Ecological Informatics 41 (2017) 33–39

Contents lists available at ScienceDirect

Ecological Informatics

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

A comparison of two sensitivity analysis techniques based on four bayesian models representing ecosystem services provision in the Argentine Pampas

Florencia Rositano[a,b,*], Gervasio Piñeiro[a,c], Federico E. Bert[b], Diego O. Ferraro[a,b]

[a] IFEVA, Universidad de Buenos Aires, CONICET, Facultad de Agronomía, Av. San Martín 4453 (C1417DSE), Buenos Aires, Argentina
[b] Universidad de Buenos Aires, Facultad de Agronomía, Departamento de Producción Vegetal, Cátedra de Cerealicultura, Av. San Martín 4453 (C1417DSE), Buenos Aires, Argentina
[c] Universidad de Buenos Aires, Facultad de Agronomía, Departamento de Recursos Naturales y Ambiente, Cátedra de Ecología, Av. San Martín 4453 (C1417DSE), Buenos Aires, Argentina

ARTICLE INFO

Keywords:
Sensitivity analyses
Probabilistic models
Bayesian Networks
One-at-a-time
Sensitivity to findings

ABSTRACT

Sensitivity analyses (SAs) identify how an output variable of a model is modified by changes in the input variables. These analyses are a good way for assessing the performance of probabilistic models, like Bayesian Networks (BN). However, there are several commonly used SAs in BN literature, and formal comparisons about their outcomes are scarce. We used four previously developed BNs which represent ecosystem services provision in Pampean agroecosystems (Argentina) in order to test two local sensitivity approaches widely used. These SAs were: 1) One-at-a-time, used in BNs but more commonly in linear modelling; and 2) Sensitivity to findings, specific to BN modelling. Results showed that both analyses provided an adequate overview of BN behaviour. Furthermore, analyses produced a similar influence ranking of input variables over each output variable. Even though their interchangeably application could be an alternative in our bayesian models, we believe that OAT is the suitable one to implement here because of its capacity to demonstrate the relation (positive or negative) between input and output variables. In summary, we provided insights about two sensitivity techniques in BNs based on a case study which may be useful for ecological modellers.

1. Introduction

Bayesian Networks (BN) consist on a set of variables with a probabilistic distribution, and their outcome assesses how likely events are and how these probabilities change with external interventions (Jensen and Nielsen, 2007; Korb and Nicholson, 2004). A BN can be represented visually as a set of nodes connected by direct links (Fig. 1). Nodes represent variables and the probability distribution of their possible states, while links represent causal relationships between nodes (Kristensen and Rasmussen, 2002). Nodes with no incoming arrows are parent nodes (i.e. input variables); while nodes with incoming arrows are child nodes (i.e. parameters) (McCann et al., 2006). Each node can take different states (e.g. high/medium/low) which are clusters delimited by intervals or ranges (Fig. 1). The number of states is dependent on the information conveyed and the possible values that they can get (Dlamini, 2010). Parent nodes have marginal probabilistic distributions that represent the frequency of each state, while child nodes are characterized by a conditional probability table that represents a factorial combination of its parent nodes along with their probabilistic values (Chen and Pollino, 2012).

Currently, BNs are an increasingly accepted method for modelling uncertain and complex domains, such as ecosystems (Uusitalo, 2007). The conceptual representation of BN results (i.e. graphical networks) is very useful for an intuitive presentation of functional relationships within complex systems. Their advantages are commonly related to the flexibility for dealing with both expert knowledge and system uncertainty (Borsuk et al., 2004; Castelletti and Soncini-Sessa, 2007). BNs have been used for modelling in a wide range of disciplines like psychology (López Puga et al., 2007), education (García et al., 2007), ecological risk assessment (Pollino et al., 2007), agroecosystems sustainability (Ticehurst et al., 2007) and ecosystem services provision (Rositano and Ferraro, 2014), among others. Regarding natural resource management, BNs are able to both capture the influence of management decisions on key ecological variables, and to help decision makers on selecting the best course of action (McCann et al., 2006).

As in other modelling methodologies, BNs require the assessment of their performance. Validation is “a demonstration that a model within its domain of applicability has a satisfactory range of accuracy consistent with the intended application of the model” (Rykiel, 1996). Model validation is not an easy process and, as a consequence, should

* Corresponding author at: Cátedra de Cerealicultura, Facultad de Agronomía, Universidad de Buenos Aires, Av. San Martín 4453 (C1417DSE), Buenos Aires, Argentina.
E-mail address: rositano@agro.uba.ar (F. Rositano).

http://dx.doi.org/10.1016/j.ecoinf.2017.07.005
Received 3 November 2016; Received in revised form 17 April 2017; Accepted 29 July 2017
Available online 31 July 2017
1574-9541/

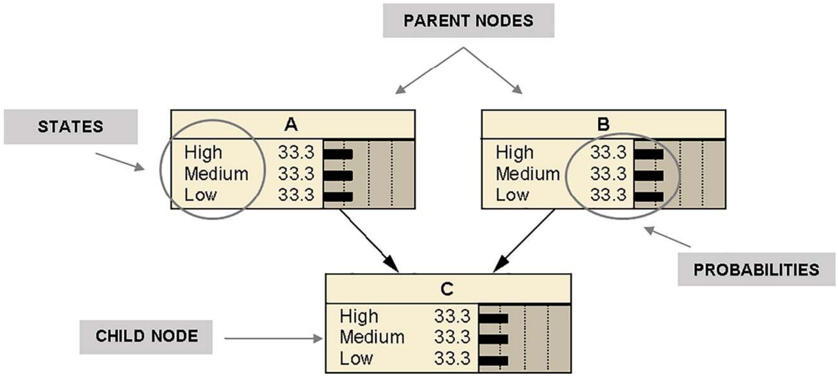

Fig. 1. Example of a Bayesian Network with three variables or nodes (A, B and C). Nodes A and B are parent nodes, while node C is a child node. Each node has three states (High, Medium and Low) with a uniform probability distribution.

be done with multiple strategies (Bert et al., 2014). Rykiel (1996) stated that sensitivity analyses (SAs) could be considered a strategy of model validation. Results of SAs are able to highlight the critical aspects of model development and data collection by identifying the impact of a change in input variables over the output variable (Newham et al., 2003; Thogmartin, 2010). Two groups of SAs are recognized: local and global (Saltelli et al., 2000). In local SA, parameter values are changed one at a time, while fixing all other variables. These SAs are not able to capture potential interactions among input variables as well as they partially explore parametric aspects (Cariboni et al., 2007; Hu et al., 2015; Saltelli and Annoni, 2010). Global SA involves varying all or several input variables at the same time, thus allowing identification of non-linear interactions among parameters (Confalonieri et al., 2010; Mackler-Pick et al., 2011). Lee et al. (2015) describe many techniques to carry out global SA. Taking this into account, environmental modellers need to be aware about the particularities of sensitivity methodologies in order to conduct a proper validation process (Cariboni et al., 2007).

Validating BNs is not simple to carry out (Payraudeau and van der Werf, 2005). In current practice, if a user has sufficient data on the phenomenon of interest, this data may be used to validate model predictions. However, BNs are commonly used to model complex systems with limited data (Chen and Pollino, 2012). Because of this, expert opinion could be an option to validate the structure, discretization and parameterization of bayesian models (Korb and Nicholson, 2004). Although expert test is quite simple, it is not sufficient to verify model validity in an independent way (Pitchforth and Mengersen, 2013). Aguilera et al. (2011) reviewed the use of BNs for environmental modelling and highlighted that ca. 40% of the studies showed no type of model validation, while only 13% of the models reviewed were validated through any kind of SA, like variance reduction (e.g. Marcot et al., 2006; Stelzenmüller et al., 2010), one-at-a-time (e.g. Bednarski et al., 2004; Chan and Darwiche, 2004; Coupé and van der Gaag, 2002; Coupé et al., 1999), sensitivity to findings (e.g. Chen and Pollino, 2012; Grêt-Regamey and Straub, 2006; Marcot, 2012; Pollino et al., 2007; Smith et al., 2007) or Latin hypercube sampling method (e.g. Borsuk et al., 2004). One-at-a-time (OAT) is the simplest methodology in order to obtain the effect of variation of parameter estimate on posterior probabilities (Coupé et al., 1999). Nonetheless, some authors have pointed out that this SA is not suitable for probabilistic methodologies (Chen and Pollino, 2012). A SA currently available in BN software packages, like *Hugin* (Madsen et al., 2005) or *Netica* (Norsys Software Corp., 2009), is "Sensitivity to findings" (STF) which is able to assess how much a finding at one variable will likely change the beliefs at another variable (Korb and Nicholson, 2004). It should be carried out with the BN previously populated since results change according to the quantitative information included into the model; therefore, this analysis is recalculated each time new information is collected. As well as OAT, this SA is only done to one variable at a time (Uusitalo, 2007). Despite conflicting opinions on which SA is the most appropriate (Saltelli and Annoni, 2010), BN modellers should be aware about advantages and disadvantages when using each approach.

In BN literature, both kinds of SAs have been used to evaluate bayesian models; however, their comparison is lacking. A case study could be useful for doing a first attempt to highlight differences and similarities between these SAs. For that reason, we used previously developed BNs originally applied for assessing four ecosystem services provision (i.e. Soil Carbon balance, Soil Nitrogen balance, N_2O emission control, and Groundwater contamination control) in the Pampa region (Argentina) (Rositano and Ferraro, 2014). Therefore, the objective of this paper was to evaluate and compare the information provided by two local SAs: one used in BNs but more commonly in linear modelling (OAT), and one specific for BN modelling (STF).

2. Methodology

2.1. Bayesian models development

Ecosystem services (ES) offer the possibility to evaluate changes in ecosystems caused by human action and to resolve conflicts arised by different land uses (Vihervaara et al., 2010). In this sense, Rositano and Ferraro (2014) developed a framework to assess changes in ES provision as a consequence of environmental variability and agricultural management practices in Pampean agroecosystems (Argentina). The framework was based on two tools capable of dealing with ecosystems complexity and uncertainty: conceptual networks and probabilistic networks (i.e. BNs).

First, a conceptual network was developed representing the set of environmental and productive variables that determine the provision of eight ES in the Pampa Region. ES selected were: 1) Soil Carbon (C) balance, 2) Soil Nitrogen (N) balance, 3) Soil structure maintenance, 4) Soil water balance, 5) N_2O emission control, 6) Biotic adversities regulation, 7) Groundwater contamination control, and 8) Species richness maintenance. This list is based on an ES concept which not only includes the attributes and processes of those ecosystems that support ES, but also strictly services. The conceptual network was the result of a bibliographic review and an expert knowledge elicitation through semi-structured interviews. Experts considered were researchers involved in several areas related to agroecosystems functioning (e.g. crop fertilization, contamination by fertilizers, nutrient dynamics, groundwater quality, soil fertility, weed ecophysiology). Researchers were selected within the academic field of Facultad de Agronomía, Universidad de Buenos Aires (FAUBA) as well as within other national universities and institutions. The expert panel was finally composed by 20 researchers.

Second, four sub-networks detached from the general conceptual network were selected in order to parameterize them with BNs. These sub-networks were: 1) Soil C balance, 2) Soil N balance, 3) N_2O emission control, and 4) Groundwater contamination control. The parameterization process consists on obtaining the conditional probabilities of child nodes (parameters) based on a conceptual network previously