



Predicting wildfire burns from big geodata using deep learning

John Ray Bergado ^{a,*}, Claudio Persello ^a, Karin Reinke ^b, Alfred Stein ^a

^a Faculty of Geo-Information Science and Earth Observation (ITC), University of Twente, the Netherlands

^b Department of Geospatial Science, Royal Melbourne Institute of Technology (RMIT), Australia



ARTICLE INFO

Keywords:

Wildfire burn prediction
Deep learning
Convolutional networks
Big geodata
Input statistical importance

ABSTRACT

Wildfire continues to be a major environmental problem in the world. To help land and fire management agencies manage and mitigate wildfire-related risks, we need to develop tools for mapping those risks. Big geodata—in the form of remotely sensed images, ground-based sensor observations, and topographical datasets—can help us characterize the dynamics of wildfire related events. In this study, we design a deep fully convolutional network, called AllConvNet, to produce daily maps of the probability of a wildfire burn over the next 7 days. We applied it to burns in Victoria, Australia for the period of 2006–2017. Fifteen factors that were extracted from six different datasets and resulted into 29 quantitative features, were selected as input to the network. We compared it with three baseline methods: SegNet, multilayer perceptron, and logistic regression. AllConvNet outperforms the other three baseline methods in four of the six quantitative metrics considered. AllConvNet and SegNet provide smoother and more regularized predicted maps, with SegNet providing greater sensitivity in discriminating less wildfire-prone locations. Input feature statistical importance was measured for all the networks and compared against logistic regression coefficients. Total precipitation, lightning flash density, and land surface temperature occur to be consistently highly weighted by all models while terrain aspect components, wind direction components, certain land cover classes (such as crop field and woodland), and distance from power lines are ranked on the lower end. We conclude that wild-fire burn prediction methods based on deep learning present quantitative and qualitative gains.

1. Introduction

Wildfire continues to be a major environmental problem in the world (Yildirim et al., 2005; Cancelo-González et al., 2014). When poorly managed, it can cause permanent and undesirable changes to certain landscape and ecosystem services (Chen et al., 2015). Although most incidents are found in remote areas and naturally occur as part of terrestrial ecosystem processes, human populations and infrastructure within wildland-urban interfaces are still exposed to wildfire related risks (Minas et al., 2014). In the context of wildfire management, we consider wildfire risk as the probability of a wildfire event that may consequently result in expected loss of lives and degradation of assets. We can characterize such risk by mapping associated risk conditions. In general, these conditions are both spatially and temporally dynamic in nature and depend on a number of factors including: fuel conditions, meteorological variables, topography, and sources of ignitions (Chowdhury and Hassan, 2015). Big geodata—in the form of multi-temporal remotely sensed images, ground-based sensor observations,

and topographical datasets—is a promising source of information on these wildfire risk factors.

Quantitative assessment of wildfire risk employs a four-stage framework (Fairbrother and Turnley, 2005; Thompson and Calkin, 2011). Firstly, problem formulation defines specific management objectives to be addressed. Secondly, exposure analysis quantifies the likely magnitude and spatiotemporal connection between the identified risk variables (Fairbrother and Turnley, 2005). Thirdly, effect analysis quantifies the response of the valuable resources at risk as a function of the fire behavior—usually flame length (Thompson et al., 2015) in combination with rate of spread and fire intensity. Finally, risk characterization integrates information from the three previous stages to come up with a complete, informative, and useful conclusion for decision-making (Sikder et al., 2006).

Exposure analysis describes how likely will valuable resources of interest interact with a wildfire (Thompson et al., 2015). Thus, it directly deals with quantifying probabilities of wildfire events (Ager et al., 2014). Probability of wildfire burn refers to how likely a geographical

* Corresponding author.

E-mail addresses: johnraybergado@gmail.com (J.R. Bergado), c.persello@utwente.nl (C. Persello), karin.reinke@rmit.edu.au (K. Reinke), a.stein@utwente.nl (A. Stein).

location will change from an unburnt state to a burnt state within a given time period, e.g. annually (Miller et al., 2008). Estimates of this probability can either serve as a proxy measure for wildfire risk or as input to probabilistic methods to quantify wildfire risk (Finney, 2005). Maps showing probability of wildfire burn can be further intersected with the asset locations (Massada et al., 2009; Atkinson et al., 2010)—to identify likely affected assets.

Formal treatment to quantify the probabilities associated to wildfire events can become complicated. Most case studies on wildfire occurrence do not adopt a consistent quantitative definition. They employ a logistic regression (LR) (de Vasconcelos et al., 2001; Catry et al., 2009; Chuvieco et al., 2010; Badia et al., 2011; Jurdao and Chuvieco, 2012; Guo et al., 2013; de Bem et al., 2018; Parisien et al., 2012) trained on a number of relevant wildfire risk factors, acquired through remotely sensed data and data products, and information on historical wildfire locations. Within this group of studies, differences exist in interpretation. For example, Chuvieco et al. (2010) predicts the number of fire incidence reclassified into low and high incidence while Badia et al. (2011) reclassifies the output of the logistic function into five ranges. Other studies have applied methods such as multilayer perceptron (MLP) (de Vasconcelos et al., 2001), weights-of-evidence Bayesian model (Jaafari et al., 2017), evidential belief function (Nami et al., 2018), and conversion from reference fuel moisture tables (Sánchez et al., 2018).

Deep learning presents a promising approach to build inference models by learning hierarchical and distributed representations of data. Hierarchical, in a way that higher-level representations are built on top of lower-level ones; and distributed, in a way that inputs are described by multiple features and each feature participates in the representation of multiple inputs (Goodfellow et al., 2016, pp. 13–19). It stands on the premise that some mapping functions may be more efficiently approximated by deeper architectures as compared to their shallower counterparts (Bengio, 2009). A wildfire event prediction problem can be seen as a function learning task, where the function maps wildfire-related input variables into probabilities of a wildfire event. Higher-level spatial and temporal association between the input variables may improve the predictive accuracy of the learned function. However, just like the mapping function, there is a knowledge gap on how to construct these higher-level features. Models based on deep learning can, therefore, be used to capture such higher-level spatial and temporal association together with learning the mapping function.

The paper presents an application of deep learning using a family of models called artificial neural networks. A variety of architectures exists such as convolutional neural networks (CNN), recurrent neural networks (RNN), autoencoders, Boltzmann machine variants—each of which, generally tailored to certain applications. For example, CNN for images and RNN for sequential data. In remote sensing context, deep artificial neural networks, such as CNN and fully convolutional network (FCN), were employed in land cover classification of very high resolution remotely sensed images (Volpi and Tuia, 2017; Bergado et al., 2018), digital terrain model extraction (Gevaert et al., 2018), and delineation of agricultural fields in smallholder farms (Persello et al., 2019).

There is limited work on employing deep neural networks and remotely sensed data for wildfire mapping. A recent relevant work presents a two-stage strategy on estimating weekly wildfire hotspots in Australia (Dutta et al., 2016) using deep belief networks for feature compression before feeding the latter to an ensemble classifier. Most relevant works produce static probability maps estimated over a certain period of time (de Vasconcelos et al., 2001; Catry et al., 2009; Badia et al., 2011; de Bem et al., 2018). Although, wildfire-driving factors are dynamic in nature, hence, producing time-series maps, either sub-hourly (Wickramasinghe et al., 2016), daily (Chuvieco et al., 2010; Guo et al., 2013) or weekly (Dutta et al., 2016; Gray et al., 2018) presents a better understanding of wildfire risk. Remotely sensed images and stationary ground-based sensors are promising sources of time-series information

that can be used to produce these maps. We combine these time-series information with historical wildfire burn records in a predictive model, e.g. a FCN, to produce dynamic probability maps of wildfire burn.

This research aims to utilize supervised deep learning techniques to estimate probabilities of wildfire burn by combining information from remotely sensed images, stationary ground-based sensors, topographical datasets, and historical wildfire data. Specifically, fully convolutional networks (FCN) are employed to produce daily predictions mapping probability of wildfire burns in Victoria, Australia for the period 2006–2017.

The major contributions of this study are:

- To design and implement a fully convolutional network for predicting daily maps of the probability of a wildfire burn over the next 7 days utilizing an extensive set of wildfire related input variables taken from various data sources such as: time series of satellite images and data products, climatological sensor observations, topographical geospatial databases, and historical wildfire burn records for Victoria, Australia over the period of 2006–2017;
- To quantitatively and qualitatively evaluate the proposed network against several baseline methods;
- To quantify the relative statistical importance of each input features used in the deep fully convolutional network.

2. Data

2.1. Study area

Victoria (37°S 144°E) is the southeastern state of Australia. It has an area of approximately 238000 km², of which roughly 16% is forest, 10% is woodland, 5% is shrubland, and 6% is grassland. Wildfire is a natural component, making the state one of the most wildfire-prone areas in the world (The State of Victoria, 2012).

The spatial extent of wildfire burns within Victoria was extracted from the “Fire History Records of Fires primarily on Public Land” dataset made publicly available by the State of Victoria, Department of Environment, Land, Water & Planning (DELWP) (The State of Victoria, 1992). This allows prescribed burns to be distinguished from wildfires, whereas the Moderate Resolution Imaging Spectroradiometer (MODIS)¹ active fires dataset does not allow for this distinction and may miss fire activity with smaller burn extents (Hantson et al., 2013) or lower intensities. Temporally, the Fire History Records dataset only provides the starting date of the fires. Fig. 1 shows a map of the study area. Planned burns conducted by DELWP were filtered out and the remaining wildfire burn extents for the period of 2006–2017 were used in this study.

2.2. Input variables

The probability of a wildfire burning a specific geographical location depends on both natural and anthropogenic factors. The spatial and temporal scale of the study required the organization of a big geodataset, not only in terms of volume and update frequency but also data heterogeneity, matching wildfire risk factors with recorded wildfire extent locations. Fifteen factors that were extracted from six different datasets and resulted into 29 quantitative features, were selected as input to our models predicting probability of wildfire burn. These features encode the factors associated to wildfire burn such as topography (elevation, slope, and aspect), weather (temperature, humidity, solar radiation, rainfall, wind speed and direction, and lightning flash density), proximity to anthropogenic interfaces (distance to roads, distance to power lines) and fuel characteristics (fuel type, fuel moisture, emissivity). Table 1 shows these input variables including the dataset from which

¹ <https://modis.gsfc.nasa.gov/>.

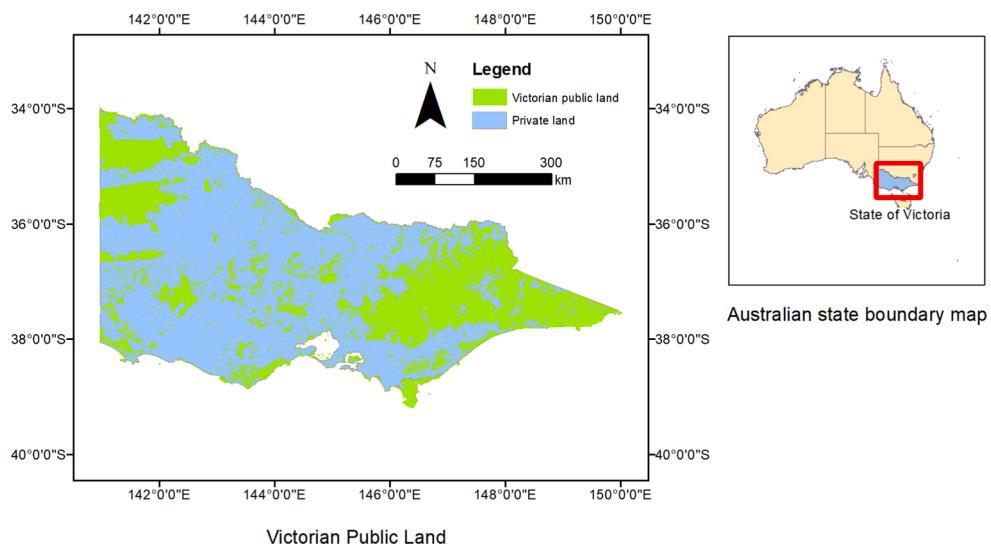


Fig. 1. Location of the state of Victoria and its public land.

Table 1
Selected input variables associated to wildfire burns for southeastern Australia and for the study period 2006–2017.

Source Dataset	Input Variable	Units	Spatial Resolution	Temporal Resolution
Digital elevation model ^a	elevation	m	250 m	stationary
	aspect	—		
	slope	—		
Topographical databases ^b	proximity measures	m	—	stationary
Land cover/use ^a	surrogate for fuel type	—	250 m	bi-annual
MODIS images ^c	fuel moisture proxy	—	500 m	8 days
	land surface temperature emissivity	—		
LIS data product ^d	lightning flash density	flashes·km ⁻²	12 km	DOY
Meteorological data ^e	temperature	°C	—	daily
	relative humidity	%		
	solar radiation	MJm ⁻²		
	precipitation	mm	—	
	wind speed	kmh ⁻¹		1 min
	wind direction	degrees		

DOY means observations are averaged based on the day of the year.

^a Obtained from Geoscience Australia web portal ([Hutchinson et al., 2008; Lymburner et al., 2011](#))

^b Obtained from Victoria's open data directory ([The State of Victoria, 2009a; The State of Victoria, 2009b](#));

^c Obtained from NASA EOSDIS Land Processes DAAC data portal ([Vermote, 2015; Wan et al., 2015](#));

^d Obtained from NASA GHRC DAAC data portal ([Cecil et al., 2014](#));

^e Sourced from Australian Bureau of Meteorology;

they are taken from and their corresponding spatial and temporal resolution before any spatial resampling. The influence of topography, weather and fuel related variables are well accepted drivers for wildfire ([Estes et al., 2017; Zhang et al., 2016; Bradstock et al., 2009](#)). Variables describing the proximity to anthropogenic infrastructure such as roads and powerlines have been added as input variables due to their ability to potentially influence the spread and intensity of wildfire ([Narayananaraj and Wimberly, 2012; Narayananaraj and Wimberly, 2011](#)).

The digital elevation model ([Hutchinson et al., 2008](#)) was

downsampled to 500 m resolution using bilinear interpolation to produce the elevation, slope, and the cosine and sine components of the aspect. Road and power line network vectors ([The State of Victoria, 2009a; The State of Victoria, 2009b](#)) were rasterized to a spatial resolution of 500 m. Euclidean distances from both the roads and the power lines resulted into two additional features extracted from this topographical vector database. These six features are temporally stationary.

Land cover/land use information were obtained from the National Dynamic Land Cover Dataset of Australia ([Lymburner et al., 2011](#)). The original dataset with a spatial resolution of 250 m and containing 22 land cover classes was downsampled using nearest neighbor into a spatial resolution of 500 m. The original 22 classes were collapsed into eight ([Table 2](#)) based on similar land use, land cover and vegetation characteristics. Compared to other land use and land cover classes, urban areas and mines and quarries will likely have the least forest fuel load. Additionally, these two classes likely have regular human activities related to fire prevention, intervention, and suppression compared to other classes. The presence of and proximity to areas with such activities and forest fuel load is intuitively tied to wildfire likelihood. The new land cover maps were transformed into binary presence maps, encoding the value one when a class is present and the value zero otherwise. This resulted in eight features extracted from the land cover dataset. The original dataset is obtained from classifying MODIS time series images spanning over two-year periods. The eight resulting features were temporally matched to other features using the year of the latest image from the two-year averaging period.

MODIS is able to observe every location on the earth's surface every 1–2 days over an array of spectral bands covering both optical, infrared,

Table 2
Aggregated land cover/use classes of the Australian Dynamic Land Cover Dataset.

New Class ID	Original Classes
1	lakes and dams, salt lakes
2	mines and quarries, urban areas
3	irrigated cropping, rain fed cropping, irrigated pasture, rain fed pasture, irrigated sugar, rain fed sugar, wetlands
4	alpine meadows
5	open hummock grassland, closed tussock grassland, open tussock grassland
6	scattered shrubs and grasses, dense shrubland, open shrubland
7	closed forest, open forest
8	woodland, open woodland

and thermal bands. With this large spatial extent and relatively short revisit time, MODIS is suitable for monitoring large-scale dynamic phenomena on the earth's surface. Hence, as an input to our models predicting the probability of wildfire burn, a proxy for fuel moisture content was obtained from spectral indices derived from a time series of a MODIS surface reflectance product (Vermote, 2015). This image product has a temporal resolution of eight days and a spatial resolution of 500 m. Four spectral indices correlated with fuel moisture content (Caccamo et al., 2012; Yebra et al., 2013) were used: (i) the normalized difference vegetation index (NDVI), (ii) visible atmospherically resistant index (VARI), (iii) normalized difference infrared index (NDII), and (iv) the normalized difference water index (NDWI). One thermal band and two emissivity bands were directly used as input features (Wan et al., 2015). This resulted into seven features extracted from MODIS images. The features are temporally matched using the latest date of the eight-day averaging period, as reported in the image product, prior to the starting date of the wildfire.

A major natural cause of wildfire ignitions is lightning. Lightning flash density information was obtained from annual lightning climatology derived from NASA's lightning imaging sensor (LIS) observations (Cecil et al., 2014). The annual lightning climatology provided estimates of lightning flash density for each day of the year averaged over more than a 16-year period with a spatial resolution of about 12 km. The dataset is spatially incomplete beyond the south of the 38°S parallel (Fig. 2). Therefore, direct sampling (Mariethoz and Renard, 2010) was used to impute locations within the study area extent with missing lightning flash density rate. The range of the empirical variogram was used as a basis for determining the search window size for the direct sampling algorithm. The imputed data were manually upsampled to 500 m spatial resolution using bilinear interpolation. This resulted into one feature that is temporally matched using the day of year (DOY) associated to the average lightning flash density rate image.

Finally, five weather variables associated with wildfire (Vasilakos et al., 2009; Sakr et al., 2011; Badia et al., 2011; Guo et al., 2013; Mundo et al., 2013; Mhawej et al., 2016) were selected, namely: temperature, humidity, solar radiation, wind speed and direction, and precipitation. The first three have a daily resolution, whereas the last two have a minutely resolution. The maximum temperature past 9 am in 24 h was used. For humidity, the lowest relative humidity from every three hours within a day was chosen. Maximum temperature and minimum relative humidity were correspondingly chosen since extreme scenario on this ends of the two variables increases the likelihood of a wildfire. The solar radiation records the total amount of daily global solar exposure. Wind speed and direction as well as the precipitation were aggregated into daily values. For both the wind speed and wind direction, the mean value was taken; and for the precipitation, the sum was taken as an input. A total wind speed and wind direction value would not have any physically interpretable meaning, meanwhile fuel moisture can be attributed to total precipitation. Observations having bad quality

indicators were removed and four days (of interest) without any qualified solar radiation observations were temporally imputed using the average of the same DOY from years with measurements having good quality indicator. The five variables were interpolated using ordinary kriging into a spatial resolution of 500 m. This resulted into five features matched using the date associated to the meteorological observation.

The raster grids were matched to have the same geographic coordinate system (WGS 84) and spatial resolution of 500 m. The organized dataset resulted into 29-band input images and one-band target images of size 1536×2304 pixels for all dates between 2006 and 2017. Fig. 3 visualizes the input images for one selected date. Fig. 4 shows the location of the wildfire burns for the whole. Directly feeding these input and target images to a predictive model would both severely limit the size of the convolutional neural network that can be trained and would make the learning process computationally slower (because of the spatial dimensions of the images) and more difficult (considerable imbalance in the number of labels). To address such computational constraints, the images were sampled into smaller input patches.

3. Methods

3.1. Deep neural networks

Deep networks can be generalized by seeing them as data-flow graphs—a graph representing how a set of input data are processed along a possibly branching chain of functions, in the end producing a set of outputs. Using such a model, the networks can be defined by three elements: the sets of data they take as an input, the operations they perform in each function block, and the intermediate and final set of outputs they produce. Aside from these three key elements of data-flow graphs, details of a unique configuration and instance of a network are defined by its *hyperparameters* and *parameters* respectively. Hyperparameters are associated with the configuration of a network architecture and are set to fixed values during training; while parameters are values associated to a specific network instance and are learned during network training. In the context of wildfire prediction, the inputs of the artificial neural networks are equivalent to the independent variables associated to a wildfire event. Identifying the outputs of the network can be challenging. Specially since there has not been any consistent formal definition of probability of burn in the literature. A generalization of common neural network operations can be found in Bergado et al. (2018).

Feedforward artificial neural networks are often employed as deterministic models. But, the outputs of these networks can be treated with probabilistic meaning. For example, in a regression problem, it can be assumed that the target variable t follows a Gaussian distribution dependant on the input x and parameters w of the network f such that

$$p(t|x, w) = \mathcal{N}(t|f(x, w), \sigma^2) \quad (1)$$

where σ is the variance of the Gaussian (Bishop, 2006, pp. 232–236). The choice of the probability distribution is not constrained by the network itself but is problem-dependent, reflecting our expected distribution of the target variable t .

3.2. Sampling of burn and non-burn locations

Non-overlapping patches of size $M \times M$, with the pixel near the center of the patch lying on a location of a wildfire burn, were systematically sampled from both the wildfire burn location (target patches) and the 29 quantitative features (input patches) accordingly matched with the recorded starting date of a wildfire. Corresponding burn locations are temporally binned together in 7-day temporal windows, hence the resulting probability is an estimate of the likelihood of a wildfire burning a specific location within the next 7 days. The temporal bins were chosen to start from the starting date of each wildfire burn, as

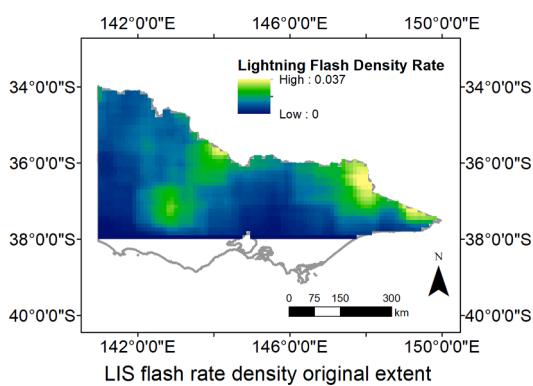


Fig. 2. LIS flash rate density original extent.

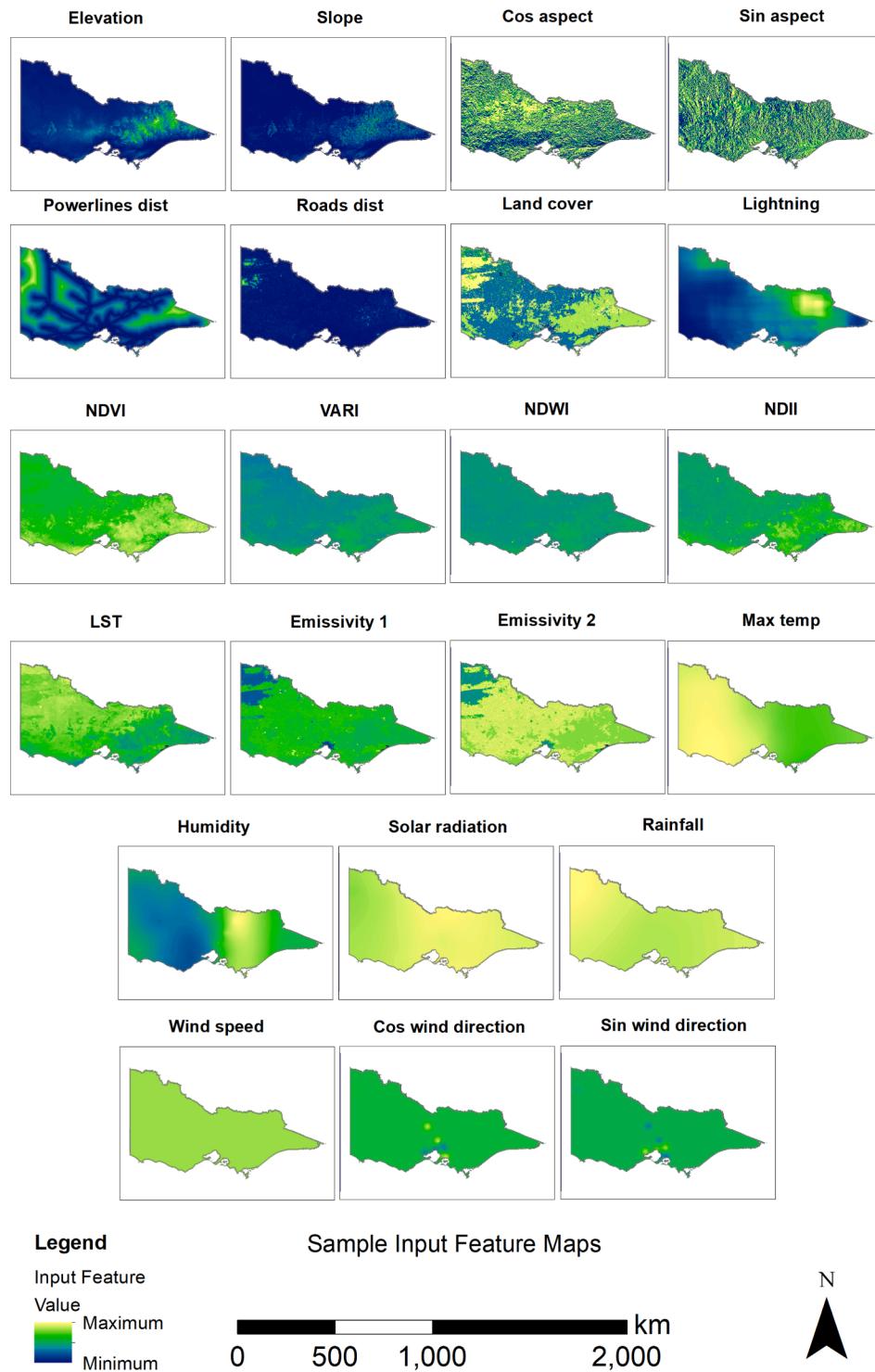


Fig. 3. Input images for one selected date (December 1, 2006). All images are continuous variables normalized to [0, 1], except for the fuel type map that shows discrete class categories from Table 2.

recorded in the Victorian Public Land Fire History Records dataset, and end 7 days after this starting date. This temporal binning of wildfire burn locations to a 7-day period transforms the prediction problem from predicting the probability of a wildfire burn occurring within the day, given we observe such values of the variables with daily temporal resolution in our input data, to predicting the probability of a wildfire burn occurring in the next 7 days. All pixels in the target patches encode a binary label, one for wildfire burn locations and zero otherwise, except for pixels lying outside the study area. Additionally, patches only having

non-burn locations were randomly collected from dates with and without any recorded wildfire.

Each of the input features was normalized to a value of [0, 1]. Training patches were sampled from images within the period 2006–2016, while all test patches are taken from an entirely separate period 2017. Data augmentation, including two flips and three 90° rotations, was applied to increase the number of training patches and help prevent the networks from overfitting.

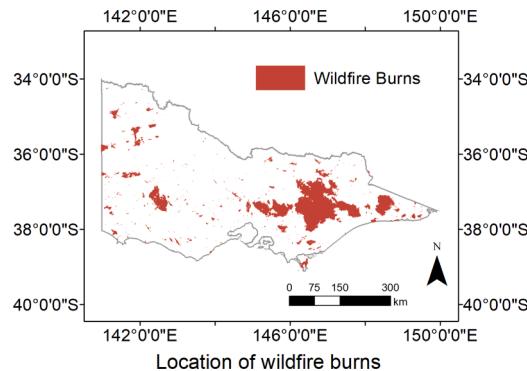


Fig. 4. Locations of wildfire burns for the whole study period obtained from the “Fire History Records of Fires primarily on Public Land” dataset made publicly available by the State of Victoria, Department of Environment, Land, Water & Planning (DELWP) ([The State of Victoria, 1992](#)).

3.3. Network architecture and learning setup

For predicting the probability of wildfire burn from our input and target patches, a variant of deep fully convolutional network (Long et al., 2015) was designed employing residual blocks with identity connections (He et al., 2016). Fig. 5 shows a simplified diagram of the network architecture used in this study. For notational purposes, this network is called AllConvNet. In contrast to the original fully convolutional network, which uses both convolutional and maximum pooling with downsampling layers, AllConvNet only uses convolutional layers (hence the shortened name for all-convolutional network). Sensitivity analysis experiments were performed on some chosen hyperparameters of the network including: the number of layers, number of filters, and input patch size of the network. For the last layer, the network applies a sigmoid activation function followed by a binary cross-entropy loss function:

$$E_N(\mathbf{w}) = -\frac{1}{N} \sum_{n=1}^N h(\mathbf{t}_n) \mathbf{t}_n \log(\mathbf{y}_n) \quad (2a)$$

$$\mathbf{y}_n = f(\mathbf{x}_n) \quad (2b)$$

$$h(\mathbf{t}_n) = \begin{cases} 1 & \text{if } \mathbf{t}_n = 0 \\ \gamma & \text{if } \mathbf{t}_n = 1 \end{cases} \quad (2c)$$

where E is the loss function value evaluated over a batch size N , h is a class-based weighting function, γ is the class-weighting value (giving more importance to the positive samples, encoded as $\mathbf{t}_n = 1$), \mathbf{t} is the target output vector, \mathbf{y} is the model output vector, \mathbf{x} is the input feature vector, and f is the series of operations performed by the network. The network is optimized using a variation of the backpropagation with stochastic gradient descent algorithm called “Adam” (Kingma and Ba, 2014). The networks are trained for 500 epochs. Aside from data augmentation as a form of regularization, an l2 weight decay term is added to the loss function and early stopping was applied by using the model with the best loss function value evaluated on a validation set. The validation set is a subset of the training set, disjoint from the test set where the accuracy metrics reported are calculated from, that is segregated to properly select the values of the network hyperparameters, and consequently not used for optimizing the loss function.

3.3.1. Accuracy assessment

Three other methods were compared against our network. Firstly, an adapted version of SegNet (Badrinarayanan et al., 2017), an encoder-decoder FCN that has been mainly used and employed in pixel-wise image classification originating from computer vision but has also been employed in remote sensing applications (Persello et al., 2019). Secondly, a multilayer perceptron. And finally, a logistic regression. Our network and SegNet accepts $M \times M$ input patches with 29 feature bands, while the MLP and LR only accepts a 1D input vector with 29 features. Both the other two artificial neural networks (SegNet and MLP) used the same optimization method as AllConvNet. Similar hyperparameter selection experimental setups and class-weighting function were used for all the three networks.

The labels of the target image predominantly consists of non-wildfire burn location pixels. Hence, using pixel-wise overall classification accuracy would be inappropriate as predicting negatives for the whole output would still yield a high accuracy. For such cases of imbalance distribution of labels, alternative accuracy metrics would be more suitable, such as a recall-biased F-beta score (Baeza-Yates and Ribeiro-Neto, 2011, pp.327–328) averaged over the two classes:

$$F_\beta = 0.5(F_\beta^+ + F_\beta^-) \quad (3a)$$

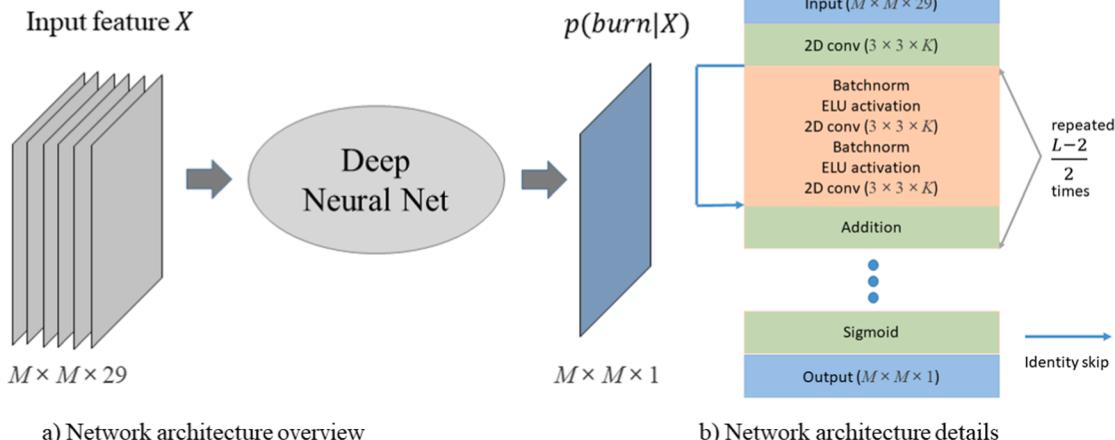


Fig. 5. The deep fully convolutional network architecture (AllConvNet) designed for predicting the probability of wildfire burn. Fig. 5.a shows the main components of an artificial neural network—input layer, hidden layer, and output layer—specifying the dimensions of the first and last layers. Fig. 5.b further shows the pattern of operations applied within the hidden layer of the network (a more detailed discussion of these operations can be found in Bergado et al. (2018)). The “2D conv” learns sets of 2D convolutional filters to extract higher-level features from the input, Batchnorm learns parameters used to normalize the values of the learned features to have zero mean and unit variance (Ioffe and Szegedy, 2015), ELU activation applies an element-wise non-linear activation function called exponential linear units (Clevert et al., 2016), and Addition performs an element-wise addition.

$$F_{\beta}^+ = \frac{(1 + \beta^2)TP}{(1 + \beta^2)TP + \beta^2FN + FP} \quad (3b)$$

$$F_{\beta}^- = \frac{(1 + \beta^{-2})TN}{(1 + \beta^{-2})TN + \beta^{-2}FP + FN}, \quad (3c)$$

the class balance accuracy (Mosley, 2013) defined as:

$$CBA = 0.5(CBA^+ + CBA^-) \quad (4a)$$

$$CBA^+ = \frac{TP}{\max(TP + FP, TP + FN)} \quad (4b)$$

$$CBA^- = \frac{TN}{\max(TN + FN, TN + FP)}, \quad (4c)$$

and the Matthews correlation coefficient (Baldi et al., 2000):

$$MCC = \frac{\frac{TP}{N} - S \times P}{\sqrt{(P \times S)(1 - S)(1 - P)}} \quad (5a)$$

$$N = TP + TN + FP + FN \quad (5b)$$

$$S = \frac{TP + FN}{N} \quad (5c)$$

$$P = \frac{TP + FP}{N} \quad (5d)$$

where F_{β} is the F-beta score parametrized by β , CBA is the class balance accuracy, MCC is the Matthews correlation coefficient, and TP , TN , FP , and FN are the true positive, true negative, false positive, and false negative counts respectively. Increasing β in F_{β} decreases the weight of the FP and TN relative to the TP and FN . Unlike the overall classification accuracy, all these measures can provide less biased estimate of the predictive accuracy on tasks involving highly imbalanced class distribution. Aside from these three measures, three important numbers in the confusion matrix when dealing with highly imbalanced binary classes were also reported, which are: the TP , FN , and FP .

All the metrics reported are calculated on test patches of size 128×128 pixels. Evaluating a pixel-wise accuracy is too restrictive to predict a very rare phenomenon that greatly varies in size from one pixel to tens of thousands. Therefore, a less-constrained measure allowing a tolerance of eight pixels was adapted. The measure is equivalent to applying 8×8 non-overlapping maximum pooling filters to both the predictions and reference before calculating the accuracy metrics. The size of the tolerance, equivalent to around 16 km^2 in ground area, is comparable to the sizes of recommended planned burned area units by several case studies reported in the Australian National Guidelines for Prescribed Burning Operations (Australasian and Forest Fire Management Group, 2014).

3.4. Measuring statistical importance of input variable

Another issue of interest is the relative statistical importance of each input feature in the prediction task. This could help future analysis and data collection efforts to choose which dataset to prioritize when considering a similar case study. For this, two measures were considered: (1) the average gradient with respect to the input feature (Vasilakos et al., 2009) for the three deep neural networks and (2) the LR coefficients.

3.5. Network hyperparameters

Experiments were performed varying hyperparameters of the three deep neural networks including: the input patch size M for both AllConvNet and SegNet; the number of layers L and the number of filters in

each convolutional layer K for both MLP and AllConvNet; and the class weighting value γ for all three networks. Values for each hyperparameter were chosen by optimizing the network accuracy on a validation set, taken from within the training period but is spatially disjoint from the training set. Thirty percent of the total training set was used as the validation set and the remaining 70% was used to optimize the loss function. Only 0.006% of the total training and test set has a positive label (wildfire location). The parameters yielding the best validation accuracy for each network are shown in Table 3.

4. Results and discussion

4.1. Burn prediction maps

We visually compare the resulting predicted maps from each of our four predictive models. Performance on contrasting days—e.g., during summer when a wildfire is recorded and winter when no fire is recorded—cannot be inferred just by observing Tables 4 and 5. We, therefore, visualize three sets of sample maps resulting from our four models' predictions. Firstly, a sampled date (date I) within the training period (December 1, 2006) having recorded a large wildfire (around 680,000 hectares); secondly, a sampled date (date II) within the test period (October 18, 2017) also with recorded, but relatively smaller than the previous date, wildfire (around 8,000 hectares); and lastly, a sampled date (date III) within the test period (June 27, 2017) without any recorded wildfire.

Fig. 6 shows the reference and predicted maps for date I. Fig. 7 shows the reference and predicted maps for date II. Fig. 8 shows the reference and predicted maps for date III. Looking at each set of results from a model across the three dates, it can be seen that the amount of predicted wildfire-prone locations (red pixels) increases together with the area of wildfire observed for a date, except for SegNet and less noticeable for LR. AllConvNet and LR perfectly predicts the locations without wildfire for a given date as shown in Fig. 8. This observation suggests that the AllConvNet and LR are relatively better at temporally distinguishing whether a day is more wildfire-prone compared to another day. Spatially comparing the two dates with recorded wildfire, the predictions of wildfire-prone locations appears to concentrate on comparably similar locations for each of the classifier, with observable size and pattern changes being more apparent in the results of AllConvNet and SegNet.

A significant portion of the maps predicted for dates I and II appear to have a high number of FP's. Since the phenomenon can be generally observed regardless of the model, then a possible explanation of this can be the lack of discriminatory information encoded in our input features. The probability of a wildfire burning can be decomposed into the probability of fuel ignition, given a causative agent triggers an ignition, and the probability of a causative agent starting an ignition. On the one hand, the information on the first probability is intuitively embedded in the features describing fuel characteristics such as topography, land cover, fuel moisture proxies, and the meteorological observation. On the other hand, only the lightning flash density rate and proximity measure (from roads and power lines) features generally embeds information on the second probability—both of which do not have daily temporal resolution. Observing the sample map predictions, all the models seem to capture the interactions within the first probability stronger, that is to say, the models are able to spatially distinguish the differences in fuel ignitability more than the differences in whether one location is more likely to experience an ignition than others. Adding features embedding

Table 3
Selected network hyperparameter values.

Model	M	L	K	γ
AllConvNet	32	21	32	100
SegNet	64	—	—	5
MLP	—	4	64	20

Table 4

Comparison of estimated predictive model accuracy based on sample counts.

Model	TP	FN	FP
AllConvNet	538	390	66881
SegNet	112	816	18873
MLP	412	516	69997
LR	167	761	24034

Table 5

Comparison of estimated predictive model accuracy based on averaged rates.

Model	F_1	F_5	F_{20}	F_{100}	CBA	MCC
AllConvNet	37.64%	56.20%	74.39%	78.57%	58.23%	0.025
SegNet	50.56%	53.45%	55.78%	56.02%	56.03%	0.027
MLP	36.63%	54.03%	64.33%	71.75%	50.48%	0.002
LR	46.52%	53.98%	58.17%	58.70%	51.54%	0.007

information on the second probability, if possible, with higher temporal resolution, might help address this challenge. Additionally, explicitly adding information on temporal weather pattern changes due to climate change as input features could help further improve the presented results.

Both the convolutional networks predict more connected fire-prone regions compared to the results of MLP and LR that shows noisier predictions—with the results from the LR and SegNet at the extreme opposing ends, SegNet producing the smoother and more regularized results. This can be explained by the fact that the two convolutional networks accept as an input $M \times M$ patches, a practice that is rarely done in similar case studies, as opposed to single pixel vectors utilized as an input by LR and MLP. By accepting input patches instead of single pixel vectors, the trainable convolutional filters of the two networks allow the latter to learn spatial-contextual dependencies that may be present between the input features.

Another notable observation is the difference on how the models assign final class score, $p(\text{burn}|x = X)$. In the one instance, AllConvNet and LR assigning more extremely high scores to most, if not all, wildfire-prone regions. Whereas conservative (in between extremely high and low) scores are more observable in the results of SegNet and MLP. Similar results were observed by de Vasconcelos et al. (2001) where the authors compared MLP and LR for wildfire ignition prediction—with the MLP producing more intermediate probability values. Extreme scores can be attributed to possible model overtraining on a dataset with limited positive samples such as recorded wildfire events. The cross-entropy loss function in Eq. 2 continues to be minimized by assigning higher scores to already correctly classified training samples instead of correctly classifying samples which are still misclassified—hence, promoting high scores assigned to wildfire-prone locations during model testing. Broadly speaking, the maps from the two convolutional networks appear to be smoother and more regularized, with the SegNet having smoother and more regularized areas and providing more information by discriminating intermediate, less wildfire-prone areas.

4.2. Predictive accuracy

Tables 4 and 5 show the results of our model compared to the three baseline models. We first report metrics based on sample counts in Table 4. Secondly, we report metrics based on class-averaged rates in Table 5. All the metrics are evaluated using results obtained by applying a threshold of 0.5 to the final class score, $p(\text{burn}|x = X)$, given by each predictive model.

AllConvNet has the highest TP and lowest FN counts but ranks 2nd lowest in terms of FP counts. On the contrary, SegNet produces the lowest FP counts but has the worst TP counts—only correctly classifying 12% of the positive samples as compared to the 58% of the AllConvNet. AllConvNet correctly classified 14% more positive samples than the MLP

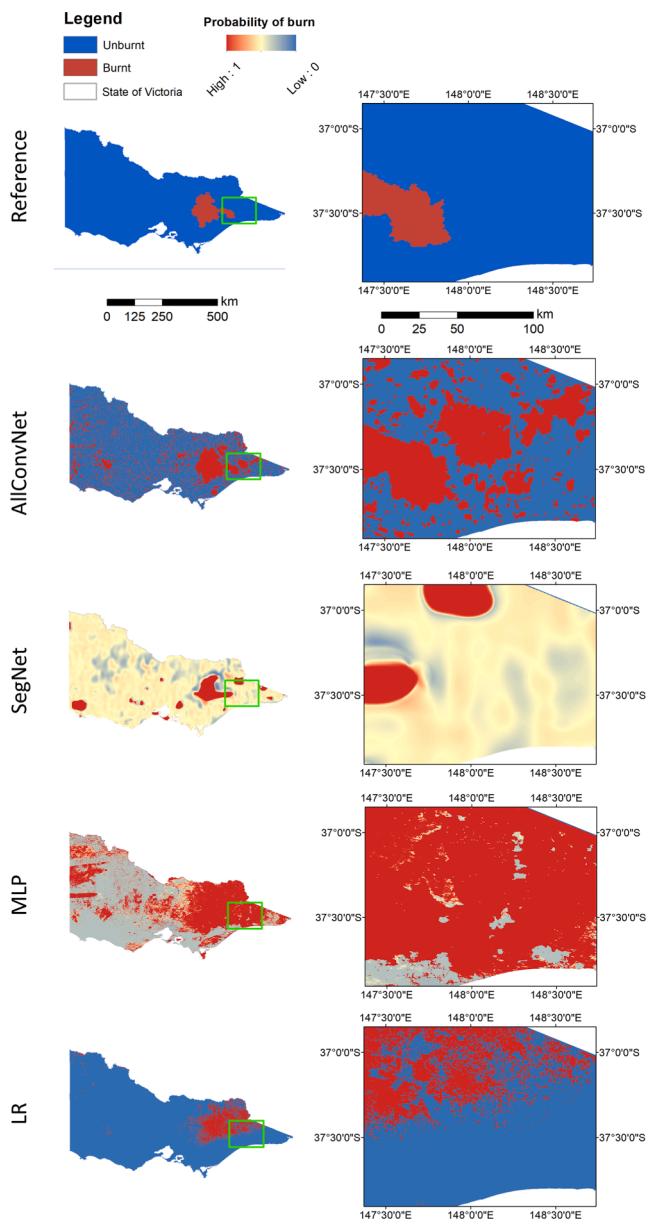


Fig. 6. Visualization of reference (topmost row) and predicted maps for a date (December 1, 2006) when a wildfire was recorded. The maps to the left shows the prediction for the whole extent of Victoria while to the right is an inset map of a zoomed location. The maps uses WGS84 as their coordinate reference system.

while still having 5% less FP samples. LR correctly classified 6% more positive samples than the SegNet but have 27% more false positives.

An ideal predictive model should naturally have high TP count and both low FP and FN counts. None of the four models outperforms the rest in all three counts with the MLP and SegNet being on the extreme ends—the MLP overpredicting and the LR underpredicting positives. Choosing the best among the four would be a matter of how much relative importance we assign to each of the three counts. We argue that putting more weight on both the TP and FN than the FP sounds more reasonable in the context of hazard-related studies such as wildfire prediction. In hazard-related studies where we predict a rare phenomenon, FP's can either be actual miscategorization or possibly genuine hazard-prone locations that are yet to develop the hazardous phenomenon (Beguería, 2006). In line with the argument of assigning less relative weight on FP (recall-biased premise), we report additional F_β measures in Table 5 aside from the conventional F_1 measure. Moreover,

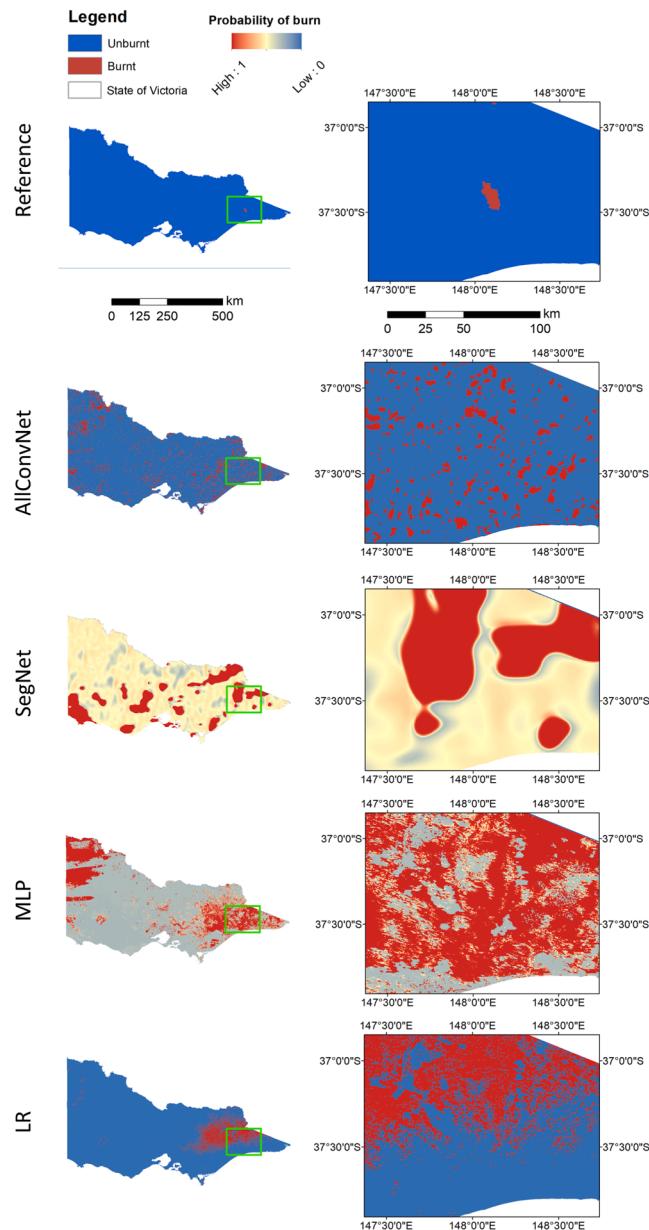


Fig. 7. Visualization of reference (topmost row) and predicted maps for a date (October 18, 2017) when a wildfire was recorded. The maps to the left shows the prediction for the whole extent of Victoria while to the right is an inset map of a zoomed location. The maps uses WGS84 as their coordinate reference system.

we report the *CBA* and *MCC* which are specifically developed for assessing problems with imbalanced class distribution.

SegNet achieves the highest F_1 and MCC while AllConvNet outperforms the rest of the models in terms of the four other measures. All F_β scores consistently increase for all the models as the β increases. The β parameter intuitively provides a way, in the model assessment step, to control how much importance we want to assign on the model's capability to predict as much as wildfire burn event as it can while being more tolerant on false alarms. However, the choice of such weighting will generally fall on the fourth stage of a wildfire risk assessment framework, risk characterization, that is better handled by land and fire management agencies in line with their management objectives.

In these experiments, we evaluated the models varying β to be equal to the optimal γ values chosen in the hyperparameter selection steps (see Table 3), since $\gamma > 1$ serves a similar purpose as β but is employed in the

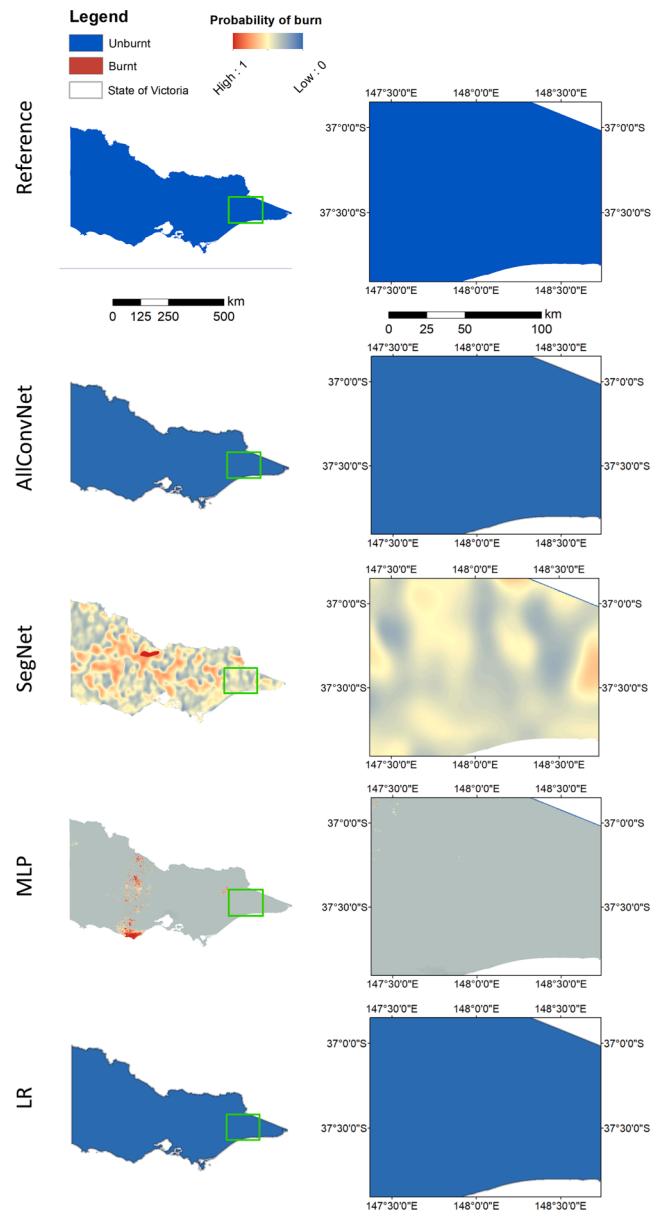


Fig. 8. Visualization of reference (topmost row) and predicted maps for a date (June 27, 2017) when no wildfire was recorded. The maps to the left shows the prediction for the whole extent of Victoria while to the right is an inset map of a zoomed location. The maps uses WGS84 as their coordinate reference system.

training phase of the models. Intuitively, assigning higher γ value would influence the model to overpredict, increasing FP. However, results presented in Table 4 demonstrates that AllConvNet, despite having higher γ than MLP, is less prone to overprediction. While scoring highest in terms of all F_β except F_1 and the parameter-free *CBA*, AllConvNet also comes out second in terms of the other parameter-free measure *MCC*. This shows that improvement in the accuracy metrics of the results from AllConvNet is independent from our recall-biased premise.

4.3. Feature statistical importance

Artificial neural networks are often criticized for being less interpretable than other models like LR, such that, it is difficult to determine the internal statistical importance of each input variable being considered in the modelling task. This apparent low model transparency can be negligible in the context of other predictive tasks, e.g. in land cover classification where there may be less significance in knowing how the

digital values assigned to pixels in a satellite image get transformed into land cover categories. In the context of wildfire prediction, understanding the relative statistical importance of each input variable as to how they affect the predictability of a wildfire event can be valuable as most fire danger ratings, such as the McArthur Forest Fire Danger Index and the Canadian Forest Fire Weather Index, are similarly based on relative importance weighting of wildfire-related variables. Fig. 9 shows the feature statistical importance measures considering two techniques: (1) gradient with respect to the input feature for AllConvNet, SegNet, and MLP and (2) normalized magnitude of the LR coefficients.

Total precipitation, lightning flash density, and land surface temperature occur to be consistently highly weighted by all models while terrain aspect components, wind direction components, certain land cover classes (such as crop field and woodland), and distance from power lines are ranked on the lower end. In a study employing MLP and LR to predict fire weather index, Vasilakos et al. (2009) compared a subset of the weather variables that we considered in our own study and found that their model also ranked the amount of precipitation as the highest among four variables they compared followed by temperature, wind speed, and finally relative humidity—a result comparable to what we can observe from the ranking provided by AllConvNet.

The slope and elevation and presence of certain land cover classes (forest, alpine meadow, and grassland) were given moderate importance by AllConvNet together with most of the bands and indices derived from MODIS except NDWI. Slope and elevation ranked consistently higher than the terrain aspect components. MODIS derived land surface temperature has higher statistical importance measure than the two emissivity bands. Notably, NDII and VARI generally ranks higher than NDVI and NDII even though Caccamo et al. (2012), regressing these

four indices against live fuel moisture content, found NDII and NDWI to have the highest and lowest coefficient of determination respectively.

The degree of agreement between the resulting set of feature statistical importance scores cannot be quantitatively observed in Fig. 9. Thus, we present Table 6 showing the correlation coefficients and corresponding p-values of the feature statistical importance scores shown in Fig. 9. Statistical importance measures from logistic regression appear to have highest correlation values with the other models, MLP generally having the lowest ones. SegNet has the lowest and highest correlation values with LR and MLP respectively. This could be due to the increased model complexity of MLP but still having the same input dimensionality as LR, possibly making the former more prone to fixating on features the other models found less statistically important.

5. Conclusion

This study presents a deep fully convolutional neural network for predicting daily maps of the probability of a wildfire burn over the next

Table 6
Correlation of the feature statistical importance measures.

Model A	Model B	Pearson's R	p-value
AllConvNet	SegNet	0.32	0.0858
AllConvNet	MLP	0.15	0.4288
AllConvNet	LR coefficients	0.45	0.0146
SegNet	MLP	0.08	0.6947
SegNet	LR coefficients	0.62	0.0002
MLP	LR coefficients	0.47	0.0097

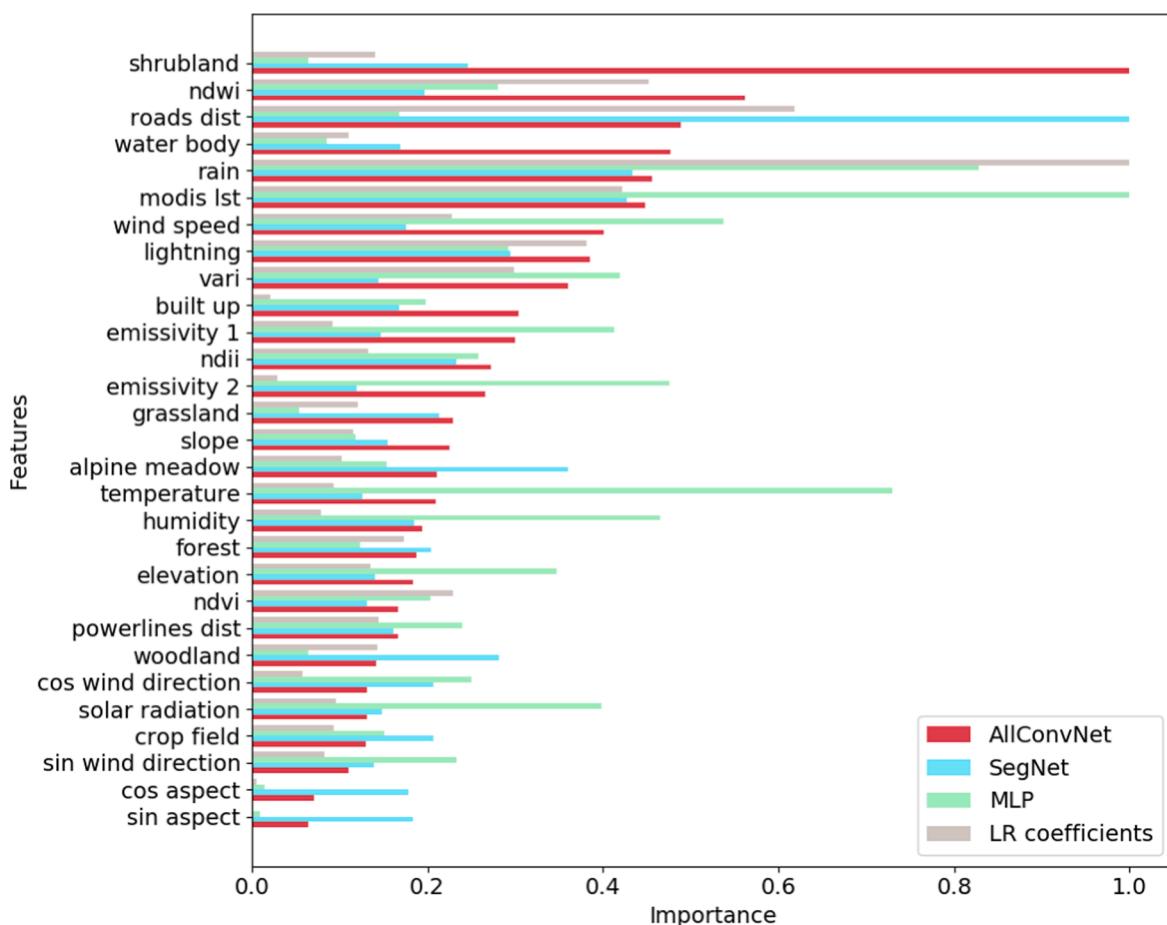


Fig. 9. Feature statistical importance measures considering gradient with respect to the input feature for AllConvNet, SegNet, and MLP and the normalized magnitude of the LR coefficients.

7 days in Victoria, Australia for the period of 2006–2017. The proposed network architecture, AllConvNet, outperforms the other three baseline methods namely: SegNet, multilayer perceptron, and logistic regression in four of the six quantitative metrics (four recall-biased F-beta scores, class balance accuracy score, and Matthews correlation coefficient) and ranks second in one of the other two parameter-free metric. Both the two convolutional networks, AllConvNet and SegNet, also provide smoother and more regularized predicted maps, with SegNet providing better visual information on discriminating less wildfire-prone locations. Input feature statistical importance was measured for the three networks and compared against logistic regression coefficients. Total precipitation, lightning flash density, and land surface temperature occurs to be consistently highly weighted by all models while terrain aspect components, wind direction components, certain land cover classes (such as crop field and woodland), and distance from power lines are ranked on the lower end.

The deep fully convolutional networks designed in this study demonstrate higher predictive accuracy and map quality than the baseline methods commonly explored in previous studies. Exploratory feature statistical importance measures presented in this work also provide a good way to better understand these rather less transparent models. Future studies on the application of deep learning for wildfire prediction may consider improving the features encoding information on the probability that a location would experience an ignition. In the context of deep learning, future works on incorporating temporal information, accepting as inputs and producing as outputs sequences instead of daily snapshots of features and predictions, and data fusion within the model design would be relevant. Resulting maps from our models can be reclassified into wildfire burn exposure indices and be intersected with information on assets-at-risk to produce alternative wildfire risk index maps.

Acknowledgements

This publication was, in part, made possible by the GEOSAFE project, an initiative to boost the progress of wildfire knowledge and the related development of innovative methods for dealing efficiently with such fires, funded by the Marie Skłodowska-Curie Research and Innovation Staff Exchange, under grant #691161.

References

- Ager, A.A., Day, M.A., McHugh, C.W., Short, K., Gilbertson-Day, J., Finney, M.A., Calkin, D.E., 2014. Wildfire exposure and fuel management on western US national forests. *J. Environ. Manage.* 145, 54–70.
- Atkinson, D., Chladil, M., Janssen, V., Lucieer, A., 2010. Implementation of quantitative bushfire risk analysis in a GIS environment. *Int. J. Wildland Fire* 19, 649–658.
- Australasian Fire Emergency Service Authorities Council, Forest Fire Management Group, 2014. National Guidelines for Prescribed Burning Operations. Australasian Fire Emergency Service Authorities Council, Level 1, 340 Albert Street East Melbourne Victoria 3002.
- Badia, A., Serra, P., Modugno, S., 2011. Identifying dynamics of fire ignition probabilities in two representative Mediterranean wildland-urban interface areas. *Appl. Geogr.* 31, 930–940.
- Badrinarayanan, V., Kendall, A., Cipolla, R., 2017. Segnet: A deep convolutional encoder-decoder architecture for image segmentation. In: *IEEE Trans. Pattern Anal. Mach. Intell.*.
- Baeza-Yates, R., Ribeiro-Neto, B., 2011. *Modern Information Retrieval: The Concepts and Technology Behind Search*. Addison Wesley.
- Baldi, P., Brunak, S., Chauvin, Y., Andersen, C.A.F., Nielsen, H., 2000. Assessing the accuracy of prediction algorithms for classification: an overview. *Bioinformatics* 16, 412–424. <https://doi.org/10.1093/bioinformatics/16.5.412>.
- Bar Massada, A., Radloff, V.C., Stewart, S.I., Hawbaker, T.J., 2009. Wildfire risk in the wildland-urban interface: A simulation study in northwestern Wisconsin. *For. Ecol. Manage.* 258, 1990–1999.
- Beguería, S., 2006. Validation and Evaluation of Predictive Models in Hazard Assessment and Risk Management. *Nat. Hazards* 37, 315–329.
- de Bem, P.P., de Carvalho Júnior, O.A., Matricardi, E.A.T., Guimarães, R.F., Gomes, R.A. T., 2018. Predicting wildfire vulnerability using logistic regression and artificial neural networks: a case study in Brazil. *Int. J. Wildland Fire* 28, 35–45.
- Bengio, Y., 2009. Learning deep architectures for AI. *Found. Trends Mach. Learn.* 2, 1–127.
- Bergado, J.R., Persello, C., Stein, A., 2018. Recurrent multiresolution convolutional networks for vhr image classification. *IEEE Trans. Geosci. Remote Sens.* 56, 6361–6374.
- Bishop, C.M., 2006. *Pattern Recognition and Machine Learning (Information Science and Statistics)*. Springer-Verlag, Berlin, Heidelberg.
- Bradstock, R., Hammill, K., Collins, L., Price, O., 2009. Effects of weather, fuel and terrain on fire severity in topographically diverse landscapes of south-eastern Australia. *Landscape Ecol.* 25, 607–619. <https://doi.org/10.1007/s10980-009-9443-8>.
- Caccamo, G., Chisholm, L.A., Bradstock, R.A., Puotinen, M.L., Pippen, B.G., 2012. Monitoring live fuel moisture content of heathland, shrubland and sclerophyll forest in south-eastern Australia using MODIS data. *Int. J. Wildland Fire* 21, 257.
- Cancelo-González, J., Cachaldora, C., Díaz-Fierros, F., Prieto, B., 2014. Colourimetric variations in burnt granitic forest soils in relation to fire severity. *Ecol. Ind.* 46, 92–100.
- Catry, F.X., Rego, F.C., Baçao, F.L., Moreira, F., 2009. Modeling and mapping wildfire ignition risk in Portugal. *Int. J. Wildland Fire* 18, 921.
- Chen, G., Metz, M.R., Rizzo, D.M., Dillon, W.W., Meentemeyer, R.K., 2015. Object-based assessment of burn severity in diseased forests using high-spatial and high-spectral resolution master airborne imagery. *ISPRS J. Photogram. Remote Sens.* 102, 38–47.
- Chowdhury, E.H., Hassan, Q.K., 2015. Operational perspective of remote sensing-based forest fire danger forecasting systems. *ISPRS J. Photogram. Remote Sens.* 104, 224–236.
- Chuvieco, E., Aguado, I., Yebra, M., Nieto, H., Salas, J., Martín, M., Vilar, L., Martínez, J., Martín, S., Ibarra, P., de la Riva, J., Baeza, J., Rodríguez, F., Molina, J., Herrera, M., Zamora, R., 2010. Development of a framework for fire risk assessment using remote sensing and geographic information system technologies. *Ecol. Model.* 221, 46–58.
- Clevert, D., Unterthiner, T., Hochreiter, S., 2016. Fast and accurate deep network learning by exponential linear units (ELUs). In: *International Conference on Learning Representations*.
- [dataset] Cecil, D.J., Buechler, D., Blakeslee, R., 2014. Lis/otd gridded lightning climatology data collection. <https://ghrc.nsstc.nasa.gov/>. Accessed: 2018-10-22.
- [dataset] Hutchinson, M.F., Stein, J.L., Stein, J.A., Anderson, H., Tickle, P.K., 2008. Geodata 9 second dem and d8: Digital elevation model version 3 and flow direction grid 2008. <https://ecat.ga.gov.au/geonetwork/srv/eng/catalog.search?node=srv#/metadata/66006>. doi:<http://pid.geoscience.gov.au/dataset/ga/66006>. accessed: 2018-09-26.
- [dataset] Lymburner, L., Tan, P., Mueller, N., Thackway, R., Thankappan, M., Islam, A., Lewis, A., Randall, L., Senarath, U., 2011. The national dynamic land cover dataset - technical report. <https://ecat.ga.gov.au/geonetwork/srv/eng/catalog.search?node=srv#/metadata/71069>. doi:<http://pid.geoscience.gov.au/dataset/ga/71069>. accessed: 2018-09-26.
- [dataset] The State of Victoria, 1992. Fire history records of fires primarily on public land. <https://www.data.vic.gov.au/data/dataset/fire-history-records-of-fires-primary-on-public-land>. doi:[ANZVI0803004741](https://doi.org/10.4225/ANZVI0803004741). accessed: 2017-02-07.
- [dataset] The State of Victoria, 2009a. Road network - vicmap transport. <https://www.data.vic.gov.au/data/dataset/road-network-vicmap-transport>. doi:[ANZVI0803002595](https://doi.org/10.4225/ANZVI0803002595). accessed: 2018-09-22.
- [dataset] The State of Victoria, 2009b. Vicmap features of interest. <https://www.data.vic.gov.au/data/dataset/vicmap-features-of-interest>. doi:[ANZVI0803003646](https://doi.org/10.4225/ANZVI0803003646). accessed: 2018-09-22.
- [dataset] Vermote, E., 2015. Mod09a1 modis/terra surface reflectance 8-day l3 global 500m sin grid v006. <https://lpdaac.usgs.gov/>. doi: 10.5067/MODIS/MOD09A1.006. accessed: 2018-07-19.
- [dataset] Wan, Z., Hook, S., Hulley, G., 2015. Mod11a2 modis/terra land surface temperature/emissivity 8-day l3 global 1km sin grid v006. <https://lpdaac.usgs.gov/>. doi: 10.5067/MODIS/MOD11B3.006. accessed: 2018-10-22.
- Dutta, R., Das, A., Aryal, J., 2016. Big data integration shows Australian bush-fire frequency is increasing significantly. *Royal Soc. Open Sci.* 3, 150241.
- Estes, B.L., Knapp, E.E., Skinner, C.N., Miller, J.D., Preisler, H.K., 2017. Factors influencing fire severity under moderate burning conditions in the Klamath mountains, Northern California, USA. *Ecosphere* 8, e01794. <https://doi.org/10.1002/ecs2.1794>.
- Fairbrother, A., Turnley, J.G., 2005. Predicting risks of uncharacteristic wildfires: Application of the risk assessment process. *For. Ecol. Manage.* 211, 28–35.
- Finney, M.A., 2005. The challenge of quantitative risk analysis for wildland fire. *For. Ecol. Manage.* 211, 97–108.
- Gevaert, C., Persello, C., Nex, F., Vosselman, G., 2018. A deep learning approach to dtm extraction from imagery using rule-based training labels. *ISPRS J. Photogram. Remote Sens.* 142, 106–123.
- Goodfellow, I., Bengio, Y., Courville, A., 2016. *Deep Learning*. MIT Press. <http://www.deeplearningbook.org>.
- Gray, M.E., Zachmann, L.J., Dickson, B.G., 2018. A weekly, continually updated dataset of the probability of large wildfires across western US forests and woodlands. *Earth Syst. Sci. Data* 10, 1715–1727.
- Guo, Z., Fang, W., Tan, J., Shi, X., 2013. A time-dependent stochastic grassland fire ignition probability model for Hulun Buir Grassland of China. *Chin. Geogr. Sci.* 23, 445–459.
- Hanton, S., Padilla, M., Corti, D., Chuvieco, E., 2013. Strengths and weaknesses of MODIS hotspots to characterize global fire occurrence. *Remote Sens. Environ.* 131, 152–159.
- He, K., Zhang, X., Ren, S., Sun, J., 2016. Deep residual learning for image recognition. In: *Computer Vision and Pattern Recognition (CVPR), 2016 IEEE Conference on*.
- Ioffe, S., Szegedy, C., 2015. Batch normalization: Accelerating deep network training by reducing internal covariate shift. *arXiv preprint arXiv:1502.03167*.
- Jaafari, A., Gholami, D.M., Zenner, E.K., 2017. A Bayesian modeling of wildfire probability in the Zagros Mountains. *Iran. Ecol. Inform.* 39, 32–44.

- Jurda, S., Chuvieco, E., 2012. Modelling Fire Ignition Probability from Satellite Estimates of Live Fuel Moisture Content. *Fire Ecol.* 7, 77–97.
- Kingma, D.P., Ba, J., 2014. Adam: A method for stochastic optimization. arXiv preprint arXiv:1412.6980 abs/1412.6980.
- Long, J., Shelhamer, E., Darrell, T., 2015. Fully convolutional networks for semantic segmentation, in: IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2015, Boston, MA, USA, June 7–12, 2015, pp. 3431–3440.
- Mariethoz, G., Renard, P., 2010. Reconstruction of Incomplete Data Sets or Images Using Direct Sampling. *Math. Geosci.* 42, 245–268.
- Mhawej, M., Fahour, G., Abdallah, C., Adjizian-Gerard, J., 2016. Towards an establishment of a wildfire risk system in a Mediterranean country. *Ecol. Inform.* 32, 167–184.
- Miller, C., Parisien, M.A., Ager, A.A., Finney, M.A., 2008. Evaluating spatially-explicit burn probabilities for strategic fire management planning, pp. 245–252.
- Minas, J.P., Hearne, J.W., Martell, D.L., 2014. A spatial optimisation model for multi-period landscape level fuel management to mitigate wildfire impacts. *Eur. J. Oper. Res.* 232, 412–422.
- Mosley, L., 2013. A balanced approach to the multi-class imbalance problem. Ph.D. thesis. Ames, Iowa.
- Mundo, I.A., Wiegand, T., Kanagaraj, R., Kitzberger, T., 2013. Environmental drivers and spatial dependency in wildfire ignition patterns of northwestern Patagonia. *J. Environ. Manage.* 123, 77–87.
- Nami, M.H., Jaafari, A., Fallah, M., Nabiuni, S., 2018. Spatial prediction of wildfire probability in the Hyrcanian ecoregion using evidential belief function model and GIS. *Int. J. Environ. Sci. Technol.* 15, 373–384.
- Narayananaraj, G., Wimberly, M., 2011. Influences of forest roads on the spatial pattern of wildfire boundaries. *Int. J. Wildland Fire* 20, 792–803. <https://doi.org/10.1071/WF10032>.
- Narayananaraj, G., Wimberly, M.C., 2012. Influences of forest roads on the spatial patterns of human- and lightning-caused wildfire ignitions. *Applied Geography* 32, 878–888. URL: <http://www.sciencedirect.com/science/article/pii/S0143622811001731>, doi: 10.1016/j.apgeog.2011.09.004.
- Parisien, M.A., Snetsinger, S., Greenberg, J.A., Nelson, C.R., Schoennagel, T., Dobrowski, S.Z., Moritz, M.A., 2012. Spatial variability in wildfire probability across the western United States. *Int. J. Wildland Fire* 21, 313.
- Persello, C., Tolpekin, V., Bergado, J.R., de By, R.A., 2019. Delineation of agricultural fields in smallholder farms from satellite images using fully convolutional networks and combinatorial grouping. *Remote Sens. Environ.* 231, 1–18.
- Sakr, G., Elhajj, I., Mitri, G., 2011. Efficient forest fire occurrence prediction for developing countries using two weather parameters. *Eng. Appl. Artif. Intell.* 24, 888–894.
- Sánchez Sánchez, Y., Martínez-Graña, A., Santos Francés, F., Mateos Picado, M., 2018. Mapping Wildfire Ignition Probability Using Sentinel 2 and LiDAR (Jerte Valley, Cáceres, Spain). *Sensors* 18, 826.
- Sikder, I., Mal-Sarkar, S., Mal, T., 2006. Knowledge-based risk assessment under uncertainty for species invasion. *Risk Anal.* 26, 239–252.
- The State of Victoria, 2012. Code of practice for bushfire management on public land. Dept. of Sustainability and Environment, Melbourne.
- Thompson, M.P., Calkin, D.E., 2011. Uncertainty and risk in wildland fire management: A review. *J. Environ. Manage.* 92, 1895–1909.
- Thompson, M.P., Haas, J.R., Gilbertson-Day, J.W., Scott, J.H., Langowski, P., Bowne, E., Calkin, D.E., 2015. Development and application of a geospatial wildfire exposure and risk calculation tool. *Environ. Model. Software* 63, 61–72.
- de Vasconcelos, M.P., Silva, S., Tome, M., Alvim, M., Pereira, J.C., 2001. Spatial prediction of fire ignition probabilities: comparing logistic regression and neural networks. *Photogram. Eng. Remote Sens.* 67, 73–81.
- Vasilakos, C., Kalabokidis, K., Hatzopoulos, J., Matsinos, I., 2009. Identifying wildland fire ignition factors through sensitivity analysis of a neural network. *Nat. Hazards* 50, 125–143.
- Volpi, M., Tuia, D., 2017. Dense semantic labeling of subdecimeter resolution images with convolutional neural networks. *IEEE Trans. Geosci. Remote Sens.* 55, 881–893.
- Wickramasinghe, C.H., Jones, S., Reinke, K., Wallace, L., 2016. Development of a multi-spatial resolution approach to the surveillance of active fire lines using himawari-8. *Remote Sensing* 8.
- Yebra, M., Dennison, P.E., Chuvieco, E., Riaño, D., Zylstra, P., Hunt, E.R., Danson, F.M., Qi, Y., Jurda, S., 2013. A global review of remote sensing of live fuel moisture content for fire danger assessment: Moving towards operational products. *Remote Sens. Environ.* 136, 455–468.
- Yıldırım, H., Özdemir, M., Güre, M., Uğurlu, B., Özel, M., 2005. Monitoring the after-effects of forest fires by satellite data. In: RAST 2005 -Proceedings of 2nd International Conference on Recent Advances in Space Technologies, pp. 626–629.
- Zhang, Y., Lim, S., Sharples, J.J., 2016. Modelling spatial patterns of wildfire occurrence in south-eastern Australia. *Geomatics, Nat. Hazards Risk* 7, 1800–1815. <https://doi.org/10.1080/19475705.2016.1155501>.