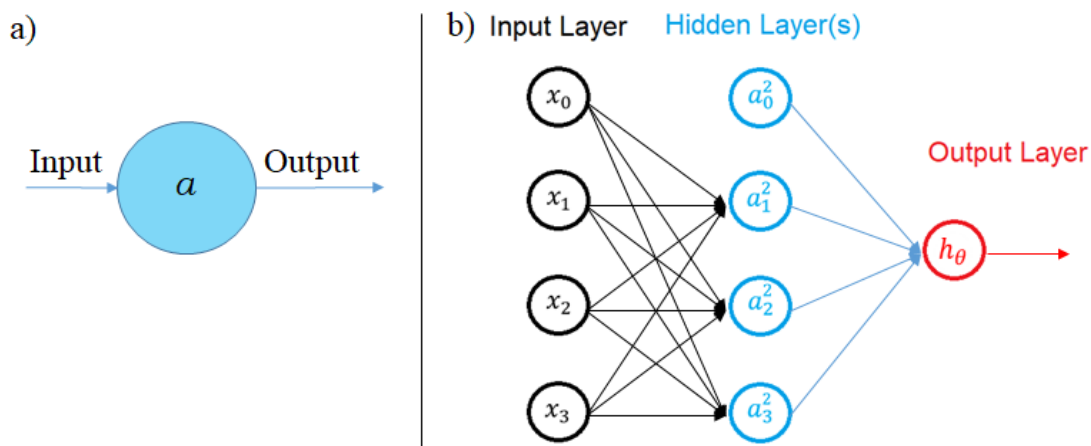


Figure 8. Schematic of: a) a neuron b) one hidden layer perceptron.

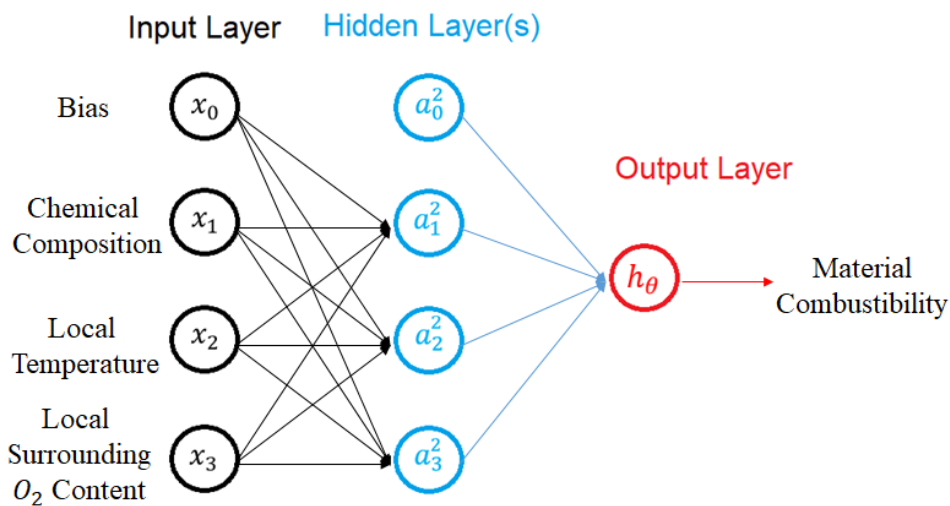


Figure 9. Example case of a one-hidden layer perceptron.

Convolutional neural networks are a common type of deep ANN that specializes in image processing. It takes an input image and learns the various features and/or objects in the image to differentiate the objects within the image. This is attributed to the convolution and pooling layers that allow the CNN to learn the features in the input images. The convolutional layer uses customizable filters to learn different features while scanning the input image. The resulting learned feature is dimensionally reduced in the pooling layer to reduce computational cost while keeping the dominant aspects of the learned features. Examples of features could be specific shapes or curves unique to the object in the image. These features are finally passed into a classification network to determine the objects in the image. Figure 10 presents the CNN architecture. For further reading and details on CNNs, refer to Krizhevsky *et al.* (2017).

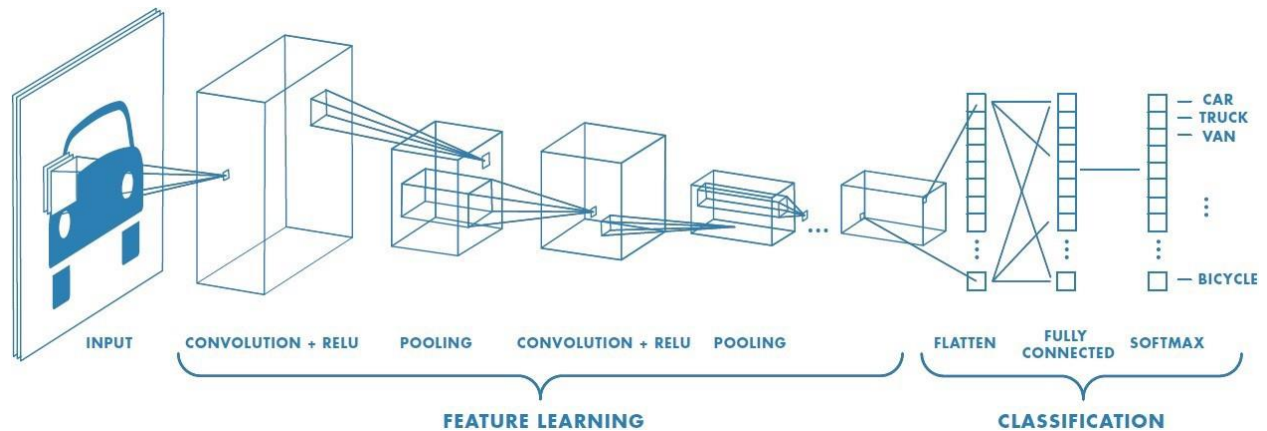


Figure 10. Convolutional neural network architecture. [towardsdatascience.com]

3 Literature Review

Machine learning has seen explosive advancements and uses in various fields such as image detection, speech recognition, and personalized profile recommendations. More recently, ML has also been applied in STEM with excellent results (Butler *et al.*, 2018; Liakos *et al.*, 2018; Carleo *et al.*, 2019; Wei *et al.*, 2019; Brunton *et al.*, 2020; Meuwly, 2021). In particular, ML has been seeing successful application in the field of fire safety and fire science (Cao *et al.*, 2019; Wu *et al.*, 2020; Park *et al.*, 2019; Hodges, 2018; Hodges *et al.* 2019a; Hodges *et al.* 2019b; Lattimer *et al.* 2020). The previous review (Duong and Gomaa, 2021) examined literature pertaining to ML used in fire dynamics and fire detection. In particular, it was shown that ML can be used to provide quick and accurate predictions of fire behaviour in compartments fires and predictions of the behaviour of fire properties such as the spread rate and heat release rate. On the detection side, it has been found that ML could be used not only to determine if a fire is present but also determine the general location of the fire. Naturally, its extension to wildfire science is expected, with several different machine learning-based methods having been tested. For a recent thorough review on the application of machine learning in wildfire science, please refer to Jain *et al.* (2020). In this literature review, the authors are interested in the wildfire literature pertaining to machine learning applied to fire susceptibility, risk and ignition, fuel characterization and mapping, and experimental work on ember induced ignition.

3.1 Machine Learning in Fire Susceptibility Mapping

The general approach of fire susceptibility mapping was to build a spatial fire susceptibility/ignition model using fire data composing of a combination of landscape, climate, structural, and anthropogenic variables as explanatory variables. Fire susceptibility can be attributed to many factors, including fuel type or land cover type, weather, climate change, landscape and for building structures, the material composition. Many works used various methods such as decision trees (DTs), random forests (RFs), artificial neural networks (ANNs), support vector machines (SVMs) and linear/logistic regression (LR) to map fire susceptibility in particular local regions. Amatuli and Camia (2007) compared decision trees against multi-variate adaptive regression splines (Friedman, 1991), a type of non-linear regression, for predicting fire occurrence in a local area of Italy. Input variables included the type of wild land cover, altitude, slope, latitude, longitude, and density and distances from primary and secondary roads, among others, while the output variable was the fire occurrence density. The authors found that the DT model outperformed the multi-variate adaptive regression spline model, with the latter model under-predicting the fire occurrence, even sometimes predicting a negative value for the fire occurrence. However, it is noted by the authors that compared to the decision tree model, the multi-variate adaptive regression spline model had a smoother transition between regions. In similar work for the northern Iranian region of Golestan, Adab (2017) investigated the influence of environmental and anthropogenic factors on land fire hazard, as well as compared ANNs with binary logistic regression on the prediction of land fire hazard. Input variable included land cover type, annual mean temperature, total solar radiation (insolation), elevation, slope, distance from roads and

settlements, normalized difference moisture index and topographic wetness index, with the output variable being fire occurrence location. The author found that the ANN model outperformed the binary logistic regression model, with the ANN model correctly locating fire spots 87% of the time, compared to the 81% accuracy of the binary logistic regression model. Additionally, the ANN model showed that anthropogenic variables described the variability of land fire frequency and that the annual precipitation, the annual mean temperature, the topographic wetness index, the type of land cover and the slope were significant variables in the prediction of land fire hazard. Guo *et al.* (2016a, 2016b) explored environmental and anthropogenic factors on forest fire occurrence in the Chinese boreal forests and forests of the Fujian province, comparing RFs to logistic regression. Similar input variables as previously mentioned works were used, and maps of likelihood of fire occurrence were produced as the output. It was found that RFs outperformed the logistic regression models, with accuracies in the mid-80%'s for the RF models compared to the mid-70%'s for the logistic regression models. Additionally, authors compared the effects of local factors, such as infrastructure, vegetation, topographic and socioeconomic factors to the effects of climate factors, such as seasonal temperature, precipitation and relative humidity; climate factors were the main driving factors of fire occurrence. In another work for the Spain region, Rodrigues and de la Riva (2014) explored the use of RFs and SVMs against logistic regression to predict wildfire ignition probabilities. Input features include various socio-economic factors (*i.e.*, forestry area in public utilities, WUI areas), traditional economic activities in rural areas (*i.e.*, wildland-agricultural interface), variables that can cause a fire by accident or by negligence (*i.e.*, power lines, railways) and factors that would hamper fires (*i.e.*, protected areas). With these wildfire ignition probabilities from different areas of Spain, a wildfire ignition probability map of Spain was produced from the models. Overall, authors found that the RF model performed the best, followed by the SVM model, then the logistic regression model. Interestingly, the authors found that the dominant factors in predicting fire occurrence was the density of agricultural machinery, negligence and accidents due to engines, machines working in or close to forest areas, the use of fire to clean up harvest wastes and crop boundaries, and the increase in human presence and pressure near wildlands in Spain, highlighting the importance of forest protection due to anthropogenic factors. In another work by Hong *et al.* (2018) applied a genetic algorithm to determine the most relevant features in predicting forest fire-related variables in Dayu County, Southeast China. A total set of thirteen variables were used, namely: land elevation, land slope angle, land aspect, land curvature, land use, soil cover, heat load index, normalized difference vegetation index, mean annual temperature, mean annual wind speed, mean annual rainfall, distance to the nearest river network and distance to the nearest road network. The important variables were found to be land elevation, land slope angle, land curvature, soil cover, normalized difference vegetation index, mean annual temperature, mean annual wind speed, and distance to the nearest road network. After prediction of the important variables, comparison of the performances of an SVM against an RF was conducted with the curated dataset. It was found that RF outperformed the SVM significantly. Another study along the same theme was performed by Jaafari and Pourghasemi (2019). Authors used an RF to determine the most important and relevant factors in wildfire occurrence from a set of factors. Afterwards, an SVM was used on the curated data to predict the wildfire probability in their study area of Central-West Iran. Similar to the Rodrigues and de la Riva (2014) study, anthropogenic factors were found the significantly

impact the occurrence of wildfire. However, unlike the Hong *et al.* (2018) study, annual rainfall was found to be an important factor in the prediction of wildfire prediction. Similar works explored other regions such as other regions of Iran (Goldarag *et al.*, 2016; Ghorbanzedehe *et al.*, 2019a, 2019b), Spain (Bisquert *et al.*, 2012), Italy (Amatulli *et al.*, 2006), Catalonia (Duane *et al.*, 2015), Serbia (Gigovic *et al.*, 2019), Portugal (De Vasoncelos *et al.*, 2001), China (Cao *et al.*, 2017), India (Sachdeva *et al.*, 2018), Vietnam (Thach *et al.*, 2018; Tehrany *et al.*, 2019) and parts of Mediterranean Europe (Oliveira *et al.*, 2012) using similar machine learning methodologies. In general, it was found that DTs, RFs and ANNs tended to outperform linear and logistic regression. This is attributed to the assumption of spatial stationarity in the relationship among the response variable and the predictors by linear regression techniques (Amatuli and Camia, 2007). Often there are non-linear and complex relationships between the factors that contribute to the occurrence of fire. Additionally, linear regression techniques generally do not allow or only allow very careful use of discrete variables, unlike in DTs, RFs and ANNs (Amatuli and Camia, 2007; Guo *et al.*, 2016a, 2016b; Goldarag *et al.*, 2016). Techniques such as binary logistic regression are also sensitive to sample distribution, *i.e.* the regression curve will incline towards the class with more abundant data (Goldarag *et al.*, 2016). Machine learning methods such as ANNs are less sensitive to distribution skewness (Goldarag *et al.*, 2016). Logistic and linear regression models are also sensitive to outliers and noise in the data, in comparison to other machine learning models such as DTs and RFs (Guo *et al.*, 2016a, 2016b). Several other authors compared and/or used other ML methods, such as max entropy such as in Vilar *et al.* (2016), Bayesian networks such as in Bashari *et al.* (2016), and neuro-fuzzy systems such as in Jaafari *et al.* (2019) and Bui *et al.* (2017). Again, it was found that these other ML models tended to outperform linear and logistic regression. Many of the listed works also compared performances of these methods to one another such as in Cao *et al.* (2017), Ghorbanzedehe *et al.* (2019a), Gigovic *et al.* (2019), Jaafari *et al.* (2018), Thach *et al.* (2018), Bui *et al.* (2017), Pourtaghi *et al.* (2016), Rodrigues and de la Riva (2014), Tehrany *et al.* (2019). The performance comparison results were mixed, as some authors found that SVM performed the best, while others found that RFs or ANNs performed best. While interesting, this introduces the possibility of a meta-analysis on the data and/or features used to train the networks, at least to perhaps explain the large differences in the performances of the various models. Nevertheless, these works showcase the capabilities of machine learning in the use of wildfire occurrence/risk prediction.

Many of the previously mentioned works analysis involved some time-averaged features. This approximation risks losing time-dependent factors that may contribute to fire occurrence/risk. Hence time-dependent factors, such as seasonality (Martin *et al.*, 2019; Vacchiano *et al.*, 2018), year-by-year (Parisien *et al.*, 2014) and month of year (Peters *et al.*, 2013; Peters and Iverson, 2017) were examined. Martin *et al.* (2019) specifically examined working days and non-working days throughout the four seasons in Spain. This produced eight different data scenarios that used to train their max entropy model. The input features to their model were the same as those in Rodrigues and de la Riva (2014). Their model outperformed time-averaged models (Rodrigues and de la Riva, 2014), for all seasons. In particular, the summer models significantly outperformed the time-averaged model. Authors also examined the input variable contribution for each of the eight scenarios, showing that the significant variables changed, depending on the time of year

and type of day (working or non-working). Parisien *et al.* (2014) produced year-by-year generalized linear models and compared them to a longer-time-averaged generalized model (31 years) for the Canadian boreal forest. It was found that the annual models slightly outperformed the multi-decade model. Peters *et al.* (2013) explored temporal factors on the monthly resolution scale using max entropy and achieving similar findings; that is, the finer temporal resolution (monthly) model outperformed the longer time-averaged (annual) model and that which of the significant input features of the model are temporally dependent. These works highlight the importance of accounting for time-dependent features to produce an accurate prediction model.

Recently the application of deep learning models, such as convolutional neural networks and recurrent neural networks, have been applied to wildfire science. One of the earliest studies applying a convolutional neural network to fire susceptibility mapping, by Zhang *et al.* (2019), compared convolutional neural network performance to previously benchmark methods, including RFs, SVMs, multi-layer perceptrons, and kernel logistic regression. Input features in the study included land elevation, land slope, land aspect, average temperature, average precipitation, surface roughness, average wind speed, maximum temperature, specific humidity, precipitation rate, forest coverage ratio, normalized difference vegetation index, distance to roads and distance to rivers. Their convolutional neural network model had an overall validation accuracy of 87.92%, compared to the accuracies of 84.36%, 80.04%, 78.47%, and 81.23% accuracies of the RF model, SVM model, multi-layer perceptron model, and kernel logistic regression model, respectively. In a follow up study, Zhang *et al.* (2021) used a convolutional neural network to create a global seasonal wildfire susceptibility map and comparing its performance to a multi-layer perceptron, with their convolutional neural network model outperforming the multi-layer perceptron. Authors split the data into four groups, each group containing three consecutive months starting with March, April, and May. The feature variables in the study were the average temperature, wind speed, specific humidity, accumulated precipitation, soil moisture, soil temperature, monthly maximum temperature, monthly minimum temperature, standard precipitation index, leaf area index, and normalized difference vegetation index. For the first group of data (March, April, and May), their convolutional neural network resulted in an overall accuracy of 91.08% compared to the 87.74% of the multi-layer perceptron. Their model has an accuracy of 89.61%, 93.18%, and 94.88% compared to the multi-layer perceptron's accuracy of 88.7%, 90.27%, and 93.50% for the second, third and fourth groups of data, respectively. Zhang *et al.* (2022) further developed a spatio-temporal model, a convolutional neural network combined with a long-short-term-memory network to predict future wildfire burned area. Similar to their previous work (Zhang *et al.*, 2021), they split the data into monthly data to be fed into the model sequentially, with similar input features. It was found that their combined model outperformed the convolutional neural network model. It is noted by the authors that previously, traditional models were often used due to limited data; however, with the recent growth and diversity of satellite-based data, the use of deep learning is becoming very possible (Zhang *et al.*, 2022).

3.2 Machine Learning in Fuel Characterization/Land Cover Mapping

Machine learning-based approaches in fuel characterization/land cover has been examined in two different approaches; prediction of fuel quantities such as the crown biomass of single trees through height and diameter and prediction of fuel quantities/types over a landscape. For example, in Riano *et al.* (2005), an ANN was used to predict and map the water thickness and dry matter content of wet and dry leaf samples using reflectance and transmittance values. Other studies identified important fuel variables such as the canopy cover, height, base height, bulk density and other biophysical variables to produce fuel maps using RFs (Pierce *et al.*, 2012; Riley *et al.*, 2014). Garcia *et al.* (2011) used an SVM to classify LiDAR and multispectral data, mapping fuel types in Spain. The authors classified the fuel types into seven groups, mainly separating tree and shrub ratios and sizes. Their model produced an overall accuracy of 88.24%. Chirici *et al.* (2013) compared the use of DTs, RFs and stochastic gradient boosting for mapping forest fuel types in Italy. The authors classified the fuel types in nine groups, mainly based on fuel bed structure, splitting grasslands from shrublands from ground litter. The gradient boosted trees were shown to outperform both the DT and RF models significantly, with an overall accuracy of 84% compared to the accuracies of 47% and 45% of the other two models. López-Serrano *et al.* (2016) compared the performances of SVM, KNN, RF and linear regression in estimating aboveground biomass in Mexico. Similar to many previous studies, the authors found that the SVM, KNN and RF outperformed the traditional regression methods. With the growth, abundance and variety of mapping data, such as LiDAR, Landsat, multi-spectral and UAV image mapping data, deep learning has seen recent success in the classification of fuel types/land cover types. For example, Wang *et al.* (2018) used LiDAR data and compared the performances of convolutional neural networks against SVMs for classification of land classes. Authors had two datasets, the first with seven land classes, the second with fifteen land classes. Overall it was found that convolutional neural networks outperformed SVMs significantly. A similar study (Maltezos *et al.*, 2019) compared the performances of a convolutional neural network to ANNs and SVMs, with similar results. Since then, several other studies have applied convolutional neural networks to fuel/land type mapping data (Xu *et al.*, 2018; Zorzi *et al.*, 2019; Yu *et al.*, 2020; Mayra *et al.*, 2021; Morchhale *et al.*, 2016; Pan *et al.*, 2020; Kwan *et al.*, 2020; Carranza-Garcia *et al.*, 2019; Arief *et al.*, 2018; Al-Najjar *et al.*, 2019; Du and Du, 2019). Different deep learning models were also developed and compared against previous benchmark convolutional neural network models, such as a 3D convolutional neural network (Xu *et al.*, 2018; Mayra *et al.*, 2021), a hybrid capsule network (Yu *et al.*, 2020), and a stochastic atrous network (Arief *et al.*, 2018). These works again highlight the improvement upon previous models using deep learning, owing to the growth and diversity of mapping data.

3.3 Experimental Ember-Driven Ignition

Relevant literature to this project are bench-scale and intermediate-scale experiments with WUI wildfires. In particular, ignition of fuel beds through hot embers and radiative flux are of interest. Simple bench-scale experiments of ignition of fuel beds from hot embers are fundamental in studying the effects of various factors in ember-driven ignition. Many experiments studied various factors such as ember temperature, ember size, ember material, number of ember particles in different fuel beds under varying conditions such as moisture content. Several studies examined different metal ember particles, such as steel, aluminum, brass and copper with varying temperatures and sizes (Fernandez-Pello *et al.* 2015; Urban *et al.*, 2015; Urban *et al.*, 2018; Zak *et al.*, 2014), while many others examined various firebrand ember particles (Hadden and Scott, 2011; Manzello *et al.*, 2006; Manzello *et al.*, 2008; Manzello *et al.*, 2009; Guijarro *et al.*, 2002; Sun *et al.*, 2018; Urban *et al.*, 2019; Viegas *et al.*, 2014; Viegas *et al.*, 2021; Yang *et al.*, 2022). Various fuel beds were also examined, such as cellulose (Hadden and Scott, 2011; Fernandez-Pello *et al.* 2015; Zak *et al.*, 2014; Urban *et al.*, 2015; Urban *et al.*, 2018; Urban *et al.*, 2019), sawdust (Urban *et al.*, 2019), grasses, barks and leaves (Viegas *et al.*, 2014; Urban *et al.*, 2018; Sun *et al.*, 2018; Manzello *et al.*, 2005; Manzello *et al.*, 2008; Guijarro *et al.*, 2002; Yang *et al.*, 2021) and paper and straw (Viegas *et al.*, 2021, Manzello *et al.*, 2005; Manzello *et al.*, 2008). Additionally, some of these studies examined building materials such as plywood, OSB, cedar shingles and several types of insulation (Manzello *et al.*, 2005; Manzello *et al.*, 2009; Wessies *et al.*, 2019). For metal particles, mapping models estimating the regions of no flaming ignition, possible flaming ignition and flaming ignition and regions of no smouldering ignition, possible smouldering ignition and smouldering ignition mapped by ember temperature and temperature diameter have been determined. As one would expect, higher ember temperatures with larger particles (higher ember particle thermal energy) resulted in higher likelihood of ignition. Another important factor that was determined was the moisture content; higher moisture content reduced that likelihood of ignition (Viegas *et al.*, 2014; Yang *et al.*, 2021). A few studies examined the effect of the number of firebrands with very few tests (Manzello *et al.*, 2005; Manzello *et al.*, 2008; Manzello *et al.*, 2009). Overall it was found that the larger number of particles increased the likelihood of ignition. These studies and more (Kasymov *et al.*, 2016; Yang *et al.*, 2021) also examined the effects of cross-wind flow speed, firebrand size and fuel bed density. As expected, increasing flow speed (without blowing the embers off the material), increasing the ember size and decreasing the fuel bed density increased the likelihood of ignition as well as increasing the likelihood of transition from smouldering ignition to flaming ignition. All of the previously mentioned studies examined solely ember-driven WUI fires; however, WUI fires are driven by a multitude of factors. A recent study documented the coupled effects of a hot metal particle with thermal radiation (Fang *et al.*, 2021). As expected, the coupled effects significantly increased ignition than either of the factors alone. Intermediate scale experiments focused on WUI wildfires mainly involved showering fuel beds and/or construction material with many ember particles (Suzuki *et al.*, 2015; Suzuki and Johnsson, 2016; Manzello *et al.*, 2017; Suzuki and Manzello, 2020). These studies documented the behaviour of the ignition, as well as identified key factors that produced ignition. Some key factors included the accumulations of the firebrands, the fuel moisture content and the wind speed. A recent study documented the coupled effects of firebrand

showers with thermal radiation (Suzuki and Manzello, 2021). An interesting observation from the study was that under slower wind speeds, thermal radiation played a larger role in ignition, while at higher speed the firebrand shower played a larger role in ignition. For a comprehensive review pre-2018, refer to Fernandez-Pello (2017). Very recently, machine learning approaches have been applied to the prediction of ember/firebrand related properties. Jha *et al.* (2022) compared a KNN model against several machine learning models, including logistic regression, Naïve Bayes, SVM, DT and RF, in the prediction of firebrand areal mass density and firebrand areal number density. The input features included the firebrand travel distance, the type of structural component, the construction material, the vegetation type and species, and the wind speed. The authors found that the KNN model outperformed the other models significantly. It should be noted that the dataset was produced in previous in-house studies (Hedayati *et al.*, 2015; Zhou *et al.*, 2019; Hedayati *et al.*, n.d.; Bahrani *et al.*, n.d.), highlighting further the need and scarcity of experimental data. In light of the available literature, only one study has examined the use of ML in the prediction of ignition probabilities based on ember-driven ignition or ember properties. This literature review serves as a foundation to explore this topic further. However, some gaps in the literature are still present, which also serves as a guide to which additional experimental work can be performed. The effects of the number of particles are examined, but with few data points. Additionally, the studies examined limited cross-wind speeds. Finally, additional investigations on the coupled effects of thermal radiation and ember are paramount to understanding real WUI fires.

4 Conclusions, Recommendations and Future Work

Machine learning has recently seen applications in various fields with impressive results. With this in mind, several ML-based studies in WUI fire safety have been performed. In particular, machine learning has seen great success in fire susceptibility, risk and ignition, fuel characterization and mapping. This literature review explores the work in these topics. Previous studies have shown that machine learning significantly outperforms traditional method such as linear and logistic regression. More recently, deep learning has been applied to these problems as well because of the ever-growing amount of diverse mapping data. Deep learning methods such as deep convolutional neural networks have also been shown to outperform previous benchmark methods, including SVMs, ANNs, DTs and RFs.

This literature review also explored experimental works in ember-driven ignition, as a guide to the data required to develop the ML-based tool. A majority of the experimental work addressed ember-only ignition, with very little work coupling thermal radiation- and ember ignition. This suggests that additional experimental work coupling these effects is a necessary endeavor. This review also shows that the use of ML in the prediction of ignition probabilities/risk based on ember-driven ignition is limited. With the available experimental ember-driven ignition/risk data from the literature, machine learning methods, such as an artificial neural network or a support vector machine can be used to create a data-driven model. However, some challenges need to be addressed. The first challenge is the collection and concatenation of the available experimental data in the literature. The next challenge is the curation of the dataset. Across the experimental literature, different features of ember ignition are explored, with gaps between features and some

features not reported in some works. This poses the challenge of filling in the gaps, either through additional experimental work or through reduction of the features in the dataset. A thorough exploration of the input features is necessary. Finally, once the dataset has been explored, the creation of a framework for the machine learning is required.

References

Abdalhaq, B., Cortés, A., Margalef, T., & Luque, E. "Enhancing wildland fire prediction on cluster systems applying evolutionary optimization techniques." *Future Gener. Comp. Sy.* 21.1 (2005): 61-67.

Adab, H. "Landfire hazard assessment in the Caspian Hyrcanian forest ecoregion with the long-term MODIS active fire data." *Nat. Hazards.* 87.3 (2017): 1807-1825.

Al-Najjar, H. A., Kalantar, B., Pradhan, B., Saeidi, V., Halin, A. A., Ueda, N., & Mansor, S. "Land cover classification from fused DSM and UAV images using convolutional neural networks." *Remote Sens-Basel* 11.12 (2019): 1461.

Amatulli, G., Rodrigues, M. J., Trombetti, M., & Lovreglio, R. "Assessing long-term fire risk at local scale by means of decision tree technique." *J. Geophys. Res:-Biogeo.* 111.G4 (2006).

Amatulli, G., & Camia, A. "Exploring the relationships of fire occurrence variables by means of CART and MARS models." *Proc. 4th Int. Wildland Fire Conf., May 13-17, 2007, Sevilla, Spain*, pp. 1–11.

Arief, H. A. A., Strand, G. H., Tveite, H., & Indahl, U. G. "Land cover segmentation of airborne LiDAR data using stochastic atrous network." *Remote Sens-Basel* 10.6 (2018): 973.

Bahrani B., Hedayati F., Zhou A., Quarles S. L., & Weise D. R. (n.d.) Data for firebrands generated from selected vegetative fuels: *Joint Fire Science Program project (15-1-04-4)*. Technical report. Forest Service Research Data Archive. Type: dataset

Bashari, H., Naghipour, A. A., Khajeddin, S. J., Sangoony, H., & Tahmasebi, P. "Risk of fire occurrence in arid and semi-arid ecosystems of Iran: an investigation using Bayesian belief networks." *Environ. Monit. Assess.* 188.9 (2016): 1-15.

Bengio, Y., Goodfellow I., and Courville A. *Deep learning*, vol. 1. Cambridge, MA, USA: MIT press, 2017.

Bisquert, M., Caselles, E., Sánchez, J. M., & Caselles, V. "Application of artificial neural networks and logistic regression to the prediction of forest fire danger in Galicia using MODIS data." *Int. J. Wildland Fire.* 21.8 (2012): 1025-1029.

Brunton, S. L., Noack, B. R., & Koumoutsakos, P. "Machine learning for fluid mechanics." *Annu. Rev. Fluid Mech.* 52 (2020): 477-508.

Bui, D. T., Bui, Q. T., Nguyen, Q. P., Pradhan, B., Nampak, H., & Trinh, P. T. "A hybrid artificial intelligence approach using GIS-based neural-fuzzy inference system and particle swarm optimization for forest fire susceptibility modeling at a tropical area." *Agr. For. Meteorol.* 233 (2017): 32-44.

Butler, K. T., Davies, D. W., Cartwright, H., Isayev, O., & Walsh, A. "Machine learning for molecular and materials science." *Nature.* 559.7715 (2018): 547-555.

Cao, Y., Wang, M., & Liu, K. "Wildfire susceptibility assessment in Southern China: a comparison of multiple methods." *Int. J. Disast. Risk Sc.* 8.2 (2017): 164-181.

- Cao, Y., Yang, F., Tang, Q., & Lu, X. "An attention enhanced bidirectional LSTM for early forest fire smoke recognition." *IEEE Access* 7 (2019): 154732-154742.
- Carleo, G., Cirac, I., Cranmer, K., Daudet, L., Schuld, M., Tishby, N., Vogt-Maranto, L., & Zdeborová, L. "Machine learning and the physical sciences." *Rev. Mod. Phys.* 91.4 (2019): 045002.
- Carranza-García, M., García-Gutiérrez, J., & Riquelme, J. C. "A framework for evaluating land use and land cover classification using convolutional neural networks." *Remote Sens-Basel* 11.3 (2019): 274.
- Chang, M. K. *Development of a testing protocol for insulation ignition by wildland fire embers*. 2017. University of Texas, Austin, PhD dissertation.
- Chicas, S. D., & Østergaard N. J. "Who are the actors and what are the factors that are used in models to map forest fire susceptibility? A systematic review." *Nat. Hazards*. (2022): 1-18.
- Chirici, G., Scotti, R., Montagni, A., Barbati, A., Cartisano, R., Lopez, G., Marchetti, M., McRoberts, R.E., Olsson, H., & Corona, P. "Stochastic gradient boosting classification trees for forest fuel types mapping through airborne laser scanning and IRS LISS-III imagery." *Int. J. Appl. Earth Obs.* 25 (2013): 87-97.
- De Vasconcelos, M. P., Silva, S., Tome, M., Alvim, M., & Pereira, J. C. "Spatial prediction of fire ignition probabilities: comparing logistic regression and neural networks." *Photogramm. Eng. Rem. S.* 67.1 (2001): 73-81.
- Dlamini, W. M. "Application of Bayesian networks for fire risk mapping using GIS and remote sensing data." *GeoJ* 76.3 (2011): 283-296.
- Du, Shouji, & Shihong Du. "Land cover classification using remote sensing images and lidar data." *IGARSS 2019-2019 IEEE Int. Geosci. Remote Sens. Symp.* IEEE, 2019.
- Duong, D. & Gomaa, I. "A Literature Review on Machine Learning and Deep Learning in Fire Dynamics and Detection," A1-017985, National Research Council of Canada, Internal Report, Apr. 2021.
- Duane, A., Pique, M., Castellnou, M., & Brotons, L. "Predictive modelling of fire occurrences from different fire spread patterns in Mediterranean landscapes." *Int. J. Wildland Fire* 24.3 (2015): 407-418.
- Fang, W., Peng, Z., & Chen, H. "Ignition of pine needle fuel bed by the coupled effects of a hot metal particle and thermal radiation." *Proc. Combust. Inst.* 38.3 (2021): 5101-5108.
- Fernandez-Pello, A. C., Lautenberger, C., Rich, D., Zak, C., Urban, J., Hadden, R., Scott, S., & Fereres, S. "Spot fire ignition of natural fuel beds by hot metal particles, embers, and sparks." *Combust. Sci. Technol.* 187.1-2 (2015): 269-295.
- Fernandez-Pello, A. C. "Wildland fire spot ignition by sparks and firebrands." *Fire Safety J.* 91 (2017): 2-10.
- García, M., Riaño, D., Chuvieco, E., Salas, J., & Danson, F. M. "Multispectral and LiDAR data fusion for fuel type mapping using Support Vector Machine and decision rules." *Remote Sens. Environ.* 115.6 (2011): 1369-1379.
- Ghorbanzadeh, O., Valizadeh Kamran, K., Blaschke, T., Aryal, J., Naboureh, A., Einali, J., & Bian, J. "Spatial prediction of wildfire susceptibility using field survey gps data and machine learning approaches." *Fire* 2.3 (2019): 43.
- Ghorbanzadeh, O., Blaschke, T., Gholamnia, K., & Aryal, J. "Forest fire susceptibility and risk mapping using social/infrastructural vulnerability and environmental variables." *Fire* 2.3 (2019): 50.

- Gigović, L., Pourghasemi, H. R., Drobnjak, S., & Bai, S. "Testing a new ensemble model based on SVM and random forest in forest fire susceptibility assessment and its mapping in Serbia's Tara National Park." *Forests* 10.5 (2019): 408.
- Goldarag, Y. J., Mohammadzadeh, A., & Ardakani, A. S. "Fire risk assessment using neural network and logistic regression." *J. Indian Soc. Remote Sens.* 44.6 (2016): 885-894.
- Guijarro, M., Hernando, C., Díez, C., Martínez, E., Madrigal, J., Lampin-Cabaret, C., Blanc, L., Colin, P.Y., Pérez-Gorostiaga, P., Vega, J.A., & Fonturbel, M.T. "Flammability of some fuel beds common in the South-European ecosystems." *4th Int. Conf. Forest Fire Res.* 2002.
- Guo, F., Wang, G., Su, Z., Liang, H., Wang, W., Lin, F., & Liu, A. "What drives forest fire in Fujian, China? Evidence from logistic regression and Random Forests." *Int. J. Wildland Fire* 25.5 (2016): 505-519.
- Guo, F., Zhang, L., Jin, S., Tigabu, M., Su, Z., & Wang, W. "Modeling anthropogenic fire occurrence in the boreal forest of China using logistic regression and random forests." *Forests* 7.11 (2016): 250.
- Hadden, R. M., Scott, S., Lautenberger, C., & Fernandez-Pello, A. C. "Ignition of combustible fuel beds by hot particles: an experimental and theoretical study." *Fire Technol.* 47.2 (2011): 341-355.
- Hedayati, F., Bahrani, B., Zhou, A., Quarles, S. L., & Gorham, D. J. "A framework to facilitate firebrand characterization." *Front. Mech. Eng.* 5 (2019): 43.
- Hedayati F., Bahrani B., Zhou A., Quarles S. L., & Weise D. R. Data firebrands generated from selected structural fuels: *Joint Fire Science Program Project* (15-1-04-4). Technical report. (n.d.) Type: dataset
- Hodges, J. L. *Predicting Large Domain Multi-Physics Fire Behavior Using Artificial Neural Networks*. PhD Thesis. Virginia Tech, 2018.
- Hodges, J. L., Lattimer, B. Y., & Luxbacher, K. D. "Compartment fire predictions using transpose convolutional neural networks." *Fire Safety J.* 108 (2019): 102854.
- Hodges, J. L., & Lattimer, B. Y. "Wildland fire spread modeling using convolutional neural networks." *Fire Technol.* 55.6 (2019): 2115-2142.
- Hong, H., Tsangaratos, P., Ilia, I., Liu, J., Zhu, A. X., & Xu, C. "Applying genetic algorithms to set the optimal combination of forest fire related variables and model forest fire susceptibility based on data mining models. The case of Dayu County, China." *Sci. Total Environ.* 630 (2018): 1044-1056.
- Jaafari, A., Zenner, E. K., & Pham, B. T. "Wildfire spatial pattern analysis in the Zagros Mountains, Iran: A comparative study of decision tree based classifiers." *Ecol. Inform.* 43 (2018): 200-211.
- Jaafari, A., & Pourghasemi, H. R. "Factors influencing regional-scale wildfire probability in Iran: an application of random forest and support vector machine." *Spatial modeling in GIS and R for Earth and environmental sciences*. Elsevier (2019): 607-619.
- Jaafari, A., Zenner, E. K., Panahi, M., & Shahabi, H. "Hybrid artificial intelligence models based on a neuro-fuzzy system and metaheuristic optimization algorithms for spatial prediction of wildfire probability." *Agr. Forest Meteorol.* 266 (2019): 198-207.
- Jain, P., Coogan, S. C., Subramanian, S. G., Crowley, M., Taylor, S., & Flannigan, M. D. "A review of machine learning applications in wildfire science and management." *Environ. Rev.* 28.4 (2020): 478-505.
- Jha, A., & Zhou, A. "Applying Machine Learning for Firebrand Production Prediction." *Fire Technol.* (2022): 1-30.

- Kasymov, D. P., Filkov, A. I., Baydarov, D. A., & Sharypov, O. V. "Interaction of smoldering branches and pine bark firebrands with fuel bed at different ambient conditions." *22nd International Symposium on Atmospheric and Ocean Optics: Atmospheric Physics*. Vol. 10035. International Society for Optics and Photonics, 2016.
- Krizhevsky, A., Sutskever, I., & Hinton, G. E. "Imagenet classification with deep convolutional neural networks." *Commun. ACM*. 60.6 (2017): 84-90.
- Kwan, C., Ayhan, B., Budavari, B., Lu, Y., Perez, D., Li, J., Bernabe, S., & Plaza, A. "Deep learning for land cover classification using only a few bands." *Remote Sens-Basel* 12.12 (2020): 2000.
- Lattimer, B. Y., Hodges, J. L., & Lattimer, A. M. "Using machine learning in physics-based simulation of fire." *Fire Safety J.* 114 (2020): 102991.
- LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. *Nature*, 521(7553), 436-444.
- Liakos, K. G., Busato, P., Moshou, D., Pearson, S., & Bochtis, D. (2018). Machine learning in agriculture: A review. *Sensors*, 18(8), 2674.
- López-Serrano, P. M., López-Sánchez, C. A., Álvarez-González, J. G., & García-Gutiérrez, J. "A comparison of machine learning techniques applied to Landsat-5 TM spectral data for biomass estimation." *Can. J. Remote Sens.* 42.6 (2016): 690-705.
- Manzello, S. L., Cleary, T. G., Shields, J. R., & Yang, J. C. "On the ignition of fuel beds by firebrands." *Fire and Mater.* 30.1 (2006): 77-87.
- Manzello, S. L., Cleary, T. G., Shields, J. R., Maranghides, A., Mell, W., & Yang, J. C. "Experimental investigation of firebrands: generation and ignition of fuel beds." *Fire Safety J.* 43.3 (2008): 226-233.
- Manzello, S. L., Park, S. H., & Cleary, T. G. "Investigation on the ability of glowing firebrands deposited within crevices to ignite common building materials." *Fire Safety J.* 44.6 (2009): 894-900.
- Manzello, S. L., Suzuki, S., & Nii, D. "Full-scale experimental investigation to quantify building component ignition vulnerability from mulch beds attacked by firebrand showers." *Fire Technol.* 53.2 (2017): 535-551.
- Martín, Y., Zúñiga-Antón, M., & Rodrigues Mimbreno, M. "Modelling temporal variation of fire-occurrence towards the dynamic prediction of human wildfire ignition danger in northeast Spain." *Geomat. Nat. Haz. Risk* 10.1 (2019): 385-411.
- Maltezos, E., Doulamis, A., Doulamis, N., & Ioannidis, C. "Building extraction from LiDAR data applying deep convolutional neural networks." *IEEE Geosci. Remote S.* 16.1 (2018): 155-159.
- Mäyrä, J., Keski-Saari, S., Kivinen, S., Tanhuanpää, T., Hurskainen, P., Kullberg, P., Poikolainen, L., Viinikka, A., Tuominen, S., Kumpula, T., & Vihervaara, P.. "Tree species classification from airborne hyperspectral and LiDAR data using 3D convolutional neural networks." *Remote Sens. Environ.* 256 (2021): 112322.
- Mell, W. E., Manzello, S. L., Maranghides, A., Butry, D., & Rehm, R. G. "The wildland–urban interface fire problem—current approaches and research needs." *Int. J. Wildland Fire* 19.2 (2010): 238-251.
- Meuwly, M. "Machine learning for chemical reactions." *Chem. Rev.* 121.16 (2021): 10218-10239.
- Morchhale, S., Pauca, V. P., Plemmons, R. J., & Torgersen, T. C. "Classification of pixel-level fused hyperspectral and lidar data using deep convolutional neural networks." *2016 8th workshop on hyperspectral image and signal processing: evolution in remote sensing (WHISPERS)*. IEEE, 2016.

Murphy, K. P. *Machine learning: a probabilistic perspective*. MIT press, 2012.

Oliveira, S., Oehler, F., San-Miguel-Ayanz, J., Camia, A., & Pereira, J. M. "Modeling spatial patterns of fire occurrence in Mediterranean Europe using Multiple Regression and Random Forest." *Forest Ecol. Manag.* 275 (2012): 117-129.

Pan, S., Guan, H., Chen, Y., Yu, Y., Gonçalves, W. N., Junior, J. M., & Li, J. "Land-cover classification of multispectral LiDAR data using CNN with optimized hyper-parameters." *ISPRS J. Photogramm.* 166 (2020): 241-254.

Parisien, M. A., Parks, S. A., Krawchuk, M. A., Little, J. M., Flannigan, M. D., Gowman, L. M., & Moritz, M. A. "An analysis of controls on fire activity in boreal Canada: comparing models built with different temporal resolutions." *Ecol. Appl.* 24.6 (2014): 1341-1356.

Park, J. H., Lee, S., Yun, S., Kim, H., & Kim, W. T. "Dependable fire detection system with multifunctional artificial intelligence framework." *Sensors* 19.9 (2019): 2025.

Peters, M. P., Iverson, L. R., Matthews, S. N., & Prasad, A. M. "Wildfire hazard mapping: exploring site conditions in eastern US wildland–urban interfaces." *Int. J. Wildland Fire* 22.5 (2013): 567-578.

Peters, M. P., & Iverson, L. R. "Incorporating fine-scale drought information into an eastern US wildfire hazard model." *Int. J. Wildland Fire* 26.5 (2017): 393-398.

Pierce, A. D., Farris, C. A., & Taylor, A. H. "Use of random forests for modeling and mapping forest canopy fuels for fire behavior analysis in Lassen Volcanic National Park, California, USA." *Forest Ecol. Manag.* 279 (2012): 77-89.

Pourtaghi, Z. S., Pourghasemi, H. R., Aretano, R., & Semeraro, T. "Investigation of general indicators influencing on forest fire and its susceptibility modeling using different data mining techniques." *Ecol. Indic.* 64 (2016): 72-84.

Ray, S. "A quick review of machine learning algorithms." *2019 International conference on machine learning, big data, cloud and parallel computing (COMITCon)*. IEEE, 2019.

Riaño, D., Ustin, S. L., Usero, L., & Patricio, M. A. "Estimation of fuel moisture content using neural networks." *International Work-Conference on the Interplay Between Natural and Artificial Computation*. Springer, Berlin, Heidelberg, 2005.

Riley, K. L., Grenfell, I. C., Finney, M. A., & Crookston, N. L. "Utilizing random forests imputation of forest plot data for landscape-level wildfire analyses." In: *Viegas, Domingos Xavier, ed. Advances in Forest Fire Research. Coimbra, Portugal: Imprensa da Universidade de Coimbra. p. 583-590.* (2014): 583-590.

Rodrigues, M., & De la Riva, J. "An insight into machine-learning algorithms to model human-caused wildfire occurrence." *Environ. Modell. Softw.* 57 (2014): 192-201.

Sachdeva, S., Bhatia, T., & Verma, A. K. "GIS-based evolutionary optimized Gradient Boosted Decision Trees for forest fire susceptibility mapping." *Nat. Hazards* 92.3 (2018): 1399-1418.

Scott, S., Hadden, R., Yun, A., Lautenberger, C., & Fernandez-Pello, A. C. "Ignition of cellulose fuel beds by hot metal particles." *Proceedings of the Fourth International Conference on Safety and Security Engineering, C. Brebbia (Ed.), WIT Press, London.* 2011.

Smyth, P. "Clustering sequences with hidden Markov models." *Adv. Neur. In.* (1996).

- Sun, P., Zhang, Y., Sun, L., Hu, H., Guo, F., Wang, G., & Zhang, H. "Influence of fuel moisture content, packing ratio and wind velocity on the ignition probability of fuel beds composed of Mongolian oak leaves via cigarette butts." *Forests* 9.9 (2018): 507.
- Suzuki, S., Manzello, S. L., Kagiya, K., Suzuki, J., & Hayashi, Y. "Ignition of mulch beds exposed to continuous wind-driven firebrand showers." *Fire Technol.* 51.4 (2015): 905-922.
- Suzuki, S., Johnsson, E., Maranghides, A., & Manzello, S. L. "Ignition of wood fencing assemblies exposed to continuous wind-driven firebrand showers." *Fire Technol.* 52.4 (2016): 1051-1067.
- Suzuki, S., & Manzello, S. L. "Role of accumulation for ignition of fuel beds by firebrands." *Appl. Energy Combust.* 1 (2020): 100002.
- Suzuki, S., & Manzello, S. L. "Investigating coupled effect of radiative heat flux and firebrand showers on ignition of fuel beds." *Fire Technol.* 57.2 (2021): 683-697.
- Tehrany, M. S., Jones, S., Shabani, F., Martínez-Álvarez, F., & Tien Bui, D. "A novel ensemble modeling approach for the spatial prediction of tropical forest fire susceptibility using LogitBoost machine learning classifier and multi-source geospatial data." *Theor. Appl. Climatol.* 137.1 (2019): 637-653.
- Thach, N. N., Ngo, D. B. T., Xuan-Canh, P., Hong-Thi, N., Thi, B. H., Nhat-Duc, H., & Dieu, T. B. "Spatial pattern assessment of tropical forest fire danger at Thuan Chau area (Vietnam) using GIS-based advanced machine learning algorithms: A comparative study." *Ecol. Inform.* 46 (2018): 74-85.
- Tymstra, C., Stocks, B. J., Cai, X., & Flannigan, M. D. "Wildfire management in Canada: Review, challenges and opportunities." *Prog. Disaster Sci.* 5 (2020): 100045.
- Urban, J. L., Zak, C. D., & Fernandez-Pello, C. "Cellulose spot fire ignition by hot metal particles." *P. Combust. Inst.* 35.3 (2015): 2707-2714.
- Urban, J. L., Zak, C. D., & Fernandez-Pello, C. "Spot fire ignition of natural fuels by hot aluminum particles." *Fire Technol.* 54.3 (2018): 797-808.
- Urban, J. L., Song, J., Santamaria, S., & Fernandez-Pello, C. "Ignition of a spot smolder in a moist fuel bed by a firebrand." *Fire Safety J.* 108 (2019): 102833.
- Vacchiano, G., Foderi, C., Berretti, R., Marchi, E., & Motta, R. "Modeling anthropogenic and natural fire ignitions in an inner-alpine valley." *Nat Hazard Earth Sys.* 18.3 (2018): 935-948.
- Viegas, D. X., Almeida, M., Raposo, J., Oliveira, R., & Viegas, C. X. "Ignition of mediterranean fuel beds by several types of firebrands." *Fire Technol.* 50.1 (2014): 61-77.
- Viegas, D. X., Oliveira, R., Almeida, M., & Kim, D. "Ignition of Fuel Beds by Cigarettes: A Conceptual Model to Assess Fuel Bed Moisture Content and Wind Velocity Effect on the Ignition Time and Probability." *Fire* 4.3 (2021): 35.
- Vilar, L., Gómez, I., Martínez-Vega, J., Echavarría, P., Riaño, D., & Martín, M. P. "Multitemporal modelling of socio-economic wildfire drivers in central Spain between the 1980s and the 2000s: comparing generalized linear models to machine learning algorithms." *PLoS One* 11.8 (2016): e0161344.
- Wang, A., He, X., Ghamisi, P., & Chen, Y. "LiDAR data classification using morphological profiles and convolutional neural networks." *IEEE Geosci. Remote S.* 15.5 (2018): 774-778.
- Wei, J., Chu, X., Sun, X. Y., Xu, K., Deng, H. X., Chen, J., Wei, Z., & Lei, M. (2019). Machine learning in materials science. *InfoMat.* 1(3), 338-358.

- Wessies, S. S., Chang, M. K., Marr, K. C., & Ezekoye, O. A. "Experimental and analytical characterization of firebrand ignition of home insulation materials." *Fire Technol.* 55.3 (2019): 1027-1056.
- Wu, X., Park, Y., Li, A., Huang, X., Xiao, F., & Usmani, A. "Smart detection of fire source in tunnel based on the numerical database and artificial intelligence." *Fire Technol.* 57.2 (2021): 657-682.
- Xu, Z., Guan, K., Casler, N., Peng, B., & Wang, S. "A 3D convolutional neural network method for land cover classification using LiDAR and multi-temporal Landsat imagery." *ISPRS J. Photogramm.* 144 (2018): 423-434.
- Yang, G., Ning, J., Shu, L., Zhang, J., Yu, H., & Di, X. "Spotting ignition of larch (*Larix gmelinii*) fuel bed by different firebrands." *J. Forestry Res.* (2022): 1-11.
- Yu, Y., Guan, H., Li, D., Gu, T., Wang, L., Ma, L., & Li, J. "A hybrid capsule network for land cover classification using multispectral LiDAR data." *IEEE Geosci. Remote S.* 17.7 (2019): 1263-1267.
- Zak, C. D., Urban, J. L., & Fernandez-Pello, C. "Characterizing the flaming ignition of cellulose fuel beds by hot steel spheres." *Combust. Sci. Technol.* 186.10-11 (2014): 1618-1631.
- Zhang, G., Wang M., & Liu K. "Forest fire susceptibility modeling using a convolutional neural network for Yunnan province of China." *Int. J. Disast. Risk Sc.* 10.3 (2019): 386-403.
- Zhang, G., Wang M., & Liu K. "Deep neural networks for global wildfire susceptibility modelling." *Ecol. Indic.* 127 (2021): 107735.
- Zhang, G., Wang M., & Liu K. "Dynamic prediction of global monthly burned area with hybrid deep neural networks." *Ecol. Appl.* (2022): e2610.
- Zhou A., Quarles S., & Weise D. (2019) Fire ember production from wildland and structural fuels. Technical report 08.
- Zorzi, S., Maset, E., Fusiello, A., & Crosilla, F. "Full-waveform airborne LiDAR data classification using convolutional neural networks." *IEEE T Geosci. Remote* 57.10 (2019): 8255-8261.
- "Ember Exposure Characterization in WUI Fires". *National Institute of Standards and Technology (NIST)*, May 2021, <https://www.nist.gov/programs-projects/ember-exposure-characterization-wui-fires>.