Environmental Modelling & Software 82 (2016) 108–120

Contents lists available at ScienceDirect

Environmental Modelling & Software

journal homepage: www.elsevier.com/locate/envsoft

Modelling spatial and temporal changes with GIS and Spatial and Dynamic Bayesian Networks

Yung En Chee [a, b, *], Lauchlin Wilkinson [c], Ann E. Nicholson [c], Pedro F. Quintana-Ascencio [d], John E. Fauth [d], Dianne Hall [e], Kimberli J. Ponzio [e], Libby Rumpff [a]

[a] School of BioSciences, The University of Melbourne, Parkville, VIC 3010, Australia
[b] School of Ecosystem and Forest Sciences, The University of Melbourne, 500 Yarra Boulevard, Burnley, VIC 3121, Australia
[c] Faculty of Information Technology, Monash University, PO Box 197, Caulfield East, VIC 3145, Australia
[d] Department of Biology, University of Central Florida, 4000 Central Florida Blvd, Orlando, FL 32816-2368, USA
[e] Division of Water Resources, St. Johns River Water Management District, PO Box 1429, Palatka, FL 32178-1429, USA

ARTICLE INFO

Article history:
Received 28 January 2016
Received in revised form
11 April 2016
Accepted 12 April 2016

Keywords:
Probabilistic graphical models
State-and-transition models
Object-oriented
Adaptive management
Willow

ABSTRACT

State-and-transition models (STMs) have been successfully combined with Dynamic Bayesian Networks (DBNs) to model temporal changes in managed ecosystems. Such models are useful for exploring *when* and *how* to intervene to achieve the desired management outcomes. However, knowing *where* to intervene is often equally critical. We describe an approach to extend state-and-transition dynamic Bayesian networks (ST-DBNs) — incorporating spatial context via GIS data and explicitly modelling spatial processes using spatial Bayesian networks (SBNs). Our approach uses object-oriented (OO) concepts and exploits the fact that ecological systems are hierarchically structured. This allows key phenomena and ecological processes to be represented by hierarchies of components that include similar, repetitive structures. We demonstrate the generality and power of our approach using two models — one developed for adaptive management of eucalypt woodland restoration in south-eastern Australia, and another developed to manage the encroachment of invasive willows into marsh ecosystems in east-central Florida.

1. Introduction

Bayesian networks (Pearl, 1988) are increasingly popular for ecological and environmental modelling, decision support and adaptive management (Nyberg et al., 2006; Korb and Nicholson, 2010; Aguilera et al., 2011). Ecosystem management problems characteristically involve variable, complex and imperfectly understood biophysical, social and economic interactions. The iterative knowledge-engineering process of developing BNs is invaluable for: a) clarifying objectives; b) identifying and articulating alternatives; c) synthesising available knowledge; d) quantifying uncertainties and d) pinpointing critical assumptions to be tested by purposeful monitoring. When fully parameterised, such models help us explore and (where possible) resolve uncertainty about the consequences of management decisions. This is integral to adaptive management (*sensu* Holling, 1978; Walters and Hilborn, 1978) which supplies the broader framework for evaluating the performance of decision actions and updating our knowledge base to improve future management (Nichols and Williams, 2006; Duncan and Wintle, 2008).

Despite the obvious value of using BNs to support learning over time for adaptive management (see e.g., Ames et al., 2005; Chee et al., 2005), most published examples of BNs for environmental applications have focused on formalising static conceptual models of the system in question, and do not explicitly represent ongoing dynamics (e.g. multiple time steps and sequential decisions) (Barton et al., 2012). Examples that incorporate spatiality explicitly are even rarer. Yet it is critical to address these gaps because the ability to understand change over time, and to account for spatial context and interactions is often necessary for meaningful decision support.

For instance, in our eucalypt woodlands case study, restoring species composition, ecosystem structure and function is a long-

* Corresponding author.

http://dx.doi.org/10.1016/j.envsoft.2016.04.012
1364-8152/

term undertaking that needs to effectively manage threats like weed establishment, so that the recovery process can build upon successive gains. In our invasive willows management case study, control efforts are long-term because adult willows have become firmly established within the catchment. In both cases, spatial considerations are crucial because the encroachment of weeds (in woodlands) and willow seedlings (in marsh ecosystems) depends on seed production and dispersal from surrounding areas, and spatial characteristics also determine the applicability and effectiveness of management actions.

State-and-transition dynamic Bayesian networks (ST-DBNs) as described by Nicholson and Flores (2011) provide a viable approach for explicitly modelling change over time. Here, we extend the capabilities of ST-DBNs – first, coupling them to GIS data so we can harness spatially relevant data, and then explicitly modelling key spatial processes using spatial Bayesian networks (SBNs). Our approach makes use of object-oriented (OO) concepts and exploits the fact that ecological systems are hierarchically structured such that key phenomena and processes of interest can be represented by nesting components that include similar, repetitive structures.

First, we explain the 'buildings blocks' and concepts of the tools we use for modelling spatial and temporal changes with BNs. We then present and illustrate our approach using two models—one developed for adaptive management of eucalypt woodland restoration in south-eastern Australia ('Woodlands weed' model, Rumpff et al. (2011)), and another developed to manage willow spread into marsh ecosystems in east-central Florida, USA ('Willows' model, Wilkinson et al. (2013)). Of course, incorporating spatial context and processes can lead to a massive increase in the size and complexity of the networks, which in turn generates computational issues and difficulties with the probabilistic updating—we discuss our approach to handling these challenges and provide a generic system architecture, templates and algorithms for combining GIS, object-oriented spatial BNs and object-oriented state-transition DBNs.

To our knowledge, this is the first demonstration of the integration of these three tools. This novel and powerful approach allows the incorporation of spatial context where it is critical for decision-making.

2. Background: building blocks and OO concepts

State-and-transition models (**STMs**) are management-focused, qualitative conceptual models that synthesise knowledge about an ecological system, in the form of observed and/or hypothesised system states and transitions that are of management interest (Westoby et al., 1989; Jackson et al., 2002). STMs are a popular tool for modelling changes over time in ecological systems that have clear transitions between distinct states. They combine graphical depiction of transitions and their causal factors with tables of qualitative descriptions of the transitions. They have been widely applied both to understand and help manage vegetation change in ecosystems such as rangelands (e.g., Westoby et al., 1989; Bestelmeyer et al., 2003; Bashari et al., 2009), grasslands (e.g., Sadler et al., 2010) and woodlands (e.g., Yates and Hobbs, 1997b; Rumpff et al., 2011).

Bayesian networks (**BNs**) are graphical models of cause-effect relationships used for reasoning under uncertainty. More formally, a Bayesian network (Pearl, 1988) is a directed, acyclic graph whose nodes represent the random variables in the problem. A set of directed arcs connect pairs of vertices, representing the direct dependencies of variables. The set of nodes pointing to X are called its parents and is denoted $pa(X)$. BNs display key variables in the system succinctly, show which variables are linked and how the causal chain or argument links events to outcomes of interest. The relationship between variables is quantified by conditional probability tables (CPTs) associated with each node, namely $P(X|pa(X))$. The CPTs together compactly represent the full joint distribution. Users can set the values of any combination of nodes in the network that they have observed. This evidence, e, propagates through the network, producing a new posterior probability distribution $P(X|e)$ for each variable in the network. There are a number of efficient exact and approximate inference algorithms for performing this probabilistic updating, providing a powerful combination of predictive, diagnostic and explanatory reasoning.

Dynamic Bayesian Networks (**DBNs**) are a variant of ordinary BNs (Dean and Kanazawa, 1989; Kjærulff, 1992; Nicholson, 1992) that explicitly model changes over time and can be used to model feedback functions in problem contexts where this is important. A typical DBN has nodes for N variables of interest and for each domain variable X_1, there is one copy for each *time slice* for interest: X_i^T, X_i^{T+1}, X_i^{T+2} etc. Links in a DBN include those between nodes in the same time slice, and those in the next time slice. Of the latter, temporal arcs may link the same variable over time, $X_i^T \rightarrow X_i^{T+1}$, and different variables over time, $X_i^T \rightarrow X_j^{T+1}$. Environmental applications employing DBNs are scarce (e.g., Shihab and Chalabi, 2007; Dawsey et al., 2007; Shihab, 2008). This may be because they are perceived to be "very tedious" (Uusitalo, 2007), or because DBN algorithms are available only in software resulting from research projects[1], with DBN functionality less well supported in popular commercial products.

State-and-transition Dynamic Bayesian Networks (**ST-DBNs**) combine the advantages of graphical visualisation of transitions and their influencing factors with quantitative representation of dependencies and uncertainty, along with explicit representation of time. Our example models are based on Nicholson and Flores (2011)'s template.

S^T represents the state of the system, has n possible values $s_1 \ldots s_n$, and may directly influence any of the environmental and management factors, which are divided into m main factors, $F_1, \ldots, F_m$ (which directly influence transitions) and other sub-factors, $X_1, \ldots, X_r$ (which influence the main factors).

Transition nodes, $ST_1, \ldots, ST_n$, represent the transitions from each state s_i. Each has at most $n+1$ values (usually fewer), one for each "next" state plus "impossible", giving explicit modelling of impossible transitions. Like ordinary DBNs, there is an implied δT, which can be included explicitly as a parent of all the ST nodes, if the time step varies. Each transition node ST has only some of the causal factors as parents. The CPT for the ST node is just a partition of the corresponding CPT if the problem was represented as an ordinary DBN, without the transition nodes. The next state node, S^{T+1}, has to combine the results of all the different transition nodes, given the starting state S, and thus has $n+1$ parents. However, the relationship between the transition nodes and S^{T+1} is deterministic, so the CPT can be generated from a straightforward equation.

It is important to note that ST-BNs that explicitly model all the transitions, only remain tractable when there are natural constraints in the domain; that is, if the number of transitions from each state is limited and only influenced by a small number of causal factors such that the underlying state transition matrix for S is sparse (Nicholson and Flores, 2011).

2.1. How does object-oriented (OO) thinking help?

The complexity of ecological systems is such that representing even a moderate degree of ecological realism tends to lead to large networks. The resulting visual 'clutter' of large networks makes

[1] e.g. BNT, code.google.com/p/bnt.