Predicting sub-continental fuel hazard under future climate and rising [CO_2] - combining ground-based observations with optimal plant behaviour in a machine learning approach

Jinyan Yang[1], Lina Teckentrup[1], Assaf Inbar[1], Juergen Knauer[1], Mingkai Jiang[2], Belinda Medlyn[1], Owen Price[3], Ross Bradstock [3], and Matthias Boer[1]

[1]Western Sydney University Hawkesbury Institute for the Environment
[2]Zhejiang University
[3]University of Wollongong

March 1, 2023

Abstract

Bushfire fuel hazard is determined by fuel hazard that represents the type, amount, density, and three-dimensional distribution of plant biomass and litter. The fuel hazard represents a biological control on fire danger and may change in future with plant growth patterns. Rising atmospheric CO2 concentration (Ca) tends to increase plant productivity ('fertilisation effect') but also alters climate, leading to a 'climatic effect'. Both effects will impact on future vegetation and thus fuel hazard. Quantifying these effects is an important component of predicting future fire regimes and evaluating fire management options. Here, by combining a machine learning algorithm that incorporates the power of large fine-resolution datasets with a novel optimality model that accounts for the climatic and fertilisation effects on vegetation cover, we developed a random forest model to predict fuel hazard at fine spatial resolution across the state of Victoria in Australia. We fitted and evaluated model performance with long-term (i.e., 20 years), ground-based fuel observations. The model achieved strong agreement with observations across the fuel hazard range (accuracy >65%). We found fuel hazard increased more in dry environments to future climate and Ca. The contribution of the 'fertilisation effect' to future fuel hazard varied spatially by up to 12%. The predictions of future fuel hazard are directly useful to inform fire mitigation policies and as a reference for climate model projections to account for fire impacts.

Introduction

Recent catastrophic fires around the world have drawn attention to the need for improved fire risk assessments (Bowman et al., 2020; Duane et al., 2021). The likelihood of fire in terrestrial ecosystems is a function of: (i) the fuel hazard (i.e., the amount, density and three-dimensional distribution of plant biomass, both dead and alive), (ii) fuel dryness, (iii) the weather conditions, and (iv) the availability of ignition sources (Bradstock, 2010; Boer et al., 2017). Existing fire behaviour models can capture the impacts of weather on fire spread rate and intensity but require spatially explicit information about a range of fuel attributes as input (Tolhurst et al., 2008). Although current fuel hazard can be mapped using a combination of ground-based and remote-sensed observations (Pierce et al., 2012), quantification of future changes in these patterns in response to climate change requires predictive models.

Fuel hazard is closely related to variation in the composition and structure of the vegetation, which in turn are shaped by plant responses to long-term environmental conditions and disturbance regimes (Kelley, et al., 2019). Consequently, predictions of future fuel hazard need to incorporate the potential impacts of climate

change. There are two major ways that climate change affects fuel hazard. First, the rising atmospheric CO_2 concentration (C_a) fertilises plants via an enhancement of photosynthesis (Ainsworth and Rogers, 2007), potentially resulting in an increase in plant biomass (Ainsworth and Long, 2005; Norby et al., 2005; Zhu et al., 2016; Walker et al., 2019). Any increase in plant biomass is likely to result in higher fuel loads, but the magnitude of change and how it will interact with other environmental factors remains uncertain (Bradstock, 2010). Second, rising air temperatures and altered rainfall patterns have distinct effects on plant productivity and species composition, both of which could lead to altered fuel hazard (Archibald et al., 2013). It is thus critical to account for plant responses to climate change when projecting future fuel hazard.

Changes in plant biomass under future climate can be predicted with a range of modelling approaches, which have been used to estimate fuel loads. For example, Clarke et al. (2016) projected future fuel loads using the net primary productivity (NPP) predictions from a land surface model, assuming a linear relationship between NPP and fuel load. However, existing evidence suggests that fire regimes (i.e., fire frequency, intensity, season, type, and extent) could varied within a biome with similar NPP - a single biome could have more than one fire regime while the same fire regime can be observed in different biomes (Archibald et al., 2013). Consequently, predictions of the spatial variation of fuel hazard under climate change need to be constrained by and evaluated against the fuel hazard observations under different environmental controls (i.e., climate, soil and topography).

Ground-based fuel observations have been routinely collected by fire management agencies in Victoria, Australia since 1995 (e.g., Hines et al. 2010), and have provided valuable insight on the spatial variation of fuel hazard at landscape to regional scales (e.g., Jenkins et al., 2020; McColl-Gausden et al., 2020). At each survey site, an ordinal score is assigned to each fuel stratum based on visual estimates of fuel hazard. The potential of these fine-resolution data to help inform process-based model predictions of future fuel hazard remains under-utilised. Assuming the spatial variation of fuel hazard along climatic gradients is indicative of how fuel hazard may change with climate over time (i.e., space-for-time substitution; Picket, 1989), these ground-based fuel surveys contain possibly the best information about the potential for changes in fuel hazard across Victoria in response to projected climate conditions.

Random forest models have been used to synthesise field-based fire observations and environmental drivers with demonstrated success (Pierce et al., 2012; Jenkins et al., 2017; McColl-Gausden et al., 2020). However, previous machine learning approaches have generally ignored plant responses to climate change (e.g., McColl-Gausden et al., 2020), due to the this 'space-for-time' approach needing additional process-based information on vegetation responses to novel climate.

The past developments of empirical approaches thus exposed the limitation of pure statistical analysis and advocate novel ways to combine strengths of process-based plant biomass predictions with data-driven approaches (e.g., Jenkins et al., 2020). Yang et al. (2018) modelled the change of leaf area index (LAI; an indicator of plant foliage biomass) under changing climate and rising C_a. Incorporating this LAI model and random forest models could help address the lack of plant responses to future climate in current machine learning frameworks. This combined framework could unite the strength of fine-resolution empirical observations and the process-based plant responses to climate changes, addressing the weakness of previous regression and process-based models.

Here, we used random forest models to predict spatial variation in fuel hazard at fine spatial resolution across the state of Victoria as a function of climate, soil and topographic attributes as well as modelled plant responses to climate change. The goal was to assess the potential of change in fuel hazard in response to projected future climate conditions and C_a. Although the training and evaluation of the models focused on a specific region, the methods and conclusions built a quantitative understanding of anthropogenic impacts on future fuel hazard, which is applicable to similar regions across the world.

Material and methods

We used ground-based fuel hazard score observations and gridded information on 15 environmental predictors related to climate (6), topography (5), soil (3) and potential vegetation cover (1). We used the ground-based fuel data to train and evaluate random forest models. We then made projections of future fuel hazard using climate model projections. The details of the data and models are outlined below.

Fuel hazard observations

The study area was the state of Victoria in Southeast Australia, where extensive fine resolution data were available to test our novel machine-learning approach (Figure 1). The Victorian Department of Environment, Land, Water and Planning (DELWP) provided field observations of fuel hazard ratings over the period 1995 to 2017 with a total of 47,245 individual records. The data contain categorical fuel hazard scores for surface, near-surface, elevated and bark fuel strata, with five levels: low (1), moderate (2), high (3), very high (4) and extreme (5) (Hines et al. 2010). The records were mostly one-off assessments for georeferenced plots of ca. 20m x 20 m.

We focused the analysis on the elevated stratum because of its high consistency among different surveys (Watson et al., 2012) and high impact on fire regimes (Hines et al. 2010). The elevated fuel hazard score is based on the cover and horizontal connectivity of dead and live plants that is close for biomass that may not be consumed by a flame height of 0.5 m. The near-surface fuel hazard score is based on cover and horizontal connectivity of dead and live plants that is close (<0.5m) but not lying on the ground. The surface fuel hazard score is based on litter depth, cover and horizontal connectivity lying on the ground. Bark fuel hazard is specific to certain tree species and alone does not contribute to the fire regime at landscape scale. Therefore, we do not specifically model bark fuel stratum in this study.

Since we are interested in predicting potential fuel hazard score, we used the observations made at least 10 years after the last fire; this length of time allows the fuel to accumulate beyond 95% of capacity (Peet et al., 1971; Fox et al., 1979; Burrows et al, 1994; Zazali et al., 2021). This filtering resulted in a total of 27,799 observations covering Victoria. Figure 1 shows the spatial distribution of the field observations, with the temporal distribution shown in Figure S1.

Climate predictors

We chose a set of climate, soil and topographic predictors that have been shown to drive fuel hazard in previous studies (e.g., Jenkins et al., 2020; McColl-Gausden et al., 2020). We used gridded daily climate data at 0.05° x 0.05° resolution from 1994-2018 obtained from the SILO project (Jeffrey et al., 2001; accessed at http://www.longpaddock.qld.gov.au/silo/). We extracted daily precipitation (PPT), maximum air temperature (T_{max}) and minimum relative humidity (RH) for the grid cells corresponding to the locations of the field observations. We aggregated the climate data to monthly time steps by taking the sum of precipitation and means for other variables. The PPT, T_{max}, minimum RH of the month before the fuel hazard observations were used as predictors to represent the impacts of short-term climate conditions. We calculated the mean annual precipitation (MAP) and T_{max} for 1994-2018 to represent long-term climate conditions. Finally, a rainfall seasonality index presented by Feng et al. (2013) was used to capture long-term variation in the temporal distribution of precipitation over the twelve months of the year; the index varies from 0 (even distribution over all months) to 2.48 (all rainfall concentrated in a single month). The details of meteorological data are in Table S1. In addition to the climatic predictors used in the model, we obtained potential evapotranspiration (PET) from SILO and averaged over 1994-2018. We then calculated an Aridity index (AI) as PET/MAP. AI is not used as a predictor in the models but rather as an indicator of long-term water balance among the pixels in the following analysis.

Soil and terrain attributes

Soil attributes, in particular proxies for water holding capacity and soil fertility, are expected to constrain spatial variation of fuel loads and fuel properties via their effect on vegetation composition, density, and structure (McColl-Gausden et al, 2020). We obtained gridded bulk density, clay content for topsoil (0-5 cm) and Available Volumetric Water Capacity (AWC; 0-200 cm) at 90 m resolution from the Whole Earth product of the Soil and Landscape Grid of Australia (Malone, 2022).

We used fine scale topographic products containing wetness index (Gallant and Austin, 2012a), adjusted monthly solar radiation in January and July (Gallant et al., 2014), and plan/profile curvature (Gallant and Austin, 2012 b,c) at 3 arcsecond resolution (~90 m). The topographic wetness index, calculated as specific catchment area divided by slope, is commonly used as an indicator of soil water availability (Gallant and Austin, 2012a). Mean monthly shortwave radiation is the mean shortwave radiation (MJ $m^{-2}d^{-1}$) received by a surface accounting for latitude, day of year, average atmospheric conditions, and terrain effects (i.e., slope, aspect and topographic shading). We chose the shortwave radiation in January and July to account for different energy inputs for summer and winter. Plan and profile curvature, derived from the Smoothed Digital Elevation Model (Geoscience Australia, 2015), added further constraints on soil moisture availability and other variation in other soil attributes (e.g., soil depth). The soil and topographic data used are summarised in Table S1.

Future climate change projections

The climate change projections are based on the downscaled output of nine General Circulation Models (Clark et al., 2021): ACCESS1-0, BNU-ESM, CSIRO-Mk3-6-0, GFDL-CM3, GFDL-ESM2G, GFDL-ESM2M, INM-CM4, IPSL-CM5A-LR, MRI-CGCM3. The chosen projections were from a full list of 12 models in Clark et al. (2021), but we excluded three models because they do not cover both Representative Concentration Pathways (RCP) 4.5 and 8.5. We used projections of PPT, T_{max} and minimum RH for the intermediate RCP 4.5 and high emission RCP 8.5 for the period 2000-2100. We obtained projected climate data for time periods of 16 years at the beginning (2000-2015), middle (2045-2060), and end (2085-2100) of the 21st century. We calculated the mean monthly climate for those three time periods for each of the nine climate models. The data were then aggregated in the same way as the historical climate data while the spatial resolution was kept at 0.036 degree (~3.6 km). The MAP and mean T_{max} were the mean of each 16-year period. The mean of current and future climate among the projections are shown in Figure S2. The future C_a used in the study is shown in Table S2.

Optimal leaf area index

Vegetation structural response to climate change was based on gridded LAI layers simulated by an optimisation model (Yang et al., 2018). Briefly, the model uses long-term mean precipitation, maximum air temperature, vapour pressure deficit (calculated based on T_{max} and RH) and photosynthetically active radiation (converted from monthly solar radiation) to predict LAI based on the concept of ecohydrological equilibrium (Woodward,1987). The long-term mean PPT is used to estimate the amount of water available to support evapotranspiration. The optimal LAI is then calculated as the LAI that maximises canopy carbon export (gross photosynthesis less leaf construction and respiration costs) subject to this constraint on evapotranspiration.

There are four advantages of the optimal LAI model (Yang et al., 2018) which makes it suitable to incorporate into machine learning: (i) it has detailed photosynthetic and stomatal processes to capture plant responses to climate change and rising C_a; (ii) the optimisation process accounts for potential changes in plant strategies under future conditions; (iii) it showed a good agreement to satellite-derived and ground-based observations over Australia during 2000-2011; (iv) it is parsimonious with minimum computational requirements and high interpretability. The optimal LAI model therefore allows us to capture essential aspects of the 'fertilisation effect' on future fuel hazard. Current and future LAI are shown in Figure S2.

Random forest models

We constructed three random forest models: one for each fuel stratum (elevated, near-surface and surface). Each model predicted a fuel hazard score of a stratum using 15 predictors including climate, soil, topography, and LAI as shown in Table S1. The field-based fuel hazard score data set was split into a training subset (random sample of 70% of observations) and an evaluation data subset (remaining 30% of observations). Due to the importance of elevated fuels for variation in fire intensity and rate of spread, we focused our evaluation and prediction on this stratum (Cheney et al., 2012). All of the data process and analyses were conducted in R (4.1, R Core Team), with the 'randomforest' function from the 'randomforest' package used for model fitting.

Model performance and importance of inputs

We used observed accuracy and Fleiss's Kappa coefficient (Fleiss, 1971) to assess model performance. Briefly, observed accuracy captures the agreement between model predictions and observations. However, due to the imbalance in each score, accuracy could be driven by a single score that has high frequency in the data. The Kappa coefficient addresses this issue by balancing the frequency of score in the data (expected accuracy) with the observed accuracy. A Kappa coefficient over 0.4 is generally considered as good (McHugh, 2012). We also used 'no-information rate' (i.e., accuracy achievable by random sampling) as a baseline of model performance. The more model accuracy exceeds the 'no information accuracy rate' the better the model performance.

To assess the importance of each predictor, we used the Gini index (also known as "Gini importance" or "mean decrease impurity") defined as the total decrease in node impurity (the proportion of a sample in each node) averaged over all trees of the ensemble (Breiman, 2001). Since the Gini index is in favour of variables with high categorical frequency, we also reported the percent change of mean square error of each variable.

Future projections

We explored the applicability of the field data to the whole region (Note S1) and found that the sample sites covered most environmental variation of the region (Figure S3). We thus predicted future fuel hazard scores for each of the nine models and two RCPs, a total of 18 climate projections. The soil and topography layers were resampled to the resolution of the climate model outputs (3.6 km) to make the projection at state scale. A LAI layer was predicted for each climate projection and period. Note that we also predicted fuel hazard scores for 2000-2015 with climate projections to avoid potential discrepancies between SILO and climate models. The projected fuel hazard scores varied by month because of monthly climate predictors. Since past intense fires in this region occurred mostly around January, we only report the fuel hazard score for January. The future fuel projections targeted potential fuel with the assumption of the land covered by natural vegetation. The projections contained the probability of each fuel hazard score on an ordinal categorical scale from 1 (low) to 5 (high). We predicted the probability of a hazard score of 4 or 5 (P_{4_5}) for the elevated fuel stratum rather than the change in categories, since high scores in this stratum have shown to strongly affect fire behaviour (Hines et al., 2010).

Predictor impact assessment

Although the importance of each predictor was reported by the Gini index and increase of error, these metrics do not directly translate into the magnitude of changes in the probability of a particular future fuel hazard score. We therefore explored the impacts of two individual predictors by making projections with the target predictor from 2000-2015 instead of the projected predictor while holding all other predictors as projected ('manipulated run'). For clarity, the projections with the full setup are referred to as 'projections'. Comparing the 'projections' with 'manipulated run' allowed us to separate the contributions of each predictor.

The two individual predictors that were assessed were the rainfall seasonality and the optimal LAI. We assessed rainfall seasonality because it showed high variability in the future and ranked high in the importance list. The difference between 'projection' and 'manipulated run' is referred to as the 'rainfall seasonality effect'.

The optimal LAI was investigated because it represents the 'fertilisation effect' of rising C_a, a key innovation of this study. The difference between 'projections' and 'manipulated run' is referred to as the 'fertilisation effect', hereafter. The fractional contribution of the 'fertilisation effect' was calculated as the 'fertilisation effect' divided by the 'projection'.

High resolution projections

To explore model performance at fine spatial resolution, we focused on a topographically heterogeneous region of about 20 km^2near Dargo, Victoria (-37.33588, -37.38554,147.27566,147.32566). This mountainous region features strong heterogeneity in environmental conditions as indicated by the fine-scale patterns of the selected terrain attributes (Figure S4). Models relying on climate only would predict no variation in fuel hazard scores due to the resolution of climate data (3.6-5.0 km). Here, we used topographic and soil data at 90 m resolution to demonstrate the capability of the model to predict fuel hazard scores at this fine resolution. We used climate data from the ACCESS 1-0 climate projection only, which was deemed sufficient for this analysis. because the goal is to highlight the capability of the model rather than predicting future change.

Results

We first compared the predictions of the three random forest models, one for each fuel stratum, against the evaluation dataset. The models agreed well with the field-based observations in the evaluation data sets in terms of accuracy, no information rate and Kappa coefficient (Table S3). For all three fuel strata, the models achieved an accuracy over 0.65, which is much higher than the 'no information rate' (0.3-0.4). The models also had Kappa coefficients over 0.5, indicating good model fits for all five score values despite the imbalanced sample size. Overall, the evaluation suggested good and consistent model performance for all three fuel strata and all five fuel hazard scores.

The random forest models also helped identify the key predictors of fuel hazard score among the 15 layers of gridded input. The importance of each predictor, which is measured in terms of the contribution to the overall model accuracy, is shown in Figure 2. Climate predictors in general ranked higher than other predictors in all models. However, solar radiation for January consistently ranked high in all strata, suggesting the importance of slope gradient and aspect for long-term climate controls on vegetation productivity and fuel hazard. Rainfall seasonality is among the most important predictors for the hazard score of the elevated fuel stratum, indicating the temporal distribution of rainfall may be more important than the absolute total. Minimum RH is a climate factor that is important for all three models, but its importance is much smaller for predicting the hazard score of the elevated fuel stratum than for the other two fuel hazard scores. The importance of optimal LAI was intermediate. Plan and profile curvature as well as clay fraction in the topsoil and available volumetric water capacity are consistently ranked low in importance for predicting hazards cores of all strata. For hazard scores of surface and near surface fuel strata, T_{max} and PPT also ranked high in importance, indicating importance of short-term topoclimatic drivers in these two strata. It is notable that the importance based on prediction accuracy and Gini Index are different for all strata. For example, for the hazard score of the elevated fuel stratum, MAP is ranked high in importance according to the Gini Index but not when importance is measured by model accuracy.

After model evaluation, we used the fitted models to predict changes in fuel hazard scores for each fuel stratum under projected climate change and increasing C_a. The mean predicted change in the P_{4_5}, averaged across the nine climate models, showed a distinct spatial distribution (Figure 3). For mid 21^{st}century (2045-2060), the model predicted an increased P_{4_5} in the north and southeast of Victoria under both the RCP4.5

scenario (Figure 3a) and the RCP8.5 scenario (Figure 3 c, d) with more severe climate change resulting in larger changes. This pattern remained consistent later in the 21st century (Figure 2b). The same projection for hazard scores of surface and near-surface strata are shown in Figures S5 and S6. The changes in the P_{4_5} (-0.2, 0.2) in surface and near-surface strata were much smaller than that of elevated stratum.

We explored the model projection of P_{4_5} of elevated fuel stratum along a climate aridity gradient (Figure 4). For current conditions, the results show a clear decrease in the median probability of high fuel hazard score in elevated fuel stratum as aridity increases, reflecting the pattern seen in the input data (Figure 4a) during 2000-2015. For future conditions, the model predicts an increased P_{4_5} of elevated stratum in dry region (AI >3.5) but not in more mesic regions under RCP8.5 by 2085-2100 (Figure 4b).

The strong spatial pattern in predicted change of the probability of high hazard scores for the elevated fuel stratum encouraged a further investigation of the predictor of that spatial variation. We first assessed the impact of changes in rainfall seasonality. Comparing the 'manipulated run' (i.e., projections with rainfall seasonality from 2000-2015) to the 'projection' showed that the spatial variation in predictions is driven by rainfall seasonality (Figure 5). With the manipulated run, the model predicted no change in $P_{4_5}$of the elevated fuel stratum, with predictions forming a narrow distribution with a single maximum (blue bars in Figure 5a). With the actual projections, the model predicted a bimodal distribution with around half of the pixels having little change in probability while the other half with a clear increase by up to 0.4 in the P_{4_5} of the elevated fuel stratum (red bars in Figure 5; Figure 3d compared with Figure 3a).

We also assessed the impact of changes in CO_2, via the change in optimal LAI. The 'fertilisation effect', varied from -9% to 12% in the region under RCP8.5 by 2085-2100 (Figure 5b). Notably, the 'fertilisation effect' is larger in regions with current low LAI (Figure S4g).

We chose a mountainous region with complex topography (Dargo, Australia) to explore the potential of the model to describe fine scale (90 m) variation in fuel hazard (Figure 6). Dargo has large small-scale variations in topography (Figure S4). A model driven solely by coarse resolution predictors, such as gridded climate, would predict no variation in fuel hazard score within this region of ca. 4 km x 5 km. The random forest model with key terrain attributes as predictors, however, predicted strong variations in the probability of high hazard scores in the elevated fuel stratum. The P_{4_5} of the elevated fuel stratum in the valleys was relatively low compared to ridges (Figure 6 a,b). Notably, the predicted response of P_{4_5} of the elevated fuel stratum to climate change is not uniform throughout the landscape (Figure 6 c,d). Wet valleys (Dark green in Figure 6d) saw higher increases in probability of high fuel hazard scores in elevated fuel stratum during 2085-2100 under RCP8.5 (Figure 6c) compared to the rest of the landscape.

Discussion

With catastrophic fire events gaining global concern, realistic fire danger assessment tools are needed more than ever. Here we present a framework to predict fuel hazard at a fine spatial resolution that is directly useful for operational fire management. It is based on mechanism-informed random forest models that make use of field-based observations of fuel hazard, gridded soil and topographic attributes, long-term climate trends, as well as plant responses to changing environment. Our modelling approach provides an important step towards a mechanism-informed fire risk assessment system. The predictions could also be used in fire behaviour models as well as to evaluate other vegetation model projections.

Fire management decision support

The modelling framework presented here has several features that are useful at regional scale for fire managers. It can incorporate long-term field-based fuel hazard surveys. It can use a comprehensive selection of predictors including climate, terrain, soil and vegetation. It can produce output that is interpretable and useful to fire managers and at a resolution that is relevant for operational management (Penman et al., 2022). Despite the high resolution, it has relatively low computational demand for regional projections

(i.e., does not require high-performance computing). The comprehensiveness separated this approach from previous empirical models that only consider a subset of predictors (e.g., Pierce et al., 2012; Jenkins et al., 2017; McColl-Gausden et al., 2020) while the high resolution (90 m) is what current process-based modelling cannot provide (Rabin et al., 2017).

The random forest model predicted increased P_{4_5} in the future for areas currently under semiarid climates (Figure 4). This prediction is consistent with a previous analysis on historical burnt area and aridity (Kelley et al., 2019). Combining the finding with current knowledge on key limiting factors in different regions (Archibald et al., 2013; Bradstock et al., 2014; Boer et al., 2016), our predictions indicate changes in future fire regime as current semiarid regions showing increasing fuel load in elevated stratum. Specific fire management measures should be thus based on the actual fuel hazard and its change in all fuel strata (Figure 3; Figure S4 and S5).

Predictors of fuel hazard

Based on our analysis we recommend using a combination of climate, fine scale topographic attributes and plant response to climate change for fuel projections. We found CO_2 fertilisation to contribute up to 12% of P_{4_5} in Victoria (Figure 4). Empirical models that do not include rainfall seasonality are unlikely to capture the actual change of fuel hazard score under future climate in Victoria where rainfall seasonality is a key driver of plant productivity (Figure 5) despite the optimal LAI ranked only in the middle in the importance list (Figure 2). Although this finding does not directly apply to other regions, the approach is generalisable and could help extract knowledge from field-based observations across the world. In contrast, land surface models running on coarse resolution (>5 km) generally cannot resolve terrain-driven variation in plant growth (e.g., Clark et al., 2016; Wu et al. 2022; but see Fiddes et al., 2022) and are unable to capture fine-scale variations in fuel hazard scores (Figure 6). Although spatial resolution might not be critical in capturing global trends in fire risk, fine scale predictions are crucial for operational fire management on regional scales (Bale et al., 1998; Nyman et al., 2015; Inbar et al., 2018).

Despite the good overall predictive performance, this machine learning approach has four major limitations: 1. It cannot provide information about the transient response of vegetation structural change due to gradual climate change; 2. It does not mechanistically model the impacts of plant composition and range shifts on fuel structure; 3. The ML models do not explicitly consider climate driven changes in fire regimes and the associated feedbacks with vegetation, which potentially affects vegetation distribution and fuel accumulation (e.g., Murphy and Bowman, 2012). 4. Human activities (e.g., land use change and fire hazard reduction efforts) are not included in the models but could result in substantial changes in future fire characteristics (e.g., Wu et al., 2022). Our approach aimed to quantify the potential for changes in fuel hazard as set by environmental and biological constrains. The shortcomings of this study could be addressed by process-based models which require significant developments in the computational capacity and the understanding of climate-vegetation-fire interactions.

Providing a benchmark for process-based models

Although many process-based models have a fire component, the evaluation of those models focused on the carbon and water components (Luo et al., 2012; Hantson et al., 2020). Existing evaluations of the fire modules relied on satellite derived burned area (e.g., Arora and Boer 2005; Hantson et al., 2020). Our predicted fuel hazard represented the potential to use machine learning approaches to upscale field-based observations to the resolution of land surface models and thus could be used as a reference for benchmarking process-based models alternative to satellite-derived burned areas.

Current process-based models run on coarse horizontal grids (Eyring et al., 2016; Friedlingstein et al., 2022) which cannot capture the fine scale variation of fuel shown in field observations. Although fine spatial resolution simulations accounting for fuel variation are possible at regional scales (Fiddes et al., 2022), the computational demand prevents such implementation at the regional and global scales. Our machine learning approach could be used in hybrid modelling framework to improve the model behaviour at fine spatial resolution (Reichstein et al., 2019).

Conclusion

Our random forest models predicted that the responses of future fuel hazard to climate change depend on climate aridity as well as local topographic attributes. We reported possible fuel hazard shifts because of changing climate and C_a. These findings highlight the fact that fuel hazard patterns are the product of the interaction among climate, vegetation, and topography. Predictions based on a subset of these factors are thus unlikely to be reliable. Our framework provides a useful decision support tool for fire risk management as well as a reference for evaluating process-based model predictions.

References

Abram, N. J. et al. 2021. Connections of climate change and variability to large and extreme forest fires in southeast Australia. - Commun Earth Environ 2: 8. Ainsworth, E. A. And Rogers, A. 2007. The response of photosynthesis and stomatal conductance to rising [CO2]: mechanisms and environmental interactions. - Plant, Cell & Environment 30: 258–270. Archibald, S. et al. 2013. Defining pyromes and global syndromes of fire regimes. - Proceedings of the National Academy of Sciences of the United States of America 110: 6442–6447. Arora, V. K. and Boer, G. J. 2005. Fire as an interactive component of dynamic vegetation models. - Journal of Geophysical Research: Biogeosciences in press. Bale, C. L., Williams, J. B., & Charley, J. L. (1998). The impact of aspect on forest structure and floristics in some Eastern Australian sites. Forest Ecology and Management, 110(1–3), 363–377. https://doi.org/0.1016/S0378-1127(98)00300-4 Beigaitė, R. et al. 2022. Identifying climate thresholds for dominant natural vegetation types at the global scale using machine learning: Average climate versus extremes. - Global Change Biology: 1–23. Boer, M. M. et al. 2016. Future changes in climatic water balance determine potential for transformational shifts in Australian fire regimes. - Environmental Research Letters in press. Boer, M. M. et al. 2017. Changing Weather Extremes Call for Early Warning of Potential for Catastrophic Fire. - Earth's Future 5: 1196–1202. Bowman, D. M. J. S. et al. 2020. Vegetation fires in the Anthropocene. - Nature Reviews Earth and Environment 1: 500–515. Bradstock, R. A. 2010. A biogeographic model of fire regimes in Australia: Current and future implications. - Global Ecology and Biogeography 19: 145–158. Bradstock, R. et al. 2014. Divergent responses of fire to recent warming and drying across south-eastern Australia. - Global Change Biology 20: 1412–1428. Breiman, L. 2001. Random Forests. - Machine Learning 45: 5–32. Burrows, N. D. 1994. Experimental development of a fire management model for Jarrah (Eucalyptus Marginata Donn ex Sm.) forest. Cheney, N. P. et al. 2012. Predicting fire behaviour in dry eucalypt forest in southern Australia. - Forest Ecology and Management 280: 120–131. Clark, S. et al. 2021a. Downscaled GCM climate projections of fire weather over Victoria, Australia. Part 2: a multi-model ensemble of 21st century trends. - Int. J. Wildland Fire 30: 596. Clark, S. et al. 2021b. Downscaled GCM climate projections of fire weather over Victoria, Australia. Part 1: evaluation of the MACA technique. - Int. J. Wildland Fire 30: 585. Clarke, H. et al. 2016. An investigation of future fuel load and fire weather in Australia. - Climatic Change 139: 591–605. Duane, A. et al. 2021. Towards a comprehensive look at global drivers of novel extreme wildfire events. - Climatic Change 165: 1–21. Emmett, K. D. et al. 2021. Adapting a dynamic vegetation model for regional biomass, plant biogeography, and fire modeling in the Greater Yellowstone Ecosystem: Evaluating LPJ-GUESS-LMfireCF. - Ecological Modelling 440: 109417. Eyring, V. et al. 2016. Overview of the Coupled Model Intercomparison Project Phase 6 (CMIP6) experimental design and organization. - Geoscientific Model Development 9: 1937–1958. Fleiss, J. L. 1971. Measuring nominal scale agreement among many raters. - Psychological Bulletin 76: 378–382. Fox, B. et al. 1979. Litter Accumulation After Fire in a Eucalypt Forest. - Aust. J. Bot. 27: 157–165. Friedlingstein, P. et al. 2022. Global Carbon Budget 2021. - Earth System Science Data 14: 1917–2005. Furlaud, J. M. et al. 2021. Bioclimatic drivers of fire severity across the Australian geographical range of giant Eucalyptus forests (G Battipaglia, Ed.). - Journal of Ecology 109: 2514–2536. Gao, X. et al. 2021. Emulation of community land model version 5 (Clm5) to quantify sensitivity of soil moisture to uncertain parameters. - Journal of Hydrometeorology 22: 259–278. Geoscience Australia, 2015. Digital Elevation Model (DEM) of Australia derived from LiDAR 5 Metre Grid.

Geoscience Australia, Canberra. Hantson, S. et al. 2020. Quantitative assessment of fire and vegetation properties in historical simulations with fire-enabled vegetation models from the Fire Model Intercomparison Project. Harris, R. M. B., Beaumont, L. J., Vance, T. R., Tozer, C. R., Remenyi, T. A., Perkins-Kirkpatrick, S. E., Mitchell, P. J., Nicotra, A. B., McGregor, S., Andrew, N. R., Letnic, M., Kearney, M. R., Wernberg, T., Hutley, L. B., Chambers, L. E., Fletcher, M.-S., Keatley, M. R., Woodward, C. A., Williamson, G., ... Bowman, D. M. J. S. (2018). Biological responses to the press and pulse of climate trends and extreme events. *Nature Climate Change*, 8(7), 579–587. https://doi.org/10.1038/s41558-018-0187-9 Hines, F. et al. 2010. Overall fuel hazard assessment guide. – Fire and Adaptive Management, Victorian Government Dept of Sustainability and Environment. Inbar, A., Nyman, P., Rengers, F. K., Lane, P. N. J., & Sheridan, G. J. (2018). Climate Dictates Magnitude of Asymmetry in Soil Depth and Hillslope Gradient. Geophysical Research Letters, 45(13), 6514–6522. https://doi.org/10.1029/2018GL077629 Jain, P. et al. 2020. A review of machine learning applications in wildfire science and management. - Environmental Reviews 28: 478–505. Jeffrey, S. J., Carter, J. O., Moodie, K. B., & Beswick, A. R. (2001). Using spatial interpolation to construct a comprehensive archive of Australian climate data. *Environmental Modelling & Software*,*16*(4), 309–330. https://doi.org/10.1016/S1364-8152(01)00008-1 Jenkins, M. E. et al. 2020. Modelling Bushfire Fuel Hazard Using Biophysical Parameters. - Forests 11: 925. Jl, F. 1971. Measuring nominal scale agreement among many raters. - Psychological Bulletin 76: 378–382. Kelley, D. I. et al. 2019. How contemporary bioclimatic and human controls change global fire regimes. - Nature Climate Change 9: 690–696. Leite, R. V. et al. 2022. Large scale multi-layer fuel load characterization in tropical savanna using GEDI spaceborne lidar data. - Remote Sensing of Environment in press. Luo, Y. Q. et al. 2012. A framework for benchmarking land models. - Biogeosciences 9: 3857–3874. Malone, Brendan (2022): Soil and Landscape Grid National Soil Attribute Maps - Soil Colour (3" resolution) - Release 1. v1. CSIRO. Data Collection. https://doi.org/10.25919/h5g4-qm95 McColl-Gausden, S. C. et al. 2020. Climatic and edaphic gradients predict variation in wildland fuel hazard in south-eastern Australia. - Ecography 43: 443–455. McColl-Gausden, S. C. et al. 2022. The fuel–climate–fire conundrum: How will fire regimes change in temperate eucalypt forests under climate change? - Global Change Biology in press. McHugh, M. L. 2012. Interrater reliability: the kappa statistic. - Biochem Med (Zagreb) 22: 276–282. Murphy, B. P. and Bowman, D. M. J. S. 2012. What controls the distribution of tropical forest and savanna? - Ecology Letters 15: 748–758. Norby, R. J. et al. 2005. Forest response to elevated CO2 is conserved across a broad range of productivity. - Proceedings of the National Academy of Sciences 102: 18052–18056. Parmesan, C. and Yohe, G. 2003. A globally coherent fingerprint of climate change. - Nature 421: 37–42. Peet, G. B. 1971. Litter Accumulation in Jarrah and Karri Forests. - Australian Forestry 35: 258–262. Penman, T. D. et al. 2022. Improved accuracy of wildfire simulations using fuel hazard estimates based on environmental data. - Journal of Environmental Management 301: 113789. Rabin, S. S. et al. 2017. The Fire Modeling Intercomparison Project (FireMIP), phase 1: experimental and analytical protocols with detailed model descriptions. - Geosci. Model Dev. 10: 1175–1197. Reichstein, M. et al. 2019. Deep learning and process understanding for data-driven Earth system science. - Nature 566: 195–204. Tolhurst, K. G. et al. 2008. Phoenix: Development and Application of a Bushfire Risk Management Tool. - The Australian journal of emergency management 23: 47–54. Viscarra Rossel, R. A. et al. 2015. The Australian three-dimensional soil grid: Australia's contribution to the GlobalSoilMap project. - Soil Res. 53: 845. Walker, A. P. et al. 2019. Decadal biomass increment in early secondary succession woody ecosystems is increased by CO2 enrichment. - Nature Communications 10: 454. Watson, P. J. et al. 2012. A comparison of bushfire fuel hazard assessors and assessment methods in dry sclerophyll forest near Sydney, Australia. - Int. J. Wildland Fire 21: 755. Wu, C. et al. 2022. Reduced global fire activity due to human demography slows global warming by enhanced land carbon uptake. - Proceedings of the National Academy of Sciences in press. Yang, J. et al. 2018. Applying the Concept of Ecohydrological Equilibrium to Predict Steady State Leaf Area Index. - Journal of Advances in Modeling Earth Systems 10: 1740–1758. Zazali, H. H. et al. 2021. A critical review of fuel accumulation models used in Australian fire management. - International Journal of Wildland Fire 30: 42–56. Zhu, Z. et al. 2016. Greening of the Earth and its drivers. - Nature Climate Change: 1–6.

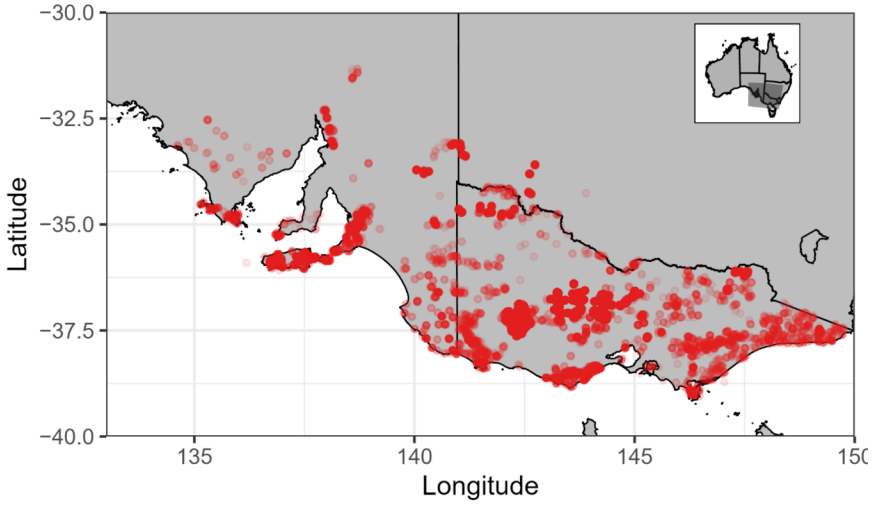

Figure 1. Locations of the field fuel observations in southeast Australia. Inset showing the sampling region relative to Australia. Darker colour means more samples in the pixel.

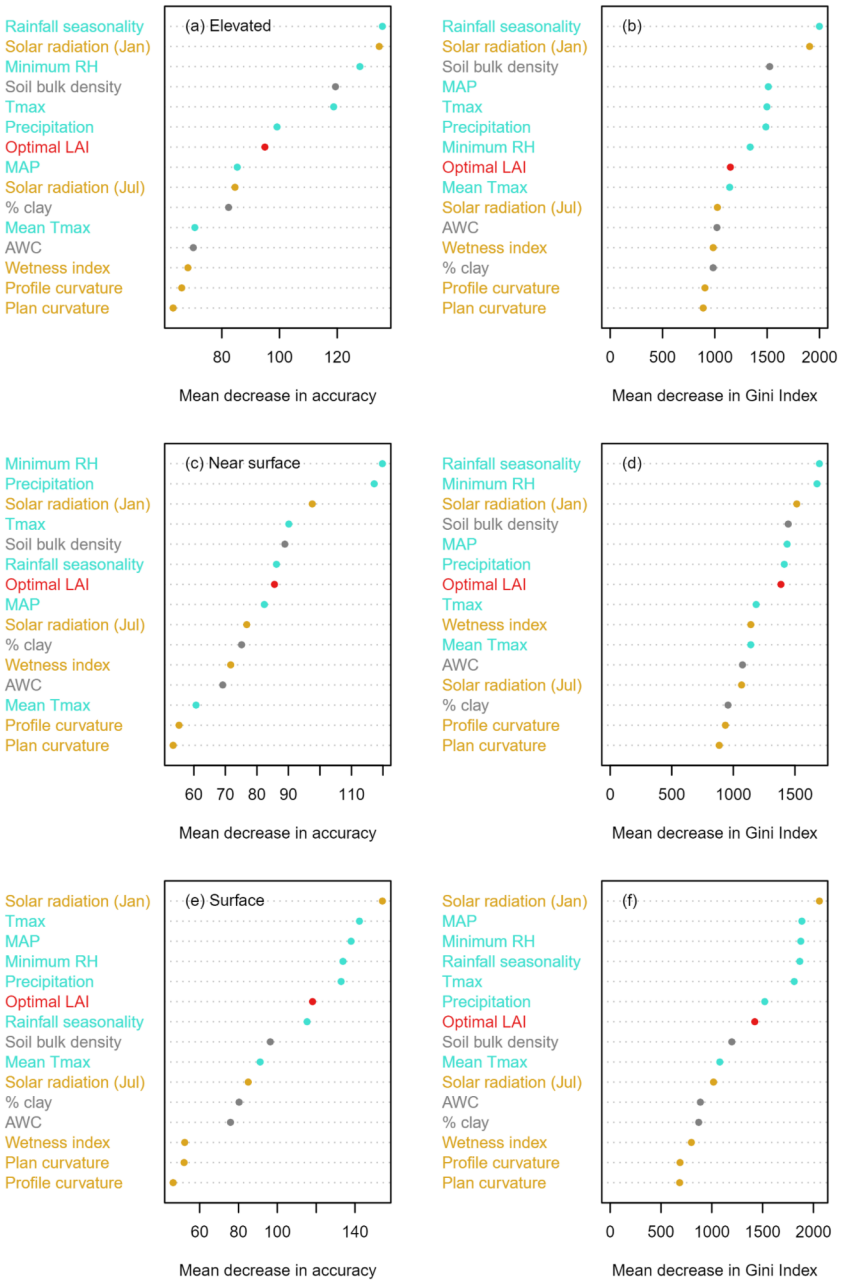

Figure 2. The importance ranking of predictors. The importance of each predictor in each model was evaluated based on the random forest model fits to training datasets. (a) and (b) show the importance based on decrease in accuracy and Gini Index for elevated fuel stratum, respectively. (c) and (d) show the same but for near-surface fuel stratum, while (e) and (f) for surface fuel stratum. Colour marks different predictor types: climate- cyan; topography-brown; soil- grey; plant-red.

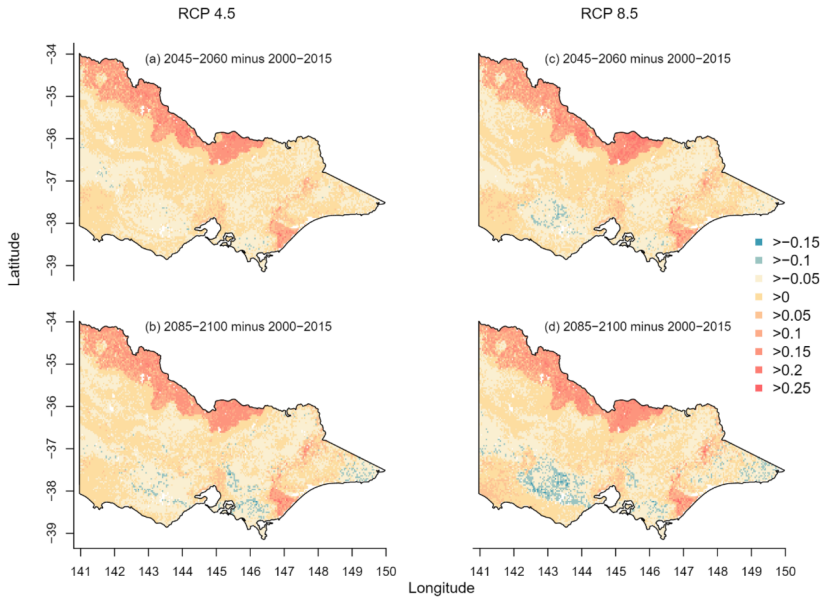

Figure 3. The projected change in P_{4_5} of elevated fuel stratum in the future in January. (a) shows the change by 2045-2060 compared to 2000-2015 under RCP4.5. (b) shows the change by 2085-2100 compared to 2000-2015 under RCP 4.5. (c) and (d) show the same as (a) and (b) but under RCP 8.5. Note that the projections ignored land use type (Figure 2).

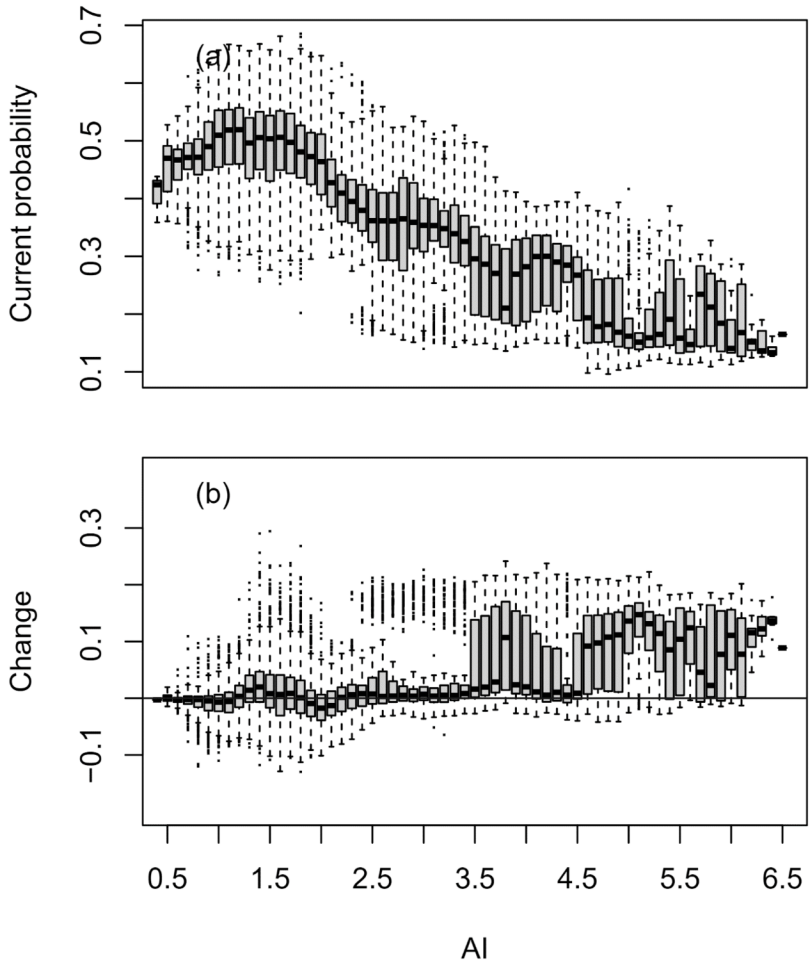

Figure 4. Predicted change in P_{4_5} for the elevated fuel stratum during 2000-2015 (a) and under RCP8.5 by 2085-2100 (b). X axis is the Aridity Index (AI) calculated as PET/MAP with long-term average SILO climate data. The changes are binned by 0.1 AI. The boxes are the predicted range of changes in probability considering both climate and LAI. The horizontal bars from top to bottom of each box show the maximum, 75% quantile, median, 25% quantile and minimum respectively. The black dots are outliers.**Figure 5.** The impacts of rainfall seasonality and CO_2 fertilisation on P_{45} of elevated stratum. (a) The change in the P_{45} of the elevated stratum using either constant or predicted rainfall seasonality under RCP8.5 by 2085-2100. The constant rainfall seasonality is the mean of all nine climate models during 2000-2015. The predicted rainfall seasonality is the mean of all nine climate models during 2085-2100 (Figure S1). The change of P_{45} is calculated as the difference of the probabilities between 2085-2100 and 2000-2015 with positive values mean increase in probability. (b) shows the fractional contribution of 'fertilisation effect' to the predicted P_{45} of the elevated fuel stratum under RCP8.5 by 2085-2100. Negative values mean the 'fertilisation effect' is in opposite direction and of smaller magnitude than the 'climatic effect'.

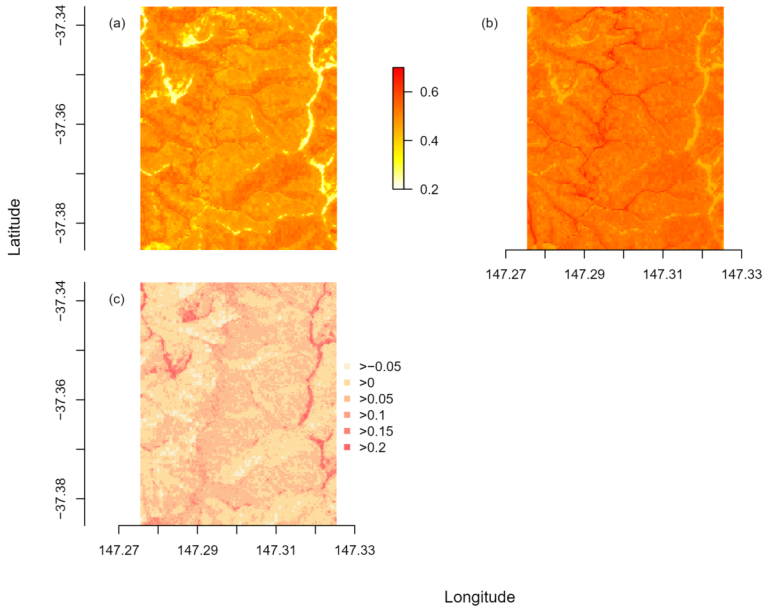

Figure 6. The projected change of probability of summer (January) high score of elevated fuel stratum in the future in Dargo region Victoria, Australia. (a) shows the probability of a high score of elevated fuel stratum during 2000-2015. (b) shows the probability during 2085-2100 under RCP 8.5. (c) shows the difference between (b) and (a).