CrossMark

ORIGINAL ARTICLE

Analysis of deforestation patterns and drivers in Swaziland using efficient Bayesian multivariate classifiers

Wisdom M. Dlamini[1]

Abstract Deforestation is a global phenomenon that threatens not only biodiversity but also the livelihoods of people who depend on them. Recent global analyses indicates that this problem is growing, hence the growing need to identify the underlying drivers in order to develop more responsive policies. We present a machine learning-based method which automatically identifies key drivers and makes predictions from available spatial data in Swaziland during the post millennium period. The efficient Bayesian multivariate classifiers (EBMC) are used to learn feature-selected Bayesian network (BN) models of deforestation from multisource data. The EBMC models, learned using the K2 and BDeu algorithms, were also used to predict the probability or risk of deforestation in addition to providing a directed acyclic graphical view of the key interacting factors. These were compared with constraint and knowledge-based BNs developed using the common EBMC-selected variables. All the models performed consistently well (log loss <0.3, AUC >0.8) when evaluated against observed deforestation patterns. The knowledge-based and constrained-based BNs performed well highlighting the need for developing a causal structure of interactions between variables. The findings indicate that deforestation patterns are determined by an interaction of proximate and underlying factors primarily fuelwood use, human population density, human settlements, protection and land ownership status. The findings indicate that, unless robust conservation measures are put in place, deforestation is likely to continue as more areas become vulnerable. The models produced plausible results that can be used for preventive planning and policy making.

Keywords Bayesian network · Deforestation · Efficient Bayesian multivariate classifier · Forest · Swaziland

✉ Wisdom M. Dlamini
mwdlamini@gmail.com

[1] Directorate of Parks, Swaziland National Trust Commission, Lobamba, Swaziland

Introduction

Deforestation and forest degradation are the major drivers of biodiversity loss in developing countries (Rolland et al. 2014). Furthermore, this results in loss of crucial ecosystem services and compromised livelihoods for people who depend on forest and woodlands for their survival. Ecosystem services such as biodiversity protection, climate regulation, carbon storage, sediment retention, water supply and pollination are some of those provided by tropical forests (Foley et al. 2005) and required for the same agricultural activities that tend to drive deforestation (Porro et al. 2015). Deforestation has also been observed to be the key contributor to carbon emissions (van der Werf et al. 2009). Hence, programmes such as the united nations framework convention on climate change's (UNFCCC) reducing emissions from deforestation and forest degradation and related activities (REDD+) mechanism have been spearheaded (UNFCCC 2014). Addressing this problem requires explicit characterization of the patterns and drivers of deforestation so as to formulate mitigation actions and to minimize the resultant carbon emissions.

Assessing the driving forces behind land change is key for reducing the uncertainty regarding the spatial and temporal occurrence of future deforestation and forest

carbon loss (Bax et al. 2016). The interactions between the causes of forest loss affect forest cover, and subsequently have impacts on ecosystems services and the livelihoods of the people who depend on tropical forests to survive (Foley et al. 2005). This highlights the importance of studying the causes of deforestation through spatial analysis so as to better understand the dynamics of land use change. However, understanding of the causes of changes in the landscape is not a simplistic representations of a few driving forces but involves complex analyses of area-specific interactions among a large number of explanatory factors (Kumar et al. 2015). Such drivers may include biophysical, socioeconomic, political and/or institutional factors interacting at different spatial and temporal scales (Geist and Lambin 2002; Rudel 2007). The increasing availability of earth observation data, open data and open-source analysis tools present an opportunity to understand forest cover change more efficiently than ever-before. Nonetheless, the increasing volume and complexity of data requires the use of techniques that can handle large multi-source and multi-scale datasets.

Over the years researchers have employed various techniques to estimate deforestation risk from possible underlying drivers. These include including statistical approaches, machine learning (ML) and spatial modelling (Mayfield 2015). Traditional parametric approaches, usually in the form of logistic regression, generalized linear or generalized additive models, are widely used compared to machine learning techniques such as Maxent (e.g., Aguilar-Amuchastegui et al. 2014). Conventional statistical approaches and tools for estimation of deforestation risk, although useful, have largely depended on technical expertise, which is often lacking in developing countries (Romijn et al. 2015). Furthermore, it may not be possible to simultaneously know all the drivers and their relative influences on deforestation. Machine learning algorithms, in particular, have the ability to represent and generalize relationships in data, thereby offering significant potential for dealing efficiently with large datasets (Committee on needs and research requirements for land change modeling et al. 2014).

Machine learning approaches have been found appropriate for situations where data concerning observed pattern are available but theory concerning process is scant (Committee on needs and research requirements for land change modeling et al. 2014). Compared to statistical approaches, such algorithms do not require strong mathematical assumptions to express a relationship between the target variable (deforestation) and predictor variable(s). Furthermore, interpretation of the model structure and performance of some traditional and machine learning approaches is difficult. Therefore, methods that are readily available, can be easily interpreted, can be run iteratively and updated are required for understanding deforestation.

In this study, we propose and test a new approach for identifying high deforestation risk areas and the underlying causes. This approach is tested in Swaziland where, to date, the drivers of deforestation have not been spatially and explicitly studied. We employ an approach based on Bayesian networks (Pearl 1988) which has largely been used in the analysis of high-dimensional datasets and can automatically select the relevant variables in addition to estimating deforestation probabilities. Although BNs have been used for land change modelling (e.g., Aalders 2008; McCloskey et al. 2011; Krüger and Lakes 2015), most of the studies used a minimal set of expert selected predictor variables. Using available landsat satellite imagery-derived deforestation and other ancillary data for Swaziland, we test the machine learning approach to BN model development from data.

The objectives of this study are to use machine learned Bayesian networks to (1) identify factors that are associated with direct deforestation drivers, on deforestation in Swaziland, and to (2) predict the risk of deforestation status of forests and woodlands. Within the context of this study, we define deforestation as the removal of primary forests and woodlands with tree cover of more than 30 %.

Methods

Study area

The study was carried out in Swaziland, southern Africa (centred on 26.5 °S and 31.5 °E) (Fig. 1) with an area totalling 17,365 km^2. The country is characterized by a combination of mountainous areas to east, undulating hills in the central parts, and relatively flat terrain and an undulating plateau to the east. The altitude of the area ranges from a lot of 40 m rising to a peak of 1862 m above sea level resulting in relatively dry to humid climates characterized by distinct wet warm summers and dry cold winters. Mean annual rainfall varies considerably from year to year, averaging at about 1500 mm in the western part of the country decreasing to 500 mm in the southeast where drought is prevalent. Conversely, mean annual temperature varies from 17 °C in the northwest rising up to 22 °C in the southeast with some localized variations caused by topographic features. The country's natural ecosystems also tend to follow the altitudinal and climatic gradients.

The underlying geology, coupled with the topography and climate, invariably influences the various land capabilities and agro-ecological zones with differentiated suitability for varying land uses such as human settlement, grazing, wildlife, irrigation agriculture, livestock ranching, and subsistence agriculture, amongst others (Remmelzwaal and Dlamini 1994). These land uses are practiced under

Fig. 1 Location map of the study area
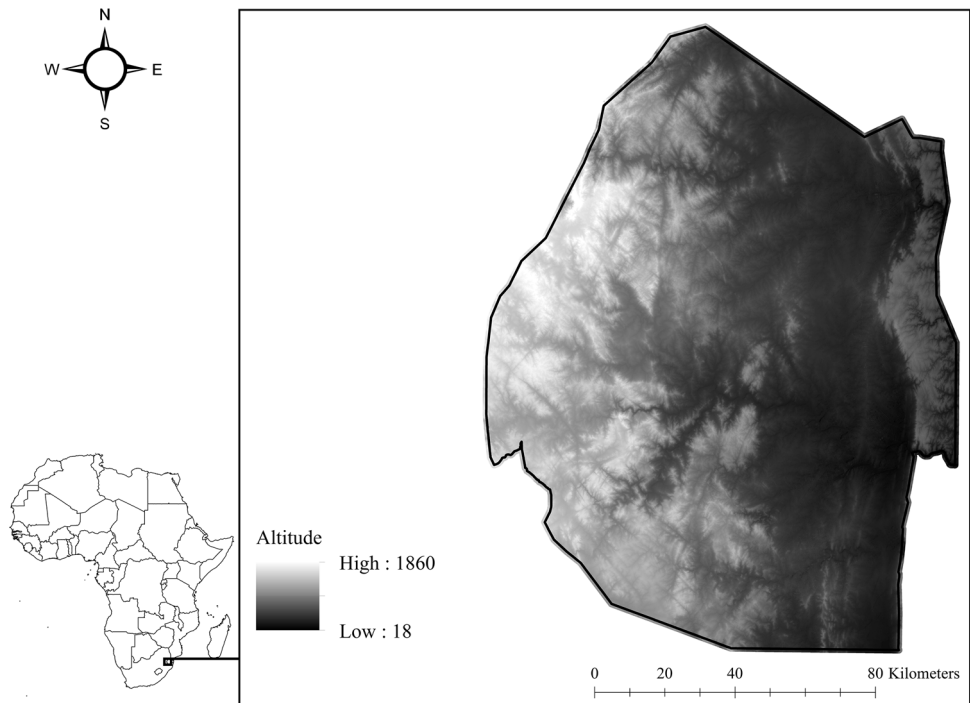
communal Swazi National Land (approximately 52 % of total land area held in trust by the King) and Title Deed Land, which constitutes about 47 % of total land area (Remmelzwaal and Vilakati 1994). Crown land (government land) and concession land are two other minor categories, which cover less than a percent of the country's surface area. Such a heterogeneous landscape with steep environmental gradients and varying land uses provides a good opportunity investigate their effects on forest cover dynamics.

Data

The data on forest cover loss was obtained from the global forest change database (GFCD, obtained through the Google Earth Engine (http://earthenginepartners.appspot.com/science-2013-global-forest). This forest loss dataset, which is the only publicly available dataset available for the country, is based on Hansen et al.'s (2013) analysis of Landsat satellite images for the period between 2000 and 2014. The procedure from the GFCD is based on tree canopy cover in the year 2000, defined as percent canopy closure for all vegetation taller than 5 m in height. Binary maps of forest loss are produced by identifying pixels with a change from non-zero to zero percent tree cover. Conversely, forest gain during the same period is defined as the inverse of forest loss. We used the R package gfcanalysis (Zvoleff 2015) to download, extract and analyses the forest loss data for Swaziland. A threshold of 30 % tree cover was set to differentiate forest and non-forest areas.

All forest loss pixels which had forest gain during the period under review were not regarded as deforestation particularly because nearly all of those were within forest plantations where there is a cycle of harvesting and replanting. However, a few areas within the plantations experienced nett loss due to factors such as fires and those pixels were identifiable and were retained. Through a review of literature on deforestation processes and with our knowledge of the local environment, we identified a total of 120 variables as possible explanatory factors at the national level (see Supplementary material 1 for datasets used and their characteristics). All the variables were then rescaled to a uniform 30 arc-second ($\sim$ 800 m) resolution grid size for the analysis. All the spatial analysis and visualization was done using the QGIS software (QGIS Development Team 2012).

Bayesian networks

In modelling deforestation with a BN, the aim is to derive a directed acyclic graph with the target (deforestation) and the explanatory variables represented as nodes connected by edges or arcs following dependencies. Bayesian thinking lies in the interpretation of Bayes' theorem or law (Eq. 1): given deforestation D and evidence e such that $P(e) \neq 0$ and $P(D) \neq 0$, then:

$$P(D|e) = \frac{P(e|D)P(D)}{P(e)}. \tag{1}$$

The theorem (Eq. 1) asserts that the probability of an deforestation event (or hypothesis) D conditioned upon

some evidence e is equal to its likelihood $P(e|D)$ times its probability prior to any evidence $P(D)$, normalized by dividing by $P(e)$ so that the conditional probabilities of all hypotheses sum to 1. The term $P(D|e)$ is often known as the *conditional probability* or posterior (or a posteriori) probability of deforestation D given e while $P(e|D)$ is the likelihood of e given a deforestation event D. The term $P(D)$ is the prior or marginal probability of deforestation. Given the conditional probability formulation, it is now possible to define what it means for deforestation and an explanatory variable to be conditionally independent. The events D and e are independent if $P(D|e) = P(D)$ and $P(e|D) = P(e)$. It follows then that if both $P(D)$ and $P(e)$ are positive, then both $P(D|e)$ and $P(e|D)$ imply the other. This notion of conditional (in)dependence is fundamental to BNs and the interpretation of probabilistic relationships.

A BN is essentially a graphical representation of a probability distribution over a set of variables $X = \{X_1, X_2,...,X_n\}$, n ≥ 1. A BN consists of two parts, $B = \langle G, \Theta \rangle$, where G is a directed network structure in the form of a directed acyclic graph (DAG), and Θ is a set of the local probability distributions for each node/variable, conditional on each value combination of the parent nodes. The graphical component of BNs, G makes them a class of probabilistic graphical models for reasoning under uncertainty, where the nodes represent variables (which can be discrete and/or continuous) and the arcs represent direct (and sometimes causal) connections and (in)dependencies between the linked variables. Those variables that are not linked directly in the graph are considered conditionally independent of each other. The second part of the BN, Θ represents the conditional probability distributions which models the quantitative strength of the connections or dependencies between variables. These are represented as conditional probability tables (CPTs), allowing probabilistic beliefs to be updated automatically as new information becomes available.

BN learning from consists of learning both the parameters and the structure or DAG. Hence, BNs provide a flexible method for probabilistic graphical modelling of highly interacting variables. Daly et al. (2011), Koski and Noble (2009) and Bielza and Larrañaga (2014) provide detailed reviews of contemporary approaches to and issues on learning BNs from data, the former focusing on the widely used discrete BN classifiers. In order to learn the Bayesian network structure from data, the algorithm approximates the likely graphical model by searching the space of possible networks via single-arc changes that improves some score. Alternatively, the structure can be based on some constraint put on the relationship between the selected variables.

The EBMC approach

Since there are many possible predictors of the complex deforestation pattern in Swaziland, it becomes difficult to ascertain which of them are better predictors. Hence, we sought for a method that can simultaneously select highly predictive variables whilst also modelling the interactions between them. Hence, the efficient Bayesian multivariate classifier (EBMC) algorithm (Cooper et al. 2010; Jiang et al. 2014) was chosen because it selects predictors from high dimensional data, and then uses the selected predictors in a BN classifier to perform inference and prediction.

This algorithm searches over a number of possible BN models to find one that is highly plausible given the training data and prior probabilities (Cooper et al. 2010). It starts by scoring all models in which a single variable is the parent of the deforestation node D, using either a K2 or a BDeu score. The model containing the highest scoring variable becomes the initial model. The algorithm then determines which variable when added as a parent of D to the first model yields the highest scoring second model. If the second model has a higher score than the first model, the new BN model is retained. This process is repeated using a greedy search by adding variables to the model as long as there is an increase in the K2 or BDeu score. If there is no further increase in the score, the algorithm then searches for a variable which when removed increases the score, and the variable whose removal increases the score the most is excluded. More variables are removed until removal of a variable does not increase the score. The algorithm then makes the variables in the final model children of D and create edges or links between them. The resultant BN produced by the EBMC algorithm is an augmented naïve BN which shows interactions or probabilistic relationships amongst the selected variables (Cooper et al. 2010).

The EBMC, therefore, searches over hybrids of Bayesian rules and augmented naïve BNs using prequential scoring in conjunction with either the BDeu or K2 scoring. The algorithm can be thought of as always searching for an additional rule that improves the prediction of the target variable in light of the selected predictor variables. The EBMC can also be viewed as a greedy search strategy that maximizes the score of a Markov blanket (Pearl 1988) of the target node D. The Markov blanket of the target node D consists of its direct parents, its direct successors, and all of its direct parents' direct successors within the given BN (Pearl 1988). Markov blankets have been shown to yield highly effective probabilistic predictions (Aliferis et al. 2010). This method differs from other BN-learning algorithms in that ordering of the features is not required in addition to an augmented Naïve Bayes representation of the learned BN structure from a subset of selected features.

We also tested a constraint-based approach, the inductive causation (ICS) algorithms (Verma and Pearl 1992) which takes into account latent variables to produce a BN with undirected, unidirected and bidirected arcs (Daly et al. 2011). This algorithm is a variant of the Spirtes, Glymour, Scheines (SGS) algorithm (Spirtes et al. 1993), which first generates an undirected graph that models dependencies between variables. Hence, the ICS algorithm is optimized for recovering the causal structure as opposed to finding an optimal classifier. The Bayesian score metric was used to learn the ICS algorithm using the variables selected by the EBMC algorithm. Finally, we developed a domain knowledge-supported BN using information from literature to determine the causal linkages. All the analyses were done using the waikato environment for knowledge analysis (WEKA) open-source data mining package version 3.9 (Hall et al. 2009).

Model validation and evaluation

A total of ten runs of tenfold cross-validation were performed for each of the K2 and BDeu scoring approaches and the ICS and knowledge-supported BN models thereby ensuring that the final calibration of every model used all of the data available. K-fold testing is more reliable with large data sets and is one of the recommended approaches for evaluating BN model prediction performance (Marcot 2012). Moreover, tenfold cross-validation has been found to be the right number to get the best estimate of error in addition to supporting theoretical evidence (Witten et al. 2011). Cross-validation is used to provide an out of sample evaluation method which repeatedly splits the data in training and validation sets. A BN structure is evaluated by estimating the network's parameters from the training set and the resulting BN's performance determined against the validation set. The average performance of the BN over the validation sets in turn provides a metric for the quality of each network.

The logarithmic loss (or log loss) was selected to evaluate model performance because of its suitability and reliability for tasks where posterior probability values are an important consideration (Marcot 2012). The log loss is also an evaluation metric whose value is only determined by the probability of the outcome that actually occurs (Cowell et al. 1993). This was calculated using the Eq. 2 (Pearl 1988; Morgan and Henrion 1990).

$$\text{Logarithmic loss} = \frac{1}{n} \sum_{i=1}^{n} \sum_{j=1}^{m} y_{ij} \big[-\ln P_{ij} \big], \qquad (2)$$

where n is the number of cases or instances in the test set, m is the number of class labels or states, ln is the natural logarithm, y_{ij} is 1 if observation i is in class j and 0 otherwise, and p_{ij} is the predicted probability that observation i belongs to class j.

The log loss, which has scores between 0 and infinity (or 1 if using a logarithm base of 2), is a cross-entropy estimate which measures the additional penalty for using an approximation instead of the true model. A value of zero indicates the lowest penalty whereby the network's probability distribution totally matches the true distribution. In addition to the log loss, we also estimated the area under the receiver operating characteristic (ROC) curve (AUC) (Hand 1997) of each model as a performance evaluation criteria.

Variable importance

The relative importance of each of the selected predictor variables was measured by computing its mutual information or entropy reduction (Pearl 1988) with the target node (deforestation). Given a probability distribution p defined over two sets of variables X and Y, the mutual information between X and Y, which is measured in bits, is given as:

$$I(X, Y) = \sum_{x,y} p(x, y) \log\left(\frac{p(x, y)}{p(x)p(y)}\right), \qquad (3)$$

where $p(\text{x})$ and $p(\text{y})$ are the probability densities of X and Y, and $p(\text{x, y})$ is the joint probability density; which can also be expressed in terms of entropy as:

$$I(X; Y) = H(X) + H(Y) - H(X, Y), \qquad (4)$$

where $H(X)$ and $H(Y)$, are the entropies of X and Y, respectively, and $H(X,Y)$ is the joint entropy of X and Y.

The mutual information extracts the relevant information that variable Y contains about X and is a very good measure of the average number of bits needed to convey the information X contains about Y and vice versa. As such, the mutual information is able to detect non-linear dependencies among variables that are undetectable using conventional measures (Guyon and Elisseeff 2003). It ranges from zero when the variables are independent and attains a maximum of 1 when one variable is a deterministic function of the other. The sensitivity analysis helps in validating the obtained relationships with the observed spatial patterns and domain knowledge. Furthermore, the mutual information indicates the potential of an explanatory variable to reduce the uncertainty in the target variable (Krüger and Lakes 2015).

Results

The 2000–2014 deforestation map of Swaziland as derived from the GFC dataset is as shown in Fig. 2. It is evident that the deforestation is complex and widespread in the

country covering most parts of the country albeit with localized hotspots. Fig 3 shows the BN models learned using K2 and BDeu scoring of the EBMC algorithm in addition to the ICS algorithm and the domain knowledge. In total, 17 variables were included in the final models of both scoring metrics from the original 120 variables. The K2 and BDeu scoring approaches produced slightly different structures although both selected 14 similar variables. The performance of all the BN learning approaches also produced accurate models as demonstrated by the AUC higher than 0.8 and logarithmic loss values lower than 0.3 (Fig. 4). The knowledge-supported BN, developed using the variables commonly selected by both the BDeu and K2 scoring approaches, outperformed the machine-learned BNs. However, the performance of the score-based algorithms was very similar with marginally better AUC and log loss values for the BDeu model.

Both proximate and underlying causes of deforestation are revealed through the selected variables from the BN models. The sensitivity analyses indicate that the percentage of people using fuelwood for cooking is the strongest deforestation risk factor followed by human population density, land use, human settlement density,

protection (conservation) status and land tenure (Fig. 5). All the BN models generally produced a similar trend regarding the relative strengths of these drivers in determining the spatial patterns of deforestation in Swaziland. Notably, both EBMC-learned BNs were predominantly augmented naïve Bayes in nature showing only a few interacting factors thereby indicating strong independence relationships between the features. For instance, the K2-derived BN indicated that land tenure interacts with cattle density and slope aspect whilst protection status interacts with wildfire frequency to determine forest cover loss. Similarly, the BDeu-derived BN also indicates the interaction of protection status and wildfire frequency, whilst additionally showing an interaction between cation exchange capacity at higher soil depths and proximity to major (perennial) rivers. The ICS algorithm produced the most complex network with feedback loops and bidirected arcs, highlighting the complex interaction of the selected drivers.

Despite differences in performance between the different modelling approaches, there are no discernible differences in the BDeu and K2 prediction maps (Fig. 6). However, the results from the ICS and knowledge-based models are visibly different from those of the BDeu and K2 algorithms. Notably, large areas have high posterior deforestation probabilities in the BDeu and K2 prediction maps compared to the ICS and knowledge-based predictions. The highly vulnerable areas tend to cluster around observed deforestation sites, which suggests that the causal BNs have better generalization capability compared to the EBMC-derived augmented naive Bayes classifiers. This also indicates that the generalization capability of the EBMC algorithm is improved by recognizing and accounting for the significant influence of all dependencies amongst causal factors.

Nevertheless, a mean deforestation map (Fig. 7) derived from the mean of all four models shows that large areas of forest are at risk of possible loss. This map can be interpreted as the probability of reduction of tree cover and height to below 30 % and 5 m, respectively. Areas that are highly vulnerable are as driven by the selected key factors particularly areas outside protected areas including forests that are in close proximity to major rivers and near existing human settlements and sugarcane plantations particularly in the central and eastern parts of the country. Protected (gazetted) areas have conspicuously low risks of deforestation as well as areas under conservation management albeit with some risk for certain areas closer to existing sugarcane plantations. The elevated risk in plantation forests and other wattle forests to the west of the country can be attributed to the cycle of harvesting and replanting which the models indicate as possible forest removal.

Fig. 2 Deforestation map derived from the GFC dataset for Swaziland. *Black pixels* represent areas where deforestation has occurred

Fig. 3 Bean plots of
logarithmic loss (**a**) and AUC
(**b**) for the four BN models
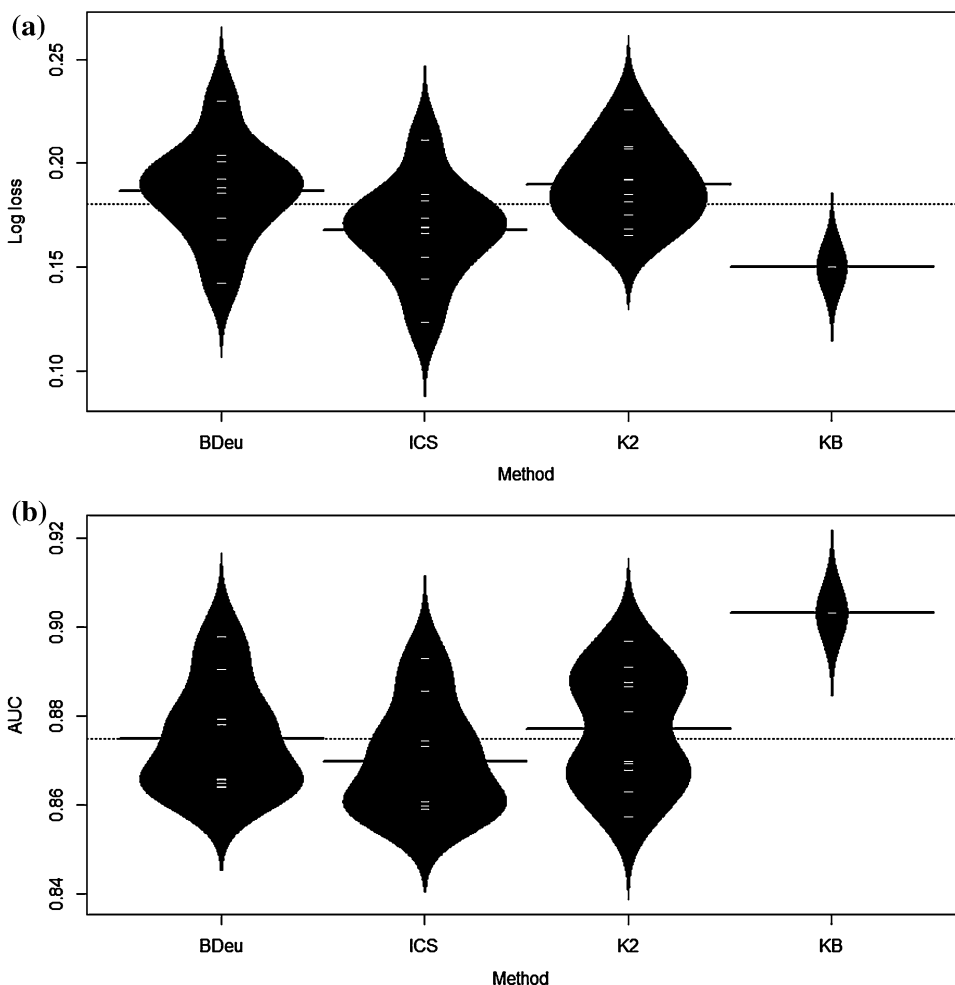
Discussion

Model performance and prediction maps

A spatially explicit analysis of deforestation and the analysis of proximate and underlying drivers is required for addressing local characteristics in forest cover change modelling. The EBMC technique was able to use high dimensional spatial data without a separate feature selection pre-processing step to select variables that are key drivers of deforestation. Although BNs do not allow for cyclic links, the ICS algorithm produced a causal BN with bidirected arcs resulting in feedback loops. This, we believe, represents a more realistic depiction of the interaction of the selected causal factors in determining deforestation in Swaziland. The performance of the ICS and knowledge-based BN highlights the importance of causality or a causal structure in modelling deforestation. The derived BNs form the basis for efforts to communicate ideas about the complexity of the observed deforestation patterns and processes. They also act as graphical

illustrations of what is both previously known and unknown about the drivers of deforestation and their interactions.

The derived deforestation risk maps provide a useful tool for proactive planning and policy making. Although there were differences in the prediction maps, these were largely differences in the probability distributions while the spatial patterns were generally similar. These differences can be attributed to the differences in the network structures. The predominantly naïve nature of the EBMC-derived maps resulted in possible overestimation of deforestation risks for large areas. On the other hand, the causal structure-based ICS and knowledge-based models were more constrained and predicted relatively lower risks in some areas with the exception of specific hotspots such as areas in close proximity to major rivers and near major water reservoirs. Hence, including more interactions among the drivers of deforestation resulted in probabilities that closely match to the observed forest loss. This observation is also supported by the log loss values which reflect lower prediction penalties for the ICS and
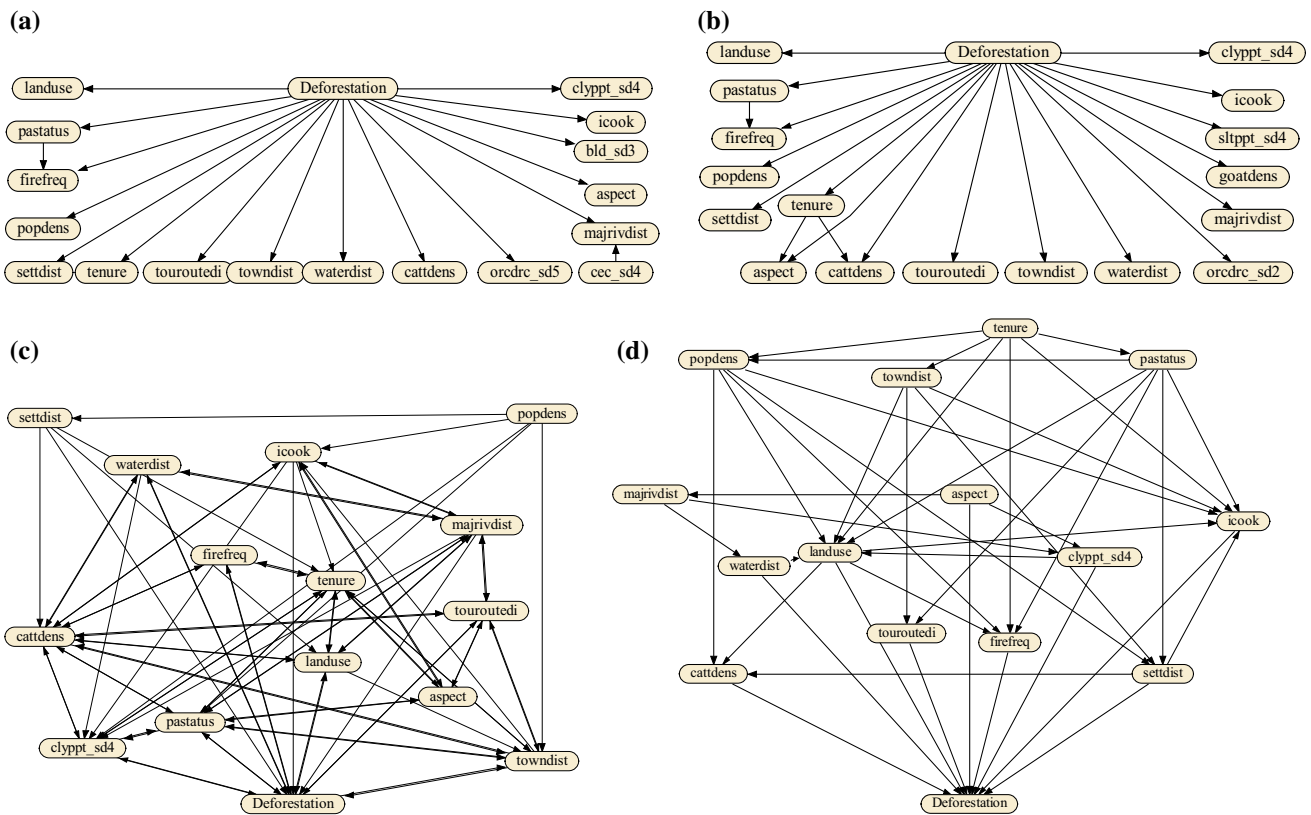
(a)

(b)

(c)

(d)

Fig. 4 Bayesian networks for the EBMC learned using the BDeu (**a**) and K2 scoring methods (**b**) together with the BNs learned using ICS algorithm (**c**) and domain knowledge (KB) (**d**). Full variables names and abbreviations are shown in Supplementary material 1
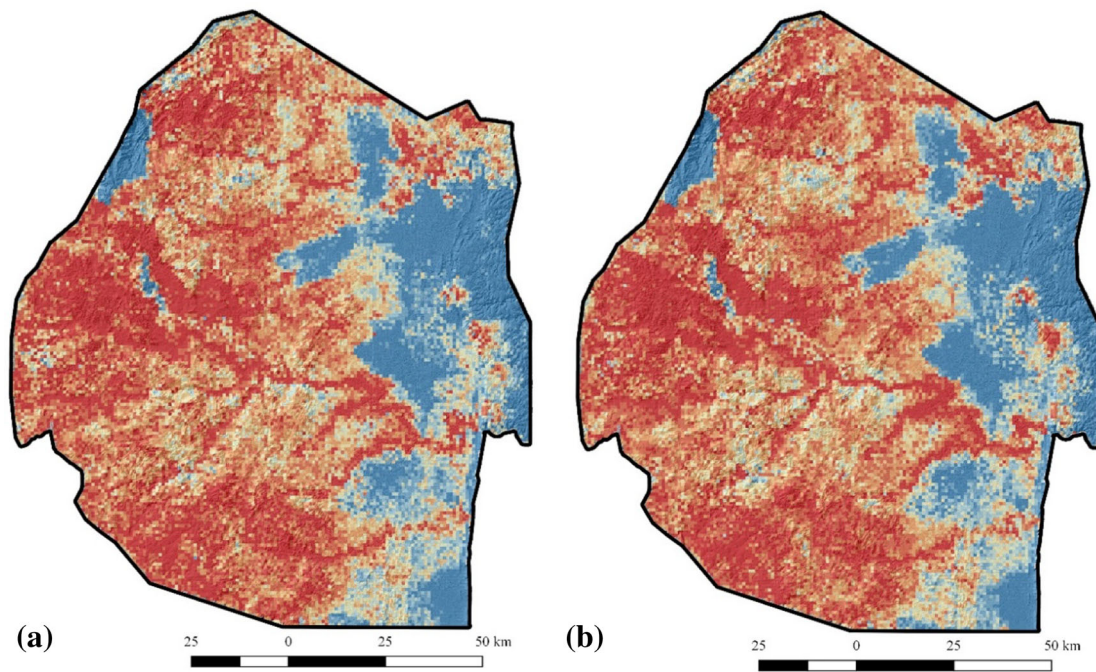
Fig. 5 Relative influence (mutual information) of selected variables on deforestation patterns. Full variables names and abbreviations are shown in Supplementary material 1
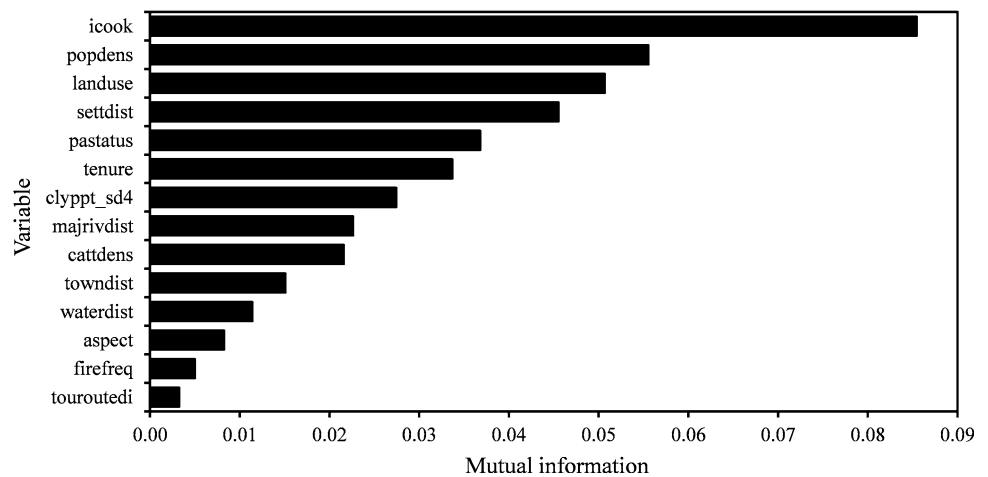
knowledge-based models. In a study using BNs for mineral prediction and species distribution modelling, respectively, Porwal et al. (2006) and Aguilera et al. (2010) also observed that including more arcs between variables improved the spatial prediction performance. The BN-based probabilistic approach, therefore, enables the communication of deforestation vulnerability in different parts of the country resulting from the various driving factors. The posterior probability maps also show
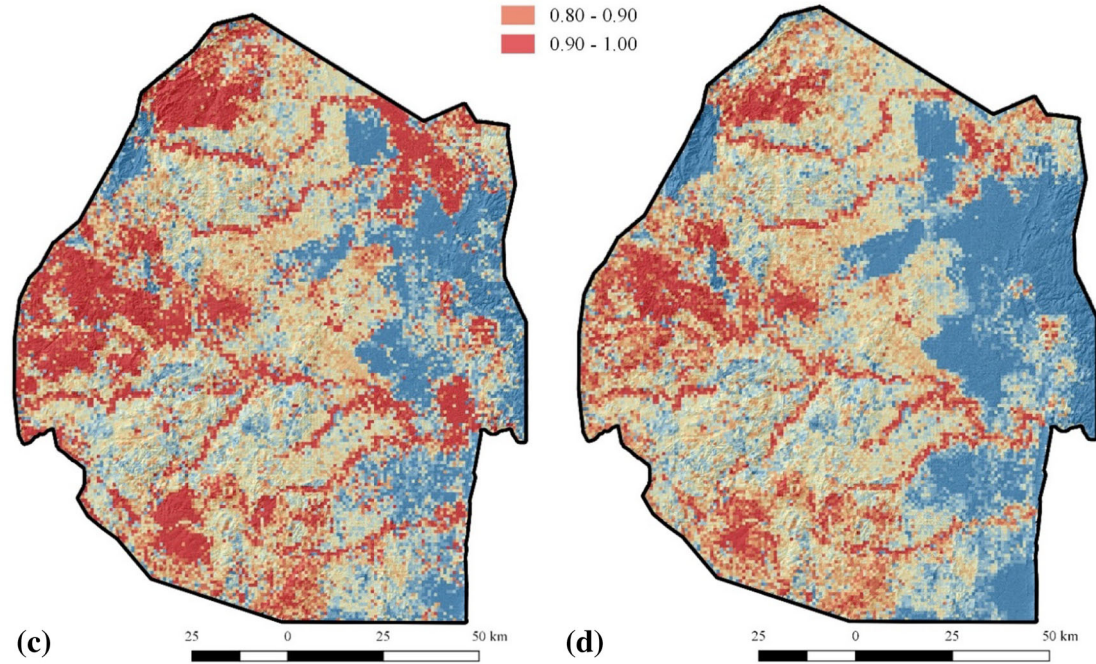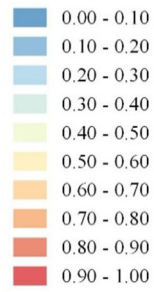
possible development paths, including the uncertainties related to the prediction models. The deforestation risk has serious negative implications for the majority of the Swazi population who direct and indirectly depend on forests for most basic necessities such as food, medicine

Fig. 6 Deforestation posterior probability (risk) maps from the ▶ EBMC learned using the BDeu (**a**) and K2 scoring methods (**b**) the ICS algorithm (**c**) and domain knowledge (**d**)

Legend

Posterior probability

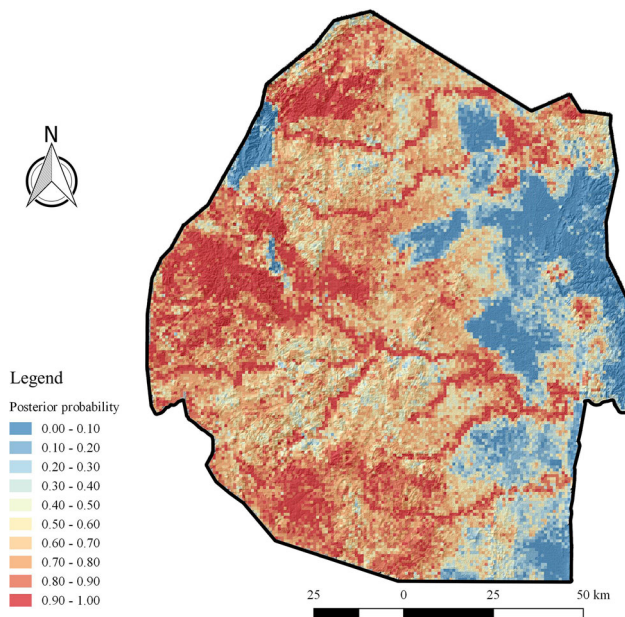|	
■	0.00 - 0.10
■	0.10 - 0.20
■	0.20 - 0.30
■	0.30 - 0.40
■	0.40 - 0.50
■	0.50 - 0.60
■	0.60 - 0.70
■	0.70 - 0.80
■	0.80 - 0.90
■	0.90 - 1.00

Fig. 7 Mean deforestation risk map derived from averaging the posterior probability all the Bayesian network models

and shelter let alone the broader suite of ecosystem services and carbon stocks that are under threat.

Whilst the findings present new evidence to the deforestation processes in Swaziland, the selected factors are consistent with previous observations elsewhere in the world. The complexity of the drivers of deforestation makes them location-specific because a variable affecting deforestation in an area may not do so in another. While we developed and compared data-driven and knowledge-based BNs, the findings indicate the added value of combining data-driven models with expert or domain knowledge to improve the deforestation models and their predictions. The usefulness of BNs is largely in their ability to facilitate the integration of expert knowledge and empirical data whilst also graphically showing the (e.g., sampling data). In addition to the improved performance of the knowledge-supported model, the model was also easy to interpret and important relationships between proximate and underlying causes and deforestation could be easily visualised.

Deforestation drivers

The key drivers of deforestation in Swaziland are revealed by the models. Fuelwood harvesting is the primary driver in the country and is known to causes forest degradation in human-dominated landscapes (Specht et al. 2015). Fuelwood collection is a key driver that can also be seen along major (tourist) roads where indigenous trees are cut to sell to passers-by (mainly city dwellers) as firewood (Stringer 2009; Manyatsi and Hlophe 2010). The 2012 National Energy Balance showed that 53 % of the energy

consumption was biomass comprising mainly of fuelwood (30 %) used in households (Government of Swaziland (GOS) 2012). Ngwenya and Hassan (2005) estimated that an average of 376 kg/annum was being extracted per person and is exacerbated by the large populations in smaller areas of natural forest and woodland that have very low regeneration capacity. Earlier findings by Wheldon (1990) and Lasschuit (1994) indicated that there is a general fuelwood deficit in the country whilst Allen et al. (1988) projected that the central part of the country would experience fuelwood deficits as early as the 1990s.

Land use is another major driver as indicated by the analysis. A multitude of policies and legislations concerning land resources management exists in the country and fragmented within various institutions. Among others, laws and regulations connected to the use and management of land has had a significant catalytic effect on deforestation processes in Swaziland. These have significant implications on major proximate drivers such as human settlements encroachments, infrastructure development projects and agricultural developments. Recently, Bailey et al. (2015) noted an increase in cropland area in areas outside protected areas in the north-eastern part of the country. The almost non-existence of forest loss within protected areas highlights their role, especially through legal gazetting and enforcement, in the avoidance of deforestation in areas that would have otherwise been deforested. This important role of protected areas as a buffer against deforestation has been observed in many parts of the world (Spracklen et al. 2015). It is therefore necessary that protected area expansion and other area based conservation strategies are urgently targeted towards the identified localities with higher deforestation risk. This is also important for policy- and decision-making especially within the framework of biodiversity conservation and international climate change mitigation policies, such as REDD+.

The land tenure system influences the spatial patterns and types of human population density, land use and subsequently the forest utilization processes in the country. A majority of the country is designated as Swazi National Land (SNL) held in trust for the nation by the King (Mavimbela et al. 2010) and carries approximately three quarters of the population through communal customary tenure (Xaba and Masuku 2013). Such land is administered by chiefs and governed through customary rules. The rest of the land is title deed land (TDL) where exclusive access rights are defined and typically allocated to corporate actors (Mushala et al. 1994) and affluent individuals. The decisions on the allocation or distribution of land, usually taken based on human population pressure and subsistence requirements, influenced the landscape. Robinson et al. (2013) found that tenure security improves forest cover in the tropics. On the other hand, tenure insecurity,

particularly in communal lands, might increase deforestation. Dlamini and Geldenhuys (2011) observed that the land tenure system in Swaziland invariably affects the use and management of edible and medicinal non-timber and timber forest resources particularly in communal areas. This has implications for the majority of the Swazi population who directly and indirectly depend on forests and woodlands for most basic necessities such as food, medicine and shelter.

The expansion of the human population and settlements especially in close proximity to major roads is another key deforestation driver. With an increase in urbanisation, deforestation would occur and these would also be associated with rural population growth. Such deforestation also occurs in peri-urban areas and within main transportation corridors between major cities and towns. The growing urban population encourages people to make increasing use of fuelwood for cooking and heating (Rudel 2013). With increasing improvement in the country's road and other linear infrastructure network, accessibility to these areas is also gradually improving, thereby opening up forests to illegal and unregulated activities such as land speculation and destructive exploitation (Laurance et al. 2009). Such disturbances can increase the vulnerability of forests to more anthropogenic and natural disturbance. For instance, new roads facilitate access to previously intact forests as more settlements are established on mountainous terrain and in the more fertile riverine ecosystems. Human settlements are also often accompanied by the construction of electrical power and telephone lines, the maintenance of which requires clearcutting of 10 to 30 m strips of vegetation.

It is important to also note that human settlements, especially in rural areas, is often accompanied by the creation of slash-and-burn agricultural fields within soils that are more fertile and have low clay content thereby exacerbating the deforestation problem. This is particularly important to note because over 70 % of the Swazi population and households rely on agricultural output as a major source of income and food security (Masuku et al. 2015). Another growing problem is the illegal cultivation of marijuana (predominantly *Cannabis sativa*) in various parts of the country during which patches of forest are cleared particularly within perennial river valleys (pers. obs.).

Uncontrolled fires are also an increasing problem which results in the loss of large tracts of forests, both indigenous and man-made (Dlamini 2010). Forests that have been burned before are more likely to be deforested because the initial fires tend to thin out the canopy and add combustible material. Such forests are also often adjacent to fire-maintained rangelands and therefore frequently exposed to ignition sources. As a result, subsequent forest fires burn with increased intensity resulting in net loss of forest cover in the affected areas. Even though our results indicate high

deforestation risks in man-made forests, we do believe that these are both sustainably managed and have higher regeneration rates after harvesting. Black wattle (*Acacia mearnsii*) and some *Eucalyptus* species, for instance, have even been particularly observed to be highly invasive due to their rapid growth rate and uncontrolled spread into other natural vegetation (Kotzé et al. 2010). Also associated with fires is the slash-and-burn practice through which patches of forest are cleared for both subsistence crops and illegal marijuana or cannabis cultivation.

Agricultural activities are the dominant land uses driving deforestation in sub-Saharan Africa (Angonese and Grau 2014). Livestock production is also one of the main agricultural activities in Swaziland, with small farmers owning about more than three quarters of the total cattle population in the country (Swaziland Environment Authority 2012). This high density of cattle, particularly in communal areas, results in overgrazing and land degradation. High consumption rates of plant cover by livestock also reduces protective plant vigour and regrowth capacity. Other indirect effects include soil compaction due to trampling which, when excessive (particularly along cattle trails and near diptanks, homesteads and water points), may cause run-off and gully erosion (Tefera 2013), thereby accentuating the problem. Roques et al. (2001) also observed that high browsing pressure by goats in communal land may suppress the recruitment of some tree species, resulting in an unsustainable population structure with limited regeneration of the population.

The importance of clay content at 30–60 cm depth point to the significance of soil characteristics in influencing land use particularly agricultural practices such as subsistence crop farming and commercial agriculture. Areas with low clay content have relatively higher deforestation probabilities. The role of slope aspect is also notable and probably links to its influence on the topo-edaphic and hydrological properties of the landscape thereby influencing both agricultural suitability and vulnerability to deforestation. For instance, Murdoch (1968) and Nixon (2006) note that clay loams to clay soils with moderate organic matter content, which usually occur in mid-slope positions on well-draining gentle slopes, are one of the best soils which produce higher sugarcane yields. Tefera (2013) also notes that human settlement-associated livestock density and land form/aspect may affect vegetation changes and subsequent vulnerability to degradation.

The importance of proximity to major water sources is primarily an indicator of the effect of the growth of the export sugarcane plantation economy. Conversion of forest land to agriculture is primarily attributed to clearings for sugar cane, business and residential structures and water supply projects (DANCED 2000). This is driven by the demand for land resources required for establishment of

sugarcane to sustain the global agricultural commodity trade. A visual analysis of the maps indicates that the construction of the Maguga dam to the north of the country and the Lubovane dam in the eastern central part of the country have induced further clearing of natural forests for sugarcane and resettlement of people previously residing in the inundated areas. These government-led developments are implemented through the Swaziland water and agricultural development enterprise (SWADE) through support to local communities to engage in commercial small-holder agricultural activities, particularly sugarcane. Such expansive irrigation programmes are supported by sectoral policies such as the Comprehensive Agriculture Sector Policy (2005), the National Food Security Policy (2005) and the Swaziland National Irrigation Policy (2005). As more dams and sugarcane development projects are anticipated, it is expected that the situation will worsen particularly in the predicted high risk areas in close proximity to major rivers where perennially flowing water can be easily abstracted or dammed and soils and topography are suitable.

Although least important of the selected variables, proximity to major tourism routes was another key factor in determining deforestation in the country. This could be an indirect indicator of the effect of traffic flows on such routes which increase both the demand and supply of fuelwood to major cities and tourism activity centres. In addition, the development of the tourism industry helps to provide job opportunities, and the steady population growth coincides with a continuous increase in the built-up area.

Even though forests provide a number of valuable goods and services to the Swazi populace, the high returns from non-forest land uses sets the protection of forest ecosystems at a disadvantage and act as incentives for deforestation as similarly observed by Kanninen et al. (2007). The situation is challenging particularly in Swaziland given the country's climatic and phytogeographic settings which exposes the central and eastern regions to the impacts of drought and desertification. Nevertheless, the observed deforestation trends should not be confused with the observed increase in woody plant cover in some parts of the country and the region. Such a phenomenon, which is observed for certain woody plant species, is a result of both bush encroachment and alien plant invasion. Deforestation can create gaps and disturbances that facilitate invasion by prevalent species such as *Dichrostachys cinerea*, *Chromolaena odorata* and *Lantana camara* which have high-growth rates.

Conclusion and recommendations

Understanding the process of forest trajectories is vital for informing forest management and conservation policy and for an efficient targeting of interventions. This is particularly true for REDD+ and climate change adaptation initiatives that rely on the formulation of effective and equitable mechanisms to reduce forest carbon loss. Deforestation in Swaziland is a result of a complex interplay of proximate drivers that are triggered by underlying drivers such as demand generated for land by agricultural activities and infrastructure development; demand for forest products, government policies for ensuring food security and meeting the rising demand for energy and infrastructure. Accentuating the problem are underlying drivers such as permissive government (land use) policies, population growth, and land tenure. Encroachment by peripheral communities, catalysed by the population size and limited enforcement of environmental laws have encouraged deforestation.

The control or reversal of deforestation can, therefore, be achieved by addressing the drivers identified to be currently contributing to deforestation in the country. The promotion of alternative energy efficient and renewable sources should be encouraged to reduce the dependence on the use of firewood. Reducing deforestation would also require creating and strengthening inhibitors of deforestation such as protected areas and forest reserves as well as strengthening participatory forest restoration and protected area expansion programmes. It is imperative, therefore, that the country enhances the land use planning process in addition to identifying and implementing appropriate market-based instruments to mitigate harmful effects of development projects on forest resources. Other proposed measures include strengthening the existing procedures of environmental impact assessments and strategic environmental assessments particularly for developments targeted at the vulnerable areas. This should be accompanied by strengthening the monitoring and enforcement capacity of relevant conservation, environmental and land management agencies. Furthermore, efforts should be made to remove incentives such as subsidies which enhance the drivers of deforestation.

References

Aalders I (2008) Modeling land-use decision behavior with Bayesian belief networks. Ecol Soc 13(1):16–37

Aguilar-Amuchastegui N, Riveros JC, Forrest JL (2014) Identifying areas of deforestation risk for REDD + using a species modeling tool. Carbon Balance Manag 9:10. doi:10.1186/s13021-014-0010-5

Aguilera PA, Fernández A, Reche F, Rumí R (2010) Hybrid Bayesian network classifiers: application to species distribution models. Environ Model Softw 25:1630–1639

Aliferis CF, Statnikov A, Tsamardinos I, Mani S, Koutsoukos XD (2010) Local causal and markov blanket induction for causal discovery and feature selection for classification—Part I: algorithms and empirical evaluation. J Mach Learn Res 11:171–234

Allen JA, Pimentel DP, Lasoie JP (1988) Fuel wood production and use in rural Swaziland: a case study of two communities. For Ecol Manag 25:239–254

Angonese JG, Grau HR (2014) Assessment of swaps and persistence in land cover changes in a subtropical periurban region, NW Argentina. Landsc Urban Policy 127:83–93

Bailey KM, McCleery RA, Binford MW, Zweig C (2015) Land-cover change within and around protected areas in a biodiversity hotspot. J Land Use Sci 11:154–176. doi:10.1080/1747423X.2015.1086905

Bax V, Francesconi W, Quintero M (2016) Spatial modeling of deforestation processes in the Central Peruvian Amazon. J Nat Conserv 29:79–88

Beyene ST (2014) Rangeland degradation in Swaziland: dip tank use effects on range condition and grazing capacity in three soil types. Afr J Ecol 52:514–523

Bielza C, Larrañaga P (2014) Discrete Bayesian network classifiers: a survey. ACM computing surveys, 47(1), Article 60, doi:http://dx.doi.org/10.1145/2576868

Central Bank of Swaziland (2015) Annual report 2014–2015. Central Bank of Swaziland, Mbabane

Committee on needs and research requirements for land change modeling, geographical sciences committee; board on earth sciences and resources, division on earth and life studies; national research council (2014) Advancing land change modeling: opportunities and research requirements. National Academies Press, Washington

Cooper GF, Hennings-Yeomans P, Visweswaran S, Barmada M (2010) An efficient Bayesian method for predicting clinical outcomes from genome-wide data. AMIA Annual Symposium Proceedings, 127–131. http://www.ncbi.nlm.nih.gov/pubmed/21346954 Accessed 17 November 2015

Daly R, Shen Q, Aitken S (2011) Learning Bayesian networks: approaches and issues. Knowl Eng Rev 26:99–157

DANCED (2000) Identification of protection worthy areas in Swaziland. Danish Corporation for International Development/Ministry of Agriculture and Co-operatives, Mbabne

Dlamini WM (2010) A Bayesian belief network analysis of factors influencing wildfire occurrences in Swaziland. Environ Model Softw 5:19–208. doi:10.1016/j.envsoft.2009.08.002

Dlamini CS, Geldenhuys CJ (2011) A resource survey for medicinal and edible plant species in the four ecological zones of rural Swaziland. J Geogr Reg Plan 4(9):557–565

Foley JA, Defries R, Asner GP et al (2005) Global consequences of land use. Science 309:570–574

Geist HJ, Lambin EF (2002) Proximate causes and underlying driving forces of tropical deforestation. Bioscience 52(2):143–150

Government of Swaziland (GOS) (2012) Swaziland's second National Communication to the United Nations framework convention on climate change. Ministry of Tourism and Environmental Affairs, Mbabane

Guyon I, Elisseeff A (2003) An introduction to variable and feature selection. J Mach Learn Res 3:1157–1182

Hall M, Frank E, Holmes G, Pfahringer B, Reutemann P, Witten IH (2009) The WEKA data mining software: an update. SIGKDD Explor 11:10–18

Hand D (1997) Construction and assessment of classification rules. Wiley, New York

Hansen MC, Potapov PV, Moore R, Hancher M, Turubanova SA, Tyukavina A, Thau D, Stehman SV, Goetz SJ, Loveland TR, Kommareddy A, Egorov A, Chini L, Justice CO, Townshend JRG (2013) High-resolution global maps of 21st-century forest cover change. Science 342:850–853

Hassan R, Mbuli P, Dlamini C (2002) Natural resource accounts for the state and economic contribution of forests and woodland resources in Swaziland. CEEPA discussion paper series. University of Pretoria, South Africa

Jiang X, Cai B, Xue D, Lu X, Cooper GF, Neapolitan RE (2014) A comparative analysis of methods for predicting clinical outcomes using high-dimensional genomic datasets. J Am Med Inform Assoc 21(2):312–319. doi:10.1136/amiajnl-2013-002358

Kanninen M, Murdiyarso D, Seymour F, Angelsen A, Wunder S, German L (2007) Do trees grow on money? The implication of deforestation research for policies to promote REDD. Center for International Forestry Research (CIFOR), Borgor

Koski T, Noble J (2009) Bayesian networks: an introduction, Wiley Series in Probability and Statistics. Wiley, Chichester

Kotzé I, Sibandze P, Beukes H, van den Berg E, Weepener H, Newby T (2010) Surveying and mapping the distribution and intensity of infestation of selected category 1 invasive alien plant species in Swaziland. Agricultural Research Council—Institute for Soil, Climate and Water, Pretoria

Krüger C, Lakes T (2015) Bayesian belief networks as a versatile method for assessing uncertainty in land-change modelling. Int J Geogr Inf Sci 29(1):111–131. doi:10.1080/13658816.2014.949265

Kumar R, Nandy S, Agarwal R, Kushwaha SPS (2015) Forest cover dynamics analysis and prediction modeling using logistic regression model. Ecol Ind 45:444–455

Lasschuit PE (1994) Rural household energy strategies in Swaziland. Universiteit van Amsterdam, Amsterdam

Laurance WF, Goosem M, Laurance SGW (2009) Impacts of roads and linear clearings on tropical forests. Trends Ecol Evol 24:659–669

Manyatsi AM, Hlophe ET (2010) Contribution of sale of firewood towards rural livelihood in Swaziland, and its environmental sustainability. Curr Res J Soc Sci 2:226–232

Marcot BG (2012) Metrics for evaluating performance and uncertainty of Bayesian network models. Ecol Model 230:50–62

Masuku MB, Kibirige D, Singh AS (2015) Impact of HIV and AIDS on agricultural production in Swaziland: strategies for mitigation. Int J Econ Commer Manag III:1–14

Mavimbela P, Masuku MB, Belete A (2010) Contribution of savings and credit cooperatives to food crop production in Swaziland: a case study of smallholder farmers. Afr J Agric Res 5:2868–2874

Mayfield H (2015) Making the most of machine learning and freely available datasets: a deforestation case study. PhD Thesis, University of Queensland. doi:10.14264/uql.2015.1018

McCloskey JT, Lilieholm RJ, Cronan C (2011) Using Bayesian belief networks to identify potential compatibilities and conflicts between development and landscape conservation. Landsc Urban Plan 101(2):190–203

Morgan MG, Henrion M (1990) Uncertainty: a guide to dealing with uncertainty in quantitative risk and policy analysis. Cambridge University Press, New York

Murdoch G (1968) Soils and land capability in Swaziland (Bulletin 24). Ministry of Agriculture, Mbabane

Mushala HM, Kanduza AM, Simelane NO, Rwelamira JK, Dlamini NF (1994) Comparative and multi-dimensional analysis of communal and private resources (Land) tenure in Africa: the case of Swaziland. Food Agric Organ, Mbabane

Ngwenya P, Hassan R (2005) An environmental accounting approach to valuing the services of natural forests and woodlands in Swaziland. Agrekon 44:264–283

Nixon DJ (2006) Guide to the soils of the Swaziland sugarcane industry. Swaziland sugar association and the South African Sugarcane Research Institute. http://www.wossac.com/downloads/19837_Soils_of_the_Swaziand_Sugar_Industry.pdf Accessed 21 February 2016

Pearl J (1988) Probabilistic reasoning in intelligent systems. Morgan Kaufmann, San Mateo

Porro R, Lopez-Feldman A, Vela-Alvarado JW (2015) Forest use and agriculture in Ucayali, Peru: livelihood strategies, poverty and wealth in an Amazon frontier. For Policy Econ 51:47–56

Porwal A, Carranza EJM, Hale M (2006) Bayesian network classifiers for mineral potential mapping. Comput Geosci 32(1):1–16

QGIS Development Team (2012) QGIS geographic information system. Open source geospatial foundation project. http://qgis.osgeo.org Accessed 7 September 2014

Remmelzwaal A, Dlamini WS (1994) Present land use map of Swaziland, scale 1:250,000. Food and Agriculture Organization/United Nations Development Programme/Government of Swaziland, Mbabane

Remmelzwaal A, Vilakati JD (1994) Land Tenure Map of Swaziland, Scale 1:250,000. Food and Agriculture Organization/United Nations Development Programme/Government of Swaziland, Mbabane

Robinson BE, Holland MB, Naughton-Treves L (2013) Does secure land tenure save forest? A meta-analysis of the relationship between land tenure and tropical deforestation. Glob Environ Change 29:281–293

Rolland J, Condamine FL, Jiguet F, Morlon H (2014) Faster speciation and reduced extinction in the tropics contribute to the mammalian latitudinal diversity gradient. PLoS Biol 12(1):e1001775

Romijn E, Lantican CB, Herold M, Lindquist E, Ochieng R, Wijaya A, Murdiyarso D, Verchot L (2015) Assessing change in national forest monitoring capacities of 99 tropical countries. For Ecol Manag 352:109–123

Roques KG, O' Connor TG, Watkinson AR (2001) Dynamic of shrubs encroachment in an African savanna: relative influences of fire, herb ivory, rainfall and density-dependence. J Appl Ecol 38:268–280

Rudel TK (2007) Changing agents of deforestation: from state-initiated to enterprise driven processes, 1970–2000. Land Use Policy 24:35–41

Rudel TK (2013) The national determinants of deforestation in sub-Saharan Africa. Philos Trans R Soc B 368:20120405. doi:10.1098/rstb.2012.0405

Specht MJ, Pinto SRR, Albuqueque UP, Tabarelli M, Melo FPL (2015) Burning biodiversity: fuelwood harvesting causes forest degradation in human-dominated tropical landscapes. Glob Ecol Conserv 3:200–209

Spirtes P, Glymour C, Scheines R (1993) Causation, prediction and search. Lecture notes in statistics No. 81. Springer-Verlag, New York

Spracklen BD, Kalamandeen M, Galbraith D, Gloor E, Spracklen DV (2015) A global analysis of deforestation in moist tropical forest protected areas. PLoS One 10(12):e0143886. doi:10.1371/journal.pone.0143886

Stringer LC (2009) Testing the orthodoxies of land degradation policy in Swaziland. Land Use Policy 26:157–168

Swaziland Environment Authority (2012) State of environment report for Swaziland. Swaziland Environment Authority, Mbabane

Tefera S (2013) Rangeland degradation in semi-arid Swaziland: effects of dip-tanks on herbaceous vegetation and soil properties. Afr J Range Forage Sci 30(3):127–140. doi:10.2989/10220119.2013.772539

United Nations Framework Convention on Climate Change UNFCCC (2014) Report of the Conference of the Parties on its nineteenth session, held in Warsaw from 11 to 23 November 2013. Part two: action taken by the Conference of the parties at its nineteenth session. Warsaw, Poland. http://unfccc.int/resource/docs/2015/cop21/eng/10a01.pdf Accessed 10 November 2015

van der Werf GR, Morton DC, DeFries RS, Olivier JGJ, Kasibhatla PS, Jackson RB, Collatz GJ, Randerson JT (2009) CO2 emissions from forest loss. Nat Geosci 2:737–738

Verma T, Pearl J (1992) An algorithm for deciding if a set of observed independencies has a causal explanation. In: Dubois D, Wellman MP (eds) Proceedings of the eighth conference on uncertainty in artificial intelligence. Morgan Kaufmann, Stanford, pp 323–330

Wheldon AE (1990) Energy in Swaziland: options for increased self-reliance at individual and national level. Solar Wind Technol 7:83–88

Witten IH, Frank E, Hall MA (2011) Data mining: practical machine learning tools and techniques, 3rd edn. Morgan Kaufmann, San Francisco

Xaba BG, Masuku MB (2013) Factors affecting the choice of marketing channel by vegetable farmers in Swaziland. Sustain Agric Res 2:112–123. doi:10.5539/sar.v2n1p112

Zvoleff A (2015) Gfcanalysis: tools for working with Hansen et al. global forest change dataset. R package version 1.4. https://CRAN.R-project.org/package=gfcanalysis. Accessed 17 March 2016