Evaluating wildfire simulators using historical fire data

George J Milne, Joel K Kelso, Drew Mellor and Mary E Murphy

University of Western Australia. Perth, Australia. george.milne@uwa.edu.au

Abstract

High-performance wildfire simulators allow the future location of a wildfire to be rapidly predicted. The accuracy of such simulators needs to be evaluated; this can be achieved by comparing simulated and observed spread for documented historical fires. A key issue relates to the accuracy of data obtained from historical fires, such as the time-varying fire location, fire-ground weather and accuracy of fuel type, load and structure data. A methodology used to evaluate the accuracy of wildfire simulators using historical fire data is presented and applied to the AUSTRALIS wildfire simulator using the four distinct phases of a large-scale wildfire occurring in Western Australian sand-plain heathlands and a fire reconstruction report on this fire produced by a wildfire expert. Challenges encountered in performing this validation exercise are highlighted.

Keywords: wildfire simulation, GIS, fire behaviour models, simulator testing

1. Introduction

A methodology used to evaluate the accuracy of wildfire simulators using historical fire data is presented. Application of the methodology was examined using the four phases of a large-scale wildfire occurring in Western Australian sand-plain heathlands and a fire reconstruction report on this fire produced by a wildfire expert. The spatio-temporal dynamics estimated from the reconstruction report was compared with simulated fire behaviour, as produced by the AUSTRALIS wildfire simulator. The availability of rapid automated fire prediction permits the many variables which influence fire spread to be quickly examined by changing simulator input parameters, such as forecast wind speed and direction, to determine how such changes may impact on the spread characteristics of the fire. While simulators such as the AUSTRALIS wildfire simulator allow the future location of a wildfire to be rapidly predicted, and geographical information systems (GIS) maps with forecast fire-lines overlayed on them to be quickly made available to fire managers, the accuracy of such simulators needs to be examined by application to high-quality datasets from prior fires. A key issue relates to the accuracy of data obtained from historical fires, such as time-varying fire location, fire-ground weather and accuracy of fuel type, load and structure data, which are necessary if meaningful comparisons are to be made.

2. Methods

Simulating the spread of wildfire across a real landscape may, like simulation of other complex natural phenomena, be impacted by multiple sources of inaccuracy. First, the input data used for simulation will be subject to inaccuracy. For example, spatial boundaries in vegetation maps have limited precision and may have changed since the map was generated; initial fire perimeters are generally approximate; the closest meteorological observations may have been taken tens of kilometres from the fire-site. Second, predictive models relevant to fire behaviour, such as fire behaviour models for predicting rate of spread, slope correction, two-dimensional fire shape models, and fuel accumulation models, are all idealised models that approximate real phenomena. Third, the simulation methodology itself can introduce inaccuracy. For example, the discrete event simulation approach of AUSTRALIS (Johnston et al. 2008) relies on spatial discretisation, where the landscape is partitioned into cells that are assumed to have homogeneous attributes, such as vegetation, slope and aspect. When the spatial
resolution of the cell grid is coarse relative to the features being modelled, then the assumption of homogeneity is likely to be inaccurate for many cells. Given the need for accurate wildfire spread prediction these issues need to be overcome, and this provides the rationale for the reported study.

The general validation technique used in this paper is as follows:

- Obtain topographic, meteorological and fuel data for a historical fire event. Also obtain reconstructed fire spread perimeters and initial ignition/fire front locations.
- Simulate the fire using the obtained data and generate a progression of fire spread perimeters.
- Compare the level of agreement between the simulated and reconstructed fire progression perimeters.
- Assess the impact of uncertainty in the input data, fire behaviour models, and simulation algorithm on simulation accuracy by extensive sensitivity analyses.

In this paper, a case study of the above methodology is presented in which the AUSTRALIS wildfire simulator (Johnston et al. 2008) is applied to a large-scale historical fire that occurred in the vicinity of the Boorabbin National Park, Western Australia (WA) in December 2007 and January 2008.

This fire burned a total area of approximately 18,000 hectares over the 4 phases that were simulated. The topography consisted of gently undulating sand-plains and broad, shallow valleys. Two types of vegetation were present: Eucalypt woodland characterised by a very sparse understorey layer and a lack of fuel continuity, and semi-arid sand-plain heath (see Figure 1). Two government reports into the Boorabbin fire were produced as part of a coronial inquest following deaths which occurred on the fourth phase of the fire. These reports provided (i) a comprehensive assessment of the fuel and meteorological conditions occurring during the course of the fire (Bureau of Meteorology 2008; de Mar 2008), and (ii) a reconstruction of the fire perimeters over several phases of the fire at time steps ranging from 15 minutes to 3.5 hours (de Mar 2008). The existence of this detailed fire reconstruction data facilitated use of the following method.

Comparison between the simulator-generated, time-varying progression of the fire-line and an independently produced fire behaviour reconstruction produced as part of a coronial inquiry into deaths resulting from one of the four constituent fire phases was made. The accuracy of the simulated fire against the fire reconstruction contained in the report was determined using a number of measures, such as agreement of the headfire rate-of-spread (RoS) and the fit of the time-varying fire perimeters. The reconstruction report made use of high-resolution (less than 20m) satellite imagery to establish the final fire perimeter; however very limited data (i.e. firefighter recollection and limited aerial and ground photography) was available for the estimation of intermediate perimeters. Four distinct component fire phases were simulated, with the maximum Fire Danger Index ranging from 28 (high) to 104 (very extreme) across the four fire phases.
Simulated fire progression perimeters were generated using the AUSTRALIS simulator, using fuel type and coverage data from the reports, a national vegetation mapping data set (Australian Government Department of the Environment 2014), meteorological data from the nearest Bureau of Meteorology Automatic Weather Station (AWS) at Southern Cross, and a fire behaviour (rate-of-spread) model developed for arid heathland (Cruz et al. 2010). Simulator accuracy was assessed via a statistical comparison of the spatial extent of the simulated perimeters against the estimated perimeters taken from the reconstruction report, at corresponding time-steps.

AUSTRALIS employs a discrete event simulation technique (Zeigler et al. 2000) that is based on partitioning the landscape into a collection of two dimensional cells and calculating the propagation delay between an ‘ignited’ cell and each of its ‘unburnt’ neighbours. Each cell contains state information (‘unburnt’ and ‘ignited’) and a number of attributes relevant for calculating propagation delay, including location, elevation, and fuel characteristics such as vegetation type and fuel load. In contrast to other cell-based approaches to wildfire simulation, the cell locations are distributed randomly, rather than regularly, across the landscape. This is done to avoid a form of fire shape distortion that results from using a regular partition, such as with a rectangular or square grid. For all simulations in this study, the average distance between cell centroids was 50 m. Other fire spread simulation systems that use a discrete cellular representation of the landscape include FSPro, part of the U.S. Forest Service WFDSS system (Finney 2002; Finney et al. 2011), PYROCART (Perry et al. 1999), and FireStation (Lopes et al. 2002).

The validation technique presented in this paper may be used to validate any fire spread simulation system that predicts fire-front time of arrival across the landscape. Systems that do not use an underlying cellular landscape structure, such as FARSITE (Finney and Ryan 1995), Prometheus (Tymstra et al. 2010), and SIROFire/Phoenix (Coleman and Sullivan 1996; Tolhurst et al. 2008), may be accommodated by first running fire spread simulations and then rasterising both the estimated and simulated fire arrival time maps to generate cells for the purpose of accuracy assessment (see below). We note that in order to conduct extensive sensitivity analyses of the type presented in this paper, the ability to run many (hundreds) of fire spread simulations rapidly is beneficial, and it may not be practical to apply this kind of sensitivity analysis to computationally intensive, physics-based simulation systems such as FIRETEC (Linn et al. 2002).

The accuracy of the AUSTRALIS simulator at predicting the fire spread progression of the Boorabbin fire was measured as follows. At the conclusion of a simulation, AUSTRALIS output the fire arrival time for each cell that ignited. The accuracy of the simulated fire spread was assessed by comparison to a...
detailed reconstruction of the estimated fire progression which had mapped intermediate fire perimeters throughout the course of the fire at time step intervals of at most 3.5 hours over the four phases of interest. Accuracy was determined using Cohen’s kappa coefficient \( (K) \), a statistical measure of agreement between two geo-spatial datasets, which has been used previously for assessing the accuracy of fire spread simulation (Arca et al. 2007). Cohen’s kappa is given by:

\[
K = \frac{N \sum_{i=1}^{k} x_{ii} - \sum_{i=1}^{k} (x_{i+} - x_{+i})}{N^2 - \sum_{i=1}^{k} (x_{i+} - x_{+i})}
\]

where \( x \) is the error matrix, i.e. \( x_{ij} \) is the number of simulation cells where the simulated and reconstructed fires arrive in time period \( i \) and \( j \) respectively; \( x_{i+} \) and \( x_{+i} \) are the marginal totals of row \( i \) and column \( i \) respectively, and \( N \) is the total number of samples. Kappa typically varies over \([0,1]\), where \( K = 0 \) indicates that agreement is due to chance alone, and \( K = 1 \) indicates perfect agreement.

3. Results

![Figure 2](image)

Figure 2. Final fire perimeters estimated by the reconstruction report (shaded) and simulated by AUSTRALIS (black line) at the end of each phase. The agreement statistic kappa is given for each phase, which takes into account agreement between intermediate estimated and simulated fire perimeters (not shown). Spread under-predictions in Phase 2 and 3a marked Y are due to vegetation mapping inaccuracies; spread over-prediction in Phase 3b (marked Z) is due to weather data inaccuracy.

Using the available data AUSTRALIS were able to approximately reproduce the observed fire behaviour in each of the four phases, as shown in Figure 2.

Following the initial simulations of the four phases, an extensive series of sensitivity analysis experiments were conducted in order to determine the factors which limited the accuracy of the simulation system. The simulation input parameters that were varied and the range of values used are given in Table 1. In the case of the weather time series, adjustments were made by adding or subtracting a fixed amount from the value of the variable at each time step. For vegetation cover, all landscape cells containing heath vegetation were set to an alternative percent cover score (PCS) value. For fire behaviour models, each sensitivity analysis simulation used an alternative fire behaviour model to calculate rate of spread occurring in heath vegetation.
Table 1. Simulation parameters varied in sensitivity analysis simulations and the series of parameter values examined for each. Abbreviations are as follows. AWS – automatic weather station; U10 – 10 m wind speed recorded at the Southern Cross AWS; WD – wind direction (degrees clockwise from North); WS – wind speed (kilometres per hour); T – temperature (degrees Celsius); RH – relative humidity (percentage); PCS – percentage cover score; HE – semi-arid heath model (Cruz et al. 2010); MH1 – mallee heath (McCaw 1997); MH2, MH3 – semi-arid mallee heath (Cruz et al. 2010); SH – shrubland (Catchpole et al. 1998); HG - (Burrows et al. 2009).

<table>
<thead>
<tr>
<th>Simulation parameters and models</th>
<th>Baseline value/model</th>
<th>Sensitivity analysis range</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Cell grid</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cell size (m)</td>
<td>50</td>
<td>50, 100, 250, 500, 750</td>
</tr>
<tr>
<td><strong>Meteorological variables</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Wind speed measured at 10 m (km/h)</td>
<td>Southern Cross AWS (U10)</td>
<td>U10 ± 5, U10 ± 10, U10 ± 15, U10 ± 20</td>
</tr>
<tr>
<td>Wind direction (°)</td>
<td>As above (WD)</td>
<td>WD ± 5, WD ± 10, WD ± 15</td>
</tr>
<tr>
<td>Temperature (°C)</td>
<td>As above (T)</td>
<td>T ± 5, T ± 10, T ± 15</td>
</tr>
<tr>
<td>Relative Humidity (%)</td>
<td>As above (RH)</td>
<td>RH ± 5, RH ± 10, RH ± 15</td>
</tr>
<tr>
<td><strong>Fuel variables</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PCS of elevated fuel layer (0–4)</td>
<td>1.5</td>
<td>1, 1.5, 2, 2.5, 3, 3.5, 4</td>
</tr>
<tr>
<td><strong>Fire behaviour models</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fire behaviour model for heath vegetation</td>
<td>HE</td>
<td>HE, MH1, MH2, MH3, SH, HG</td>
</tr>
</tbody>
</table>

In the following subsections we summarise the results of these sensitivity analyses. We describe the outcome of the analyses in two parts: the effect of simulation algorithm cell size, and results showing a dichotomy between the first two phases of the fire and the second two phases, which occurred under extreme fire weather conditions.

3.1. **Impact of the cell grid on simulation accuracy**

As described previously, the AUSTRALIS simulator discretises the landscape into a set of randomly placed cells. This technique potentially introduces two sources of inaccuracy into the simulation results. Firstly, as with any simulation algorithm based on a cellular landscape discretisation, if cells are too large, it will not be possible for any set of simulation-generated cells to represent realistic fire shapes without either under- or over-fitting. Secondly, since the cell locations used by AUSTRALIS are randomly generated, simulation outputs can vary from one simulation run to another, even with identical simulation inputs. Both of these potential sources of inaccuracy decrease with decreasing cell size, as illustrated in Figure 3. In Figure 3A it can be seen that with a large cell size, the simulation of an ideal elliptical fire spread shape is considerably distorted, whereas in Figure 3B the elliptical fire shape is closely approximated. Comparing Figure 3B with 3C, it can be seen that simulations using two different randomly chosen grids with the same small cell size give rise to very similar fire shapes.
In order to quantify these effects of cell size in the context the Boorabbin simulation study, each of the four phases of the fire were simulated using cell grids generated with cell spacing ranging between 50–750 m, with five random cell grids being generated for each cell size. Two important results were apparent. Firstly, accuracy (i.e. agreement between estimated and simulated fire spread) increased as cell size decreased from 750 to 250 m, and remained constant below 250 m, indicating that below this cell size inaccuracy in simulation output was not due to discretisation error but was due to other sources (as discussed below). Secondly, the variance in simulation outputs due to the random nature of the cell grid diminished with cell size. For the 50m cell size used in the rest of the study, the results indicated that random grid variance artefacts do not significantly influence our results (95% confidence intervals for Kappa values were +/- 0.016 or smaller).

### 3.2. Impact of extreme fire weather on simulation accuracy

In this section we describe a clear distinction in the simulation accuracy sensitivity analysis results between the first two phases of the fire, during which the fire danger index (McArthur 1966) ranged from 20 – 28 (High), and the third and fourth phases, during which the fire danger index ranged from 47 to 104 (Extreme). For each of 4 key simulation parameters, Table 2 shows which parameter value resulted in the most accurate simulation i.e. the highest kappa value \( k \). If this value was not the baseline parameter value (baseline values are given in Table 1), the variation from the baseline is given along with the improvement in kappa over the baseline.

Table 2. Summarised results of sensitivity analysis simulations, showing the best accuracy value kappa (k) achieved by varying each simulation parameter. Where the best \( k \) value is greater than that of the baseline simulation, the baseline \( k \) value is given in parenthesis for comparison. The simulation parameter value that gave the best accuracy is also shown: where the notation “baseline” is used, this value was the baseline value shown in Table 1; otherwise, the deviation from the baseline value is given.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Phase 1</th>
<th>Phase 2</th>
<th>Phase 3A</th>
<th>Phase 3B</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>best ( k ) value for best ( k )</td>
<td>best ( k ) value for best ( k )</td>
<td>best ( k ) value for best ( k )</td>
<td>best ( k ) value for best ( k )</td>
</tr>
<tr>
<td>wind speed</td>
<td>0.63 (0.62) – 5kph</td>
<td>0.49 baseline</td>
<td>0.44 (0.33) + 20km/h</td>
<td>0.58 (0.42) + 20km/h</td>
</tr>
<tr>
<td>wind direction</td>
<td>0.62 baseline</td>
<td>0.49 baseline</td>
<td>0.47 (0.33) – 15°</td>
<td>0.42 baseline</td>
</tr>
<tr>
<td>PCS</td>
<td>0.62 baseline</td>
<td>0.49 baseline</td>
<td>0.54 (0.33) + 3.5</td>
<td>0.58 (0.42) + 1.5</td>
</tr>
<tr>
<td>FBM</td>
<td>0.62 baseline</td>
<td>0.49 baseline</td>
<td>0.49 (0.33) MH1</td>
<td>0.56 (0.42) MH1</td>
</tr>
</tbody>
</table>
The sensitivity analysis simulations revealed that for Phases 1 and 2, the most accurate simulation results were given by the baseline parameter settings. In other words, the weather time series from the Southern Cross AWS, the estimate of vegetation cover (PCS) as having a value of 1.5, and the selection of the heathland fire behaviour model yielded the most accurate simulation results. In the case of the wind speed during Phase 1, a very slight improvement in simulation accuracy was given by reducing the Southern Cross AWS wind speed by 5 km/h. Note that this is not a claim that the baseline parameter values were “correct” – rather, these results show that simulation inaccuracies are not simply explained by errors in the weather data, vegetation coverage estimate, or selection of an inappropriate fire behaviour model. For example, one apparent source of inaccuracy was due to the resolution of the vegetation map. Inspection of the Landsat imagery of the area showed that some areas of the map marked as Eucalypt woodland were interspersed with heath and carried fire; simulation in these areas under-predicted the extent of fire-spread (see areas marked Y in Figure 2).

The situation for Phases 3A and 3B is quite different. In terms of wind direction, in Phase 3A the simulation with the greatest accuracy occurred for a wind direction series in which all wind readings were rotated 15° counter-clockwise from the (northerly) winds recorded at the Southern Cross AWS. This result is consistent with de Mar’s (2008) analysis that transient westerly winds occurred during this phase at the fire-site and impacted the shape of the perimeters. In Phase 3B, the wind direction was clearly inaccurate at the beginning of the period (see Figure 4, up to 2100), resulting in over-prediction of westerly spread in the area marked ‘Z’ in Figures 2 and 4. Unlike Phase 3A, no simple uniform alteration of wind direction improved accuracy.

Table 2 shows that simulation accuracy is improved if the wind speed during the Phases 3A and 3B were 20km/h faster than the value recorded at Southern Cross AWS. Simulation accuracy was also improved if it was assumed that the vegetation cover of the heath vegetation burned during Phases 3A and 3B was higher than estimated, or if a Mallee Heath (MH1) fire behaviour model was used for heath vegetation. These three sensitivity analysis results are actually manifestations of a single factor: that the baseline simulation systematically under-predicted the rate of spread during Phase 3B. This is illustrated in Figure 4, which shows both estimated and simulated fire perimeters at 2030, 2045, 2100 and 2359. In each case the simulated fire front lags the fire front position estimated from the fire reconstruction report. The under-prediction of the rate of spread may thus be due to:

1. The actual wind speed being higher than the value recorded at Southern Cross. Given that the Southern Cross AWS was located 75km from the fire ground, and the clear discrepancies in wind direction noted above, discrepancies in wind speed are plausible. The HE fire behaviour model would then predict a faster rate of spread, more closely matching the estimated fire behaviour. Or,
2. The heath vegetation burned during Phase 3B having a higher vegetation cover than in the previous phases. Although the available vegetation input map data did not distinguish between heath scrub types, according to (de Mar 2008) the fire site contained both ‘heath-scrub’ and ‘Tamma scrub’. Of these two types of heath vegetation Tamma scrub, which was present in some areas burnt during Phase 3B, had characteristically higher levels of cover. With a higher vegetation cover, the HE fire behaviour model predicts a faster rate of spread. Or,
3. The HE fire behaviour model under-predicting the rate of spread in extreme fire weather conditions. This is plausible, given that the heathland experimental fires on which the HE model was based were conducted under less severe weather conditions: the maximum experimental fire danger rating was Very High compared to the Extreme conditions during Phase 3B, and maximum experimental wind speeds were approximately 18 km/h, compared to the average Phase 3B wind speed of 37 km/h. Or,
4. Some combination of the above.
Using the data available for our validation analysis, we were not able to distinguish between the possibilities 1-3 above.

4. Discussion

Several previous studies have sought to validate fire spread simulation systems by comparing simulated fire spread against historical fire data, for example the studies of Finney (Finney 2000), Fujioka (Fujioka 2002), Arca et al (Arca et al. 2007), and Fillipi et al (Filippi et al. 2014). In common with previous validation exercises, this study found that the ability to perform validation was limited by the reliability of available data, with fireground weather presenting the largest obstacle. While this study identified that the AUSTRALIS simulation system under-predicted the rate of spread in arid heathland vegetation under extreme fire conditions, it was not possible to further diagnose either wind data, fuel mapping data, or fire behaviour models as the cause of the under-prediction.

The very features which this study highlighted as presenting difficulties in conducting a validation of the accuracy of computer simulation of wildfire spread (using high quality data) are exactly those which impact on the use of such simulation technology “in the field”. As well as conducting simulation model validation using historical fire data, there is a pressing need to collect accurate fire data during active wildfires, rather than conducting analysis after the event. Such data-gathering efforts include regular fire line mapping at hourly intervals and the recording of fire ground weather conditions. Together, accumulated GIS data on fuel types, fuel load, and the development of fire behaviour rate-of-spread models which are experimentally calibrated for extreme fires, these data will facilitate: (1) higher fidelity simulator validation studies and, (2) more accurate prediction of “live” wildfires, which currently may be compromised by source data quality. Fire agencies and fire personnel organisations are to be encouraged to address these data issues.

5. Acknowledgements

The authors acknowledge funding for the study from the Australian Research Council, Fire and Emergency Services WA, and Landgate WA.

6. References


McArthur AG (1966) 'Weather and grassland fire behaviour.' (Forest Research Institute, Forestry and Timber Bureau of Australia: Canberra)


