Exploratory analysis of lightning-ignited wildfires in the Warren Region, Western Australia

Bryson C. Bates\textsuperscript{1}, Lachlan McCaw\textsuperscript{2}, Andrew J. Dowdy\textsuperscript{3,*}

\textsuperscript{1} School of Agriculture and Environment, University of Western Australia and CSIRO Oceans and Atmosphere, Perth, Western Australia, Australia
\textsuperscript{2} Department of Biodiversity, Conservation and Attractions, Manjimup, Western Australia, Australia
\textsuperscript{3} Bureau of Meteorology, Melbourne, Victoria, Australia

\textit{Accepted 28/7/2018}

\textit{Journal of Environmental Management}

Virtual special issue on “Fire in the Environment: Prevention, causes, impacts and management”

* Corresponding author. Bureau of Meteorology, GPO Box 1289, Melbourne, VIC 3001, Australia.

\textit{E-mail address:} Andrew.Dowdy@bom.gov.au
Abstract

An exploratory analysis of lightning-ignited wildfire data for the Warren Region of Western Australia was carried out for the period from April 1976 to December 2016. Temporal patterns in the series were examined in terms of characterizing the seasonal cycle, and detecting long-term trends and changes in seasonality over time. A generalized additive modelling approach was used to ensure that temporal features were determined by the data rather than a priori assumed mathematical forms (e.g. linear or low-order polynomial functions). The spatial organization of the data was evaluated using concepts from the theory of stochastic point processes. Results indicate a strong seasonality in the monthly lightning ignition series, the presence of a long-term trend and an interaction between trend and seasonality. There is also strong evidence of spatial variation in the number of ignitions per unit area in terms of location and distance from nearest ignition. Within the Warren Region, observation platforms for fire detection and reporting protocols have remained stable over the period of record, and changes in land use are unlikely to have altered the pattern of lightning ignition. Thus, the above results might reflect an interplay between: landscape attributes (e.g. vegetation classes, elevation, slope, aspect); changes in rainfall and fuel moisture; changes in fuel management practices; and, perhaps, an increase in the frequency of dry thunderstorms and fire weather conditions.

Keywords:

Lightning
Wildfire
Temporal trend analysis
Point pattern analysis
1. Introduction

Lightning is an important natural cause of wildfire ignition in temperate and high latitudes during summer months (see, e.g., Weber and Stocks, 1998; Dowdy and Mills, 2009; Price, 2013; Yang et al., 2015). In some regions of the world, wildfires ignited by lightning have been reported as typically burning a larger area (on average per fire) than fires ignited by other means, which has been attributed to lightning occurrence in remote locations and in large spatial and temporal clusters (see, e.g., Vazquez and Moreno, 1998; Wotton and Martell, 2005; Dowdy and Mills, 2012a; McCarthy et al., 2012).

Lightning-ignited wildfires in the northern hemisphere, particularly in Canada, the United States and Europe, have been examined in extensive detail in terms of: the influence of climate and weather (see, e.g., Price and Rind, 1994; Hess et al., 2001; Rorig and Ferguson, 2002; Krawchuk et al., 2006; Hall, 2007; Lutz et al., 2009; García-Otegi et al., 2011; Kraaij et al., 2013); spatial and temporal patterns (see, e.g., Vázquez and Moreno, 1998; Wierzchowski et al., 2002; Podur et al., 2003; Larjavaara et al., 2005a; Amatulli et al., 2007; Duncan et al., 2010; Wang and Anderson, 2010; Veraverbeke et al., 2017); and statistical modelling of lightning-fire occurrence (see, e.g., Wotton and Martell, 2005; Reineking et al., 2010; Magnussen and Taylor, 2012; Nieto et al., 2012; Vícen-Arias et al., 2016). In comparison, the literature focused on lightning-ignited wildfires in Australia is relatively sparse (see, e.g., McRae, 1992; Kilinc and Beringer, 2007; Dowdy and Mills, 2009, 2012ab; McCaw and Read, 2012). This is despite the fire-prone nature of Australia and its impact on public health and safety, ecosystems and infrastructure (Yelland et al., 2010; Bradstock et al., 2012; Productivity Commission, 2014).

The objective of this paper is to present and demonstrate an exploratory analysis approach that can be used to inform the development of fire-response and fire-suppression strategies in Australia and elsewhere. The approach is based on the use of generalized additive modelling.
(Wood, 2006) to detect and characterize temporal trends, and concepts from stochastic point
process theory to characterize the spatial pattern of lightning ignitions (Baddeley et al., 2015).
This is because quantitative analyses of changes in seasonality and temporal trends in
lightning ignition data have not received much attention to date, and spatial pattern
evaluations have not always been carried out with due consideration of the assumptions
underpinning the methods used (see, e.g., Hering et al., 2009). An Australian case study is
used to illustrate the application of the approach and demonstrate its usefulness.

2. Study area

The following description of the study area parallels that of McCaw and Read (2012), and
the text immediately below is derived from there with minor modification. The Warren
region (1.6 M ha) is in the southwest corner of the State of Western Australia (Fig. 1). The
southern and southwestern edges of the region define the extent of the coastal zone. The
region has a warm-summer Mediterranean climate (type Csb in the Koppen climate
classification system) with cool moist winters and an extended dry summer and autumn.
Annual rainfall varies from 1400 mm near the Southern Ocean to 700 mm in the northeast.
Winter rainfall across the southwest has been declining since the 1960s, particularly for the
late autumn to mid-winter months (May-July) (see, e.g., Bates et al., 2008; IOCI, 2012).
Historically, June and July have been the wettest months of the year.

Sixty percent of the Warren Region is public land managed by the Department of
Biodiversity, Conservation and Attractions (DBCA). Public land includes State forest, forest
plantations, national parks, and other conservation land tenures. Vegetation types include
(Christensen, 1992): open forests consisting of karri (Eucalyptus diversicolor), jarrah
(Eucalyptus marginata), red tingle (Eucalyptus jacksonii), wandoo (Eucalyptus wandoo) and
yellow tingle (Eucalyptus guilfoylei); and woodland and coastal shrubland. The region
consists of 26 recognized forest-ecosystem types, 19 of which have been impacted by lightning ignitions over the study period (Table 1).

Fig. 1. Location map for Warren region and spatial pattern of lightning-ignited wildfires. Coloured circles indicate the forest-ecosystem type associated with each ignition: JS = Jarrah South, KMB = Karri Main Belt, SHS = Shrubland, Herbland and Sedgelands and VCU = Vegetation Cleared or Unknown, and Other = the remaining types listed in Table 1. Black numerals indicate townships with populations of 300 or more in Australian Bureau of Statistics (2016): 1 = Manjimup (pop. 4349); 2 = Pemberton (pop. 974); 3 = Northcliffe (pop. 300); 4 = Walpole (pop. 439); 5 = Denmark (pop. 2637). The inset shows the State of Western Australia (which is bounded by latitudes 13° 44’ and 35° 08’ south and longitudes 113° 09’ and 129° east), and the
boundaries of the Department of Biodiversity, Conservation and Attractions (DBCA) regions. Boundary of the Warren region is coloured in black. [NB: colour should be used for this figure.]

Table 1
Details of forest-ecosystem types with one or more lightning ignitions during the period April 1976 to December 2016.

<table>
<thead>
<tr>
<th>Forest-ecosystem</th>
<th>Abbreviation</th>
<th>Area (ha)</th>
<th>Number of lightning-ignitions</th>
</tr>
</thead>
<tbody>
<tr>
<td>Darling Scarp Vegetation</td>
<td>DSV</td>
<td>527</td>
<td>1</td>
</tr>
<tr>
<td>Jarrah Blackwood Plateau</td>
<td>JBP</td>
<td>6976</td>
<td>1</td>
</tr>
<tr>
<td>Jarrah Mount Lindesay</td>
<td>JMT</td>
<td>26457</td>
<td>3</td>
</tr>
<tr>
<td>Jarrah Red Tingle</td>
<td>JRT</td>
<td>221</td>
<td>1</td>
</tr>
<tr>
<td>Jarrah South</td>
<td>JS</td>
<td>396194</td>
<td>169</td>
</tr>
<tr>
<td>Jarrah Unicup</td>
<td>JU</td>
<td>17720</td>
<td>8</td>
</tr>
<tr>
<td>Jarrah Woodland</td>
<td>JW</td>
<td>18246</td>
<td>8</td>
</tr>
<tr>
<td>Jarrah Yellow Tingle</td>
<td>JYT</td>
<td>8277</td>
<td>4</td>
</tr>
<tr>
<td>Karri Main Belt</td>
<td>KMB</td>
<td>152193</td>
<td>55</td>
</tr>
<tr>
<td>Karri Rates Tingle</td>
<td>KRaT</td>
<td>788</td>
<td>2</td>
</tr>
<tr>
<td>Karri Red Tingle</td>
<td>KRT</td>
<td>5209</td>
<td>6</td>
</tr>
<tr>
<td>Karri Yellow Tingle</td>
<td>KYT</td>
<td>11853</td>
<td>7</td>
</tr>
<tr>
<td>Peppermint and Coastal Heathland</td>
<td>PCH</td>
<td>45447</td>
<td>11</td>
</tr>
<tr>
<td>Rocky Outcrops</td>
<td>RO</td>
<td>2330</td>
<td>2</td>
</tr>
<tr>
<td>Shrubland, Herbland and Sedgelands</td>
<td>SHS</td>
<td>212854</td>
<td>30</td>
</tr>
<tr>
<td>Swamps</td>
<td>SWP</td>
<td>4954</td>
<td>3</td>
</tr>
<tr>
<td>Vegetation Cleared or Unknown*</td>
<td>VCU</td>
<td>503652</td>
<td>54</td>
</tr>
<tr>
<td>Western Wandoo Woodland</td>
<td>WWW</td>
<td>1047</td>
<td>2</td>
</tr>
<tr>
<td>Western Wandoo Forest</td>
<td>WWF</td>
<td>9202</td>
<td>4</td>
</tr>
</tbody>
</table>

* Includes areas not covered by DBCA’s forest-ecosystem data set.
Fig. 2 displays a topographic map of the Warren region at 1:250000 scale. The topographic data were retrieved from Geoscience Australia's GeoData Topo 250K Series 3 data product (https://researchdata.ands.org.au/geodata-topo-250k-file-format/1278865, accessed 6/20/2018). The terrain is undulating, with the highest elevations consisting of lateritic ridges in the northern part of the region and monadnocks of igneous rock in the south. Most of the public land is relatively remote and offers limited access to fire-fighting operations.

Fig. 2. Topographic map of the Warren region. Contour levels are marked in metres above mean sea level. Key to black numerals is given in caption for Fig. 1. [NB: colour should be used for this figure.]
Fig. 3 displays information about the lightning climatology of this region. Lightning ground flash data are presented, based on satellite observations, as described by Dowdy and Kuleshov (2014). The period of available data is from July 1995 to June 2013, representing 18 complete years spanning the austral summer. The number of flashes per square km is shown averaged for individual months of the year for these 18 years, as well as shown as the total number of flashes for the months from November to April over each individual summer. It shows that although lightning can occur at any time of the year, it mostly occurs over the warmer months of the (roughly from about November to April) year spanning the austral summer, but with considerable interannual variability.

**Fig. 3.** Lightning ground flash data, based on satellite observations. The number of flashes per square km is shown averaged for individual months of the year (upper panel) and for the months from November to April (lower panel), based on observations from July 1995 to June 2013.
3. Data

Lightning ignition data for the period from April 1976 to December 2016 were compiled from reports prepared by the DBCA and its predecessor agencies. The record consists of 371 observations from lookout towers, spotter aircraft, volunteer fire-fighting brigades, land owners, and the public. For each ignition, the data set specifies: date of occurrence; geographical position (easting and northing) with a spatial resolution of ±160 m; burnt area (ha) caused by the ignition; and the forest-ecosystem type involved. An auxiliary data file supplied by DBCA contained the coverage areas of the forest-ecosystem types (ha). There were zero lightning ignitions in the years 1982, 1985 and 1996 and, throughout the period of record, during the months of July, August, and September.

4. Methods

4.1. Temporal patterns

Results of preliminary analyses (sec. 5) indicated that the monthly lightning ignition series may exhibit a long-term trend and a change in seasonality. Consequently, the generalized additive modelling approach described by Bates et al. (2015) was applied to the monthly ignition series to unravel the individual and joint effects of seasonality and long-term trend. The following description parallels that found in Bates et al. (2015), and the text immediately below is derived from there with minor modifications.

Generalized additive models (GAMs) (Wood, 2006) were used to extract temporal patterns from the data. GAMs permit the seasonal cycle to be irregular and not perfectly harmonic, the long-term trend to be nonlinear and non-monotonic, and any two-dimensional dependency between seasonality and long-term trend to be judged without the imposition of a rigid functional form. In this framework, time \((month, date)\) is modelled as a smooth function
of seasonality \( (\text{month} = 1, \ldots, 12) \) and \( \text{date} = \text{year} + \text{month}/12 - 1/24 \) where \( \text{year} \) denotes the calendar year (1988, for example). The fraction 1/24 corresponds to half a month and ensures that calculated values of \( \text{date} \) correspond to the mid-points of the respective values of \( \text{month} \). This ensures that observations and model predictions are properly aligned in time series plots. For the monthly lightning ignition series \( Y = Y_1, \ldots, Y_n \), the GAMs considered herein specify a distribution for \( Y_i \) with mean \( \mu_i \), linked to one or more vectors of covariates via an equation of one of the following forms

\[
\begin{align*}
\text{(1)} & \quad g(\mu_i) = \beta_0 + f_1(\text{month}_i) \\
\text{(2)} & \quad g(\mu_i) = \beta_0 + f_1(\text{month}_i) + f_2(\text{date}_i) \\
\text{(3)} & \quad g(\mu_i) = \beta_0 + f_3(\text{month}_i, \text{date}_i)
\end{align*}
\]

where: \( \mu_i \equiv E(Y_i) \) in which \( E(\cdot) \) is the expectation operator; \( g(\cdot) \) is some monotonic function known as the link function; \( \beta_0 \) is the intercept term; \( f_1(\cdot) \) and \( f_2(\cdot) \) are centred smooth functions of the covariates (i.e. they are constrained to sum to zero over the data); and \( f_3(\cdot) \) is the contribution of the smooth to the modelled means on the scale specified by the link function. The smooth (potentially non-monotonic) functions are estimated using penalized maximum likelihood, where a penalty term is added to the usual log-likelihood criterion to avoid overfitting the data with functions that are too ‘wiggly’. The size of the penalty is controlled via a smoothing parameter that is chosen using leave-one-out cross-validation; larger values of the smoothing parameter produce smoother estimates. Several
diagnostics were used to check the fit of the GAMs; for details of these diagnostics, see Wood (2006).

Let Model 1, Model 2 and Model 3 denote the GAMs defined by Eqs. (1), (2) and (3), respectively. Model 1 presumes that the distribution of observed series is dependent on the seasonal cycle alone. Model 2 is the traditional long-term trend plus seasonality model where the dependence on the seasonal cycle is assumed to remain unchanged throughout the period of observation. Model 3 is a bivariate model that allows the seasonal cycle to change through time along with the overall mean. A formal comparison of Models 1 to 3 can be undertaken by noting that Model 1 is a special case of Model 2 in which \( f_2(\cdot) \) is set to zero, and Model 2 is a special case of Model 3. A ‘null’ model containing only an intercept can also be considered.

The above models were fitted using the “mgcv” package in the R programming environment (Wood, 2006; R Core Team, 2016). This package represents the smooth functions in the models using flexible collections of spline bases and reports the estimated degrees of freedom EDF (or effective number of parameters) as a measure of model complexity. Cyclic penalized cubic regression splines were used for \( f_1(\cdot) \) as they connect the beginning and end points of the seasonal cycle, penalized cubic regression splines for \( f_2(\cdot) \), and tensor product smooths for \( f_3(\cdot) \). The tensor product smooths use bivariate splines constructed from the individual basis functions for each variable, analogous to the technique described in Chandler (2005, sec. 4.3). Further details can be found in Wood (2006).

A formal comparison of the competing models was undertaken using an analysis of deviance. The calculation of \( p \) values for the analysis is based on large sample approximations that are most accurate for GAMs based on the normal distribution, with an identity link \( g(\mu) = E(Y) \) and a fixed smoothing parameter rather than one estimated from
The Warren lightning ignition series has a highly non-normal distribution. Thus, the log link function was used to ensure that the estimated means are all positive, and the GAMs were successively fitted in order of decreasing complexity with the amount of smoothing not allowed to decrease at each step.

4.2. Spatial patterns

The lightning ignition locations $x_i$, $i = 1, \ldots, n$ were considered as events in a planar region $A$, and several functions from the “spatstat” package of R (Baddeley et al., 2015) were used to characterize their spatial structure. A traditional starting point is to assume that the $n$ events are a realization of a homogeneous Poisson process which characterizes point patterns that exhibit complete spatial randomness (CSR). For fixed $n$, this means that each event is likely to occur randomly within $A$ and that the $n$ events are located independently in space (i.e. there is no interaction between them). This process defines the number of points in region $A$ to be Poisson distributed with constant intensity (mean number of events per unit area, $\lambda$):

$$E(\text{locations in } A) = \lambda |A|$$

where $|A|$ denotes the area of region $A$. Traditional tests of CSR are based on the derived distances defined by

$$s_{ij} = \|x_i - x_j\|, \quad i \neq j$$

$$t_i = \min_{j \neq i} s_{ij}, \quad i = 1, \ldots, n$$
\[ d(u) = \min_{i} \| u - x_i \| \]  

(5c)

where \( \| \| \) denotes the Euclidean distance, \( s_{ij} \) is the distance between neighbours \( x_i \) and \( x_j \).

Eqs. (5a), (5b) and (5c) are the pairwise, nearest neighbour and empty space distances, respectively and \( u \) denotes an arbitrary spatial location within \( A \).

It is considered good practice to examine several different summary functions, tests and diagnostics for a point pattern data set (Baddeley et al., 2015). Examples include: kernel density estimation; the empty space function \( F(r) \) where \( r \) denotes a distance of interest; the nearest neighbour distance distribution function \( G(r) \); the reduced second moment function \( K(r) \); the inhomogeneous \( K \) function \( K_i(r) \); Berman’s (1986) \( Z \) test; and Stienen’s (1982) diagram. In previous lightning-ignition studies, only limited subsets of these techniques have been used (see, e.g., Podur et al., 2003; Amatulli et al., 2007; Wang and Anderson, 2010).

Given that the boundary of the study region is somewhat irregular (Fig. 1), the lattice-based density estimator of Barry and McIntyre (2011) was used to produce an intensity map. For this method, estimation begins by overlaying the region with nodes, linking the nodes together in a lattice and then computing the density of random walks of length \( k \) on the lattice. The length of the random walk controls the smoothness of the intensity map. An approximation to the unbiased cross validation criterion is used to find the optimal value of \( k \). The resulting density and intensity functions differ by a constant of proportionality.

The \( F \) function is the empirical cumulative distribution function (CDF) of the shortest distance between events in a point pattern and a set of \( m \) randomly placed locations in the study region. Its raw estimator is defined by
\[ \hat{F}(r) = m^{-1} \sum_{j=1}^{m} I \left( d(u_j) \leq r \right) \] (6)

where \( I(\cdot) \) denotes the indicator function. The \( G \) function is the empirical CDF of the nearest neighbour distance between events. Its raw estimator is defined by

\[ \hat{G}(r) = n^{-1} \sum_{j=1}^{n} I \left( t_i \leq r \right) \] (7)

Estimation of the \( F \) and \( G \) functions is hindered by edge effects arising from the unobservability of events external to the study region. Edge corrections are routinely applied as raw estimates are negatively biased, and the magnitude of the bias increases with \( r \) (Baddeley et al., 2015, sec. 8.11). The \( K \) function is based on all inter-event distances for a point pattern. In the “spatstat” package, the estimator for the \( K \) function is defined by

\[ \hat{K}(r) = \frac{|A|}{n(n-1)} \sum_{i} \sum_{j} I \left( s_{ij} \leq r \right) e_{ij} \] (8)

where \( e_{ij} \) is a weight that corrects for edge effects.

Estimates of \( F(r) \), \( G(r) \) and \( K(r) \) are useful statistics for summarizing aspects of the spatial organization of the data. They are usually compared to theoretical values assuming CSR. The theoretical curves are defined by \( F_p(r) = 1 - \exp(-\lambda \pi r^2) \), \( G_p(r) = 1 - \exp(-\lambda \pi r^2) \) and \( K_p(r) = \pi r^2 \) where the subscript \( p \) indicates a homogeneous Poisson process. Deviations between the empirical and theoretical curves may suggest spatial regularity (inter-point
dependence) or spatial clustering (aggregation). The goodness-of-fit of the theoretical curves can be assessed using a Monte Carlo approach. Pointwise envelopes under the assumption of CSR can be obtained by sampling from a homogeneous Poisson process with intensity \( \hat{\lambda} \).

The envelopes can be used to check whether the empirical summary functions fall within the boundary of the envelopes or deviate from it.

If the point pattern does not follow the CSR assumption, interpretations of the \( F \), \( G \) and \( K \) functions are open to question (see, e.g., Hering et al., 2009; Baddeley et al., 2015). Indeed, perusal of Fig. 1 suggests that the point pattern may not be spatially homogeneous given the higher concentration of data points in the northern half of the region. Therefore, the inhomogeneous \( K \) function (\( K_i \)) described by Baddeley et al. (2015, sec. 7.10.2) was also used. The \( K_i \) function allows the intensity to have any form but assumes that the correlation structure is stationary. Each location \( x_i \) is weighted by \( 1/\hat{\lambda}(x_i) \) (i.e. the reciprocal of the intensity at \( x_i \)) and each location pair by \( 1/(\hat{\lambda}(x_i)\hat{\lambda}(x_j)) \). The estimated intensity at each data point must be nonzero. The theoretical curve for the inhomogeneous \( K \) function is defined by \( K_{i,p}(r) = \pi r^2 \), which is identical to that for \( K_p(r) \).

Berman’s (1986) \( Z_i \) test was used to assess the degree of spatial association between a point process and a continuous covariate (which in this study is the distance from the nearest lightning ignition point). The null hypothesis \( H_0 \) is that the pattern was generated by a homogeneous Poisson process model that is independent of the spatial covariate. The alternative hypothesis \( H_1 \) is that the pattern is an inhomogeneous Poisson point process with an intensity function depending on the spatial covariate. The test statistic is defined by \( Z_i = (S - \mu)/\sigma \) where \( S \) is the sum of the covariate values at the points in the pattern, \( \mu \) in this instance is the predicted mean value of \( S \) under \( n \) realizations of the null model and \( \sigma^2 \) is
the corresponding variance. The null distribution of the test statistic is approximately the standard normal distribution. The Berman diagnostic plot displays the CDF of the covariate at the data points and the predicted CDF of the model assumed under $H_0$. If the model is correct, the two CDFs and the mean values of the distributions (depicted by vertical lines in the plot) should be close. Another useful approach is to use kernel smoothing to fit the model defined by

$$\lambda(u) = \rho(Z(u))$$  \hfill (9)

where $\lambda(u)$ is the intensity at location $u$, $Z(u)$ is the value of the spatial covariate at $u$, and $\rho$ is an unknown function to be estimated. If $\rho$ is constant, it follows that $\lambda(u) = \lambda$.

Stienen’s (1982) diagram is useful for detecting inhomogeneity in the scale of the pattern (e.g., in the spacing between points). For each point in the diagram $x_i$, the distance $s_i$ to its nearest neighbour is computed, and a circle of radius $s_i/2$ is drawn around $x_i$. By construction, the resulting circles never overlap. Pairs of touching circles represent points that are mutual nearest neighbours, and filled circles represent observations that are not biased by edge effects ($s_i$ is less than the distance from $x_i$ to the boundary of $A$). Inhomogeneity is indicated by a trend in the radii of the Stienen circles.

5. Results

5.1. Preliminary analyses

A dot chart representation of monthly lightning ignitions is displayed in Fig. 4. In the figure, black dots indicate tied sets of monthly ignitions for two or more years. Non-black dots indicate monthly ignitions for a single year. There is a distinct seasonal cycle with peak
ignition activity (77 percent of total ignitions) in January to March (mid-summer to early autumn). This is when forest fuels are at their driest (McCaw and Hanstrum, 2003). Ignitions were detected as early as October (in 2006, 2009 and 2014) and as late as June (in 2006). The non-black dots for November represent 1, 2, 4, 5 and 7 ignitions in the years 2006, 2009, 2010, 2015 and 2013, respectively. Thirteen out of the fifteen non-black dots for January, February and March represent monthly ignitions that occurred post-1999 (i.e. in the second half of the record). The black dot for April represents ignitions in 1978, 1984, and 2007 to 2012, and the black dot for May represents ignitions in 2003, 2012 and 2015. Overall, these results suggest that the amplitude and length of the seasonal cycle have increased, particularly in the second half of the record.
Fig. 4. Seasonal cycle of lightning ignitions. Numerals indicate number of years with zero lightning ignitions in the period from April 1976 to December 2016. Black dots indicate tied sets of monthly ignitions for two or more years. Non-black dots indicate monthly ignitions for a single year. As depicted in the legend, the colours of these dots are graduated from green (1976) to red (2016). [NB: colour should be used for this figure.]

A barplot of the lightning ignition intensities for the four forest-ecosystem types that have coverages greater than 1000 km² is shown in Fig. 5. These types were selected on the bases that their coverage areas are of similar size, they account for 83 percent of all lightning ignitions, and that the use of other types with small coverage areas can lead to inflated intensity estimates. (For example, the intensity for the Jarrah Red Tingle type is 0.45
The intensities for the Jarrah South (JS) and Karri Main Belt (KMB) types are noticeably higher than those for Shrubland, Herbland and Sedgeland (SHS), and Vegetation Cleared or Unknown (VCU). Together, the JS and KMB types account for 60 percent of all lightning ignitions. This, combined with the high number of ignitions for other types consisting of jarrah, karri, red tingle, wandoo, yellow tingle and peppermint, suggests that a sizeable proportion of lightning ignitions (76 percent) originated from strikes on or near trees.
Barplot of lightning ignition intensities (number of ignitions per km\(^2\)) for the study period and forest-ecosystem types covering more than 1000 km\(^2\). Key to forest-ecosystem types is given in Table 1. Numerals indicate the corresponding coverage areas in km\(^2\).

5.2. Temporal patterns

Fig. 6a displays a time series plot of the lightning ignition data. The dispersion index \(D\), the ratio of the sample variance and sample mean) for the data is \(D = 8.23\) which means that the data are ‘overdispersed’ relative to the Poisson distribution \((D = 1)\). Furthermore, there is a distinct linear relationship between mean and variance (Fig. 6b). While this suggests that the use of quasi-Poisson regression can be justified, it does not preclude the use of a negative binomial model (which would be indicated by a quadratic relationship between the mean and variance). The latter was chosen since the quasi-Poisson model does not define a full distribution and thus it does not have a true likelihood. It is therefore unusable with likelihood-based smoothness selection and the standard likelihood ratio test for comparing nested models is not available. Use of the negative binomial distribution adds a complication in that it requires the estimation of a shape parameter that must be held fixed for the approximate \(p\) values to be valid. Thus, initial estimates of the shape parameter were obtained separately for Models 1, 2 and 3, and the models refitted using with the shape parameter fixed at the mean value (Bates et al., 2015).

Figs. 6c and 6d display the estimated smooth functions of date and month in Model 2. The plots include variability bands indicating the size of two standard deviations above and below the estimated functions. The long-term trend (which is linear on the logarithmic scale of the link function) and the seasonal cycle are particularly strong.
Fig. 6. Sample results from analysis of lightning ignition data for the Warren Region: (a) raw time series (dashed horizontal line depicts the mean), (b) variance-mean relationship by month (solid line is a linear regression fit), (c) contribution to the log link function \( g(\mu) = \log(\mu) \) by penalized cubic regression spline fit to lightning ignitions versus date data for Model (2) with variability band (cyan shading), and (d) contribution to the log link function \( g(\mu) = \log(\mu) \) by penalized cyclic cubic regression spline fit to lightning ignitions versus month data for Model (2) with variability band (cyan shading). [NB: colour should be used for this figure.]

Tables 2 and 3 summarize the results of GAM fits to the monthly lightning ignition series. The percentages of deviance (i.e. the percentages of the variability in the data) explained by the models ranges from 50 to 58% (Table 2), which is reasonably high given the variability of
the data (Fig. 6a). Perusal of Table 3 indicates that there is: overwhelming evidence against the null hypothesis of no seasonality (Model 1 versus the null model); overwhelming evidence against the null hypothesis that the series does not contain a temporal trend (Model 2 versus Model 1); and moderate to strong evidence against the null hypothesis that the seasonal cycle is constant over time (Model 3 versus Model 2). The last result indicates the presence of an interaction between date and month.

Table 2
Summary of generalized additive models for the monthly series of lightning ignitions in the Warren region.

<table>
<thead>
<tr>
<th>Model</th>
<th>EDF</th>
<th>Deviance</th>
<th>Explained (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>4.862</td>
<td>229.76</td>
<td>49.8</td>
</tr>
<tr>
<td>2</td>
<td>5.820</td>
<td>204.33</td>
<td>55.3</td>
</tr>
<tr>
<td>3</td>
<td>9.268</td>
<td>192.04</td>
<td>58.0</td>
</tr>
</tbody>
</table>

Table 3 Comparisons of fitted models for the monthly series of lightning ignitions in the Warren region.

<table>
<thead>
<tr>
<th>Analysis of deviance</th>
<th>Approximate p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model 1 versus null</td>
<td>&lt;2.2 × 10^{-16}</td>
</tr>
<tr>
<td>Model 2 versus Model 1</td>
<td>1.62 × 10^{-6}</td>
</tr>
<tr>
<td>Model 3 versus Model 2</td>
<td>0.0169</td>
</tr>
</tbody>
</table>

Fig. 7 displays a panel of plots of the logarithms of contributions to the modelled means $\mu$ for date = 1980, 1995 and 2010 obtained from Model 3; and the superimposed profiles for
these values of date. There is a marked change in the amplitude of the seasonal cycle with 

date, caused by increases in the number of ignitions for all months other than July to 

September. This is also reflected in the decreasing sizes of the variability bands for the 

months April to June, and October to November.

Fig. 7. Logarithms of contributions to modelled means $\mu$ for the years: a) 1980; b) 1995; c) 2010; and d) 1980, 


2010. [NB: colour should be used for this figure.]
A natural question is whether a simpler modelling approach would produce similar findings. Perusal of Fig. 6c and 6d suggests that the smooths of date and the seasonal cycle could be well represented by a straight line and very low-order harmonics, respectively. Generalized linear models provide a comprehensive framework for regression analysis with counted responses. Following the model building principles set out in Chandler and Scott (2011, Secs. 3.2 and 3.5), the model $g(\mu) = \beta_0 + \beta_1 date + \beta_2 \cos(2\pi \text{month}/12) + \beta_3 \sin(2\pi \text{month}/12) + \beta_4 date \times \sin(2\pi \text{month}/12)$ provided a reasonable fit to the lightning ignition data. The $p$-values for the regression coefficients $\beta_0, \ldots, \beta_4$ were $2.29 \times 10^{-8}, 3.29 \times 10^{-8}, <2 \times 10^{-16}, 0.019$ and $0.022$. Thus, there is moderately strong evidence against the null hypothesis $H_0: \beta_4 = 0$ and the product $date \times \sin(2\pi \text{month}/12)$ can be interpreted as indicating that the sine component of the seasonal cycle changes with time. This result is consistent with that obtained using GAMs.

5.3. Spatial patterns

The intensity map for the Warren lightning ignition data is shown in Fig. 8. The intensity varies by an order of magnitude across the region, suggesting an inhomogeneous point process. Comparison of Figs. 1, 2 and 8 indicates a well-defined ‘hotspot’ located to the east of the town of Manjimup that appears to be highly associated with the Jarrah South forest-ecosystem type and moderate to high elevation areas. Vegetation in the Jarrah South type consists mainly of dry sclerophyll forest and woodland. It is more prone to ignitions in the earlier months of the wildfire season than the karri forest (sec. 6).
Fig. 8. Contour map of the lattice-based estimator of intensity for lightning ignitions in the Warren region. Contour levels are marked in units of number of ignitions per km$^2$. Key to black numerals is given in caption for Fig. 1. [NB: colour should be used for this figure.]

The empirical and theoretical curves of the $F(r)$, $G(r)$ and $K_i(r)$ functions are displayed in Figs. 9a, 9b and 9c. The empirical curve of $F(r)$ dips below the pointwise envelope for theoretical values for $r>2$ km, suggesting that clustering is present over that range of distance. The empirical curve for $G(r)$ lies above the theoretical curves for almost the entire range of $r$ shown. Again, this suggests that the data are clustered. However, these interpretations assume CSR. Inspection of Fig. 9c suggests that the empirical and theoretical curves of $K_i(r)$ do not
deviate substantially from each other, with the deviation occurring at distances $r > 3.5$ km, approximately. The Berman diagnostic plot is shown in Fig. 9d. The mean and CDF of the distances to the nearest lightning ignition point are well separated from those predicted under the assumption of a homogeneous Poisson process that is independent of the distances. The value of the test statistic is -12.62 which has a $p$-value $< 2.2 \times 10^{-16}$. Thus, there is overwhelming evidence against the null hypothesis that the observed point pattern exhibits CSR and an indication that the assumption of an inhomogeneous Poisson process would be a good starting point for future modelling efforts.
Fig. 9. a) $F$ function; b) $G$ function; c) $K_i$ function, and d) Berman diagnostic plot. Red dashed lines are theoretical curves for complete spatial randomness in a) and b), and the theoretical curve for an inhomogeneous Poisson process in c). Cyan bands depict simulation envelopes and solid black lines depict empirical curves in a), b) and c). In d), red dashed curve is cumulative distribution function (CDF) of the covariate at the data points and black curve the predicted CDF of the model assumed under $H_0$. Mean values of the distributions are depicted by vertical red and black lines, respectively. [NB: colour should be used for this figure.]

Fig. 10 displays the kernel-smoothed estimate of the intensity function $\rho(Z(u))$ in Eq. (9). It is not constant (again suggesting that the data do not conform with the assumption of CSR), but a concave function with a pronounced peak when the distance from the nearest ignition is about 850 m. This most likely reflects the spike in intensity in the central-northern part of the region (Fig. 7).
Fig. 10. Kernel-smoothed estimate of the intensity of the point process of lightning ignitions, as a function of the spatial covariate $u$. Cyan band depicts a simulation envelope. [NB: colour should be used for this figure.]

The Stienen diagram for the lightning ignition data is shown in Fig. 11. There is little or no broad trend in the radii of the Stienen circles. However, there are several instances where closely-spaced circles appear to form thread-like structures. A possible explanation for these structures is the position and variability of storm tracks across the region. Perusal of the data revealed that there are 12 episodes of severe lightning-fire occurrence (i.e. consecutive days with more than seven lightning ignitions). A superimposed plot of the lightning ignition
locations for these episodes (not shown) revealed that they only provide a partial explanation for the structures. Thus, the thread-like structures might be regarded as preferred strike points built up over many storms, and that their locations might reflect landscape attributes such as: private-public land boundaries; vegetation type and height; and wooded ridges, riparian zones or road verges. This possibility will be investigated in future work.

**Fig. 11.** Stienen diagram for Warren lightning ignition data. For each data point, circle radius indicates half-distance to its nearest neighbour. By construction circles never overlap, and pairs of touching circles represent pairs of points that are mutual nearest neighbours. Filled circles represent data points that are not biased by edge effects. [NB: colour should be used for this figure.]
6. Discussion

Overall, the above analysis has revealed that the lightning ignition series for the Warren Region exhibits a pronounced long-term trend that interacts with seasonality over time. Thus, it follows that intensity of the point process of lightning ignitions (i.e., the number of ignitions per unit area) varies with time. Possible causative factors include changes in: observation platforms; detection efficiency; the areal coverage under surveillance; reporting practices; weather and climatic conditions; and fuel management.

According to McCaw and Read (2012): it is unlikely that the increased number of lightning ignitions since circa 2002 is due to improved detection or reporting practices as the resulting wildfires would increase to sizes where detection becomes inevitable; the proportion of land subject to surveillance has been relatively constant over the study period; and reporting protocols have been relatively consistent over the study period. Moreover, light aircraft and the same lookout towers have been used throughout the period of record. While most reports were for fires on public land, some relate to ignitions on adjoining private land. While there is the possibility that private landowners are more likely to report lightning-ignited wildfires on their land, the frequency of their reports is unlikely to have changed over time.

Nevertheless, the proportion of the landscape that has not been subjected to prescribed burns has increased over time (Boer et al., 2009; Burrows and McCaw, 2013), which could plausibly lead to an increased ignition frequency (e.g., due to increased fuel age and loads). Decreasing fuel and soil moisture in the pre- and early-fire season due to the well-documented winter rainfall decline in southwest Australia (sec. 2) could be a contributing factor, as could changes in the frequency of dry thunderstorms and fire weather conditions.
The analysis of the spatial organization of the lightning ignition data has shown that intensity varies across the region. It has also been revealed that the point pattern is consistent with an inhomogeneous Poisson process, but the exact nature of the process requires further investigation. As well as a possible interplay between forest-ecosystem type and elevation, a second possible explanation for the hotspot near Manjimup (Fig. 8) is that a surface pressure pattern known as the west coast trough anchors itself near Manjimup. During the warmer months (November to March), the trough is a semi-permanent feature that affects temperatures, winds and thunderstorm development in southwestern Australia. The trough develops at the boundary between warm-dry north-easterly winds driven by a strong high-pressure system in or south of the Great Australian Bight, and cool-moist winds from the Indian Ocean. Typically, each trough lasts from a few days to a week, before moving eastwards due to the approach of a cold front to the southwest of Australian continent (http://www.bom.gov.au/watl/about-weather-and-climate/australian-climate-influences.shtml?bookmark=westtrough, accessed 11/10/2017). Given that intensity also varies with time, future research efforts will involve the development and application of a spatial-temporal point process model. Further investigation is also needed to identify influential covariates such as fuel state, landscape attributes, and weather or climatic conditions.

Changes in lightning ignition activity have important implications for fire management in the Warren Region. As identified earlier, fires ignited by lightning have been shown to account for a dis-proportionately large percentage of area burnt because of their occurrence in large spatial and temporal clusters, and in remote areas not otherwise prone to ignition by human causes. Clusters of lightning ignition have caused unusually large fires in the Warren Region in February 2015 (98 000 ha), February 2012 (28 000 ha) and the period from
February to April 2003 (36 300 ha). The scale of these fires reflects the difficulty in suppressing multiple ignitions at a time of year when fuels are very dry and there is a high level of connectivity between different vegetation types that allows fires to spread extensively through the landscape. The development of fires ignited by lightning storms earlier in the season (November, December) may be constrained temporarily by damp fuels in tall karri and tingle forest, and by moist vegetation associated with creeks and swamps. Such areas lose effectiveness as barriers to fire spread as they dry progressively over the austral summer. Lightning strikes that ignite dead wood in standing trees and fallen logs on the forest floor can also result in fires that persist for many weeks, subsequently becoming active during periods of dry windy weather. Our findings showing increased level of lightning ignition and a lengthening of the period prone to ignition therefore suggest that since 2002 lightning has played a larger role in determining fire regimes for the Warren region that was the case in the earlier decades of the study period.

There does not appear to be a significant trend in lightning activity based on the data shown in Fig. 3, while noting the considerable degree of interannual variability and the relatively limited period of available data (from July 1995 to June 2013). Similarly, although there is some broad-scale spatial variation in the climatology of lightning data around this region with slightly more lightning activity to the northeast than the southwest in general (Dowdy and Kuleshov 2014), this variation is small in magnitude and not intended to explain the different spatial features in the fire ignitions within the Warren Region as presented in this study. It is also noted that there is a range of additional contributing factors that can influence the occurrence of lightning-ignited fires in this region, including variations in fuel state and in near-surface weather conditions.

The lightning data based on the satellite observations can be used to provide an estimate of the lightning ignition efficiency for fires in this region, based on the number of lightning-
ignited fires divided by the number of lightning ground flashes in this region. For the austral summer period (from November to April) during the 18 years of available lightning data, there were 225 fires in 18 summers (i.e., 12.5 fires on average) and 0.37 lightning ground flashes per square kilometre over the 1.6 M ha region (i.e., 5920 lightning ground flashes), providing an estimate of about 0.2% of lightning ground flashes resulting in a fire. It is difficult to directly compare values such as this with those from previous studies, given significant differences in data and methods, as well as noting a range of uncertainties around lightning observations (Kuleshov et al. 2009). However, a value of about 0.4% was reported for southeast Australia based on lightning stroke observations (noting that lightning flashes can sometimes contain multiple strokes) (Dowdy and Mills 2012a), with that being the only other previous estimate of this quantity for Australia. Values ranging from about 0.07% to 2% per lightning flash have been reported in different parts of Canada (Wiezchowski et al. 2002) and as low as 0.015% per lightning stroke in Finland (while noting that a significant proportion of lightning-ignited fires may be unreported in Finland) (Larjavaara et al. 2005a,b).

7. Conclusions

The temporal and spatial patterns exhibited by a 41-year record of lightning-ignited wildfires in the Warren region of Western Australia were investigated. A generalized additive modelling approach was used to detect and characterize seasonality, long-term trend and their interaction. The data were also considered as a point pattern in space, and a suite of summary functions, tests and diagnostics was used to characterize aspects of the spatial organization of the data. The main findings can be summarized as follows:
1. There are a pronounced seasonal cycle and a long-term trend in the lightning ignitions series. The presence of a trend indicates that the intensity of the point process (number of ignitions per unit area) varies with time.

2. There are marked changes in the seasonal cycle of lightning ignitions over time. The changes are expressed in terms of increases in the number of ignitions for the months October to June and a broadening of the lightning-fire season, particularly in the earlier part of the season.

3. There is evidence that a sizeable proportion of lightning ignitions (76 percent) originated from strikes on or near trees.

4. The point pattern of lightning ignitions does not exhibit complete spatial randomness. The intensity varies across the region, and the estimated intensity function (sec. 4) is concave with a pronounced peak when the distance from nearest ignition is about 850 m.

5. The point pattern is consistent with an inhomogeneous Poisson process, and this, coupled with the variation of intensity with time, will be used as a starting point in future modelling efforts. These efforts will include investigation of the dependence between intensity and covariates such as landscape attributes, fuel load, fuel and soil moisture, and weather and climate variables.

Acknowledgments

The lightning ignition data, boundary coordinates and areal coverages for forest-ecosystem types for the Warren Region were provided by the Department of Biodiversity, Conservation and Attractions, Western Australia. Financial support for AJD was provided by the Australian Government’s National Environmental Science Programme. The R computing environment (http://cran.r-project.org/) was used extensively throughout the study. We also thank two
anonymous referees and Managing Guest Editor, Paulo Pereira, for their thoughtful and constructive comments on the original manuscript.

References


http://dx.doi.org/10.1071/WF11167


Service Authorities Council and the Bushfire Cooperative Research Centre, Perth, 28–31 August 2012


