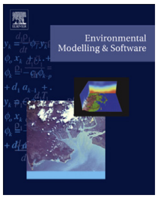
Environmental Modelling & Software xxx (2014) 1–11

Contents lists available at ScienceDirect

Environmental Modelling & Software

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



Modelling uncertainty in social–natural interactions☆

R.F. Ropero [a], R. Rumí [b,] *, P.A. Aguilera [a]

[a] *Informatics and Environment Laboratory, Dept. of Biology and Geology, University of Almería, Spain*
[b] *Dept. of Mathematics, University of Almería, Spain*

ARTICLE INFO

Article history:
Received 7 June 2013
Received in revised form
16 June 2014
Accepted 14 July 2014
Available online xxx

Keywords:
Systemic change
Socio-ecological system
Water flows
Hybrid Bayesian networks
Mixtures of truncated exponentials

ABSTRACT

Socio-ecological systems can be represented as a complex network of causal interactions. Modelling such systems requires methodologies that are able to take uncertainty into account. Due to their probabilistic nature, Bayesian networks are a powerful tool for representing complex systems where interactions between variables are subject to uncertainty. In this paper, we study the interactions between social and natural subsystems (land use and water flow components) using hybrid Bayesian networks based on the Mixture of Truncated Exponentials model. This study aims to provide a new methodology to model systemic change in a socio-ecological context. Two endogenous changes – agricultural intensification and the maintenance of traditional cropland – are proposed. Intensification of the agricultural practices leads to a rise in the rate of immigration to the area, as well as to greater water losses through evaporation. By contrast, maintenance of traditional cropland hardly changes the social structure, while increasing evapotranspiration rates and improving the control over runoff water. These results indicate that hybrid Bayesian networks are an excellent tool for modelling social–natural interactions.

1. Introduction

Nature and society are clearly related and so any delimitation between natural and social systems is artificial and arbitrary (Berkes and Folke, 1998). Instead, they should be considered as a complex system of interactions operating on different scales; this is referred to as a Socio-Ecological System (SES) (Anderies et al., 2004; Cadenasso et al., 2006; Folke, 2006). In the context of SES, a systemic change can be defined as a fundamental change in the interactions within a system, arising either from an external hazard event or from gradual endogenous change, which leads to a shift in the state of the system to another with new properties (Kinzig et al., 2006; Filatova and Polhill, 2012).

Graphically the SES can be represented as a network of nodes (social and natural components), with a number of links between them. When a hazard event occurs or a component undergoes gradual change, the change can be propagated through the entire system by means of cause–effect interactions between the components of the SES. These types of interactions are subject to the uncertainty inherent in the system (Clark, 2002; Refsgaard et al., 2007). This uncertainty can be modelled using probability theory (Ricci et al., 2003; Walker et al., 2003; Refsgaard et al., 2007; Warmink et al., 2010).

Bayesian networks (BNs) are considered one of the most powerful tools for representing complex systems of causal interactions between variables that are subject to uncertainty (Borsuk et al., 2004; Jensen and Nielsen, 2007; Pourret et al., 2008; Korb and Nicholson, 2011; Carmona et al., 2013; Kelly et al., 2013; Landuyt et al., 2013; Nash et al., 2013). Their graphical structure allows stakeholders to understand the relationships easily. In addition, they provide stakeholders with a participatory framework because the learnt model can be refined manually by adding or removing arcs (or even variables) from the graph to better represent reality (Voinov and Bousquet, 2010).

BNs models have already been successfully applied in environmental modelling (Aguilera et al., 2011; Kelly et al., 2013; Landuyt et al., 2013; Dyer et al., 2014). Whereas BNs usually estimate models using discrete domains, most environmental and social variables are continuous. A common solution is to discretize continuous variables, but this involves some loss of statistical information (Uusitalo, 2007). Estimation of a model directly from the original discrete and continuous (hybrid) data returns a more accurate model. In turn, this model is able to report more specific answers to the proposed scenarios. The main problem when dealing with hybrid BNs is that, initially, there is no common structure to represent the distribution of the variables. The *Mixtures*

☆ Special Issue on Modelling Systemic Change.

* Corresponding author.

E-mail addresses: rosa.ropero@ual.es (R.F. Ropero), rrumi@ual.es (R. Rumí), aguilera@ual.es (P.A. Aguilera).

http://dx.doi.org/10.1016/j.envsoft.2014.07.008
1364-8152/© 2014 Elsevier Ltd. All rights reserved.

of Truncated Exponentials (MTE) model (Moral et al., 2001) provides us with a common structure to represent both the discrete and continuous variables simultaneously, in such a way that all the computations needed to perform probability propagation in the model can be done using the same structure (Moral et al., 2001). The versatility of BNs allows any statistic of interest to be calculated from the variables, including the probability of extreme values.

Two of the main challenges in SES research (Filatova and Polhill, 2012) are: (i) to accommodate the study of systemic change while taking uncertainty into account (Clark, 2002), and (ii) to represent the new state of the system after systemic change has been propagated (Filatova and Polhill, 2012). Since BNs are modelled by means of probability distributions, risk and uncertainty can be estimated more accurately than by using models which only consider mean values (Uusitalo, 2007). They allow a system to be represented both in its current state (*a priori*), and *a posteriori*, once the change has been propagated through the system, using the probability distribution functions of the variables. Their main purpose is to provide a framework for efficient reasoning about the system they represent, in terms of updating information about unobserved variables, when new information (changes to a single or several observed variables) is incorporated to the system. This is known as probability propagation or probabilistic inference. However, not every change included into a component of the system (one or more variables) will lead to systemic change because some components may be conditionally independent. This property is expressed in the graph by means of the d-separation concept (see Section 2 for a more detailed explanation).

1.1. *Outline of the paper*

Global socio-economic changes affect regional and local socio-economic structures (Lambin et al., 2001; Foley et al., 2005) and lead to changes in land uses in the landscape (Schmitz et al., 2005; Caillault et al., 2013) and in the structure and functionality of natural ecosystems (Matson et al., 1997; Foley et al., 2005; Rudel et al., 2009). One of the main effects of these changes relates to the behaviour of water flows (Scanlon et al., 2005; Maes et al., 2009; Toda et al., 2010; Park et al., 2014). The concepts of green and blue water flows were defined to introduce the whole water cycle into water management plans (Falkenmark, 1997; Rockstroem, 2000). Blue water is the amount of rainfall that exceeds the soil's storage capacity and feeds rivers, lakes and aquifers. Green water refers to the rainfall that infiltrates into the root zone of the soil to support the primary productivity of natural and agricultural systems through evapotranspiration (Falkenmark, 1997; Falkenmark and Folke, 2002). Green and blue water flow through natural subsystems across the landscape, participating in several ecological processes; as a result, there is a clear interaction between land use and green and blue water flows (Willaarts et al., 2012). The characteristics of soil and the type and cover of vegetation determine the amount of water that evaporates back to the atmosphere, infiltrates into the soil or flows away as runoff (Falkenmark, 2003; Willaarts et al., 2012). For example, tropical forest and Mediterranean pasture have high green water flows, whilst urban land and irrigated herbaceous croplands have high blue water flows (Rockstroem and Gordon, 2001; Willaarts, 2009).

In this paper, we study a Spanish catchment as a SES. Three variables were selected to represent the socioeconomic subsystem while land use and green and blue water flows were selected to represent the natural subsystem.

The aim of the study is to demonstrate the ability of hybrid BNs to model systemic change. We develop a new methodology, which considers the tails of the probability distribution functions to identify systemic change, and we carry out statistical tests to differentiate between different states of the system. By this means, we provide the expert with a set of tools to help assess systemic change.

2. Methodology

2.1. *Study area*

The study area comprises the catchment of the river Adra in south-eastern Spain (Fig. 2). It is bounded to the north by the Sierra Nevada, to the south by the Mediterranean Sea, to the aast by the Sierra de Gádor, and to the west by the Sierra Filabres. It occupies 74.400 ha, and supports an estimated population of 124.000 people distributed over fourteen municipalities.

The landscape of the Sierra Nevada mountain range is characterized by dense woodland, mainly oaks and conifers species with Mediterranean scrubland. In the upper reaches, the original Mediterranean forest with oaks remains, whilst the scrubland is the result of several episodes of deforestation (García-Latorre and Sánchez-Picón, 2001). The socioeconomy is characterized by several small municipalities accommodating and ageing population with a high rate of migration. In the foothills of Sierra Nevada, mixed and irrigated crops replace woodland and the social structure indicates a slightly younger population, though still with a high migration rates.

In the east of the area, in the Sierra de Gádor foothills, land uses comprise traditional croplands including olive and almond groves with patches of woodland and scrub, creating a complex and heterogeneous landscape. The socioeconomy is characterized by depopulation and an older population.

In the middle and west of the study area, the landscape is composed by scrubs and some patches of woodland whose configuration was determined by historical trends in the 19th century (mining and the deforestation of natural forest) (García-Latorre and Sánchez-Picón, 2001).

In the lower reaches, the land uses are intensive agriculture with greenhouses and irrigated crops, mixed with scrubland. Immigration rate is significant given the incoming of a new workforce to the greenhouses.

2.2. *Data collection*

Taking into account socio-economic characteristics of the study area (CCA, 2007; Camarero et al., 2009), three representative variables (ageing, emigration and immigration rates) were selected. Data on these variables were obtained for each municipality from the Andalusian Statistical Institute (Fig. 1 *i)*). The ageing component was summarized by calculating the percentage of people older than 65 years old, while emigration and immigration rates were calculated as percentages of the total population.

The BalanceMEd model (Willaarts, 2009; Willaarts et al., 2012) was applied to calculate the water flows described above (Fig. 1 *i)*). This is a semi-deterministic model developed to quantify hydrological functioning in Mediterranean catchments using long time series of monthly rainfall and potential evapotranspiration data. The model assumes that a fraction from the total precipitation is intercepted by vegetation or soil and evaporates directly as a Non Productive Green Water (NPGW). Another fraction from the total precipitation can be intercepted on impermeable surfaces and is returned to the atmosphere as Consumptive Blue Water (CBW). The remaining precipitation reaches the soil and is taken up by plants and transpired, this portion is termed Productive Green Water flow (PGW). When the infiltrated water exceeds the soil storage capacity, it can either percolate or drain as Runoff Blue Water (RBW). In the specific case of greenhouse crops, we consider that the concept of PGW is not applicable since the crops are irrigated from groundwater flows rather than from direct precipitation. Moreover, evaporative flows are difficult to evaluate under a greenhouses cover. For that reason, in this specific case, we focus on CBW when considering greenhouse crops as the land use.

Nine land uses representative of the study area landscape were selected (Table 2). These data were obtained from the Land Use and Land Cover shape file from Andalusian Regional Government using ArcGis v.9.3.1 (ESRI, 2006) (Fig. 1 *i)*). They are expressed as a discrete variable which represents the presence of each land uses as a percentage.

Table 1 shows the main statistics of the continuous socioeconomic and water flow variables in the data set. The land use variable is discrete and the description of each category and its percentage in the data set are shown in Table 2.

2.3. *Model description*

A BN is a statistical multivariate model for a set of variables $\mathbf{X} = (X_1, \ldots, X_n)$, which is defined in terms of two parts (Jensen and Nielsen, 2007)[1]:

1. Qualitative part: A directed acyclic graph (Fig. 3(a)) where each vertex represents a variable in the model and each edge, linking two variables, represents the

---

[1] Capital letters e.g. X_i represent random variables, whilst lower case letter, e.g. x_i represent values of the corresponding variables.

Please cite this article in press as: Ropero, R.F., et al., Modelling uncertainty in social–natural interactions, Environmental Modelling & Software (2014), http://dx.doi.org/10.1016/j.envsoft.2014.07.008