Ecological Modelling 371 (2018) 101–118

Contents lists available at ScienceDirect

Ecological Modelling

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

Research Paper

Reprint of "Validation approaches of an expert-based Bayesian Belief Network in northern Ghana, West Africa"☆

Janina Kleemann[a,b,c,*], Enrico Celio[d], Christine Fürst[c]

[a] Karlsruhe Institute of Technology, Institute of Meteorology and Climate Research, Atmospheric Environmental Research (IMK-IFU), Kreuzeckbahnstraße 19, 82467 Garmisch-Partenkirchen, Germany
[b] University of Bonn, Center for Development Research (ZEF), Genscherallee 3, 53113 Bonn, Germany
[c] University of Halle-Wittenberg, Institute for Geosciences and Geography, Dept. Sustainable Landscape Development, Von-Seckendorff-Platz 4, 06120 Halle, Germany
[d] Institute for Spatial and Landscape Development (IRL), Planning of Landscape and Urban Systems (PLUS), ETH Zürich, Stefano-Franscini-Platz 5, 8093 Zürich, Switzerland

ARTICLE INFO

Article history:
Available online 31 January 2018

Keywords:
Conditional probabilities
Expert knowledge
Extreme-condition test
Uncertainty
Predictive power
Sensitivity analysis

ABSTRACT

Model validation is a precondition for credibility and acceptance of a model. However, it appears that there is no scientific standard for validation of Bayesian Belief Networks (BBNs). In this paper, we present a novel combination of BBN validation approaches. A set of qualitative and quantitative validation approaches for the BBN structure, the Conditional Probability Tables and the BBN output is presented and discussed. The validation approaches were tested for a BBN on food provision under land use and land cover changes and different weather scenarios in rural northern Ghana. Experts played an important role in developing and validating the BBN due to data scarcity. Furthermore, selected nodes and the BBN output were compared to existing data. A sensitivity analysis was conducted. Validation approaches show that structural model uncertainties are still high and reliability of input data is low. However, the extreme-condition test shows that the BBN works according to the assumed system understanding that food provision decreases under floods, droughts, land pressure and poverty. Therefore, the BBN can provide general trends for output nodes but lacks reliability if detailed results of single system components are required.

1. Introduction

The validation of data and a model is key to provide confidence in the findings of a study and to ensure compliance with specific requirements (IEEE, 1990; Oreskes et al., 1994). However, a model can only to a certain degree be an accurate representation of the real world (AIAA, 2002). Hence, in complex human-environmental interactions, we need to recognize assumptions and over-simplification (Banks, 1999). Furthermore, model validity is closely linked to the character and context of the problem, the model objective, and the background of the user and modeler, among others (Barlas and Carpenter, 1990). Therefore, a model cannot be an absolute and objective representation. Model validation even ends in a philosophical question between logical empiricist philosophy and relativist philosophy (Ibid.).

Mankin et al. (1977) suggested that models should be evaluated according to their usefulness rather than to their validity because of the imperfect representation of reality. Furthermore, the use of a model without validation is legitimate (Caswell, 1988), for example when the model is used to systematize knowledge or to develop a theory. But validation is often essential for user acceptance (Rykiel, 1996).

Bayesian Belief Networks (BBN) are regarded as a suitable tool for overcoming data gaps, estimating uncertainties, and visualizing complex causal relationships, e.g. socio-ecological interactions (Varis, 1997). Due to the explicit consideration of uncertainty, predictions are closer to the (uncertain) reality (Reichert and Omlin, 1997). They are graphical non-spatial statistical models that represent a set of variables and their conditional dependencies through directed acyclic graphs (for further reading, see Jensen, 2001; Kjærulff and Madsen, 2005; Reckhow, 1999).

DOI of original article: http://dx.doi.org/10.1016/j.ecolmodel.2017.09.018.
☆ This article is a reprint of a previously published article. For citation purposes, please use the original publication details "*Ecological Modelling*" 365 (2017) 10–29.
* Corresponding author. Present address: Helmholtz-Centre for Environmental Research (UFZ)/German Centre for Integrative Biodiversity Research (iDiv), Deutscher Platz 5e, 04104 Leipzig, Germany.
E-mail addresses: jkleemann@uni-bonn.de (J. Kleemann), ecelio@ethz.ch (E. Celio), christine.fuerst@geo.uni-halle.de (C. Fürst).

https://doi.org/10.1016/j.ecolmodel.2017.12.021
0304-3800/

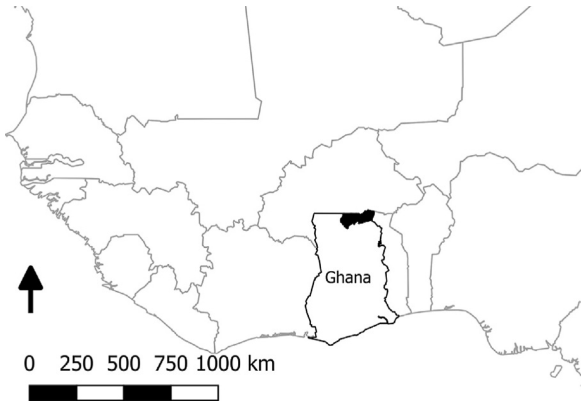

Fig. 1. Location of the Upper East Region (in black) in northern Ghana. National and administrative boundaries from OpenStreetMap (http://www.openstreetmap.org).

However, many BBNs are not validated. Aguilera et al. (2011) found that 37.7% of 114 reviewed Bayesian Belief Networks (BBNs) were not validated. Landuyt et al. (2013) found that only one-third of 47 papers on modeling ecosystem services with BBNs used data for model validation.

In this paper, we apply a range of BBN validation methods to the example of an expert-based BBN for the rural agricultural area in northern Ghana. This study is a follow-up of the BBN presented in Kleemann et al. (2017b). The BBN was developed to estimate the likelihood of food security in the dry and rainy season under different land use and weather scenarios. First, we briefly present the model study area and then describe the participatory expert-based approach for developing the structure of the BBN. Experts were an important information source in the data-scarce context of this study. Second, we present a range of validation approaches. We split the methodological description and results of our analysis in quantitative and qualitative validation methods.

2. Material and methods

2.1. Study area

The Upper East Region (UER) is located in northern Ghana close to the border of Burkina Faso and Togo and covers 8842 km^2 (3.7% of Ghana, GSS 2008; Fig. 1). The majority of the population is engaged in small-scale rain-fed subsistence farming (Birner et al., 2005) of maize, sorghum, and millet often intercropped with groundnuts or beans. Vegetables and rice are grown in irrigated areas or rain-fed lowlands.

The region is characterized by a dry season from November to April and a rainy season from May to October. All rain-fed crops are grown and harvested in the rainy season. During the dry season, only irrigated crops can be cultivated and this land use comprises only a small part of the area. During the dry season, food is stored and consumed until the beginning of the next rainy season. Migration, especially of the youth, is one of the strategies for coping with food shortage in the dry season (Hjelm and Dasori 2012; Quaye 2008; Van der Geest et al., 2010).

The people living in the study area are the poorest in Ghana and have a low educational level with high illiteracy (GSS, 2008; GSS, 2014). Malnutrition and famines are triggered by high population densities (103 people/km^2), degraded soils, and low fertilizer input (Antwi-Agyei et al., 2012; Assan et al., 2009; Dietz et al., 2004; Songsore 1996; Quaye 2008). In addition, erratic rainfall, droughts, floods, extreme temperatures and the shifting onset of the rainy season cause crop failure (Armah et al., 2011; Dietz et al., 2004; Ofori-Sapong 2001; Yengo et al., 2010).

2.2. Overview of the participatory development of the Bayesian Belief Network (BBN)

Due to manifold data gaps, missing data, incomplete knowledge, and the computationally very demanding task, an automatic generation of the BBN structure, also known as structural learning (Jensen, 2001; Steck and Tresp, 1999), was not possible, hence we derived the model structure using expert knowledge (Morgan and Henrion, 1990; Uusitalo 2007). Expert knowledge is a common approach with regard to the development of BBNs. Aguilera et al. (2011) revealed that expert knowledge was the most often used method in the discretization, model learning, and validation processes of BBNs.

In our case, scientific experts with a background in West African land management were selected because they represent the most reliable and comprehensive knowledge source for reflecting on superordinate links of this specific socio-ecological system (Cooke 1991; Weible et al., 2010). Scientists were selected based on published papers, recommendations by other scientists and/or consultation of directors of scientific institutions such as university departments and institutes of the Council for Scientific and Industrial Research (CSIR). The majority of these scientists were involved in WASCAL (West African Science Service Center on Climate Change and Adapted Land Use), which is a West African-German scientific collaboration established to improve the resilience of human-environmental systems with respect to climate variability and other environmental changes (WASCAL, 2016). They were chosen due to their regional and thematically relevant knowledge. Agricultural extension officers were consulted in addition to the scientists in order to reflect on locally relevant interlinkages of the socio-ecological system. Overall, 58 scientists and 37 representatives from agricultural planning (field officers) were consulted for the BBN development (Fig. 2).

The main methods were focus group discussions, interviews and questionnaires. The development of the structure of the BBN was based on a focus group discussion with 11 scientists from WASCAL. After a general introduction to the purpose of the study, the participants were split into groups to work on three sub-BBNs because each of the sub-models was complex. Experts were allocated to the groups according to their field of expertise. One sub-BBN dealt with land use changes, one with parameters related to crop yield, and one with water in relation to food provision. The nodes had been predefined to maintain the overall theme of three sub-BBNs but it was left open to the group to add more nodes. The group decided on major links between the nodes. Finally, all three sub-BBNs were presented and discussed.

The modeler merged the three sub-BBNs into one, and three representatives of each sub-BBN were asked for their feedback. It turned out that one BBN part for the dry season and one part for the rainy season better reflected the differences in land use during an annual cycle. More details are presented in Appendix A. Details on the focus group discussions to validate the BBN are provided in Section 2.3.5.

The Conditional Probability Tables (CPTs) are based on data derived from experts (34%), through calculations (32%), literature (32%), remote sensing/maps (2%) and other models (2%) due to varying data availability. Specifications of The CPTs based on literature, calculations and crop model output are provided in Kleemann et al. (2017b). Details on the expert-based CPTs are provided in Section 2.3.1.

2.3. Validation approaches

In contrast to verification, validation does not imply the presentation of the truth but is rather characterized by the establishment