A Bayesian Belief Network to assess rate of changes in coral reef ecosystems

Chiara Franco, Leanne Appleby Hepburn, David J. Smith, Stephen Nimrod, Allan Tucker

A Coral Reef Research Unit, University of Essex, United Kingdom
b St. George's University, Grenada
c Department of Computer Science, Brunel University, United Kingdom

ARTICLE INFO

Article history:
Received 8 January 2016
Received in revised form 3 February 2016
Accepted 23 February 2016
Available online 5 March 2016

Keywords:
Coral reef
Calcium carbonate budget
Bayesian network
Anthropogenic disturbances
Climate change
Environmental management

ABSTRACT

It is crucial to identify sources of impacts and degradation to maintain functions and services that the physical structure of coral reef provides. Here, a Bayesian Network approach is used to evaluate effects that anthropogenic and climate change disturbances have on coral reef structure. The network was constructed on knowledge derived from the literature and elicited from experts, and parameterised on independent data.

Evaluation of the model was conducted through sensitivity analyses and data integration was fundamental to obtain a balanced dataset. Scenario analyses, conducted to assess the effects of stressors on the reef framework state, suggested that calcifying organisms and carbonate production, rather than bioerosion, had the largest influence on the reef carbonate budgetary state. Despite the overall budget remaining positive, anthropogenic pressures, particularly deterioration of water quality, affected reef carbonate production, representing a warning signal for potential changes in the reef state.

1. Introduction

Coral reefs are important ecosystems that support biodiversity and provide ecological, social and economic benefits for many communities (Moberg and Folke, 1999; Cesar et al., 2003; Burke et al., 2011).

The extent to which services (e.g. shore protection) and functions (e.g. biodiversity) are maintained by coral reef ecosystems is associated with the persistence of their three-dimensional structure (framework; Perry et al., 2008). Unfortunately, coral reefs have suffered, and continue to suffer, significant framework degradation and loss (Alvarez-Filip et al., 2009; Perry et al., 2013). Anthropogenic disturbances and pressures, such as urban and industrial developments, destructive fishing activities, catchment misuse and coastal and inland deforestation (Burke et al., 2002, 2011; Edinger et al., 2000), have increased the vulnerability of these systems to climate variability (Hoegh-Guldberg et al., 2007; Anthony et al., 2008; Baker et al., 2008; Eakin et al., 2010), overall threatening reefs' functionality (Kennedy et al., 2013).

Rate of changes of the reef framework have been largely investigated through carbonate budget assessments that estimate contribution from reef-building (e.g. hermatypic corals, crustose coralline algae) and bioerosive (e.g. sea urchins, sponges, parrotfish) organisms (Eakin, 1996, 2001; Edinger et al., 2000; Hubbard et al., 1990; Mallela and Perry, 2007; Perry et al., 2013; Stearn and Scoffin, 1977). Coral reef structural integrity is associated with positive budgets that occur when calcium carbonate production exceeds the rate of erosion, whereas negative budgets occur generally as a result of changes in the natural reef processes (Perry et al., 2008; Kennedy et al., 2013).

Despite carbonate budgets being valuable in determining the ‘state’ of a reef system, they do not always provide a full picture of disturbances and pressures responsible for changes in the framework state, and are therefore limited in their application for long-term management. In addition, varied and often incomplete datasets, as well as limited knowledge on the relationships between coral reef abiotic and biotic factors, can result in considerable uncertainty in the parameters of resulting models. In coral reef ecosystems uncertainty may be associated with ecological and biological processes (e.g. coral reef framework growth and erosion)
and with changes triggered by climatic and anthropogenic disturbances and pressures (e.g., changes in ecosystem state due to extreme sea water temperatures, sedimentation, water pollution). This has the potential to limit the identification of management priorities and the definition of effective management actions (Olsson et al., 2004; Smith et al., 2011).

A comprehensive approach that integrates uncertainty, can aid sustainable coral reef management and prevent further decline. Although, it is impossible to predict with certainty the result of management decisions, it is important to provide decision-makers with models that consider the impacts of implementing management interventions or decision options in order to maximize their benefit (Uusitalo et al., 2015). Therefore, to evaluate the effects of anthropogenic and climatic disturbances on the reef framework, we propose a Bayesian Belief Network (BBN) approach, which offers a methodological framework to address uncertainty (Bennett et al., 2013; Kelly et al., 2013).

BBNs associate variables via conditional probability distributions and use inference algorithms to calculate posterior probabilities of the outcome states (Jensen and Nielsen, 2007). They consist of two structural components: (1) a direct acyclic graph (DAG), where each vertex represents one of the variables in the model; (2) conditional probability tables (CPTs), indicating the strengths of the links in the graph by denoting the likelihood of the state of a ‘child’ node given the states of its ‘parent’ nodes (those from which edges entering the node originated) (Renken and Mumb, 2009; Landuyt et al., 2013). The DAG consists of a set of variables or nodes that can take on a number of pre-defined discrete “states”, which are mutually exclusive and exhaustive (Borsuk et al., 2004). The presence of an edge linking two variables indicates the existence of statistical dependence between them (Aguilera et al., 2011). Inference can be used to propagate conditional probabilities through the network (Aguilera et al., 2011), whilst accounting for uncertainty.

BBNs enable the integration of empirical data and expert knowledge (Uusitalo, 2007; Aguilera et al., 2011; Chen and Pollino, 2012), can operate in a data poor environment (Uusitalo, 2007), and can be readily updated with newly available data, by combining the new information with prior probabilities such that, the network posterior probabilities are updated in response to additional observational information (Marcot et al., 2001). Although BBNS are efficient in integrating variables presented at different scales (Wooldridge et al., 2005), they are constrained in describing explicit spatial and temporal dynamics and interactions, requiring the use of different nodes to represent and incorporate information on different locations or times (Marcot et al., 2001; Kelly et al., 2013). In addition, since they do not allow for feedback loops among variables, time steps to describe such effects are needed (Marcot et al., 2001; Kelly et al., 2013), adding complexity to the model and limiting their application to systems or processes described by feedback interactions (e.g. nutrient cycle; food webs). Due to the explicit handling of uncertainty (as well as their ability to integrate different type of data and knowledge) BBNS provide the opportunity to identify key knowledge gaps in our scientific understanding of complex systems, and hence inform future research priorities (Marcot et al., 2001; Uusitalo, 2007; Renken and Mumb, 2009).

The graphical structure of BBNS is particularly relevant in ecosystem management since it facilitates a participatory approach during the development of the model and provides a user-friendly framework to communicate the results (Marcot et al., 2001; Borsuk et al., 2004; Jakeman et al., 2006; Aguilera et al., 2011; Chen and Pollino, 2012; Vilizzi et al., 2013). Extensive reviews of the use of BBN for environmental modelling can be found in Aguilera et al. (2011) and Chen and Pollino (2012).

The Carbonate Budget BBN (CARBNET) was developed to evaluate coral reef carbonate balance under changing environmental conditions and across reef bioregions. The aim was to identify those disturbances and pressures that exert the greatest influence in modifying the reef framework CaCO₃ (carbonate) budgetary state.

2. Methods

2.1. Network development process

CARBNET construction followed a well-defined procedure through which i) variables affecting and describing the state of the ‘Calcium carbonate budget’ output node were identified, ii) the relationships among these variables were identified, iii) the CPT tables were populated with data cases after discretisation of the data.

2.1.1. Identification of the variables composing CARBNET network

In CARBNET, variables contributing to coral reef framework growth and destruction were identified through a literature search with the key words ‘carbonate budget + coral reef’, ‘CaCO3 budget + coral reef’ and ‘carbonate budget + coral reef’ conducted in ISI Web of Science (Reuters) and ReefBase (http://www.reefbase.org) between November 2010 and January 2011. The variables composing the network were selected among those that defined the quantitative contribution of the reef-building and bioerosive taxa to the reef carbonate budget (see Appendix A). Reef-building organisms were identified as calcifying organisms (i.e. hard corals, crustose coralline algae and epibionts) that contribute to biogenic carbonate production and deposition. Bioeroders were identified as organisms contributing to chemical or mechanical removal of carbonate from the reef framework while grazing on (i.e. sea urchins and parrotfish) or boring into (i.e. sponges, bivalves, sipunculans, polychaetes and euendoliths) the reef substrate.

Climatic and anthropogenic disturbances and pressures were also included in the network to determine the extent to which impacts affect reef preservation and carbonate balance. In this paper we refer to disturbances as ‘actions’ (e.g., logging) that can translate into increasing pressures (e.g. sedimentation and eutrophication) on the ecosystem, leading to likely changes in the state of the reef communities (response), that as a consequence, may impact reef framework functionality. Many of these disturbances included those arising from climate change, such as sea surface temperature rise, ocean acidification and increasing occurrence of hurricane or cyclones, as well as disturbances from human activities including destructive fishing practices, inland deforestation, coastal degradation and fish farming. In light of increasing regional and global anthropogenic and climate change disturbances, features giving information on the effects of disturbances on reef ecosystem and communities, allowed for ‘what if’ analysis, providing the basis to underpin changes in the framework CaCO₃ budgetary state (Cooper et al., 2009; Burke et al., 2011), as well as illustrating the effects that implemented management interventions may have on preserving coral reef framework. Disturbance and pressure variables were identified among the carbonate budget studies that assessed changes in the CaCO₃ budgetary state, relative to climatic and anthropogenic impacts (Table A.2, Appendix A), as well as from the Reefs at Risk tool (Burke et al., 2002; Burke and Maidens, 2004; Burke et al., 2011, 2012) where threats to the world’s coral reefs are described through map-based indicators (Table A.3, Appendix A).

As part of this review, the information was conceptualised in a diagram (not shown) in which dependencies were identified among variables including the ‘Calcium carbonate budget’ response node.
2.1.2. CARBNET structure evaluation

The CARBNET conceptualisation was proposed to twenty experts in the field of coral reef management and ecology to identify flaws in the network structure and address structural bias before model parameterisation.

The number of experts formally consulted in environmental modelling studies vary (Krueger et al., 2012). BNs constructed for water quality and watershed management have been conceptualised with the inputs from three to six experts (Marcot et al., 2001; Martin de Santa Olalla et al., 2012; Lynam et al., 2010), whilst expert numbers can increase for models constructed and developed during workshops (Villizi et al., 2013). Generally, a participatory approach with workshops including experts, stakeholders and final users has been used to construct and develop BBN models (Tiechurst et al., 2007; Inman et al., 2011; Richards et al., 2013), whilst other studies have validated the structure using questionnaires proposed to experts and stakeholders (Martin de Santa Olalla et al., 2007). Here we used a mixed approach based on questionnaires with experts and interviews with the final users.

Experts were consulted through an online questionnaire comprising of 29 questions, of which 73% were closed questions (see supplemental material, Appendix C). Open ended questions were used to provide the experts with an opportunity to detail their opinions. Feedback to the questionnaires was provided by twelve experts located in the United Kingdom, Caribbean, Philippines, Thailand, Western Indian Ocean and USA. As a result, nodes considered redundant and edges indicated as incorrect or missed, were re-investigated and evaluated through a tailored literature search. The majority of the experts (84%) agreed with most of the literature-based edges in the network, whilst 13% partially agreed and 3% disagreed. All edges that experts queried as incorrect were re-investigated in the literature and in the case of conflicting information on dependency between variables, the edge was omitted to avoid introducing bias in the network structure. Similarly, the 15 edges indicated by experts as missed dependencies among variables were investigated in the literature before updating the network. For instance, the impact exerted by sediments on macroalgae cover was included in the network following feedback from the experts that indicated the importance of this dependency. Indeed, previous studies confirmed that direct deposition of sediments on frondose elongate algae can affect their physiological processes (e.g. photosynthesis or gas and nutrient exchange) reducing growth and biomass with direct effects on their cover (Schaffelke, 1999; Fabricius, 2005; Airoldi and Cinelli, 1997; Umar et al., 1998; Kawamata et al., 2012). Based on feedback from the experts, the network was updated and presented to the final users — managers and wardens of the Grenada Molinière-Beauséjour and Carriacou Marine Protected Areas (MPAs). Interviews with the final users were conducted, between January and February 2013, to assess how the model conceptualisation would be interpreted and if any final changes needed to be made. Cosmetic changes (i.e. name changes) were suggested from both experts and final users, and incorporated into the final model. For instance, managers suggested that the use of common names instead of taxa specific names (e.g. sea urchin instead of echinoids) can improve model interpretation, overall improving accessibility to the graphical representation of the model.

This participatory approach was used to refine the conceptual diagram to its final version (Fig. 1). The final CARBNET diagram (Fig. 1) is comprised of anthropogenic and climatic disturbances, the pressures they exert on the ecosystem, through variables representing reef-building and bioeroder communities and the CaCO3 budgetary state.

Nodes representing different levels of spatial resolution were used to capture changes that may occur at different spatial scales. Presence/absence of reef-building and erosive organisms or reef growth and erosion processes are captured at the smallest scale of reef depth, but also for an entire reef (‘Site’), sub-region (‘Reef type’, ‘Reef topography’) or region (‘Coral reef region’). For instance, the node ‘depth’ (states: shallow, mid and deep), relates depth-specific reef responses to reef growth and erosion, whilst the node ‘Reef topography’ relates the same processes to whole reef systems. The final model is composed of 58 nodes and 94 cause-effects links.

2.2. Data and discretisation

In CARBNET, variables representing anthropogenic and climatic disturbances were described through categorical states, while ranges were used to represent the states of most of the environmental and taxa nodes (Appendix B).

Quantitative data used to inform the biological components of the network, were collected in the Wakatobi Marine National Park (southeast Sulawesi, Indonesia; Fig. 2) and in the Molinière-Beaúséjour MPA (Grenada, Caribbean; Fig. 2). Data on benthic percentage cover, sea urchin dimensions (test size) and relative density, porrofish dimensions, porrofish life phase and relative density, macroborders density and water quality were collected for each region from four sites at three depths (shallow: 2–4 m depth; mid: 6–8 m depth; deep: 12–16 m depth). Estimates of benthic percentage cover and sea urchin density were conducted along 10 m belt transects (n = 6) in Grenada and 30 m belt transects (n = 2) in Indonesia. Porrofish density, dimensions and life stage were visually assessed along 30 m belt transect (Grenada n = 6; Indonesia n = 3) deployed parallel to the shore and following reef contours. Before the node states were defined, CaCO3 production and erosion rates (kg CaCO3 m⁻² y⁻¹) were calculated using the ReefBudget tool (Perry et al., 2012). Carbonate budgets were determined for each site as the difference between mean gross carbonate production and erosion. ReefBudget is a reproducible methodology that provides a quantitative estimation of the carbonate deposited and accumulated on the reef framework by determining net reef production. Integrating ReefBudget components and results in CARBNET allowed standardised data collection and analysis.

The effects of anthropogenic pressures on the sites were determined for sedimentation, turbidity and nutrient concentrations. Sediment rates and turbidity data were collected at all sites, whilst nutrients concentration data was collected for Grenada only due to logistical challenges. This resulted in nutrients nodes for the Wakatobi sites having uninformative uniform distributions due to missing data.

Sea surface temperature (SST) variations for Wakatobi and Grenada for the years 2011 and 2012 were obtained from satellite bi-weekly data from NOAA Coral Reef Watch virtual stations (CRW, coralreefwatch.noaa.gov/satellite/vs/docs/list_vs_group_latlon_201103.php). Wakatobi virtual station was located approximately 25 km north from the sites (5.0° S, 124° E) while Buccoo Reef (Tobago,11.5° N, 61° W), the virtual station used for Grenada’s data was located approximately 65 km south of Grenada east coast.

During the surveys bleached coral colonies were not observed at any site. Field observations were corroborated using NOAA CRW records for 2011 and 2012 on SST anomalies (defined as the difference between SST and the daily climatological SST) and degree of heating weeks (a measure of thermal stress accumulated over a 12-weeks period used to define the likelihood of bleaching). Atmospheric carbon dioxide was assumed to be equal to 380 ppm (equal to aragonite saturation above 3.3) at all sites and depths, as for coral reef scenario A (CRS-A) from Hoegh-Guldberg et al. (2007), corresponding to reefs that can sustain net carbonate accretion. The dataset comprised 192 records for each of the 58 variables, with
missing data for nutrient concentrations for the Indonesian sites and for the erosive sponges, worms and bivalve variables for one site in Indonesia. In the case of the nodes ‘Sediment infilling’, ‘Phytoplankton bloom’ and ‘Sub-lethal bleaching’ for which information were not available, CPTs were defined using uninformative uniform distributions (see Table B.1, Appendix B for details on the variables and units).

Data cases derived from the literature (LIT) on studies conducted on the state of Discovery Bay reefs (Jamaica, Caribbean) were used as independent dataset for model validation. This literature-based dataset (LIT) included data from 1980 to 2003 on hard coral and crustose coralline algae percentage covers, sea urchin density and dimension, and parrotfish density and biomass (Liddel and Ohlhorst, 1987; Morrison, 1988; Picou-Gill et al., 1991; Liddel and Ohlhorst, 1992; Steneck, 1993; Andres and Witman, 1995; Crawford, 1995; Miller et al., 1996; Sary et al., 1999; Aronson and Precht, 2000; Cho and Woodley, 2000; Munro, 2000; Grant et al., 2001; Crawford and Carpenter, 2001; Haley and Solandt, 2001; Hawkins and Roberts, 2003; Quinn and Kojis, 2005; Bechtel et al., 2006; Crabbe, 2008, 2010; Gayle et al., 2010). Carbonate production, erosion and budget were calculated from 2000 to 2003 using mean benthic percentage covers data, following Perry et al. (2012) and Mallela and Perry (2007). Reef topographic complexity (i.e. rugosity) was estimated using Alvarez-Filip et al. (2009) for the same years. Rates of carbonate erosion were estimated following Perry et al. (2012) for Diadema antillarum urchins, from individuals’ size class measurements and density data, and for parrotfish from density and size class ranges, assuming that all parrotfish individuals below 45 cm (total length) were in their initial life phase for the years between 2000 and 2003. In the LIT dataset variables were defined by 24 records, with missing data relative to the macro-invertebrate variables.

Continuous variables were discretised based on bins derived from the literature and used to define the states of the network nodes (Appendix B). The literature-based discretisation process provided a transparent method to update the nodes states based on the state of the knowledge. Node states were assigned based on the ecological and functional thresholds of the biological and environmental variables.

For instance, a sedimentation rate threshold of 10 mg cm$^{-2}$ d$^{-1}$ has been suggested as a tipping point for hermatypic coral survival (Rogers, 1990; Fabricius, 2005). However, below this threshold, chronic sedimentation can affect coral growth and fitness by increasing coral metabolic costs through the removal of settled particles or reduction of photosynthetic yield (Fabricius, 2005). Furthermore, the sedimentation threshold is lowered for coral recruits that are more sensitive to changes in sedimentation than adult colonies, with negative consequences for the reseeding of coral communities (Fabricius, 2005). The stress exerted on corals by a quantity of sediment below the coral survival threshold, is linearly related to the duration of sedimentation, so that long periods of sediment deposition exerts a similar effect to that produced by a large amount of sediment deposited in a shorter time (Rogers, 1990; Fabricius, 2005). From a management perspective, sedimentation stress to coral reefs is due to changes in sedimentation rates and consequently water quality, rather than high-ambient...
sediment levels (Risk and Edinger, 2011). Therefore, rather than define the states of the ‘Sediment rates’ node (Fig. 1) on the 10 mg cm$^{-2}$ d$^{-1}$ coral survival threshold, bins were selected to account for the amount of sediment deposited and time of deposition (Appendix B; Table B.1). Although the boundary selected for the intermediate node state (3–10 mg cm$^{-2}$ d$^{-1}$) accounted for estimates that should be regarded as a warning for a clear-water reef, it should be noted that too large boundaries may not be able to capture subtle dynamics within a system (Renken and Mumby, 2009; Chen and Pollino, 2012). Importantly, the model does not apply for coral reefs growing in turbid and sedimentary environments, where corals grow and survive under high ambient sedimentation (Roy and Smith, 1971; Perry and Larcombe, 2003), since bins were selected based on literature that evaluated the effect of sedimentation on clear-water reef settings.

Node size was maintained from a minimum of two to a maximum of four states to reduce CPT size and therefore decrease computational costs.

2.3. Parameter estimation process

The relationships between the states of parent (independent) and child (dependent) nodes were quantified within CPTs that presented the probability of a node to take on each discrete state, given the states of its parent nodes (Marcot et al., 2001; Pollino et al., 2007; Chen and Pollino, 2012). Parameter estimation was conducted using the Bayes Net Toolbox in MATLAB (Murphy, 2001). CPTs were initially compiled with dirichlet priors (Heckerman et al., 1995). These prior CPTs were updated using data from the field. Field data were combined and bootstrapped (Bennett et al., 2013). This involved re-sampling with replacement ($n = 250$) and the cases excluded from the bootstrapped sets during the bootstrapping process (BLO), were used together with the LIT dataset as independent data (testing set) for model validation. Due to the presence of missing values, the training set was created by allowing the model to learn the conditional probabilities from bootstrapped partially observed data using the expectation maximization (EM) algorithm (Lauritzen, 1995). EM estimates the CPTs based on the structure of the network and the dataset by finding the marginal posterior probability for each node that yields the greatest likelihood given the available data (Lauritzen, 1995; Chen and Pollino, 2012). The model was then tested using i) LIT and ii) LIT combined with BLO, as independent datasets (testing set). In this way none of the data used to train the model were used for testing it, hence validation was conducted using a testing set of data that resulted as ‘new’ or ‘unknown’ to the model.

2.4. Evaluation of the CARBNET

2.4.1. Accuracy

The ability of the model to correctly predict instances on independent/unseen data was investigated by using resampling/holdout analysis and inspecting the resultant confusion matrices (Bennett et al., 2013). For example, for the node ‘Calcium carbonate budget’, the predicted and actual instances were cross-tabulated for each of the four states resulting in a $4 \times 4$ matrix (positive high,
positive low, negative low, negative high; Pollino et al., 2007). Correctly classified instances (CCI) were assessed as the percentage of cases correctly predicted by the model divided by the total number of cases, providing a measure of how many instances the model predicts correctly when tested against an independent dataset. Model accuracy (Bennett et al., 2013) was estimated by averaging the CCI of all nodes obtained by testing the model against the independent sets i) LIT and ii) the LIT combined with BLO. In addition, weighted Kappa and the correctly classified instances are presented for the output node ‘Calcium carbonate budget’. Weighted Kappa measures the degree of agreement between the observed and predicted cases, and accounts for the proportion of disagreement between the actual categories, assigning less weight to the agreement as categories are further apart (Fleiss and Cohen, 1973; Viera and Garrett, 2005).

2.4.2. Sensitivity analysis

Sensitivity analyses (Norton, 2015) were conducted using the LinkStrength package (Ebert-Uphoff, 2007) implemented for MATLAB’s Bayes Net Toolbox to identify the relative influence of the variables in the network.

Since “the probability distribution (CPT) of each node is a depiction of uncertainty” (Marcot et al., 2001), the uncertainty associated with each node was measured using entropy which is a score of a variable’s “richness” (i.e. how much information is within the data for that particular variable; Pollino et al., 2007; Pearl, 1988). Therefore, as suggested by Marcot et al. (2001), the uncertainty associated with a particular node is propagated to the probability distribution of the output node when solving the network. Nodes were then ranked accordingly from the most uncertain to the least uncertain, with the most uncertain variables being less informative within the network.

The sensitivity of one node to multiple other nodes, was evaluated through the mutual information (MI) that determines if the state of a particular node is sensitive to the state of others by measuring the connection strength for any pair of nodes (adjacent or not), taking any possible path between them into account (Ebert-Uphoff, 2007). When MI is equal to zero, nodes are known to be mutually independent and therefore the condition of one node does not affect the state of another (Pearl, 1988).

One limitation of entropy-related measures is based on the uncertainty being affected by imbalanced data as uniform distributions are typically treated as the most uncertain. However, if discretisation is poorly chosen and there are not many cases of some states then a highly non-uniform distribution (with apparently low uncertainty) is generated. What is more, the scale and ordering of the intervals are not taken into account. Care must therefore be taken when interpreting these results but it should be noted that these limitations are shared by any other measure that is a function of only the probabilities of a random variable’s states to measure uncertainty. Entropy nevertheless remains by far the most popular measure for uncertainty (Pearl, 1988).

Results of these two measures are generally used to identify the most relevant variables for predicting the output node as well as in particular cases identifying gaps associated with specific nodes—and therefore the need for more data collection.

2.5. Model predictions

CARBNET was applied to scenario-based analysis to determine the alternative conditions in the state of the reefs relative to the effects of anthropogenic and climatic disturbances and pressures on the coral reef framework. Predictions are aimed at demonstrating the utility of CARBNET to quantify alternative states of the ‘Calcium carbonate budget’ node, determined by changes in the carbonate producer and bioeroder communities driven by modifications in the state of the disturbances and pressures variables. Disturbances and pressures can affect reef-building organisms and bioeroder communities, reducing carbonate production and/or increasing framework erosion with consequences to the CaCO3 budgetary state (Kennedy et al., 2013). Here we present three examples of the scenarios modelled. Scenario 1 presents the current budgetary condition and was defined by leaving the node states unaltered, i.e. the distribution with no evidence introduced. Scenario 2 assumes that changes in the budgetary state are due to changes in the environmental conditions as a consequence of the effect that disturbances have on the natural reef processes. The scenario was defined by setting the disturbance nodes (Fig. 1) to their highest state (e.g. probability of severe coastal degradation equal to 1, meaning that severe degradation was certain). Scenario 3 considers changes in the budgetary state for a system in which overexploitation of herbivore fish, defined by the certainty of low parrotfish density, and coastal waters deterioration due to eutrophication (Ammonia >10 μM, Phosphate >0.1 μM) and severe sedimentation (>10 mg cm−2 d−1) act together to increase macroalgae cover (state >75% equal to 1) and decrease calcifying organisms’ abundances (dominant state for ‘Hard coral cover’ and ‘CCA cover’ was <10%). The scenario outcomes, relative to the output node ‘Calcium Carbonate Budget’, were then evaluated by running the model in ‘most probable explanation’ (mpe) mode, which shows the most likely state of the network nodes given the evidence.

3. Results

3.1. Accuracy

The model showed a great disparity in the mean CCI values obtained during validation against the two independent datasets. Indeed, the model was able to classify correctly less than 50% of the instances (42%) when tested against LIT, whilst it performed better when tested against the combined LIT and BLO dataset (77%). However, CCI values were close when model was tested against LIT (62%) and against the combined LIT and BLO dataset (65%) in relation to the output node ‘Calcium carbonate budget’. In addition, a substantial agreement was found between the observed and the predicted cases with regard to the four states associated with the output node ‘Calcium carbonate budget’ when the model performance was investigated using weighted Cohen’s kappa. Also in this case agreement increased from 0.61 for the LIT to 0.64 for the combined testing set, corresponding to a greater number of ‘negative’ carbonate budget (‘low’ and ‘high’ negative) instances correctly classified when BLO was included in the independent set (Fig. 3B).

3.2. Sensitivity analysis

3.2.1. Entropy

Entropy results showed that variables for which the CPT was described by less informative probability distributions, such as, some of the biological and environmental variables were the most uncertain, whilst less uncertain variables were either those defined regionally, such as SST or ocean acidification, or those in which CPTs were defined by uninformative equal priors, such as ‘Sediment infilling’ (Table 1). Among the most uncertain nodes were intermediate nodes influencing the state of the carbonate budget by influencing calcium carbonate production (e.g. ‘Light availability’) or degradation (e.g. ‘Sea urchin size’). In particular, sea urchin erosion rates are size and species-specific, with larger urchins removing a greater quantity of carbonate substrate (Bak, 1990). Similarly,
hermatypic corals mean calcification rates (kgCaCO₃ m⁻² y⁻¹), relevant for the production of calcium carbonate from this group, are dependent on the light available for photosynthesis (Table 1).

### 3.2.2. Mutual information

Overall the model showed 13 out of 58 nodes being highly sensitive (MI > 0.3) to others. However, some of the relationships described through MI were inaccurate (e.g. ‘Ammonia concentration in waterways’ influencing ‘Phosphate concentration in waterways’), reducing the number of sensitive interactions occurring in the network.

The response node, ‘Calcium carbonate budget’, was highly sensitive to the ‘Calcium carbonate production’ and ‘Hard coral carbonate production’, whilst some sensitivity was observed in relation to ‘Calcium carbonate removal’ (Table 2), suggesting that reef budgetary conditions are influenced by carbonate production rather than degradation through erosion. Other variables influencing the output node were ‘Reef rugosity’, ‘Reef depth’ and ‘Light availability’, which either contribute to the budgetary state, as a consequence of biologically-driven carbonate deposition (i.e. reef rugosity) or influencing calcification rates. The quantity of carbonate removed through erosion from the reef framework was moderately sensitive to the bioerosion activity macro-invertebrates (Table 2).

Climate change variables had a negligible effect on the other network variables with the exception of ‘Sea Surface Temperature rise’, which mildly affected hard coral cover (MI = 0.02) and crustose coralline algae carbonate production (MI = 0.02). Conversely, anthropogenic disturbances (see Fig. 1a), had no effects on the network nodes (MI = 0). Catchment degradation had a mild effect on water quality variables (MI < 0.1) such as sedimentation and nutrient concentrations (‘Sediment load in waterways’, ‘Phosphate in waterways’, ‘Sediment rate’). However, ‘Sediment rates’ was sensitive to ‘Water entering coastal water’, suggesting that catchment management may be critical to reduce sediment pressures and sustain reef framework growth (Table 2). Similarly, ‘Lethal’ and ‘Sub-lethal bleaching’ nodes were sensitive to changes in water quality due to the nodes ‘Freshwater discharge’ and ‘Sediment rate’ affecting indirectly reef-building organisms (Table 2). The node ‘Turbidity’ (suspended particle in the water column) was sensitive to the load of sediment reaching the coastal area, and its child node ‘Light availability’ was also sensitive.
3.3. Model predictions

For the scenarios proposed, the model showed a mild response to the effects of anthropogenic and climatic disturbances on the CaCO₃ budgetary state, whilst a distinctive response was observed in relation to the anthropogenic pressures, with conditions also influenced by the level of grazing occurring on the reef.

Scenario 1. Predictions on the actual state of the reef based on the output node ‘Calcium carbonate budget’ showed a high probability of finding the carbonate budget in a positive state (Fig. 4). In this case mpe for this budgetary state was relative to the ‘positive high’ calcium carbonate production.

Scenario 2. Impacts from disturbances (Fig. 1a) were associated with a substantial increase in erosion with 12% of variation with respect to the previous scenario. When anthropogenic and climatic disturbances were defined to have highest impact on the system, negative budget states increased substantially, whilst the budgetary state decreased drastically (Fig. 4). However, this variation in the posterior probability did not produce a substantial shift in the budgetary condition and the mpe for the output node under this scenario remained a ‘positive high’ carbonate production state.

Scenario 3. Under the conditions of elevated sedimentation, eutrophication and overexploitation of herbivore fish, the posterior probability varied by 13%, resulting in substantial changes in the likelihood for the states of the output node (Fig. 4). This condition was associated with a shift to a relative lower budgetary state (mpe = ‘positive low’). It appeared that intermediate nodes can substantially affect the CaCO₃ budgetary state, by reducing the probability of the positive states whilst increasing the likelihood of both negative budget states (Fig. 4).

4. Discussion and conclusions

CARBNET provides a means for capturing the dependencies between climate change and human disturbances and pressures, and their influence on the reef framework state. The benefit of using Bayesian Belief Networks is that they explicitly model ecological causal influences and uncertainty, providing a better level of information that decision makers can use to interpret the ecological and biological changes occurring in a system (Marcot et al., 2001; Bennett et al., 2013).

During model construction, literature-based discretisation had the advantage of including information derived from studies conducted under different environmental conditions and in different coral reef bioregions, introducing a large range of possible conditions and reducing the bias that can occur when automatic discretisation is applied to variables with unbalanced data cases (Usisitalo, 2007; Chen and Pollino, 2012). These multiple sources of knowledge aided capturing a richer set of variable states, in order to improve the ability of the model to generalise across coral reef systems. However, the discretisation process is generally regarded
as a potential source of bias since “too large categories may not be able to capture subtle dynamics occurring in a system, whilst narrower categories may not be feasible due to the uncertainty and lack of data to define accurately the conditional probabilities for each possible combination of levels” (Renken and Mumby, 2009). For some of the CARBNET variables, a trade-off between uncertainty and accuracy may have been exacerbated by too large categories, which may have overlooked understudied or rare conditions that occur in coral reef systems. In this case, experts can be consulted to elicit rare or less studied conditions in order to improve model applicability. In general, we recognise that further involvement of experts is needed to assess flaws in discretisation and that this participatory process can be further applied to corroborate model outputs.

The use of standardised methodology for data collection and analysis (ReefBudget) was a practical approach to reduce the bias that can occur during the standardisation of data collected and analysed with different methodologies. ReefBudget is set up as an excel document, where estimates of carbonate production and erosion are automatically calculated based on benthic cover and density data collected using a consistent field methodology (Perry et al., 2012). This methodology provides information that is beyond just the hard coral-macroalgae relationship, hence adding to the knowledge relative to some of the complex relationships occurring on the reef.

Overall, sensitivity analyses showed the need to inform the knowledge gaps, which reside in data characterised by little variation and therefore high uncertainty. The biological and environmental components of the network resulted in high entropy values potentially due to a lack of data to encode the CPTs. Integrating the non-informative nodes with more varied data (data with a high mixture of combinations of variable states) can help to determine whether the variation observed is noise or based on some relationship with other variables. However, measures of uncertainty should be treated carefully since sometimes a uniformly distributed variable may be the result of intrinsic qualities of the variables (noise/external factors) or how the data has been discretised. Conversely, data that has non-uniform distributions may not necessarily be certain – for example due to lack of data for some states. Careful use of intervals surrying discretisation helps to overcome this to some degree. The sensitivity of the node ‘Calcium carbonate budget’ to the quantity of calcium carbonate deposited by reef-building organisms (e.g. hermatypic corals), rather than from carbonate degradation through bioerosion, may be due to the reduced number of data cases for some of the bioeroder taxon variables (e.g. macrorhizum density). Therefore, it is likely that knowledge gaps in the bioerosive taxa components (Fig. 1, d1) of the model have caused misclassification of the negative budgets.

Despite the overall budget remaining positive, the model predicted a likelihood increase in bioerosion when the system is subjected to extreme climatic and anthropogenic disturbances, indicating that disturbances have the potential for changes in the CaCO$_3$ budgetary state. However, climatic and anthropogenic input nodes were less important in influencing the output node state, being unable to produce a consistent shift to any of the other node states (mpe = “positive high” calcium carbonate production). Declining carbonate production was observed also in relation to poor water quality, supporting the general view that high sedimentation and nutrients levels are important drivers of change in coral reef framework (Hallock and Schlager, 1986; Hallock, 1988; Le Campion-Alsumard et al., 1993; Kennedy et al., 2013). Direct effects of sedimentation are well documented (see Fabricius, 2005); however, its effect on erosion is still poorly understood (Perry and Larcombe, 2003). In this context, depletion of herbivore fish appeared to magnify the response of the system to water quality degradation, suggesting that the effect of overfishing is additive to poor water quality in determining changes in the CaCO$_3$ budgetary state. There have been examples on how exploitation of herbivore fish may reduce the erosive pressure on the reef, overall maintaining a positive budget (Mallela and Perry, 2007; Perry et al., 2014). However, negative budgets may offset positive net production in the case of extensive loss in coral cover and low recruitment of juvenile corals also as a consequence of algal blooms (Perry et al., 2013). This complex interaction seems to not be accounted for, despite dependencies presented in the model. The consequence is that the model is perhaps limited in detecting hidden conditions although additional sources of data cases may improve this drawback. In a management context scenario 3 indicates that managing sources (e.g. catchment degradation) of pressures (e.g. sedimentation) may aid positive balance and reef framework growth (Perry et al., 2008; Kennedy et al., 2013).

In this paper we have demonstrated the potential of a BBN approach to model coral reef state, improving our knowledge into the complex interactions occurring between disturbances and reef framework dynamics. In addition, we have shown how such a model can be used to identify knowledge gaps that need to be informed to prompt comprehensive management strategies. Future work will involve the inclusion of information on marginal systems (e.g. coral reefs flourishing in turbid and sedimentary settings), especially in relation to the interaction between bioerosion and sedimentation, and integration of the model with more varied data to add knowledge.

Acknowledgements

We are grateful to the three anonymous reviewers for the supportive and constructive comments on the manuscript. We thank the Molinier-Beauséjour Marine Protected Area staff, Ronald Baldwin, Codddinton Jeffery and Operation Wallacea Conservation for facilitating permits and logistics. We also thank Jerry Mitchell and Jason Williams for providing sites maps and Bawo Teddy Ikolo for the support at the True Blue Campus, St. George’s University, Grenada. We acknowledge the assistance of Richard Huber, the Organization of the American States (OAS) ReefFix Program and Carlos Gigoux Gramenea. Finally we would like to thank the experts, wardens and managers consulted to improve the model structure.

Appendix A. Supplementary data

Supplementary data related to this article can be found at http://dx.doi.org/10.1016/j.envsoft.2016.02.029.

References

Reefs 21, 853–866.