A Bayesian approach to exploring the influence of climate variability modes on fire weather conditions and lightning-ignited wildfires

Bryson C. Bates¹ · Andrew J. Dowdy² · Lachlan McCaw³

Received: 21 September 2020 / Accepted: 8 April 2021 / Published online: 25 April 2021
© Crown 2021

Abstract

Understanding the relationships between large-scale, low-frequency climate variability modes, fire weather conditions and lightning-ignited wildfires has implications for fire-weather prediction, fire management and conservation. This article proposes a Bayesian network framework for quantifying the influence of climate modes on fire weather conditions and occurrence of lightning-ignited wildfires. The main objectives are to describe and demonstrate a probabilistic framework for identifying and quantifying the joint and individual relationships that comprise the climate-wildfire system; gain insight into potential causal mechanisms and pathways; gauge the influence of climate modes on fire weather and lightning-ignition relative to that of local-scale conditions alone; assess the predictive skill of the network; and motivate the use of techniques that are intuitive, flexible and for which user-friendly software is freely available. A case study illustrates the application of the framework to a forested region in southwest Australia. Indices for six climate variability modes are considered along with two hazard variables (observed fire weather conditions and prescribed burn area), and a 41-year record of lightning-ignited wildfire counts. Using the case study data set, we demonstrate that the proposed framework: (1) is based on reasonable assumptions provided the joint density of the variables is converted to multivariate normal; (2) generates a parsimonious and interpretable network architecture; (3) identifies known or partially known relationships between the variables; (4) has potential to be used in a predictive setting for fire weather conditions; and (5) climate modes are more directly related to fire weather conditions than to lightning-ignition counts.

Keywords Climate modes · Lightning · Wildfire · Bayesian networks

1 Introduction

Protecting human life and health, economic infrastructure, natural ecosystems and resources from wildfire is a key challenge for fire management agencies. In many parts of the world there is ongoing interest in (concern about) wildfire activity, its impacts, social perceptions of risk, and risk communication efficacy (Gordon et al. 2013; Fairman et al. 2016; McGee et al. 2015; Zhang et al. 2018; Menezes

et al. 2019; Messier et al. 2019; Silva et al. 2019). Weather and climate strongly influence wildfire activity across space and time scales including the occurrence of lightning ignition (see, e.g., Jolly et al. 2015; Earl and Simmonds 2017; Holz et al. 2017; Abatzoglou et al. 2018; Nigro and Molinari 2019).

Our knowledge of the relationships between large-scale, low-frequency climate modes (e.g., the El Niño–Southern Oscillation), fire weather conditions and the occurrence of lightning-ignited wildfires comes mainly from empirical studies. There is a large and growing body of research on fire weather including climate variations for Australia and other regions of the world (Cai et al. 2009; Clarke et al. 2012; Bedia et al. 2015; Harris et al. 2014; Williamson et al. 2016; Dupire et al. 2017; Jain et al. 2017; Dowdy 2018; Harris and Lucas 2019), lightning-ignited wildfires (Dowdy and Mills 2012; Müller et al. 2013; Pineda et al. 2014; Abatzoglou et al. 2016; Bates et al. 2018) and their relationships with climate variability modes (Trouet et al. 2010; Úbeda and

✉ Andrew J. Dowdy
Andrew.Dowdy@bom.gov.au

¹ CSIRO Oceans and Atmosphere and School of Agriculture and Environment, University of Western Australia, Perth, WA, Australia

² Bureau of Meteorology, GPO Box 1289, Melbourne, VIC 3001, Australia

³ Department of Biodiversity, Conservation and Attractions, Manjimup, WA, Australia

Sarricolea 2016; Chen et al. 2017a; Mariani and Fletcher 2016; Mariani et al. 2016, 2018; Pan et al. 2018).

Perusal of the literature cited above (and references therein) suggests that assessments of the strength of relationships between climate modes and fire weather and lightning ignitions are often based on one or more of the following approaches: (1) graphical methods; (2) simple and partial correlation analyses, (3) conventional linear regression; and (4) time series data mining techniques such as superposed epoch analysis and wavelets. While useful for data exploration such methods have several limitations. A graph can only depict a few characteristics of the data and may not sufficiently represent key interrelationships between factors. This is the case for dangerous wildfires which are typically caused by a combination of different factors (i.e., a compound event). Correlation measures the extent to which a pair of variables is linearly related. It neither explains the causal relationship between the variable pair nor accounts for the effects of additional explanatory variables. Generally, regression-based approaches relate one outcome variable of interest to one or more explanatory variables. However, there may be situations that involve multiple outcome variables, multiple dependencies between explanatory variables and variables that are both an explanatory and an outcome variable (e.g. Sect. 4). Also, it can be a difficult process to build models in the absence of complete or partial knowledge of the processes that produce the observations. Any potential interactions between variables need to be specified a priori as an input feature. Moreover, predictions made outside the range of the original variables have the tendency to be poor. Superposed epoch analysis (also known as compositing) can be used to link discrete events to continuous time or space–time processes. For large sample sizes, it can separate small signals from noise. The limitations of the method are discussed by Rao et al. (2019 and references therein). They include: one or more events might have very high leverage on the mean response; the implicit assumption that all event signals are equal rather than unique, and the same kind of events might differ due to climate variability and the influence of antecedent background states. Wavelet analysis is a useful tool for revealing time–frequency localization information in time series (Torrance and Compo 1998). However, one of its disadvantages is that it is intended for examining the frequency spectrum over time for an individual variable rather than considering more complex interrelationships between multiple variables. It also requires user-specification of the shape and width of the wavelet function so that it reflects the features present in a time series. Once the wavelet function is selected, a set of scales must also be specified for use in the wavelet transform.

This study differs from previous research in four respects. First, it seeks to identify and characterize the joint as well as the individual contributions of six large-scale, low-frequency

climate variability modes to fire weather conditions and to lightning-ignited wildfires. Moreover, emphasis is placed on the physical basis of the relationships found. A better understanding of the interactions between these modes, forest fire danger (as indicated from weather conditions and fuel moisture measures), prescribed burn area and lightning ignition counts has the potential to inform fire-weather forecasting, fire management decisions and conservation strategies. Second, we use model-based machine learning and statistical approaches to investigate these interactions. The main technique used (dynamic Bayesian network analysis) provides a graphical probabilistic framework for modeling and reasoning under uncertainty, and a useful communication device for supporting decision-making processes that can be easily updated as new data and knowledge become available. Third, an extensive assessment is made of the assumptions underlying the various steps of the analysis. Remedial methods for violations of these assumptions are described. Fourth, rolling origin evaluation is used to assess the potential predictive skill of the proposed model for lead times of interest to wildfire managers. This skill was compared to that of an alternative prediction scheme that uses local-scale information only. This assessment provides a crucial baseline to which the effect of accounting for the behavior of climate variability modes on predictive skill can be compared.

Our main objectives are to: propose a probabilistic framework for identifying and quantifying the joint and individual relationships of key factors that comprise the climate-wild-fire system; consider a larger number of relevant climate modes than in previous studies using a multivariate framework; gain insight into potential causal pathways from climate variability modes to fire weather and to lightning ignition in a forested region subjected to prescribed burning; gauge the influence of climate modes on fire weather and lightning-ignitions relative to that of local-scale conditions; and motivate the use of techniques that are intuitive, flexible and for which user-friendly software is freely available. The remainder of this article is organized as follows: Sect. 2 gives a brief introduction to Bayesian networks. Section 3 describes the design, building and testing of the proposed framework. Section 4 presents a real case study where the relationships between climate mode indices, fire weather conditions and lightning-ignitions are identified, quantified, and tested. A discussion of our findings and an examination of their physical plausibility are presented in Sect. 5. Finally, Sect. 6 provides a summary and conclusions of the study.

2 Bayesian networks

Bayesian networks (BNs) are an increasingly popular tool for representing and reasoning with uncertain knowledge. They are based on well-established theoretical foundations

of probability calculus and most often used to represent the state of a system at an instant in time. There are other applications that focus on a network of interacting processes that drive the temporal evolution of the system state. There, quantification of any underlying structure requires modeling and analysis of multiple time series. BNs are also useful for investigating and constructing probable explanations for observed events. In what follows we provide some background on Bayesian networks for the sakes of completeness and convenience. We use uppercase letters to denote random variables (e.g., Y), lowercase letters for their fixed values (y), boldfaced uppercase letters for variable sets ($\mathbf{Y}$), and boldfaced lowercase letters for column vectors ($\mathbf{y}$).

2.1 Basic concepts and terminology

BNs provide a means for combining prior (background) expert knowledge of the underlying domain with observational data. They enable the compact representation and efficient computation of the joint probability distribution of a set of random variables based on a set of conditional independence assumptions. The graphical representation of a BN is a directed acyclic graph (DAG) in which each random variable corresponds to a node (vertex) and the direct influence of one node (the parent) on another (a child) corresponds to a directed arc (edge). Thus, arc directions represent conditional dependencies (directions of statistical influence). Unconnected nodes are conditionally independent of each other in the BN. Four classes of node types exist: *root* nodes have no parent nodes; *intermediate* nodes have parent and child nodes; *leaf* or *terminal* nodes are only connected to their parents; and *orphan* nodes are not connected to any other node in the network. Thus, a DAG exhibits a topological ordering. The acyclic requirement means that it is impossible to start at any node in the DAG, follow a sequence of directed arcs, and return to the same node.

The notion of conditional independence is central to BNs. Let $\mathbf{Y} = \{Y_1, Y_2, \dots, Y_p\}$ denote the ordered set of random variables under consideration. By using the chain rule in probability theory, the full joint probability distribution can be expressed as

$$\begin{aligned} P(Y_1, Y_2, \dots, Y_p) &= P(Y_1)P(Y_2|Y_1)P(Y_3|Y_1, Y_2) \dots P(Y_p|Y_1, \dots, Y_{p-1}) \\ &= \prod_{i=1}^p P(Y_i|Y_1, \dots, Y_{i-1}). \end{aligned} \quad (1)$$

BNs provide a decomposed representation of (1) by encoding a collection of conditional independence statements to yield a more compact factorization that reduces computational requirements and the number of parameters to be estimated:

$$P(Y_1, \dots, Y_p) = \prod_{i=1}^p P(Y_i|\text{par}(Y_i)), \quad (2)$$

where $\text{par}(Y_i) \subset \{Y_1, Y_2, \dots, Y_{i-1}\}$ is called the parent function. For completeness, we note that if all random variables are mutually independent the joint distribution is simply

$$P(Y_1, \dots, Y_p) = P(Y_1)P(Y_2)P(Y_3) \dots P(Y_p). \quad (3)$$

The way BNs are defined means that a network structure must be a DAG. Otherwise, the network parameters cannot be estimated because the factorization of the joint probability distribution (Eq. (2)) is not completely known. If a network structure contains one or more undirected arcs, their direction can be set by the analyst. However, we prefer the more objective approach of letting ‘the data speak for itself’ as the strength and direction of the interdependencies between the variables are not well known and may be time variant.

Multivariate data sets may be composed of random variables that are continuous (real-valued), discrete (integer-valued) or a mixture of both. It is commonly assumed that the random variables in a BN are discrete, as many algorithms for building network structures do not handle continuous variables efficiently. However, there are many applications that require the use of both discrete and continuous data (including the study herein, in which nine of the eleven random variables considered are continuous). There are common approaches for extending the use of these algorithms to continuous variables (Cobb et al. 2007; Chen et al. 2017a, b): (1) automatic discretization; (2) modeling the conditional distributions using nonparametric methods, and (3) modeling the conditional probability distributions of each continuous variable using specific families of parametric distributions. The discretization of continuous variables leads to a loss of information. The definition of the discretization breakpoints is a major challenge as the choice of discretization policy can have a significant impact on the accuracy, speed, and interpretability of the resulting models. Automatic discretization methods exist, but the basis for selection of one method over another is not clear and their use may result in a discretization inappropriate for the

model purpose and user needs (Lara-Estrada et al. 2018). For extreme events such as wildfires the upper tail of a distribution might be most important detail to represent, a task for which automatic methods may not be ideally suited. Moreover, while a nonparametric model can often fit any

underlying distribution the number of parameters grows with the amount of data used to fit the model. This can be computationally expensive/limiting and increases the potential for overfitting which leads to the fitting of noise in the data set. In contrast, parametric models have a fixed number of parameters.

2.2 Gaussian Bayesian networks

Gaussian BNs are widely used for modeling the behavior of continuous random variables. In contrast, the construction of non-Gaussian networks remains a topic of active research. Gaussian BNs assume that the joint probability density of $\mathbf{Y}$ is a multivariate Gaussian (normal) distribution $N_p(\boldsymbol{\mu}, \mathbf{V})$ where $\boldsymbol{\mu}$ is the p -dimensional mean vector and $\mathbf{V}$ is the $p \times p$ symmetric and positive definite covariance matrix (Appendix A.1). It is well known that the marginals and conditionals of a joint Gaussian distribution are also Gaussian. Thus, the joint density function can be factorized using the conditional probability densities of Y_i ($i = 1, \dots, p$) given its parents in the DAG. These are univariate Gaussian distributions with densities:

$$f(y_i | \text{par}(y_i)) = N\left(y_i | \mu_i + \sum_{j=1}^{i-1} b_{ji}(y_j - \mu_j), \nu_i\right), \quad (4)$$

where μ_i is the unconditional mean of Y_i , b_{ji} are the unknown regression coefficients of Y_i given $Y_j \in \text{par}(Y_i)$ and ν_i is the conditional variance of Y_i given its parents in the DAG. Thus, a Gaussian BN is equivalent to a set of multivariate linear regression equations, and the distribution of each node can be viewed as a linear regression. The number of parameters in each regression is equal to the number of the parents of the node plus one (the intercept term). Importantly, the information from factorizations such as (4) can be used for making predictions of the most probable values of any node (including the terminal node).

Gaussian BNs are computationally efficient because they offer exact algorithms for prediction and inference. However, they impose the restriction that discrete nodes cannot have continuous nodes as parents. To avoid this restriction, we use an appropriate transformation to ensure that the joint probability distribution is approximately Gaussian (Appendix A.2).

2.3 Dynamic Bayesian networks

Dynamic Bayesian networks (DBNs) are a useful tool for compactly describing the structure of statistical relationships and dependencies among multiple time series (Murphy 2002; Sucar 2015, Chap. 9). DBNs involve the discretization of the temporal range of interest into a finite number of slices (time points) and creating an instance of each random

variable for each slice. Each temporal slice has the same set of nodes, and the dependency structure between the nodes at that slice is defined (the base network). This represents the system's complete state at a single time point. The same set of arcs are placed between nodes from different slices to capture temporal relations, with their directions following the direction of time (the transition network). Thus, it is possible to accommodate possible loops and feedbacks in the network over time points (Nagarajan et al. 2013, Chap. 3). Common key assumptions are that the network structure and the probability distributions and associated parameters describing the temporal dependencies are time invariant; and that the system state follows a first-order Markov process. That is, the DBN is usually specified using only two consecutive time slices $t-1$ and t . This is because DBNs involving higher order and non-stationary processes pose computational challenges in *structure learning* (see below) and parameter estimation. It can also lead to a deterioration in the performance of model selection criteria (as was found to be the case here).

2.4 Structure learning

The first task in a BN analysis is to seek a DAG that best explains a given dataset. When enough prior knowledge exists the structure of the BN can be based on current process understanding. In cases where there is limited knowledge and little information available about the physical mechanisms that affect the system under examination, it is possible to infer an appropriate network structure directly from the data. This process is called structure learning. Learning the DAG from data alone is a challenging problem. The two primary methods for structure learning are constraint-based and score-based algorithms. Constraint-based algorithms start with a fully connected DAG and rely on conditional independence tests to learn the dependence structure of the data. A score-based search assigns goodness-of-fit scores to candidate networks (DAGs) and uses an optimization algorithm to maximize the fit. Scutari et al. (2019) suggest that constraint-based algorithms are often less accurate than the score-based algorithm tabu search (TS) for both small and large sample sizes. They also found that TS is often faster than its alternatives. Briefly, TS is an iterative procedure that can be superimposed on local search methods to enable a more thorough exploration of the solution space, escape from local optima and convergence towards the global or a near global optimum. From a feasible initial solution, the algorithm explores its neighborhood while using knowledge held in its *tabu list*. This list marks previously visited solutions as either temporarily or permanently forbidden. It improves the efficiency of the search and discourages the phenomenon called *cycling* which involves the reversal or repetition

of selected moves near a local optimum. The exploration continues until a stopping criterion is met. Further details can be found in Fouskakis and Draper (2002) and Gendreau and Potvin (2019).

In many situations there are known structural constraints which if ignored can result in unrealistic networks with respect to certain characteristics such as child-parent relationships, the presence or absence of arcs and arc directions. The imposition of these constraints reduces the size of the search space and improves the performance and convergence of the search algorithm. Two common types of structural constraints relevant to this study are: 1) *ancestral*—directed arcs from local- to large-scale variables are forbidden; and 2) *temporal*—arcs directed backwards in time are forbidden since the value of a variable at time t is determined by the values of a set of variables at t or previous time points. Such constraints can be imposed via so-called *whitelist* and *blacklist* arguments in the call to the learning algorithm. A whitelist can be used to force the inclusion of specific arcs in the solution. A blacklist can be used to force the exclusion of specific arcs.

Typically, the initial network is large and may contain orphan or redundant nodes and substructures that have negligible impact on network performance. Pruning is a popular technique for eliminating regions of the search space that can be safely ignored. It can quickly find an initial feasible solution and then, given more computation time, move the solution towards optimality by improving the score. Potential benefits of pruning are reduction in network size without unduly losing predictive performance, reduction in the number of parameters (which makes overfitting less likely), and improved network robustness and predictive reliability for data not used for structure learning. Once a BN has been learned, its DAG can be used to ascertain what variables remain and gain insight into the relationships between them and their strength. Importantly, it can also be used to perform probabilistic inference for diagnostic purposes and prediction.

3 Design, construction and testing of the model

3.1 Framework design

Figure 1 illustrates the framework's components and their interfaces in the physical (climate-wildfire) system under consideration. The data sets consist of lightning-ignited wildfire counts (LICs), climate mode indices and hazard variables (such as an index of fire weather conditions and prescribed burn area). An exploratory data analysis is performed to identify the distributional and temporal characteristics of the data. The probability integral transform and the inverse transform are used to convert the distributions of continuous and discrete variables to standard normal (Appendix A.2). This step is needed because the Bayesian networks considered herein are predicated on the assumption that all variables are normally distributed (Gaussian). Structure learning is undertaken to estimate the DAG that best describes the dependencies in the data set. A consequence of the analysis is that it provides insight into how the potential causal mechanisms operate and control fire weather conditions and lightning-ignition counts over time. Finally, the predictive performances of the selected network and a reference model are assessed and compared using rolling origin evaluation (Sect. 3.5). By exploiting the factorization described in Sect. 2.2, predictions can be obtained for LICs and a measure of fire weather conditions.

3.2 Model design

Our goal is to identify a compact representation of the system dynamics without incurring an undue loss of information. Our approach represents an attempt to provide as much explanatory power and predictive power as possible. Explanatory power refers to the extent to which an existing theory or competing theories provide a description of system behavior. In the absence of a theoretical foundation for the

Fig. 1 Schematic of proposed framework. Grey boxes and accompanying labels indicate the major task sequence (left to right). Arrows depict data flow lines. Hazard variables could include, but are not limited to, a fire danger index (FDI) and a measure of the level of prescribed burning activity (e.g. prescribed burn area, PBA)

relationships that comprise the physical system, it can refer to the extent to which a learned network can provide a causal explanation for any observed impact of the system attributes.

A simplified variant of a DBN is a sparse DBN (sDBN), where the term *sparse* means that a sizeable portion of the model parameters in (4) are zero. This translates to a more compact DAG than what might have been expected. For the intended purposes of this study we make the following assumptions for our sDBN: (1) the network structure and the probability distributions and associated parameters describing the temporal dependencies are time invariant (while noting that the temporal patterns of residuals from the models are examined for trend); (2) the joint probability distribution of the random variables is Gaussian; (3) the DAG is not required to be fully connected, and (4) the system behaves like a low-order autoregressive process (i.e., the current value of a time series is dependent on one or more past observations in that series, plus noise). A DAG that is not fully connected contains independence assumptions. The more such assumptions are made and found to be reasonable, the more compact the network structure becomes.

While a focus on the optimization of the scoring function can lead to performance improvements on the training set, it can also compromise the stability of the final network (i.e., its robustness to small perturbations of the data set). As outlined below, bootstrap resampling is a useful means to assess the stability of selected networks. A learned network with high explanatory power does not necessarily possess high predictive power. Predictive power refers to the performance of the network with new data previously unseen during structure learning and parameter estimation. While the focus of predictive modeling is typically on association rather than causation, our focus is on model transparency (i.e., being able to understand the concepts captured by a learned network and the reasons behind them). In this context, the use of predictive modeling and testing can help to: uncover and confirm potential causal mechanisms and pathways, lead to the generation of new hypotheses, suggest improvements to existing explanatory models, compare competing theories and ascertain practical relevance (Shmueli 2010; Heinze et al. 2018).

DAGs are a particularly useful tool for identifying redundant variables. Given that the networks considered herein are not complex, in the sense that the number of observations is much larger than the number of nodes, we used seven heuristic pruning steps:

1. Build an initial network (fully connected DAG)
2. Remove any orphan nodes and isolated subnetworks in the DAG to obtain a candidate network. Repeat as necessary.
3. Compare the candidate network score to the local scores for the remaining nodes. Prune nodes that make a rela-

tively small fractional contribution to the network score. Repeat step two as necessary, then proceed to step four.

4. Use a nonparametric bootstrap procedure with 500 replications to assess *arc strength* and *arc direction* for each possible arc in the candidate network. Here:
 - (a) Arc strength is the probability of observing an arc between two nodes (i and j) across the bootstrapped networks, regardless of the arc's direction.
 - (b) Arc direction is the probability of an arc directed from node i to node j across the bootstrapped networks, conditional on the presence of an arc between the two nodes regardless of its direction.
5. Remove any non-terminal leaf nodes and associated arcs in the DAG that have low values of arc strength and arc direction.
6. Repeat steps two to five as necessary until no non-terminal leaf nodes remain and arc strengths exceed a specific threshold value to prevent spurious or weak node connections. A threshold value of 0.7 seemed appropriate for the task at hand.

The introduction of randomization in the data transformation step (Fig. 1 and Appendix A.2) can affect the trajectory of the pruning process. However, while different trajectories were identified in our experiments, they all led to the same compact network structure.

A key assumption underpinning our analysis is that the data are generated by underlying stochastic processes that are stationary. Therefore, it is important to investigate the robustness of the network structure over time. We adopted an ad hoc approach in which individual sDBNs are constructed for three splits of the available data. The minimum number of observations in each training set is given by $\lceil f_m * N \rceil$ where f_m is the minimum proportion of observations for model training, N is the total number of available observations and the ceiling brackets indicate rounding up to the nearest integer. We set $f_m = 0.7, 0.85$ and 1.0 . The similarity of the corresponding DBNs is judged by comparing the frequencies of the nodes and directed arcs present in the three DAGs.

3.3 Data pre-processing and exploratory data analysis

We chose a monthly time step for our calculations given most of the variables considered are available as monthly series. Therefore, the lightning-ignition series was aggregated by month. We applied the Euclidean norm to daily

continuous series to characterize the magnitude of these variables at a monthly time step.

Application of the univariate Shapiro–Wilk test to the pre-processed data indicated that all but one of the marginal distributions exhibited marked departures from normality (Sect. 4.3). This result indicates the need for data transformation. We used a dual transformation approach to convert the marginal distributions of the variables to standard normal (Appendix A.2). Normality in the marginal distributions of a multivariate population is a necessary but insufficient condition for multivariate normality. Numerous statistical tests and graphical approaches are available to check the multivariate normality (MVN) assumption. Different MVN tests may give with different results as they are based on different measures of non-normality and there are many types of departures from MVN (Thode 2002, Chap. 9). Consequently, we use the chi-square Q-Q plot (Appendix A.3) which allows for a visual assessment of multivariate normality analogous to the normal quantile–quantile plot for univariate data. It is also useful for identifying atypical points or outliers. Similarly, the validity of the linearity assumption in (4) is assessed using pairwise scatter plots of the random variables.

3.4 Software

All analyses were carried out in the R computing environment (<https://www.r-project.org/>). Our Bayesian networks were built, refined, and evaluated using the R package *bnlearn* (Scutari 2010). The package supports structure learning, parameter learning and prediction. The structure learning algorithms in *bnlearn* permit the imposition of structural constraints by specifying a whitelist and/or a blacklist. We used the tabu search algorithm and a score-equivalent Gaussian posterior density (also known as the Bayesian Gaussian equivalent score, BGe) to guide the learning procedure. The BGe can be decomposed into a sum of all local scores that depends only on each node and its parents. This is useful information for pruning the network structure (Sect. 3.2). The *bnlearn* package was also used for conditional independence tests, nonparametric bootstrapping and to obtain maximum likelihood estimates of the unknown parameters of a network given its structure (Eq. (4)). Another R package, *bnviewer*, was used for network visualization (Fernandes 2019).

3.5 Rolling origin evaluation

Assessments of predictive power are typically based on out-of-sample measures. The available sample is split into two parts: the training set (in-sample) and the test set (out-sample or holdout sample). The model is then fitted to the training set and its predictive skill is evaluated using one

or more error measures on the test data. The overall aim is to minimize the combined bias and variance by sacrificing some bias in exchange for reducing sampling variance. If network performance is noticeably better during training, overfitting might be an issue (Shumeli 2010).

Evaluating model performance based on a single split may provide a biased estimate of model performance due to the presence of outliers and changes in the mean level of the response variable. Application of more sophisticated conventional techniques (such as k -fold cross-validation) to time series can be problematic in three respects: (1) there is an implicit temporal dependence in such data (which might include seasonality and a chronic trend); (2) data leakage from the response to lagged variables is more likely to occur (i.e., information from the test set is inappropriately incorporated into the training set); and (3) the possible presence of non-stationarity in the series (due, for example, to temporal variations in the mean and variance).

Rolling origin evaluation is a resampling method in which the origin of the test set is updated successively, and model predictions are produced from each origin (Tashman 2000; Petropoulos et al. 2017). There are different options of how this can be done. Our approach is illustrated in Fig. 2. Each circle represents h observations corresponding to the length of the prediction horizon. For convenience assume that N is exactly divisible by h . Let M denote the number of observations in the training set for each iteration (red circles). It follows that there are h observations in each test set (blue circles). During the first iteration, the network structure and parameter estimates learned using the first M observations are used to make predictions for FDI_t and LIC_t , $t = M + 1, M + 2, \dots, M + h$. These predictions are stored for later retrieval. During the second iteration the first h observations are discarded, and the training set is moved forward to maintain the training set size (M). The model parameter estimates are updated using the new training set with a view to improve the accuracy of network predictions and reduce their uncertainty. Predictions are then obtained for next h observations and stored. This process is repeated until the last observation is used for testing. Model predictive skill is evaluated by applying six different measures to stored predictions and the corresponding observations (Appendix A.4). They are: mean bias error (MBE); maximum absolute error (MAE); Manhattan distance (MANH); Euclidean distance (EUC); cross-validated coefficient of determination (often written as Q^2); and the proportion of variation in the test set explained by model predictions based on cross-validation (VE_{CV} ; Li 2016).

As a benchmark, a reference model (rDBN) was constructed using (local-scale) hazard variables and their lag-one time series, and the structure learning approach described above. Including the rDBN in the rolling origin procedure permits an assessment of the additional predictive

Fig. 2 Schematic of rolling origin procedure for model evaluation. For a given iteration: interior circles each represent h time steps (observations); for the sake of convenience in maintaining an almost constant prediction horizon, the first and last circles may represent slightly smaller or slightly larger numbers of time steps; filled red circles indicate observations in the training set; filled blue circles indicate the test set, and grey circles indicate observations that are temporarily excluded from the training or the test set

skill offered by a more complicated sDBN that incorporates both local-scale and climate mode information.

4 Case study

4.1 Overview of the study region

The proposed method was applied to the Warren region (1.6 M ha) in the southwest corner of Australia (Fig. 3). The following description parallels that of Bates et al. (2018), and the text below is derived from there with minor modification. The region has a warm-summer Mediterranean climate 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. The region's terrain is undulating. 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. Most of the public land is relatively remote and offers limited access to fire-fighting operations. The peak wildfire season occurs in the austral summer (December to February), extending into autumn months.

Fig. 3 Location map. Black circles indicate locations of lightning-ignitions. The inset shows the location of the study region (red square) within the wider context of Australia. Adapted from Fig. 1 of Bates et al. (2018). Forest-ecosystem and topographic maps for the study region appear in Bates et al. (2018, Figs. 1 and 2)

4.2 Data description

Modeling should start with defensible assumptions on the possible roles of variables based on previous studies, expert knowledge, or common sense (Heinze et al. 2018). Our analysis is based on indices for six climate modes that are known to either individually or jointly influence atmospheric circulation across Australia (see, e.g., Power et al. 1999; Min et al. 2013; Marshall et al. 2014; Hendon and Abhik 2018). The indices and modes considered are: (1) the Niño3.4 index (N34) of the El Niño–Southern Oscillation (ENSO), where N34 is defined by area-mean sea surface temperature (SST) over the central equatorial Pacific Ocean bounded by (5° S–5° N, 170° W–120° W); (2) the dipole mode index (DMI) of the Indian Ocean Dipole (IOD) which is represented by an anomalous SST gradient between the western equatorial Indian Ocean bounded by (50° E–70° E, 10° S–10° N) and the south eastern equatorial Indian Ocean bounded by (90° E–110° E, 10° S–Equator) (Saji et al. 1999; Saji 2019); (3) the tripole index (TPI) of the Interdecadal Pacific Oscillation (IPO) where TPI is based on the difference between SST anomalies averaged over the central equatorial Pacific and the average of the SST anomalies in the Northwest and Southwest Pacific (Henley et al. 2015); (4) the real-time multivariate indices of the Madden–Julian Oscillation (MJO) proposed by Wheeler and Hendon (2004), herein designated by RM1 and RM2; (5) the stratospheric Quasi-Biennial Oscillation (QBO) of mean zonal winds at 30 hPa (U30) and

50 hPa (U50) over Singapore, and (6) Marshall's (2003) observation-based index of the Southern Annular Mode (SAM, which refers to the north–south movement of a belt of westerly winds that circles Antarctica, and is measured by the monthly difference in normalized zonal-mean sea level pressure at six stations close to 40° S and six stations close to 65° S).

Three out of the six climate modes listed above (ENSO, IOD, and SAM) have been considered in previous studies linking climate modes with fire weather conditions and lightning-ignited wildfire in Australia (see cited references in Sect. 1). The remaining modes (IPO, MJO and QBO) were considered here for the sake of completeness. The spatial configuration of the climate modes is displayed in supplementary Figure S1. Time series plots and boxplots showing data distributions by month appear in supplementary Figure S2(a–i) and S2(l–t), respectively. Further details are provided in Table 1.

We also consider two local-scale hazard variables: fire weather and fuel moisture conditions, and prescribed burn area. The McArthur Mark V forest fire danger index (FDI, see McArthur (1967) and Noble et al. (1980)) was used to characterize the first hazard variable. The index combines several meteorological factors with fuel moisture information and is commonly used in Australia for forecasting the influence of weather on fire behavior. It is based on an empirical relationship comprising daily maximum temperature, relative humidity, wind speed and a longer-term estimate of fuel availability based on soil dryness and recent rainfall. The construction of the FDI data set used here is described in detail by Dowdy (2018) and is primarily based on a gridded analysis of observations. The second

hazard variable (prescribed burn area, PBA) was included to control for the potential effect of fire management on lightning ignitions. Since 1961 DBCA has used prescribed fire to lower fuel loads across large areas of southwest Australia and reduce the severity and size of wildfire (Burrows and McCaw 2013 and references therein). The decision to burn is dependent on suitable daily weather conditions and fuel moisture. DBCA maintains a database of the annual prescribed burn area for entire forest region of southwest Western Australia. We included this annual series in our study since prescribed burns may mitigate the potential hazard of lightning ignition. Given the lack of detailed spatial and temporal information about prescribed burn area, expert opinion is that about one-third of that annual burn area takes place in the Warren region. Prescribed burns are typically conducted in the austral autumn and spring. Burns can also take place in winter months (June–August) in more open vegetation types subject to prevailing conditions. To disaggregate the annual series to monthly, we assumed that the intra-annual burn area can be represented by a bimodal Gaussian mixture distribution with modes at April and October, variances equal to 1, and the probabilities for mixture component equal to 0.5. While this constitutes a set of ad hoc assumptions, it is not unreasonable in the absence of data at a finer temporal scale. Time series plots for FDI and PBA, and boxplots showing data distributions by month, appear in supplementary Figure S2(j–k) and S2(u–v), respectively.

The lightning ignition data for the Warren region covers the period from April 1976 to December 2016. The monthly record consists of 489 observations, 95 of which are non-zero lightning-ignition counts. Observation platforms for fire detection and reporting protocols have

Table 1 Summary statistics and probability-values (p values) obtained from the Shapiro–Wilk test for departures from univariate normality before and after dual transformation (Sources key: BoM=Bureau of Meteorology; DBCA=Department of Biodi-

versity, Conservation and Attractions, Western Australia; NOAA/ESRL=National Oceanic and Atmospheric Administration/Earth System Research Laboratory)

Data details			Pre-processed data			Transformed data		
Monthly index series	Notation	Source	Mean	Variance	p value	Mean	Variance	p value
Lightning-ignition counts	LIC	DBCA	0.759	6.24	0.000	0.0156	0.975	0.741
Dipole Mode Index	DMI	NOAA/ESRL	0.155	0.101	0.000	0.00143	0.980	1.00
McArthur forest fire danger index	FDI	BoM, Australia	0.229	0.0263	0.000	$<1 \times 10^6$	0.980	1.00
Madden–Julian Oscillation—RMM1	RM1	BoM, Australia	5.22	3.64	0.000	0.00000	0.980	1.00
Madden–Julian Oscillation—RMM2	RM2	BoM, Australia	5.18	4.11	0.000	0.00000	0.980	1.00
Niño 3.4 SST	N34	NOAA/ESRL	27.1	0.941	0.005	0.0056	0.980	1.00
Prescribed burn area	PBA	DBCA	5.5×10^3	27.8×10^6	0.000	0.00835	0.988	1.00
Quasi-Biennial Oscillation (mean zonal wind components at 50 and 30 hPa)	U50	Freie Universität, Berlin	−0.193	163	0.000	0.00930	0.987	1.00
	U30	Freie Universität, Berlin	−6.73	330	0.000	0.00773	0.983	1.00
Southern Annular Mode	SAM	British Antarctic Survey	0.231	3.06	0.015	0.00286	0.981	1.00
Tripole Index	TPI	NOAA/ESRL	0.0275	0.555	0.088	0.611×10^3	0.980	1.00

remained stable over the period of record, and changes in land use are unlikely to have altered the pattern of lightning ignition. The spatial pattern of ignition locations is displayed in Fig. 3.

4.3 Examination of modeling assumptions and remedial action

We used a variety of techniques to maximize insight into the features of the available data. The specific tasks undertaken (and methods used) included: (1) detection of potential outliers (boxplots and dot charts); (2) calculation of summary statistics (mean and variance); (3) testing the distributional assumptions underlying the analysis described below (Shapiro–Wilk test for univariate normality and the chi-squared quantile–quantile plot to visually assess departures from multivariate normality); (4) testing the assumption of linear relationships among the variables (scatter plots and cursory linear regression modeling); and (5) characterization of the underlying temporal patterns (autocorrelation and partial autocorrelation function plots).

Visual inspections of boxplots and dot charts did not reveal any outliers in the data set. Table 1 reports summary statistics and probability-values (p values) for the Shapiro Wilk statistics for the pre-processed and transformed random variables. The values of the pre-processed variables are on different scales (different units of measurement, means and variances). Most exhibit departures from normality (p value < 0.05). In comparison, the dual transformation has achieved the outcome of converting the distributions of the variables to standard normal. The p -value for the transformed LIC data (0.741) is lower than those for the remaining transformed variables (all 1.00). This discrepancy is due to the use of randomization in the transformation (Eq. (A2)) and the effects of sample size. Nevertheless, the p value for the transformed LIC data is much closer to 1.0 than that for the pre-processed data (0.000). The chi-square Q-Q plot for the transformed data (Fig. 4) indicates only slight departure from multivariate normality in the extreme upper tail. This departure is most likely due to sample size effects and the predominance of zeros in the pre-processed LIC series. The consequences of this departure on model performance are investigated in Sect. 4.5. Pairwise scatter plots of the data (not shown) did not reveal any nonlinear associations. Except for the PBA series, inspection of autocorrelation and partial autocorrelation plots for the remaining series revealed that their temporal structures are well-described by low-order autoregressive processes. The temporal structure of the PBA series is far more complicated. We postpone consideration of this issue to Sect. 5 at which time a more informed judgement can be made about the contribution of PBA to network performance.

Fig. 4 Chi-squared quantile–quantile plot of transformed observations of LIC

4.4 sDBN structure and features

Time series plots of the residuals obtained from the rDBN and sDBN exhibited a random pattern around zero (see supplementary Figures S3a, c for the sDBN case). This suggests that the DBNs provide reasonable representations of the temporal patterns of these data including trends and seasonality (cf. Bates et al. 2018). The quantile–quantile plot of observed and fitted FDI values for the sDBN is shown in supplementary Figure S3b. Overall, the model fit is good. A disparity exists in the upper half of the corresponding quantile–quantile plot of observed and fitted LICs in the training set (see supplementary Figure S3d). This issue is discussed further below (Sect. 4.4).

The DAG for the sDBNs for the $f_m = 0.7$ and 1.0 cases is displayed in Fig. 5. It represents a generalization of the structure of the final sDBNs for $f_m = 0.85$, and corresponds to the following factorization of the joint density:

$$\begin{aligned}
 P(\mathbf{Y}) = & P(\text{DMI}_{t-1})P(\text{RM1}_{t-1})P(\text{U50}_{t-1}) \\
 & P(\text{RM2}_{t-1}|\text{RM1}_{t-1})P(\text{N34}_{t-1}|\text{DMI}_{t-1}) \\
 & P(\text{U30}_{t-1}|\text{U50}_{t-1})P(\text{TPI}_{t-1}|\text{N34}_{t-1}) \\
 & (\text{FDI}_{t-1}|\text{RM2}_{t-1}, \text{U30}_{t-1})P(\text{FDI}_t|\text{DMI}_{t-1}, \\
 & \text{N34}_{t-1}, \text{TPI}_{t-1}, \text{FDI}_{t-1})P(\text{LIC}_t|\text{FDI}_t).
 \end{aligned} \quad (5)$$

There are six main features in the DAG. First, it comprises ten nodes and eleven directed arcs whereas the DAG for the $f_m = 0.85$ case contains eight nodes and nine directed arcs. All three DAGs have no undirected arcs. Second, the

Fig. 5 DAG for the final sDBN structures learned for the $f_m=0.7$ and 1.0 cases. Labelled grey circles indicate nodes and black arrows directed arcs. The DAG layout is based on aesthetic criteria such as: connected nodes are drawn close to each other, minimal arc crossings, inherent symmetry, and containment within a plotting frame. Thus, arc lengths have no specific meaning. A lagged one-month series is indicated by $_1$ at the end of the node name (due to limitations of the plotting software)

nodes corresponding to climate mode indices are all lag-one series (cf Eq. (5)). This indicates that the interactions between climate modes, FDI and LIC are characterized not by immediate but slight delayed responses. Third, there are no nodes for PBA and SAM. During the pruning process for each sDBN, either the initial network contained the isolated arc $SAM_{t-1} \rightarrow SAM_t$ or both nodes emerged early as leaf nodes. Consequently, the SAM nodes were removed from the analysis leading to higher values of BGe. Pruning of the PBA nodes followed the same pattern as that for SAM. This

indicates that PBA is not influential in the determination of fire weather or the occurrence of lightning-ignited wildfires. (It does not follow, however, that prescribed burning is an ineffective tool for fire management.) Fourth, while the set of nodes comprising DMI_{t-1} , $N34_{t-1}$ and TPI_{t-1} are interconnected they are not individually or jointly connected to members of the set $RM1_{t-1}$, $RM2_{t-1}$, $U50_{t-1}$ and $U30_{t-1}$. Instead, these sets are connected to the FDI_t and FDI_{t-1} nodes, respectively. Also, $RM1_{t-1}$, $U50_{t-1}$ and DMI_{t-1} are all root nodes and therefore are independent of each other. Fifth, there are no direct interactions between LIC_t and the climate mode indices. There is the directed arc $FDI_t \rightarrow LIC_t$ but no arc from FDI_{t-1} to LIC_t . This suggests that the lightning-ignition process is itself conditioned on concurrent weather and fuel moisture conditions. Sixth, inspection of Fig. 5 or Eq. (5) reveals that network contains two variables that are measured simultaneously: FDI_t and LIC_t . While this might be useful for explanatory purposes, it could present a problem for predicting LIC_t in an operational environment because the FDI at the time of prediction is unknown. However, LIC_t is only conditionally dependent on FDI_t and after processing the data for FDI_t and its parents the network can be used to can compute LIC_t .

Table 2 summarizes the characteristics of the DAGs for the three final sDBNs. The table's contents are arranged as a matrix. Diagonal terms represent nodes and off-diagonal terms directed arcs. The numerals in the table cells are the number of instances of each node or directed arc in the three DAGs. The number of used table cells is 21 with 10 representing nodes. Seventeen of the 21 cells represent nodes and arcs that appear in the three DAGs. Eight out of the ten nodes shown occur in all sDBNs. The two exceptions are $RM1_{t-1}$ (appears in the DAGs for $f_m=0.7$ and 1.0) and $U50_{t-1}$ (which also appears in the DAGs for $f_m=0.7$ and 1.0). In all three DAGs, FDI_t is the only immediate parent to the terminal node LIC_t . Across the three DAGs, the number of

Table 2 Agreement between sDBN DAGs for three data splits ($f_m=0.7, 0.85, 1.0$)

Parent	Child									
Node	DMI_{t-1}	FDI_{t-1}	FDI_t	$RM1_{t-1}$	$RM2_{t-1}$	$N34_{t-1}$	TPI_{t-1}	$U30_{t-1}$	$U50_{t-1}$	LIC_t
DMI_{t-1}	3		3			3				
FDI_{t-1}		3	3							
FDI_t			3							3
$RM1_{t-1}$				2	2					
$RM2_{t-1}$		3			3					
$N34_{t-1}$			3			3	3			
TPI_{t-1}			3				3			
$U30_{t-1}$		3						3		
$U50_{t-1}$								2	2	
LIC_t										3

Node names denote dual transformed variables. Numerals indicate the number of times nodes and directed arcs appear in the three DAGs. Bold represent nodes, and italics off-diagonal cells directed arcs

directed arcs is 31. The minimum, median and maximum arc strengths are 0.498, 0.999 and 1.000. The direction probabilities for 21 of the 31 arcs are all 1.0 to four significant figures. The direction probabilities for seven of the remaining ten arcs fall in the closed interval [0.37, 0.66], but their arc strengths are all 1.0 to four significant figures. The directed arcs concerned are: $N34_{t-1} \rightarrow TPI_{t-1}$ (three instances); $RM1_{t-1} \rightarrow RM2_{t-1}$ (two instances); and $U50_{t-1} \rightarrow U30_{t-1}$ (two instances). These results indicate that while the directions of dependency are not well established for these arcs the relationships are (very) strong. The situation for the arc $DMI_{t-1} \rightarrow TPI_{t-1}$ (3 instances) is somewhat different in that while the direction probabilities are all close to 0.5 the arc strengths increase from 0.50 to 0.90 as f_m increases. Taken overall, the results suggest that the structure of the sDBN has only a slight sensitivity to temporal fluctuations in the full data set.

There are no arcs connecting any of the climate mode indices directly to the terminal node. Therefore, it follows that

$$P(LIC_t | DMI_{t-1}, FDI_{t-1}, FDI_t, MJA_{t-1}, MJP_{t-1}, N34_{t-1}, TPI_{t-1}, U30_{t-1}) = P(LIC_t | FDI_t). \quad (6)$$

Equation (6) indicates that, for the case study region, local-scale fire weather conditions influence lightning-ignition occurrence much more so than large-scale climate variability modes. It also dictates that the fitted and predicted values of LIC_t using either the rDBN or the sDBN are identical. Second, as noted above, the only nodes at time slice t are FDI_t and LIC_t . This peculiar configuration is distinctly different to that of a conventional DBN in that there is no common base network for the temporal slices $t-1$ and t . Third, the parent nodes for FDI_{t-1} are not directly connected to the parent nodes for FDI_t . This result indicates an absence of direct interaction between the parent node sets.

4.5 Rolling origin evaluation

The rolling origin evaluation is performed under the following settings: $N = 489$ (the number of months in the period April 1976 to December 2016); $h = 12$ months (covering July to June of the following year), $f_m = 0.7$, and $M = 351 \geq [f_m * N]$ which ensures that the last observation in all training sets corresponds to the month of June. The size of the complete test data set is, of course, $N - M = 138$.

Table 3 lists the values of the model performance measures for the FDI data only. Separate results for the LIC data are not reported since the rDBN and sDBN provide the same predictions (Sect. 4.3). Perusal indicates that the sDBN is superior to the rDBN for five out of the six predictive skill measures examined, with the exception being MBE. The MBE obtained for the rDBN (-0.00338) is close to zero and smaller than that for the sDBN (-0.0237). Nevertheless, the latter value of MBE is small relative to the standard deviation of FDI observations (0.169). For the sake of comparison, the MBEs obtained at the end of the structure learning task (i.e. for $f_m = 1.0$) were -0.000135 for the rDBN and -0.000507 for the sDBN. Put together, these results suggest that the rDBN and sDBN provide a good fit to the data and produce reliable predictions in terms of bias. However, the sDBN produces predictions that are more precise (cf. MAE, MANH, EUC, Q^2 and VE_{CV} values in Table 3).

Figure 6 displays quantile–quantile plots of mean predicted values of FDI_t and LIC_t obtained from the sDBN during rolling origin evaluation versus observed. The pre-

dicted distribution of FDI_t is close to the observed with a small additive bias (Fig. 6a). This indicates the sDBN predictions can be improved by a simple (and small) additive bias correction. For both DBNs, time series plots of the residuals obtained from the evaluation procedure exhibited a random pattern around zero (see supplementary Figures S4a, b for the sDBN case). These findings indicate that

Fig. 6 Quantile–quantile plots of mean predicted versus observed series obtained from the sDBN in rolling origin evaluation: **a** monthly forest fire danger index (FDI_t), with the MBE indicated by the blue dashed line, and **b** monthly lightning-ignition counts (LIC_t)

Table 3 Comparison of predictive skills offered by rDBN and sDBN for FDI

Model	MBE	MAE	MANH	EUC	Q^2	VE_{CV}
rDBN	-0.00338	0.0804	11.1	1.26	0.63	0.592
sDBN	-0.0237	0.0622	8.58	1.08	0.74	0.700

Best values of model performance measures are presented in boldface type

the DBNs provide reliable predictions for FDI despite their slight sensitivity to temporal fluctuations in the full data set (Sect. 4.3). The equivalent rDBN and sDBN predictions of LIC_t exhibit a slight underestimation for the interval $0 \leq LIC_t < 5$, where 126 of the 138 observations in the test set lie (Fig. 6b). However, both DBNs underestimate the upper tail of the LIC_t distribution. A similar but smaller disparity exists in the quantile–quantile plot of observed and predicted LIC_t in the training set (see supplementary Figure S3c). Since time series plots of the residuals obtained from the evaluation procedure exhibited a random pattern around zero (see above), this suggests that the disparity evident in Fig. 6b is due to a failure to adequately predict high ignition counts rather than the increasing trend in the data (Bates et al. 2018). Taken together, these results may be indicative of a relationship between LIC_t and one or more missing synoptic-scale variables related to dry season storminess including convective available potential energy, convective inhibition and cloud top height (e.g. Riemann-Campe et al. 2009; Clark et al. 2017).

5 Discussion

The climate-wildfire system considered herein is not well understood from either a theoretical or observational standpoint. Our analysis is based on observational data containing measurement errors, probabilistic properties, and conditional independence relationships in the presence of limited background knowledge. Therefore, it cannot be guaranteed that: (1) all the important processes and phenomena are included properly in the learned network; (2) the directed arcs in the network represent actual cause-and-effect relationships, or (3) that the network possesses an appropriate level of complexity. To some extent, the impact of such issues on a Bayesian network can be examined using a sensitivity analysis approach. Bayesian networks provide a sensitivity analysis capability in a natural way. For example, the effects of different types and levels noise in the data on network parameters, structure, and performance can be easily investigated. Also, structure transformations such as node removal, arc removal and arc reversal can also be explored. For the case study reported herein, we found that changes in the sequence of node deletions during a pruning process had no impact on the configuration of the nodes comprising the final network. Arc reversal was sometimes observed for those arcs where the directions of dependency were not well established (Sect. 4.4). However, our results suggest that such arc reversals are not critical to model performance. A comprehensive sensitivity analysis will be undertaken in future research.

A result worthy of further investigation is that PBA was not found to be influential in the determination of fire

weather conditions or the occurrence of lightning-ignited wildfire. Fuel condition and loading would be expected to more strongly influence fire intensity and resistance to containment than it would the likelihood of ignition (McCaw et al. 2012), particularly given the observation that lightning strikes may ignite dead wood in standing trees resulting in fires that persist for many weeks (Bates et al. 2018). While PBA was the only human-induced variable considered in our study, the potential exists to apply our approach to aid historical analyses in which it can be difficult to disentangle human and climate influences on wildfire processes.

Further research involving theoretical modeling and numerical simulation, together with observational data, is necessary to test the veracity of our results and findings. However, many of our findings withstand scrutiny from a variety of perspectives. All the climate mode indices included in the sDBN are lag-one time series that highlight the importance of tropical-extratropical teleconnection patterns. These patterns fluctuate over a wide range of time scales, from synoptic (e.g., regional clusters of enhanced or decreased convective processes near Australia can be associated with different phases and amplitudes of the MJO) to multidecadal (e.g., variations in Pacific Ocean conditions similar to ENSO but for low-frequency components as indicated by the IPO). The node and directed arcs depicted in Fig. 5 are broadly consistent with current knowledge about the relationships between climate modes and fire weather conditions in Australia such as for the influence of ENSO and IOD on FFDI (Harris and Lucas 2019). They also provide potential insight into the climate mode interactions that are relevant to fire weather conditions and the occurrence of lightning ignitions in the case study region. There is also scope for further investigations into physical processes such as the role of QBO and MJO on antecedent weather conditions of relevance for lightning-ignited fires. These claims do not pertain to all climate mode interactions in the global climate system.

The ENSO is an oscillation of the coupled atmosphere and ocean and is the strongest interannual signal in the global climate system (Halpert and Ropelewski 1992; McPhaden et al. 2006). It has been shown to exert a strong (moderate) influence on fire activity in southeast (southwest) Australia (Harris et al. 2014; Mariani et al. 2016; Harris and Lucas 2019, and references therein). This is associated with ENSO's influence on fire weather conditions in the southeast as well as in the southwest particularly during summer (Dowdy 2018). The Interdecadal Pacific Oscillation (IPO) has a dominant period of roughly 20–30 years whereas the El Niño and La Niña phases of ENSO tend to last less than one year. Disentangling the dependence relationship between the ENSO and IPO is a challenging and ongoing research topic as the two are intertwined (Santoso et al. 2015). Our results are broadly consistent with this current

general understanding of IPO/ENSO relationships including that the direction of dependency is not well established even though the relationship is strong.

The Indian Ocean dipole (IOD) is a major mode of inter-annual climate variability in the tropical Indian Ocean. Inspection of Fig. 5 in Harris and Lucas (2019) indicates weak to moderate correlations (0.36–0.40) between the 90th percentile of December–February FDI in southwest Australia and the DMI lagged by three and six months. The existence, strength and direction of the relationship between the IOD and ENSO has been the subject of extensive debate (Zhou and Nigam 2015; Lestari and Koh 2016 and references therein). In the context of our case study, we have found that while the relationship is strong (for $f_m = 0.7$ and 1.0) the direction of dependency is not well established. This result is consistent with the current state of knowledge about the IOD-ENSO interaction. Yamagata et al. (2004) and Dowdy (2016) found that the strongest correlation between ENSO and IOD indices occurs in the austral spring (September–November) ($r = 0.53$ and 0.68 , respectively). The strength probabilities for the arc $\text{DMI}_{t-1} \rightarrow \text{N34}_{t-1}$ were 0.50, 0.81 and 0.90 for the $f_m = 0.7$, 0.85 and 1.0 cases, respectively. This result indicates a marginal to moderate increase in the strength of the relationship in the latter part of the study period. While this result may be attributable to noise, it is also consistent with recent evidence that the strength of the relationship is not time invariant (Lim et al. 2017). Furthermore, the results depicted in Fig. 5 and listed in Table 2 indicate that the IOD and ENSO have individual and joint effects on fire weather and fuel moisture conditions, and hence the occurrence of lightning-ignited fires.

The Madden Julian Oscillation (MJO) is an eastward-moving intra-seasonal (20- to 90-day) oscillation in tropical winds and convection. It is the most dominant mode of variability in the tropical atmosphere and is strongest during the months of November to April (Madden and Julian 1994; Zhang 2013). It is known to cause variations in weather in far-reaching extratropical locations around the globe (Wheeler and Hendon 2009 and references therein). The results depicted in Fig. 5 and Table 2 suggest a lack of direct dependency between ENSO and the MJO. Zhang et al. (2001) note the lack of a strong statistical relationship between ENSO indices and conventional MJO indices based on global winds and convection. Similarly, Moon et al. (2011) state that the roles of changes in MJO convection and the mean circulation in accounting for the differences between El Niño and La Niña are difficult to find. Harris and Lucas (2019) recommend that future work should consider analyses of the role of the MJO on fire weather variability in northern Australia. Our results suggest that the MJO also influences fire weather even further afield, in southwest Australia.

The QBO is a quasi-periodic oscillation of the equatorial zonal wind in the tropical stratosphere. Between 3 and 100 hPa, zonal wind at the Equator is characterized by a pattern of alternating easterly and westerly wind regimes that propagate downward with time. The wind direction changes about every 14 months. There is increasing evidence from observations and numerical models that stratospheric anomalies (such as the QBO) affect the extratropical troposphere at or near the Earth's surface (Baldwin et al. 2001; Thompson et al. 2005; Marshall and Scaife 2009; Kuroda and Yamazaki 2010). There is also a growing body of literature on interactions between the QBO and ENSO (Schirber 2015; Hansen et al. 2016; Gao et al. 2017 and references therein). Similarly, recent studies have found evidence of an interconnection between the QBO and MJO during the austral summer, when the relationship is at its strongest (Abhik and Hendon 2019; Kim et al. 2019 and references therein). However, for the data and study region at hand, there is no clear evidence that interactions between the QBO and the other modes of climate variability affect the prediction skill of fire weather conditions and lightning-ignited wildfires.

There is a paucity of observations-based research on the relationships between climate variability modes, lightning-ignitions, and wildfire in the southern hemisphere (Mariani et al. 2018 and references therein). Typically, previous Australian studies have focused on one or more of the following climate modes: ENSO, IOD and SAM. We are neither aware of any wildfire studies that have considered climate mode interactions in a formal quantitative sense beyond correlation and regression methods, nor are we aware of any similar studies that have included both the IPO and MJO. We have also considered the contributions of the QBO to fire weather conditions and lightning-ignited wildfires.

The use of statistical or machine learning models in climate change impact assessments is often questioned on the basis that they are likely to miss emerging signals of changes in variable distributions and the temporal evolution of climate modes and local weather. This is true if the model is frozen in terms of its current structure and parameter values. As new knowledge and observations become available, it is a simple matter to add new nodes, set or reset arc directions, relearn the network and re-estimate its parameters. For example, the sparse DBN is ideally placed to examine teleconnection processes at time scales longer than one month. This is because new nodes can be added to the network for the variables and lags that are of particular interest alone. Results from rolling origin evaluations of the current and new model can be used to compare their predictive skills and assess the practical significance of changes in network structure and parameters. This updating process can be repeated multiple times in the future. With each update, the model judged to be the best representation of the data is retained and used until at least the next update.

6 Summary and conclusions

This article proposes and demonstrates a probabilistic framework for investigating the relationships between modes of climate variability, fire weather conditions and the occurrence of lightning-ignited wildfires. The sparse dynamic Bayesian network (sDBN) at the heart of the framework is directed primarily towards identifying potential causal mechanisms and pathways, assessing the nature and magnitude of the contribution of climate modes to fire weather conditions and the occurrence of lightning-ignited wildfire episodes, and evaluating the skill of short-term forecasts. Within the framework, the sDBN is assumed to follow a low-order Markov process. A dual transformation approach is used to convert the distributions of continuous and discrete random variables to standard normal. Linear regression is used to estimate the relationship between a child node and its parents. An initial sDBN structure is learned from data using a scoring function approach. This structure is pruned by a set of heuristic rules that in a systematic way leads to a simpler structure without unduly sacrificing network performance. The predictive skill of the sDBN was assessed using rolling origin evaluation and six common error measures for time series that are based on the notions of statistical distance and similarity. The proposed methodology could be applicable to a wide range of problems in pure climate science such as emergent constraint studies as well as applied studies involving the natural environment and climate sensitive industries.

The case study demonstrates the applicability of the proposed framework to a real climate-wildfire system. It illustrates how the results can provide analysts with key insights into the physical plausibility and complex nature of the causal mechanisms and pathways that influence fire weather conditions and the occurrence of lightning-ignited wildfires. All the climate mode indices included in the sDBN are lagged one-month time series. This indicates that the interactions between climate modes, FDI and LIC are characterized not by immediate but slight delayed responses. These responses could help precondition the environment so that it is more conducive to the occurrence of lightning-ignited wildfires (e.g., through influences on fuel state such as moisture content, load and availability). The structure for the sDBN and its prediction skill were found to be largely insensitive to temporal fluctuations in the available data. The Southern Annular Mode was found not to have a strong influence on either FDI or LIC. It is known to influence cool season rainfall in southern Australia. However, the focus of our study is on the warmer months of the year when antecedent moisture is low. The generality of these findings needs to be established. The causal mechanisms and pathways postulated above are tentative and should be subjected to verification by suitable theoretical and modeling studies. Further

consideration of plausible physical processes that might be associated with these findings is also needed.

Another interesting result is that the relationship between ignition counts and climate modes may be indirect, via fire weather and fuel moisture conditions. We have demonstrated that the use of information about the behavior of climate modes can reduce uncertainty in predictions of such conditions. Predictions of high ignition counts are not convincing as they exhibit a systematic departure from the corresponding observations. These are the cases of most interest to fire managers and planners. This points to the importance of including summary statistics of synoptic-scale processes and examining the interactions of these processes with atmospheric teleconnection patterns in determining the probable number of severe thunderstorm events. Such an investigation may give a wider view of the causal mechanisms and the system response, and lead to improved predictions for lightning-ignition counts. This topic will be the subject of future research.

Appendix A. Technical notes on methods

Multivariate Gaussian density function

The multivariate Gaussian joint density function f_Y of $\mathbf{Y} = \{Y_1, Y_2, \dots, Y_p\}$ is defined by

$$f_Y(\mathbf{y}) = (2\pi)^{-p/2} \det(\mathbf{V})^{1/2} \exp\left(-\frac{1}{2}(\mathbf{y} - \boldsymbol{\mu})^T \mathbf{V}^{-1}(\mathbf{y} - \boldsymbol{\mu})\right), \quad (7)$$

where $\mathbf{y} = (y_1, y_2, \dots, y_p)$ and $-\infty < y_i < \infty$ ($i = 1, \dots, p$). As noted in Sect. 4.3.2, $\boldsymbol{\mu}$ is the p -dimensional mean vector and $\mathbf{V}$ is the $p \times p$ symmetric and positive definite covariance matrix.

Dual transformation method

The dual transformation method draws directly from the concept of randomized quantile residuals (Dunn and Smyth 1996). The intent of this concept was to find residuals whose empirical distribution is close to normality in non-normal regression situations. Here we use the same approach to transform discrete data or continuous non-normal data to normality. Let $F(Y, \boldsymbol{\theta})$ be the cumulative distribution function for a random variable Y . Suppose

$$F_i^* = \begin{cases} F(y_i; \hat{\boldsymbol{\theta}}) & \text{if } Y_i \text{ is continuous} \\ F(y_i - 1; \hat{\boldsymbol{\theta}}) + u_i p(y_i; \hat{\boldsymbol{\theta}}) & \text{if } Y_i \text{ is discrete} \end{cases}, \quad (8)$$

where the subscript i denotes the data point, $\hat{\boldsymbol{\theta}}$ is the vector of estimated distributional parameters and u_i is a random variable drawn from the standard uniform distribution on the interval $(0, 1]$. (The randomization in Eq. (8) is used

to convert discrete to continuous values.) Then, the corresponding standard normal quantile is defined by

$$z_i = \phi^{-1}\{F_i^*\}, \quad (9)$$

where $\phi^{-1}\{\cdot\}$ is the cumulative distribution for the standard normal distribution. The z_i are exactly standard normal, apart from sampling variability in $\hat{\theta}$.

For thoroughness, we note Anscombe (1948) suggested two inverse hyperbolic sine transformations to convert negative binomial to more ‘nearly’ normally data. The ignition counts data in our case study are well described by a negative binomial distribution. However, the maximum likelihood estimates of the distribution’s parameters (mean = 0.759 and shape = 0.100) do not confirm with Anscombe’s requirements (mean is large, and shape is either > 2 or ≥ 1).

Chi-square Q-Q plot

When $\mathbf{Y} \sim N_p(\boldsymbol{\mu}, \mathbf{V})$, the squared Mahalanobis distances defined by

$$d_j^2 = (\mathbf{y} - \boldsymbol{\mu})^T \mathbf{V}^{-1} (\mathbf{y} - \boldsymbol{\mu}), \quad j = 1, \dots, n, \quad (10)$$

are distributed as χ_p^2 (chi-square distribution with p degrees of freedom). This result can be used to evaluate whether a data point may be an outlier and whether the observed data exhibit departures from multivariate normality. A chi-square Q-Q plot is an informal, but useful, way to assess whether d^2 is distributed as χ_p^2 . If the data indeed follow the assumed distribution, then the points in a plot of the sample quantiles (quantiles of d^2) versus the theoretical quantiles (quantiles of χ_p^2) will fall approximately on a straight line through the origin with slope equal to 1.

Statistical distance and similarity measures

Let $\mathbf{y}$ denote the n -dimensional vector of observations and $\mathbf{z}$ the corresponding n -dimensional vector of Bayesian network predictions obtained from cross validation, where in this instance n is the length of the test set.

The mean bias error (MBE) is used primarily to estimate average bias in a predictive model. It is defined by

$$\text{MBE} = \frac{1}{n} \sum_{k=1}^n (z_k - y_k). \quad (11)$$

The mean absolute error (MAE) measures the average magnitude of the errors in the set of predictions without considering their sign:

$$\text{MAE} = \frac{1}{n} \sum_{k=1}^n |y_k - z_k|. \quad (12)$$

The Manhattan (or city block) distance (MANH) is defined by

$$\text{MANH} = \sum_{k=1}^n |y_k - z_k|. \quad (13)$$

The Euclidean distance is defined by

$$\text{EUC} = \left(\sum_{k=1}^n (y_k - z_k)^2 \right)^{1/2}. \quad (14)$$

The MAE, MANH EUC have the same scale as the data.

The cross-validated coefficient of determination (Q^2) can be written as

$$Q^2 = \frac{(\sum_{k=1}^n (y_k - \bar{y})(z_k - \bar{z}))^2}{\sum_{k=1}^n (y_k - \bar{y})^2 \sum_{k=1}^n (z_k - \bar{z})^2}. \quad (15)$$

Unlike the conventional R^2 which measures goodness-of-fit, Q^2 is a reliable measure of model quality.

The proportion of variation in the test set explained by model predictions based on cross-validation (VE_{CV} , Li 2016) is defined by

$$\text{VE}_{\text{CV}} = \left(1 - \frac{\sum_{k=1}^n (y_k - z_k)^2}{\sum_{k=1}^n (y_k - \bar{y})^2} \right) \quad (16)$$

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1007/s00382-021-05764-2>.

Acknowledgements The lightning ignition data set, boundary coordinates for the Warren Region and annual prescribed burn area series were provided by the Department of Biodiversity, Conservation and Attractions, Western Australia. Time series for the McArthur forest fire danger index were supplied by the Bureau of Meteorology. The R computing environment (<http://cran.r-project.org/>) was used extensively throughout the study. We thank two anonymous referees for their thoughtful and constructive criticisms of the original manuscript, and their supportive comments and encouragement.

Author contributions BCB developed the probabilistic framework, carried out the Bayesian network analysis and drafted the original manuscript. AJD suggested the inclusion of fire weather conditions in the study and provided technical guidance on climatological and meteorological aspects. LM initiated the study design and provided technical guidance on meteorological and forest management aspects. AJD and LM were involved in critical revision of the original manuscript. All authors contributed to the refinement of the study design, project coordination and interpretation of results, and have read and approved the final manuscript.

Funding Financial support for AJD was provided by the Australian Government’s Earth Systems and Climate Change Hub (ESCC) of the National Environmental Science Program (NESP).

Availability of data and material Climate mode data used in this study was downloaded from open sources. Forest fire danger index data can be obtained from the corresponding author.

Code availability Questions about code availability can be directed to the first author at bryson.bates@csiro.au.

Declarations

Conflict of interest All authors declare that there are no conflicts of interest.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Abatzoglou JT, Kolden CA, Balch JK, Bradley BA (2016) Controls on interannual variability in lightning-caused fire activity in the western US. *Environ Res Lett* 11:045005. <https://doi.org/10.1088/1748-9326/11/4/045005>
- Abatzoglou JT, Williams AP, Boschetti L, Zubkova M, Kolden CA (2018) Global patterns of interannual climate–fire relationships. *Glob Change Biol* 24:5164–5175. <https://doi.org/10.1111/gcb.14405>
- Abhik S, Hendon HH (2019) Influence of the QBO on the MJO during coupled model multiweek forecasts. *Geophys Res Lett* 46:9213–9221. <https://doi.org/10.1029/2019GL083152>
- Anscombe, FJ (1948) The transformation of Poisson, binomial and negative-binomial data. *Biometrika* 35:246–254. <https://www.jstor.org/stable/2332343>
- Baldwin MP, Gray LJ, Dunkerton TJ et al (2001) The quasi-biennial oscillation. *Rev Geophys* 39:179–229. <https://doi.org/10.1029/1999RG000073>
- Bates BC, McCaw L, Dowdy AJ (2018) Exploratory analysis of lightning-ignited wildfires in the Warren region, Western Australia. *J Environ Manag* 225:336–345. <https://doi.org/10.1016/j.jenvman.2018.07.097>
- Bedia J, Herrera S, Gutiérrez JM, Benali A, Brands S, Mota B, Moreno JM (2015) Global patterns in the sensitivity of burned area to fire-weather: Implications for climate change. *Agr Forest Meteorol* 214–215:369–379. <https://doi.org/10.1016/j.agrformet.2015.09.002>
- Burrows N, McCaw L (2013) Prescribed burning in southwestern Australian forests. *Front Ecol Environ* 11:e25–e34. <https://doi.org/10.1890/120356>
- Cai W, Cowan T, Raupach M (2009) Positive Indian Ocean Dipole events precondition southeast Australia bushfires. *Geophys Res Lett* 36:L19710. <https://doi.org/10.1029/2009GL039902>
- Chen Y, Morton DC, Andela N, van der Werf GR, Giglio L, Rander-son JT (2017a) A pan-tropical cascade of fire driven by El Niño/Southern Oscillation. *Nat Clim Chang* 7:906–912. <https://doi.org/10.1038/s41558-017-0014-8>
- Chen Y-C, Wheeler TA, Kochenderfer MJ (2017b) Learning discrete Bayesian networks from continuous data. *J Artif Intell Res* 59:103–132. <https://doi.org/10.1613/jair.5371>
- Clark SK, Ward DS, Mahowald NM (2017) Parameterization-based uncertainty in future lightning flash density. *Geophys Res Lett* 44:2893–2901. <https://doi.org/10.1029/2018GL078294>
- Clarke H, Lucas C, Smith P (2012) Changes in Australian fire weather between 1973 and 2010. *Int J Climatol* 33:931–944. <https://doi.org/10.1002/joc.3480>
- Cobb BR, Rumf R, Salmerón A (2007) Bayesian network models with discrete and continuous variables. In: Lucas P, Gámez JA, Salmerón A (eds) *Advances in probabilistic graphical models*, Springer-Verlag, Berlin and Heidelberg, pp 81–102. https://doi.org/10.1007/978-3-540-68996-6_4
- DaSilva NTC, Fra.Paleo U, Neto JAF (2019) Conflicting discourses on wildfire risk and the role of local media in the Amazonian and temperate forests. *Int J Disaster Risk Sci* 10:529–543. <https://doi.org/10.1007/s13753-019-00243-z>
- Dowdy AJ (2016) Seasonal forecasting of lightning and thunderstorm activity in tropical and temperate regions of the world. *Sci Rep* 6:20874. <https://doi.org/10.1038/srep20874>
- Dowdy AJ (2018) Climatological variability of fire weather in Australia. *J Appl Meteor Climatol* 57:221–234. <https://doi.org/10.1175/JAMC-D-17-0167.1>
- Dowdy AJ, Mills GA (2012) Atmospheric and fuel moisture characteristics associated with lightning-attributed fires. *J Appl Meteor Climatol* 51:2025–2037. <https://doi.org/10.1175/JAMC-D-11-0219.1>
- Dunn KP, Smyth GK (1996) Randomized quantile residuals. *J Comput Graph Stat* 5:236–244. <https://doi.org/10.1080/10618600.1996.10474708>
- Dupire S, Curt T, Bigot S (2017) Spatio-temporal trends in fire weather in the French Alps. *Sci Total Environ* 595:801–817. <https://doi.org/10.1016/j.scitotenv.2017.04.027>
- Earl N, Simmonds I (2017) Variability, trends, and drivers of regional fluctuations in Australian fire activity. *J Geophys Res Atmos* 122:7445–7460. <https://doi.org/10.1002/2017JD027749>
- Fairman TA, Nitschke CR, Bennett LT (2016) Too much, too soon? A review of the effects of increasing wildfire frequency on tree mortality and regeneration in temperate eucalypt forests. *Int J Wildland Fire* 25:831–848. <https://doi.org/10.1071/WF15010>
- Fernandes R (2019) bnviewer: Interactive Visualization of Bayesian Networks. R package version 0.1.4. <https://CRAN.R-project.org/package=bnviewer>
- Fouskakis D, Draper D (2002) Stochastic optimization: a review. *Int Stat Rev* 70:315–349. <https://doi.org/10.1111/j.1751-5823.2002.tb00174.x>
- Gao P, Xu X, Zhang X (2017) On the relationship between the QBO/ENSO and atmospheric temperature using COSMIC radio occultation data. *J Atmos Sol-Terr Phys* 156:103–110. <https://doi.org/10.1016/j.jastp.2017.03.008>
- Gendreau M, Potvin J-Y (2019) Tabu search. In: Gendreau M, Potvin J-Y (eds) *Handbook of metaheuristics*, Chap. 2, Springer, Boston
- Gordon JS, Gruver JB, Flint CG, Luloff AE (2013) Perceptions of wildfire and landscape change in the Kenai Peninsula, Alaska. *Environ Manag* 52:807–820. <https://doi.org/10.1007/s00267-013-0127-4>
- Halpert MS, Ropelewski CF (1992) Surface temperature patterns associated with the Southern Oscillation. *J Climate* 5:577–593. [https://doi.org/10.1175/1520-0442\(1992\)005%3c0577:STPAWT%3e2.0.CO;2](https://doi.org/10.1175/1520-0442(1992)005%3c0577:STPAWT%3e2.0.CO;2)
- Hansen F, Matthes K, Wahl S (2016) Tropospheric QBO-ENSO interactions and differences between the Atlantic and Pacific. *J Clim* 29:1353–1368. <https://doi.org/10.1175/JCLI-D-15-0164.1>
- Harris S, Lucas C (2019) Understanding the variability of Australian fire weather between 1973 and 2017. *PLoS ONE* 14:e0222328. <https://doi.org/10.1371/journal.pone.0222328>
- Harris SL, Nicholls N, Tapper NJ (2014) Forecasting fire activity in Victoria, Australia, using antecedent climate variables and ENSO

- indices. *Int J Wildland Fire* 23:173–184. <https://doi.org/10.1071/WF13024>
- Heinze G, Wallisch C, Dunkler D (2018) Variable selection—a review and recommendations for the practicing statistician. *Biom J* 60:431–449. <https://doi.org/10.1002/bimj.201700067>
- Hendon HH, Abhik S (2018) Differences in vertical structure of the Madden–Julian oscillation associated with the Quasi–Biennial Oscillation. *Geophys Res Lett* 45:4419–4428. <https://doi.org/10.1029/2018GL077207>
- Henley BJ, Gergis J, Karoly DJ, Power S, Kennedy J, Folland CK (2015) A tripole index for the interdecadal pacific oscillation. *Clim Dyn* 45:3077–3090. <https://doi.org/10.1007/s00382-015-2525-1>
- Holz A, Paritsis J, Mundo IA et al (2017) Southern Annular Mode drives multicentury wildfire activity in southern South America. *Proc Natl Acad Sci USA* 114:9552–9557. <https://doi.org/10.1073/pnas.1705168114>
- Jain P, Wang X, Flannigan MD (2017) Trend analysis of fire season length and extreme fire weather in North America between 1979 and 2015. *Int J Wildland Fire* 26:1009–1020. <https://doi.org/10.1071/WF17008>
- Jolly WM, Cochran MA, Freeborn PH et al (2015) Climate-induced variations in global wildfire danger from 1979 to 2013. *Nat Commun* 6:7537. <https://doi.org/10.1038/ncomms8537>
- Kim H, Richter JH, Martin Z (2019) Insignificant QBO–MJO prediction skill relationship in the SubX and S25 subseasonal reforecasts. *J Geophys Res Atmos* 124:12,655–12,666. <https://doi.org/10.1029/2019JD031416>
- Kuroda Y, Yamazaki K (2010) Influence of the solar cycle and QBO modulation on the Southern Annular Mode. *Geophys Res Lett* 37:L12703. <https://doi.org/10.1029/2010GL043252>
- Lara-Estrada L, Rasche L, Sucar LE, Schneider UA (2018) Inferring missing climate data for agricultural planning using Bayesian networks. *Land* 7:4. <https://doi.org/10.3390/land7010004>
- Lestari RK, Koh T-Y (2016) Statistical evidence for asymmetry in ENSO–IOD interactions. *Atmos–Ocean* 54:498–504. <https://doi.org/10.1080/07055900.2016.1211084>
- Li J (2016) Assessing spatial predictive models in the environmental sciences: accuracy measures, data variation and variance explained. *Environ Modell Softw* 80:1–8. <https://doi.org/10.1016/j.envsoft.2016.02.004>
- Lim E, Hendon HH, Zhao M, Yin Y (2017) Inter-decadal variations in the linkages between ENSO, the IOD and south-eastern Australian springtime rainfall in the past 30 years. *Clim Dyn* 49:97–112. <https://doi.org/10.1007/s00382-016-3328-8>
- Madden RA, Julian PR (1994) Observations of the 40–50-day tropical oscillation—a review. *Mon Wea Rev* 122:814–837. [https://doi.org/10.1175/1520-0493\(1994\)122%3c0814:OOTD%3e2.0.CO;2](https://doi.org/10.1175/1520-0493(1994)122%3c0814:OOTD%3e2.0.CO;2)
- Mariani M, Fletcher M-S (2016) The Southern Annular Mode determines interannual and centennial-scale fire activity in temperate southwest Tasmania, Australia. *Geophys Res Lett* 43:1702–1709. <https://doi.org/10.1002/2016GL068082>
- Mariani M, Fletcher M-S, Holz A, Nyman P (2016) ENSO controls interannual fire activity in southeast Australia. *Geophys Res Lett* 43:10,891–10,900. <https://doi.org/10.1002/2016GL070572>
- Mariani M, Holz A, Veblen TT, Williamson G, Fletcher M-S, Bowman DMJS (2018) Climate change amplifications of climate–fire teleconnections in the Southern Hemisphere. *Geophys Res Lett* 45:5071–5081. <https://doi.org/10.1029/2018GL078294>
- Marshall GJ (2003) Trends in the southern annular mode from observations and reanalyses. *J Clim* 16:4134–4143. [https://doi.org/10.1175/1520-0442\(2003\)016%3c4134:TITSAM%3e2.0.CO;2](https://doi.org/10.1175/1520-0442(2003)016%3c4134:TITSAM%3e2.0.CO;2)
- Marshall AG, Scaife AA (2009) Impact of the QBO on surface winter climate. *J Geophys Res* 114:D18110. <https://doi.org/10.1029/2009JD011737>
- Marshall AG, Hudson D, Wheeler MC, Alves O, Hendon HH, Pook MJ, Risbey JS (2014) Intra-seasonal drivers of extreme heat over Australia in observations and POAMA-2. *Clim Dyn* 43:1915–1937. <https://doi.org/10.1007/s00382-013-2016-1>
- McArthur AG (1967) Fire Behaviour in Eucalypt Forests. Australia, Forestry and Timber Bureau, Leaflet 107, Forestry and Timber Bureau, 36 pp
- McCaw WL, Gould JS, Cheney NP, Ellis PFM, Anderson WR (2012) Changes in behaviour of fire in dry eucalypt forest as fuels increase with age. *For Ecol Manag* 271:170–181. <https://doi.org/10.1016/j.foreco.2012.02.003>
- McGee T, McFarlane B, Tymstra C (2015) Wildfire: A Canadian perspective. In: Shroder JF, Paton D (eds) Wildfire hazards, risks, and disasters, Elsevier, Amsterdam, pp 35–58. <https://doi.org/10.1016/B978-0-12-410434-1.00003-8>
- McPhaden MJ, Zebiak SE, Glantz MH (2006) ENSO as an integrating concept in earth science. *Science* 314:1740–1775. <https://doi.org/10.1126/science.1132588>
- Menezes GSC, Cazetta E, Dodonov P (2019) Vegetation structure across fire edges in a Neotropical rain forest. *For Ecol Manag* 453:117587. <https://doi.org/10.1016/j.foreco.2019.117587>
- Messier KP, Tidwell LG, Ghetu CC et al (2019) Indoor versus outdoor air quality during wildfires. *Environ Sci Technol Lett* 6:696–701. <https://doi.org/10.1021/acs.estlett.9b00599>
- Min S-K, Cai W, Whetton P (2013) Influence of climate variability on seasonal extremes over Australia. *J Geophys Res Atmos* 118:643–654. <https://doi.org/10.1002/jgrd.50164>
- Moon J-Y, Wang B, Ha K-J (2011) ENSO regulation of MJO teleconnection. *Clim Dyn* 37:1133–1149. <https://doi.org/10.1007/s00382-010-0902-3>
- Müller MM, Vacik H, Diendorfer G, Arpacı A, Formayr H, Gossow H (2013) Analysis of lightning-induced forest fires in Austria. *Theor Appl Climatol* 111:183–193. <https://doi.org/10.1007/s00704-012-0653-7>
- Murphy KP (2002) Dynamic Bayesian networks: representation, inference and learning. Dissertation, University of California, Berkeley
- Nagarajan R, Scutari M, Lèbre S (2013) Bayesian networks in R with applications in systems biology. Springer, New York
- Nigro K, Molinari N (2019) Status and trends of fire activity in southern California yellow pine and mixed conifer forests. *For Ecol Manag* 441:20–31. <https://doi.org/10.1016/j.foreco.2019.01.020>
- Noble IR, Bary GAV, Gill AM (1980) McArthur’s fire-danger meters expressed as equations. *Aust J Ecol* 5:201–203. <https://doi.org/10.1111/j.1442-9993.1980.tb01243.x>
- Pan X, Chin M, Ichoku CM, Field RD (2018) Connecting Indonesian fires and drought with the type of El Niño and phase of the Indian Ocean Dipole during 1979–2016. *J Geophys Res Atmos* 123:7974–7988. <https://doi.org/10.1029/2018JD028402>
- Petropoulos F, Goodwin P, Fildes R (2017) Using a rolling training approach to improve judgmental extrapolations elicited from forecasters with technical knowledge. *Int Forecast* 33:314–324. <https://doi.org/10.1016/j.ijforecast.2015.12.006>
- Pineda N, Montanyà J, van der Velde OA (2014) Characteristics of lightning related to wildfire ignitions in Catalonia. *Atmos Res* 135–136:380–387. <https://doi.org/10.1016/j.atmosres.2012.07.011>
- Power S, Casey T, Folland C, Colman A, Mehta V (1999) Inter-decadal modulation of the impact of ENSO on Australia. *Clim Dyn* 15:319–324. <https://doi.org/10.1007/s003820050284>
- Rao MP, Cook ER, Cook BI, Anchukaitis KJ, D’Arrigo RD, Krusic PJ, LeGrande AN (2019) A double bootstrap approach to Superposed Epoch Analysis to evaluate response uncertainty. *Dendrochronologia* 55:119–124. <https://doi.org/10.1016/j.dendro.2019.05.001>
- Riemann-Campe K, Fraedrich K, Lunkeit F (2009) Global climatology of convective available potential energy (CAPE) and convective inhibition (CIN) in ERA-40 reanalysis. *Atmos Res* 93:534–545. <https://doi.org/10.1016/j.atmosres.2008.09.037>

- Saji NH (2019) The Indian Ocean dipole. Oxford Res Encycl Clim Sci Oxford Univ Press USA. <https://doi.org/10.1093/acrefore/9780190228620.013.619>
- Saji NH, Goswami BN, Vinayachandran PN, Yamagata T (1999) A dipole mode in the tropical Indian Ocean. *Nature* 401:360–363. <https://www.nature.com/articles/43854>
- Santoso A, Cai W, Collins M, McPhaden M, Jin F-F, Guilyardi E, Vecchi G, Dommenges D, Wang G (2015) ENSO extremes and diversity: dynamics, teleconnections, and impacts. *Bull Amer Meteorol Soc* 96:1969–1972. <https://doi.org/10.1175/BAMS-D-15-00141.1>
- Schirber S (2015) Influence of ENSO on the QBO: results from an ensemble of idealized simulations. *J Geophys Res Atmos* 120:1109–1122. <https://doi.org/10.1002/2014JD022460>
- Scutari M (2010) Learning Bayesian Networks with the bnlearn R Package. *J Stat Softw* 35:1–22. <https://www.jstatsoft.org/article/view/v035i03>
- Scutari M, Graafland CE, Gutiérrez JM (2019) Who learns better Bayesian network structures: accuracy and speed of structure learning algorithms. *Int J Approx Reason* 115:235–253. <https://doi.org/10.1016/j.ijar.2019.10.003>
- Shmueli G (2010) To explain or to predict? *Stat Sci* 25:289–310. <https://doi.org/10.1214/10-STS330>
- Sucar LE (2015) Probabilistic graphical models. Springer-Verlag, London
- Tashman LJ (2000) Out-of-sample tests of forecasting accuracy: an analysis and review. *Int J Forecast* 16:437–450. [https://doi.org/10.1016/S0169-2070\(00\)00065-0](https://doi.org/10.1016/S0169-2070(00)00065-0)
- Thode HC (2002) Testing for Normality. Marcel Dekker, New York
- Thompson DWJ, Baldwin MP, Solomon S (2005) Stratosphere-troposphere coupling in the Southern Hemisphere. *J Atmos Sci* 62:708–715. <https://doi.org/10.1175/JAS-3321.1>
- Torrence C, Compo GP (1998) A practical guide to wavelet analysis. *Bull Am Meteorol Soc* 79:61–78. [https://doi.org/10.1175/1520-0477\(1998\)079%3c0061:APGTWA%3e2.0.CO;2](https://doi.org/10.1175/1520-0477(1998)079%3c0061:APGTWA%3e2.0.CO;2)
- Trouet V, Taylor AH, Wahl ER, Skinner CN, Stephens SL (2010) Fire-climate interactions in the American West since 1400 CE. *Geophys Res Lett* 37:L04702. <https://doi.org/10.1029/2009GL041695>
- Úbeda X, Sarricolea P (2016) Wildfires in Chile: a review. *Glob Planet Change* 146:152–161. <https://doi.org/10.1016/j.gloplacha.2016.10.004>
- Wheeler MC, Hendon HH (2004) An all-season real-time multivariate MJO index: development of an index for monitoring and prediction. *Mon Wea Rev* 132:1917–1932. [https://doi.org/10.1175/1520-0493\(2004\)132%3c1917:AARMMI%3e2.0.CO;2](https://doi.org/10.1175/1520-0493(2004)132%3c1917:AARMMI%3e2.0.CO;2)
- Wheeler MC, Hendon HH (2009) Impacts of the Madden-Julian Oscillation on Australian rainfall and circulation. *J clim* 22:1482–1498. <https://doi.org/10.1175/2008JCLI2595.1>
- Williamson GJ, Prior LD, Jolly WM, Cochrane MA, Murphy BP, Bowman DMJS (2016) Measurement of inter- and intra-annual variability of landscape fire activity at a continental scale: the Australian case. *Environ Res Lett* 11:1–12. <https://doi.org/10.1088/1748-9326/11/3/035003>
- Yamagata T, Behera SK, Luo J-J, Masson S, Jury MR, Rao SA (2004) Coupled ocean-atmosphere variability in the Indian Ocean. In: Wang C, Xie S-P, Carton JA (eds) *Earth's climate—The ocean-atmosphere interaction*. American Geophysical Union, Washington DC, pp 189–211
- Zhang C (2013) Madden-Julian oscillation: Bridging weather and climate. *Bull Am Meteorol Soc* 94:1849–1870. <https://doi.org/10.1175/BAMS-D-12-00026.1>
- Zhang C, Hendon HH, Kessler WS, Rosati A (2001) A workshop on the MJO and ENSO. *Bull Am Meteorol Soc* 82:971–976
- Zhang H, Chen K, Han F (2018) An increasing threat of wildfire to human health. *Curr Pollut Rep* 4:56–57. <https://doi.org/10.1007/s40726-018-0077-9>
- Zhou Y, Nigam S (2015) The Indian Ocean Dipole: a monopole in SST. *J Clim* 28:3–19. <https://doi.org/10.1175/JCLI-D-14-00047.1>

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.