Environmental Modelling & Software 42 (2013) 157–167

Contents lists available at SciVerse ScienceDirect

Environmental Modelling & Software

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

The construction of causal networks to estimate coral bleaching intensity

Lilian Anne Krug [a,b,*], Douglas Francisco Marcolino Gherardi [a], José Luís Stech [a], Zelinda Margarida Andrade Nery Leão [c], Ruy Kenji Papa Kikuchi [c], Estevam Rafael Hruschka Junior [d], David John Suggett [e]

[a] National Institute for Space Research, Remote Sensing Division, Avenida dos Astronautas, 1758, Zip Code: 12227-010 São José dos Campos, SP, Brazil
[b] University of the Algarve, Centre for Marine and Environmental Research, Campus de Gambelas, Zip Code: 8005-139 Faro, Portugal
[c] Federal University of Bahia, Institute of Geosciences, R. Barão de Jeremoabo s/n, Zip Code: 40170-115 Salvador, BA, Brazil
[d] Federal University of São Carlos, Computer Science Department, Rod. Washington Luis km 235, Zip Code: 13565-905 São Carlos, SP, Brazil
[e] University of Essex, Department of Biological Sciences, Wivenhoe Park, Colchester CO43SQ Essex, United Kingdom

ARTICLE INFO

Article history:
Received 30 January 2012
Received in revised form
8 January 2013
Accepted 9 January 2013
Available online 9 February 2013

Keywords:
Bayesian network
Coral reef
Coral bleaching
Remote sensing
Environmental variability
South Atlantic coral reefs

ABSTRACT

Current metrics for predicting bleaching episodes, e.g. NOAA's Coral Reef Watch Program, do not seem to apply well to Brazil's marginal reefs located in Bahia state and alternative predictive approaches must be sought for effective long term management. Bleaching occurrences at Abrolhos have been observed since the 1990s but with a much lower frequency/extent than for other reef systems worldwide. We constructed a Bayesian Belief Network (BN) to back-predict the intensity of bleaching events and learn how local and regional scale forcing factors interact to enhance or alleviate coral bleaching specific to Abrolhos. Bleaching intensity data were collected for several reef sites across Bahia state coast (~12°–20°S; 37°–40°W) during the austral summer 1994–2005 and compared to environmental data: sea surface temperature (SST), diffuse light attenuation coefficient at 490 nm (K_{490}), rain precipitation, wind velocities, and El Niño Southern Oscillation (ENSO) proxies. Conditional independence tests were calculated to produce four specialized BNs, each with specific factors that likely regulate bleaching intensity. All specialized BNs identified that a five-day accumulated SST proxy (SSTAc5d) was the exclusive parent node for coral bleaching producing a total predictive rate of 88% based on SSTAc5d state. When SSTAc5d was simulated as unknown, the Thermal-Eolic Resultant BN kept the total predictive rate of 88%. Our approach has produced initial means to predict beaching intensity at Abrolhos. However, the robustness of the model required for management purposes must be further (and regularly) operationally tested with new *in situ* and remote sensing data.

© 2013 Elsevier Ltd. All rights reserved.

1. Introduction

Coral bleaching can occur as a response to stressful environmental conditions induced by direct local pressures as well as indirect regional/global climatic change (Carpenter et al., 2008; Suggett and Smith, 2011). Although anomalous surface water heating is acknowledged to be the most important stressor, the interplay of additional environmental variables moderates the net bleaching response (Glynn, 1993; Fitt et al., 2001). Indeed, most observations to date demonstrate that corals are most vulnerable to bleaching during periods of clear sky, calm sea, weak winds and clear water that maximise the amount of heat and light reaching the coral soft tissue (Baker et al., 2008; Brown, 1997; Glynn, 1993). However, such patterns may also be influenced by genetic composition of the corals, symbiotic algae or hosts (see Baker et al.,

Abbreviations: AGRRA, Atlantic and Gulf Rapid Reef Assessment; BN, Bayesian networks; BNPC, Belief Network Power Constructor; BSST, best sea surface temperature; CI, conditional independence; CPT, conditional probabilities table; K_{490}, diffuse light attenuation coefficient at 490 nm; DAG, directed acyclic graph; EM, expectation–maximization; MaxSST, maximum sea surface temperature; |W|, mean surface wind fields; V, meridional wind; MEI, multivariate El Niño index; PCA, principal component analysis; PPT, rain precipitation; SST, sea surface temperature; SSTAc5d, sea surface temperature accumulated in five days; U, zonal wind.

* Corresponding author. University of the Algarve, Centre for Marine and Environmental Research, Campus de Gambelas, Zip Code: 8005-139 Faro, Portugal. Tel.: +351 289 800 900x7372; fax: +351 289 800 100.

E-mail addresses: lakrug@ualg.pt (L.A. Krug), douglas@dsr.inpe.br (D.F.M. Gherardi), stech@dsr.inpe.br (J.L. Stech), zelinda@ufba.br (Z.M.A.N. Leão), kikuchi@ufba.br (R.K.P. Kikuchi), estevam@dc.ufscar.br (E.R. Hruschka), dsuggett@essex.ac.uk (D.J. Suggett).

1364-8152/$ – see front matter © 2013 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.envsoft.2013.01.003

2008), which may, in turn, reflect the intrinsic environmental history of the populations in question (e.g. Oliver and Palumbi, 2011).

Indirect climatic forcing of coral bleaching is often observed across large regional scales and frequently attributed to El Niño driven thermal anomalies (Baker et al., 2008; Goreau and McClanahan, 2000). Globally significant mass coral-bleaching episodes occurred in 1982–83 and 1997–98 but with clearly variable effects amongst reefs within regions (e.g. Berkelmans et al., 2004; Mumby et al., 2001). Thus, accounting for regional scale variability in compounding environmental variables is vital to increasing the accuracy in predicting the likely severity of bleaching outbreaks (Maina et al., 2008; McClanahan et al., 2007).

Reefs throughout Bahia state represent South Atlantic's largest regional scale reef complex and contain 84% of the Brazilian coral and hydrocoral species, which appear to exhibit mixed responses to global patterns of coral bleaching (Leão et al., 2008). Variations in mean SST, cloud cover and turbidity induced by the 1997–98 El Niño caused partial mortality of octocorals and actiniarians at Bahia, but had limited impacts on scleractinian communities (Kelmo et al., 2003). Such observations suggest that SST anomalies alone may not explain the observed bleaching. For example, Brazil's reefs are dominated by endemic coral species, notably *Mussismilia braziliensis* is an endemic species that have adapted to high levels of turbidity (Leão et al., 2003). It is possible that environmental regulation and localised adaptation shape a specific relationship between bleaching susceptibility to anomalous environmental conditions here (see Suggett and Smith, 2011). However, whilst observations of coral bleaching for Brazil's reefs have been recorded since 1994 during the peak SSTs (February–April; Castro and Pires, 1999; Leão et al., 2008), there still remains an overall scarcity of field data (Castro and Pires, 1999). Robust relations governing local and remote atmospheric and oceanic forcing of coral bleaching have not been found and other approaches must be sought.

Modelling the complex interactions of atmosphere (e.g. winds and clouds) and oceans (e.g. SST and the underwater light field) with coral biology is a major challenge for anticipating future coral bleaching events. Site-specific variations have prompted the research community to stress the need for alternative approaches to improve coral bleaching predictions (Berkelmans et al., 2004; Donner et al., 2007; Maina et al., 2008; Maynard et al., 2008; Sheppard, 2003). One such effort (Wooldridge and Done, 2004) modelled the dependencies of a series of proxies using a Bayesian Belief Network (hereafter called Bayesian Network or BN), to predict coral mortality due to the 2002 bleaching event at the Great Barrier Reef (GBR). Specifically, the authors built a probabilistic graphical model that incorporated information on thermal stress, geographic location, ecological and topographic attributes of reef areas, and achieved a predictive success rate of 72%.

The advantages of using BNs include the possibility of (i) constructing a process oriented model domain, (ii) providing explicit representation of dependency relationships between variables, (iii) handling of uncertainty and complexities by the propagation of evidence based on conditional probabilities, and (iv) testing future scenarios to support decision-making (Charniak, 1991; Chen and Pollino, 2012; Wooldridge and Done, 2004). Furthermore, BNs allow the use of different types of data and enable continuous refining of analyses upon data addition(s) and testing of hypotheses about past or future processes/events. Such predictive power provides structure to high dimensionality models and has meant that BNs have become widespread in fundamental diagnostic applications, such as the medical industry (Kahn et al., 1997; Maskery et al., 2008), ecological studies (Kuikka et al., 1999; Little et al., 2004; Uusitalo, 2007; Varis and Kuikka, 1997) as well as environmental modelling (Aguilera et al., 2011; Alameddine et al., 2011) and management (Kragt et al., 2011; Pérez-Miñana et al., 2012; Ticehurst et al., 2011).

In light of the difficulty in constructing predictive models (Maina et al., 2008; Maynard et al., 2008; McClanahan et al., 2007) for Brazil's reefs, we have developed a BN capable of learning how regional and remote forcing factors interact to determine coral bleaching intensity for the major reef complex of Bahia State, in the tropical southwest Atlantic. The complexities involved in the bleaching scenarios that have been identified for the Bahia state reefs are highlighted. This should be the starting point for future development of an operational predictive tool to aid decision-making in reef areas.

2. Materials and methods

2.1. Study area and coral bleaching cases

Reefs of the Bahia State (Brazil's East coast) are the centre of south Atlantic coral biodiversity (Spalding et al., 2001). At the northern and central coast, reefs are distributed along a narrow continental shelf ca.15 km wide. The reefs further south are mostly onshore of the Royal Charlotte and Abrolhos banks, where the continental shelf extends up to 200 km from the coast. Seven reef areas are specifically identified for this paper: 1) Litoral Norte (LN), 2) Baía de Todos os Santos embouchure (BTS), 3) Tinharé, Boipeba and Baía de Camamu (TB), 4) Cabrália (CAB), 5) Itacolomis Reefs (ITA), 6) Abrolhos Coastal Reefs (ABC), and 7) Parcel dos Abrolhos (PAB) (Fig. 1). All these reefs typically grow in shallow water $\leq$10 m depth, except for ITA, PAB and certain areas of ABC, that extend to ca. 20 m depth.

Coral bleaching data (Leão et al., 2008) were collected in these seven areas between 1998 and 2005 during the austral summer, via 10 m line transects according to the Atlantic and Gulf Rapid Reef Assessment (AGRRA) protocol (Ginsburg et al., 1998). Two other bleaching records for Abrolhos Reefs (ABC) were made in 1994 (Castro and Pires, 1999) and 2005 (Travassos, pers. comm.) that were collected via 10 m and 50 m line intercept transects, respectively. Sampling effort was comparable across surveys and all results are given in percentage of bleached coral surface relative to the total surveyed area. Unfortunately, these assessments did not consistently include detailed information on the severity of bleaching of the surveyed coral colonies (e.g. McClanahan et al., 2007). Therefore, coral bleaching events were instead arbitrarily categorised as either weak, or strong whether the percentage bleached area of all surveyed colonies was $<$ or $>$20%, respectively. Only one case of 'no bleaching' was observed for 2004 in Cabrália (CAB) reefs (Table 1), therefore, this category was not included in the prediction of bleaching intensity.

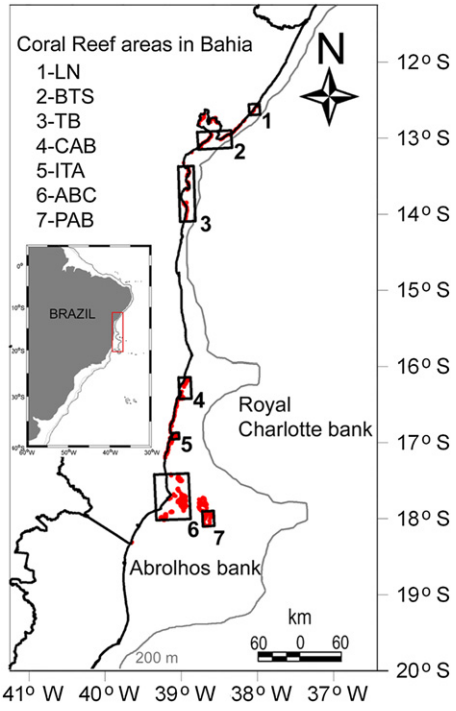

Fig. 1. The seven main reef areas at the coast of Bahia State, Brazil. The red polygons illustrate the reef patches, the black rectangles define these study reef areas limits and grey contours illustrate isobaths. In detail, the limits of the broad study area (red rectangle) used at the Principal Component Analysis. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Table 1
Environmental data and El Niño index (MEI) for months of bleaching events. Variables units are: °C for maximum sea surface temperature (MaxSST) and sea surface temperature accumulated in five days (SSTAc5d); m^{-1} for light attenuation coefficient (K_{490}); m s^{-1} for resultant wind |W|, and its components zonal (U) and meridional (V); mm d^{-1} for rain precipitation (PPT).

Area	Month/year	Intensity of bleaching	MaxSST	SSTAc5d	K_{490}	|W|	(U)	(V)	PPT	MEI	MEI_7
LN	April 1998	Strong	29.1	145.2	0.0521	7.33	−6.73	2.27	2.12	2.613	2.837
BTS	April 2003	Strong	29.4	147.0	0.0347	6.41	−5.82	2.04	6.38	0.385	0.795
T/B	March 2002	Weak	28.9	144.4	0.0635	3.39	−3.17	0.36	3.27	−0.081	0.301
T/B	April 2003	Strong	29.3	146.6	0.0943	3.18	−2.63	0.71	6.38	0.385	0.795
T/B	March 2004	Weak	28.7	143.6	0.1396	2.97	−2.71	−0.43	7.93	−0.083	0.276
ITA	April 2005	Weak	28.6	142.5	0.0705	2.03	−1.74	0.18	5.67	0.555	0.580
ABC	March 1994	Strong	28.1	140.6	0.1139[a]	5.20	−3.38	−2.47	5.49	0.137	1.064
ABC	March 2001	Weak	28.4	141.5	0.1128	5.81	−4.96	−2.45	0.86	−0.552	−0.162
ABC	March 2002	Weak	28.3	141.4	0.1053	5.55	−4.50	−1.82	2.21	−0.081	0.301
ABC	March 2003	Weak	28.7	143.6	0.1226	5.84	−4.10	−1.47	1.25	0.812	0.864
ABC	March 2005	Strong	28.4	142.1	0.1148	4.53	−3.00	−2.19	5.55	0.929	0.611
PAB	March 2000	Weak	28.0	140.3	0.0712	4.97	−3.82	−0.78	6.55	−0.927	−0.723
PAB	March 2001	Weak	28.4	141.5	0.0657	5.81	−4.96	−2.45	0.86	−0.552	−0.162
PAB	March 2002	Weak	28.2	141.0	0.0624	5.55	−4.50	−1.82	2.21	−0.081	0.301
PAB	March 2003	Weak	28.8	144.0	0.0627	5.84	−4.10	−1.47	1.25	0.812	0.864
PAB	April 2005	Strong	28.4	142.1	0.0728	3.58	−2.71	0.41	2.74	0.555	0.580
CAB	March 2004	Absent	28.4	141.8	0.0907	2.76	−2.10	−1.02	9.95	−0.083	0.276

[a] Data on K_{490} before 1997 was not available, therefore a mean value calculated for events occurring at same area and month was used instead.

2.2. Environmental data

Bleaching environments were characterized using a 13-year long time series (1993–2005). Data used from remote sensing were sea surface temperature (SST, daily and pentad) and diffuse light attenuation coefficient at 490 nm (K_{490}, 8 day mean), from analysis were rain precipitation (PPT, pentad) and from reanalysis data were zonal (U) and meridional (V) wind velocity components and resultant surface wind fields |W| (daily). All these variables were selected and calculated to cover the main forcing factors that have been reported as relevant for coral bleaching (see above, also Maina et al., 2008), in an attempt to reconstruct likely bleaching conditions for the Bahia reefs (Table 1). Association of these environmental variables with bleaching intensities was based on the mean values of each variable (except SST, see below) for the chosen reef areas (black rectangles in Fig. 1) and for the months when bleaching occurred. Summertime (February to April) maps for each year were built from the maximum values of the SST (MaxSST) and mean values of K_{490}, PPT, (U), (V) and |W|.

Daily and pentad (5-day composites) 4 km resolution SST fields (Kilpatrick et al., 2001) were obtained from the Advanced Very-High Resolution Radiometer (AVHRR) sensor (http://data.nodc.noaa.gov/pathfinder/Version5.0/). Two thermal anomaly proxies most relevant for coral bleaching were derived from the SST data. First, the MaxSST, which indicates the highest SST value in an area for each summer, obtained as the maximum SST value per pixel cell within the 18 pentads available in each summer or six pentads per month (Table 1). Second, the sea surface temperature accumulated in five days (SSTAc5d), intended to capture the persistence of high temperatures over a running five days, was calculated as follows. The daily SST in each cell was recursively summed along five consecutive days generating 18 pentads per summer (6 per month), from which the highest value was selected from each pixel to create a map with the SSTAc5d for that particular summer month (Equation (1)). A similar proxy (3-days accumulated temperature) has been reported as being highly correlated with coral bleaching in the GBR (Berkelmans et al., 2004).

$$\mathrm{SSTAc5d} = M\left(\sum_{n}^{n+4}\mathrm{SST}\right) \quad (1)$$

where M is the highest value among all the five day summations, $n = \{1, 6, 11,\ldots,86\}$ for summer and {1,6,…,26} for monthly data.

K_{490} provides a measure of water transparency and was calculated for 1998–2005 from measurements made by the Sea-viewing Wide Field of View Sensor (SeaWiFS – Mueller, 2000) (ftp://oceans.gsfc.nasa.gov/SeaWiFS/Mapped/8Day/K490/). To account for the 1994 bleaching event an average value calculated from the time series for March was used.

Rain precipitation (PPT) data were obtained from the Enhanced version of Merged Analysis of Precipitation (CMAP) developed by NOAA Climate Prediction Center (CPC) with 2.5° spatial resolution (http://www.cdc.noaa.gov/cdc/data.cmap.html). The 1.9° spatial resolution zonal (U) and meridional (V) wind data for the same period were obtained from reanalysis model developed by the National Centres for Environmental Prediction/Atmospheric Research (NCEP/NCAR). The resultant wind velocity magnitude |W| was built by the vector sum of its zonal (U) and meridional (V) components.

Two monthly time series of the Multivariate El Niño Index (MEI), one with zero lag and another with a seven months lag, were used as El Niño Southern Oscillation (ENSO) proxies. This index takes into account a number of measures of oceanographic and atmospheric parameters in the Pacific Ocean (Wolter and Timlin, 1998). The seven months negative lag relative to the bleaching period was based on the maximum cross correlation coefficient with the 20-year SST on the Abrolhos Bank ($r = 0.3$). Refer to Krug et al. (2012) for more details on the environmental data.

2.3. Statistical analyses

A standardized PCA (Eastman, 1992) was performed separately on each oceanic and atmospheric variable to obtain proxies (PCA) related to summer amplitudes. These proxies in turn provide an indication of the acclimatization extent for corals relative to their geographical location since highly variable regions inherently impose more stress but in the longer term this can carry more potential for acclimatization.

The data set consisted of a single variable time series; the first principal component (PC1) indicated the largest possible fraction of the variability contained in the original data (Eastman, 1992; Wilks, 2006). Each pixel is an element of an N-dimensional vector time series constructed from 13 summer periods, between 1993 and 2005, with the exception of K_{490}, for which the time series starts in 1998. In doing so, one can consider this vector as a continuous field sampled at N discrete points in space, according to the spatiotemporal resolution of each data set. Also referred to as the empirical orthogonal function (EOF) analysis, this statistics has the advantage of showing the mutually orthogonal spatial patterns modulated by mutually uncorrelated time series (Monahan et al., 2009).

The PCA was generated for the whole study area (red rectangle in the inset of Fig. 1) and the average of the first Principal component (PC1) loadings for each of the seven reef areas (Table 2) was used in the BNs as proxies that incorporate typical patterns of variability influencing each reef. The PC1 maps reflect the spatial patterns of summertime amplitude of variation: the higher the absolute loading value, the larger the variation for the time period (Krug et al., 2012).

Models capable of incorporating observed evidences and inherent uncertainties to both measurements and dependency relationships can be used to alert the probability of bleaching intensity for certain environmental conditions. According to Charniak (1991), given the important role played by causality on a situation or

Table 2
Mean PC1 loadings representing the variation amplitude for the seven reef areas and the maximum and minimum loadings for the broad study area (red rectangle in the inset of Fig. 1) MaxSST, K_{490}, PPT, |W|, (U) and (V) (see Table 1 for abbreviations).

Areas	MaxSST	K_{490}	PPT	|W|	(U)	(V)
PC1 eigenvalue (explained variance)	92%	37%	76%	49%	55%	54%
LN	0.0094	−0.0185	0.27	0.215	0.187	0.182
BTS	0.0093	−0.0222	0.27	0.215	0.187	0.182
TB	0.0095	−0.0126	0.27	0.216	0.186	0.233
CAB	0.0096	−0.0054	0.27	0.206	0.195	0.232
ITA	0.0095	−0.0029	0.27	0.206	0.195	0.232
ABC	0.0095	−0.0034	0.26	0.252	0.237	0.234
PAB	0.0096	−0.0067	0.26	0.252	0.237	0.234
Maximum loading	0.0110	−0.028 and 0.023	0.29	0.252	0.242	0.244
Minimum loading	0.0033	0	0.16	0.036	0.107	0.031

a phenomenon and our incomplete knowledge of its functioning, it is easier to describe things probabilistically. BNs calculate the conditional probability of an event (given all available evidences) applying the Bayes Theorem.

In a more formal approach, a BN can be described as an influence diagram represented as BN = $\langle N,A,\Theta\rangle$, where $\langle N,A\rangle$ is a directed acyclic graph (DAG); each node, $n \in N$, represents a domain variable and each arc, $a \in A$, represents the probabilistic dependence between two connected nodes. For each node, $n_i \in N$, exists a conditional probability table (CPT) whose elements are represented by $\Theta = \{\theta_i\}$ that quantifies the dependency of a child node to its parents (Neapolitan, 2004). Variables (or nodes) can represent events, objects, propositions or other entities whilst the orientation of the arcs connecting two nodes is from the parent node to the child node. These dependencies can be seen as indicative of cause–effect influences (or temporal relations) given by the conditional probabilities (Pearl, 2000). On the other hand, if two nodes X and Y are not connected given some evidence C, then they are conditionally independent given that evidence. This concept is called direction dependent separation or d-separation (Pearl, 2000).

BNs (as well as other probabilistic graphical models) can be built based on experts' knowledge, but they can also be induced from data (in an inductive learning process – Pearl, 1988). In order to start building the coral bleaching BN from data it is critical to identify which variables are 'connected'. Such a connection can be either a causal relation (e.g. heat stress and coral bleaching intensity) or a temporal sequence of events (e.g. El Niño and SST anomalies). This is achieved by deriving the network structure from the data according to two main classes of methods: those based on "search and score" (SS) algorithms (Santos et al., 2011) and those on the "conditional independence" (CI) definition (Spirtes et al., 2000). Additional methods can be based on a combination of these approaches (Neapolitan, 2004). CI methods are designed to find arc direction without the need of ordering, the attributes algorithms based on these methods can still be improved when the ordering is known (Spirtes et al., 2000). For our approach, such a node ordering is based on the domain knowledge that specifies a causal or temporal order of conditions known to induce, exacerbate or minimise a bleaching event. At this point, the actual arc orientation is not important as we wish to understand the conditional independencies avoiding any a priori expert knowledge. In addition, each environmental variable was discretized (Yang and Webb, 2003, 2009) into classes, ensuring that a balanced distribution of values within each class was maintained; such a step is key to allow the learning algorithms to separate conditionally independent nodes (variables) (Cheng et al., 2002; Pearl, 2000).

Once the BN structure (nodes and arcs) is defined, it is necessary to set the numerical parameters to determine the strength of dependencies between a parent and a child node. Thus, the probabilities of nodes states are specified for the direct dependents and then used to compute the indirect probabilities. In our approach, the mutual information and the conditional mutual information algorithms (equations (2.1) and (2.2) in Cheng et al., 2002) were used to construct a BN without expertise information. The performed learning process is intended to help determining how variables relate to each other in terms of information flow. Therefore, to further investigate the conditional dependencies expressed in this initial BN, specialized BNs were derived by adding some straightforward expert knowledge (details are given in the following sections) on the relationships between nodes ("expert guidance"). Such an approach has the advantage of allowing identification of weaker dependency signals within the data set, as well as variables more strongly related to bleaching events in Bahia.

The Belief Network Power Constructor (BNPC) software (Cheng et al., 2002) was employed in the performed experiments to learn BN networks from data. By computing the mutual information of variable pairs, BNPC constructs a BN by applying CI tests to analyse the relationships among nodes. The CI test based on the mutual information is capable of informing whether two variables are dependent and how closely related they are. The algorithm performs a drafting procedure, which is the tree construction algorithm of Chow and Liu (1968); this procedure subsequently performs a 'thickening and thinning' process to extend tree construction to general BN construction. The full description of the mutual information algorithms and the procedures involved in the construction of networks is beyond the scope of this manuscript, we refer the reader to Cheng et al. (2002) for details.

The numerical parameters (presented at the CPTs) were calculated with the expectation–maximization (EM) algorithm (Gupta and Chen, 2010) available with Netica (Norsys Software Corp., 2006). Specific implementation of this EM algorithm is achieved by iteratively taking random initial BNs to optimise the initial BN after a sequence of expectation (E) and maximization (M) steps. This procedure continues until the log likelihood numbers converge (cease to improve). Here, the EM algorithm is not used for estimating missing values but instead for updating parameters to find a maximum for the log likelihood (Needham et al., 2007). A probabilistic inference algorithm (see Lauritzen and Spiegelhalter, 1988) is then used to propagate the effects of the evidences of field, model and remote sensing data throughout the dependence-structured BN.

Both the structure and the probabilities were derived from a table consisting of rows with bleaching cases and columns with corresponding environmental variables (Table 1). Since there is only one field record of no bleaching event (March 2004 in Table 1), we excluded that category from the CPTs learning procedure. For the sake of completeness we kept the environmental information observed during that period during the CI calculations. The main focus here is in the analysis of the environmental conditions that modulate the intensity of coral bleaching, instead of focussing in the prediction of bleaching occurrence. There is an ongoing project dedicated to the creation of a model for bleaching prediction on a global scale using a significantly larger database.

2.4. Validation and sensitivity analysis

BNs were evaluated for accuracy using scoring (loss) functions and contingency tables, whilst the reliability of predictions was assessed using calibration values that express the percentage of times the predicted node state was the true state. The three scoring functions used are the logarithmic loss, ranging from zero (best performance) to infinity; the Brier score ranging from zero (best performance) to 2; and the spherical payoff, ranging from zero to 1 (best performance) (see Bickel, 2007 for comparisons between scoring rules).

A sensitivity test of the bleaching node to all other nodes was also run based on variance and entropy reduction. This test helps to identify how the probability of a specific bleaching state is influenced by using knowledge on the other nodes. Finally, to validate the results of the final BN a 'leave-one-out cross validation' procedure was used; for this, one bleaching case is excluded prior to probabilities construction to avoid its influence on the calculation of the CPTs. Then, the newly formed BN is used to predict the excluded observation. This process is then repeated for all bleaching cases so that each excluded case can be treated as independent data.

3. Results

3.1. Environmental variability

The average summer SST ranges from 27.7 °C to 28.4 °C in the north sector of the shelf, and from 27.1 °C to 28.1 °C in the south sector. The leading PC mode (PC1) of the MaxSST explained 92% of this variable total variance (Table 2) and demonstrates that heating is higher offshore of the Abrolhos Bank than elsewhere (Fig. 2a). Overall SST variability was relatively low in the northern sectors but coastal SST was highest between latitudes 13°S and 18°S, i.e. where most reefs are located. The PC1 of the K_{490} explained only 37% of the time series variance (Table 2), reflecting the spatially restricted distributions of larger PC loadings in the northern shelf break zones and some inner shelf areas of the Royal Charlotte and Abrolhos Banks (Fig. 2b). Thus K_{490} variability appeared to have limited effect on most reefs, except for ABC and PAB reefs.

The first leading mode of PPT (Table 2), seems to add little information, except for the fact that it tends to decrease towards the northern and southern limits of the Bahia coast (Fig. 2c). Variability in wind magnitude |W| for the Bahia coast (Fig. 2d) explained 49% of total variance and, as for K_{490}, with high PC loadings spatially restricted to the southern coastal areas. Higher variability concentrated south of 18°S, was coincident with areas of cooler SST over Abrolhos Bank, i.e. where ABC and PAB are located; most of the variability observed for |W| was explained (55% of total variance) by the zonal wind component (U) (Fig. 2e). Finally, the meridional wind component (V) (Fig. 2f) was spatially very similar to the observed pattern of PPT and explained 54% of total variance. A more detailed environmental description of these (bleaching) sites can be found in Krug et al. (2012).

3.2. The Bayesian networks

Our initial BN, based on information flow and without expert knowledge (Fig. 3), suggests that the direct influence of local environmental variables is likely more important than remote forcing (i.e. MEI) and characteristic variability pattern obtained from the first PCs in driving bleaching. K_{490} and its parent nodes appeared to add little in determining bleaching intensity. On the other hand, local SST proxies, wind and PPT were more closely linked with the bleaching events. Even so, it is important to highlight that the single use of CI tests was capable of producing a BN that is very close to what is currently known as the likely causal relationships amongst relevant environmental conditions for other reef systems (e.g. Maina et al., 2008, above). Therefore, it is also possible to infer that the environmental data used to produce the PCA loadings do not produce unrealistic relations. However, as stated earlier, the importance of this network relies on the

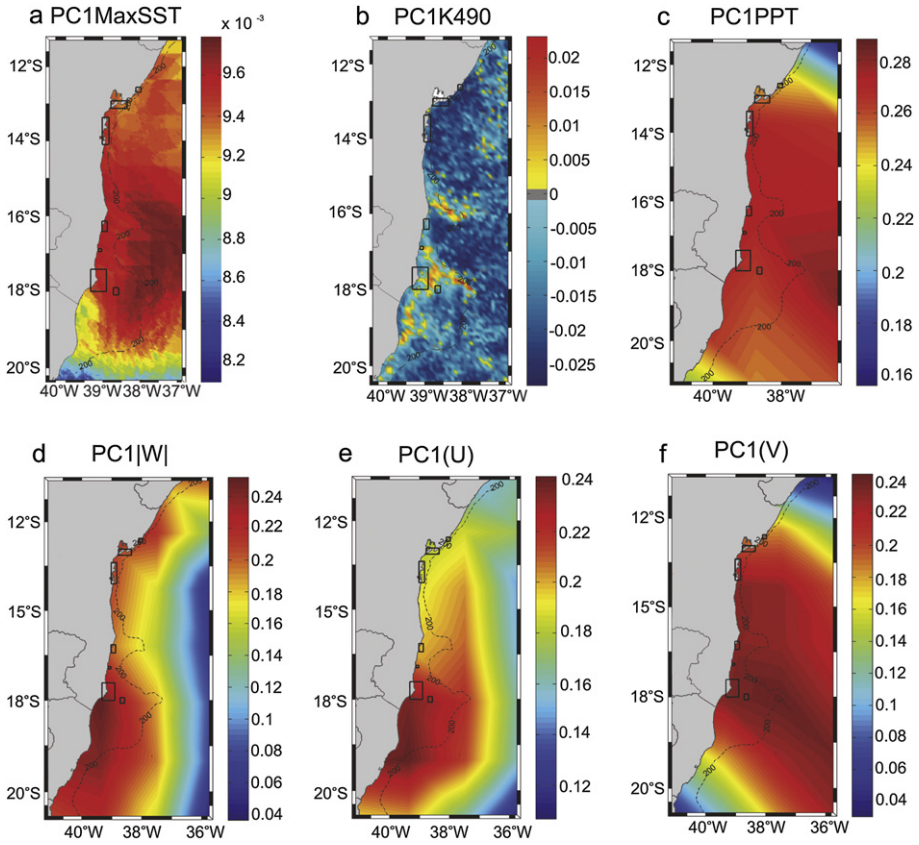

Fig. 2. First principal component (PC1) maps for a) maximum sea surface temperature (MaxSST); b) light attenuation coefficient (K_{490}) where the positive values at the edges of the banks represent an inverse environment, with high values of K_{490}, opposed to the rest of the area normally low values of K_{490}; c) rain precipitation (PPT); d) resultant wind |W|, and its components e) zonal (U) and f) meridional (V). The non-dimensional values represent the variation amplitude.

determination of the conditional dependencies as the direction of arcs is easily achieved by later adding expert knowledge.

The CI tests resulted in four alternative BNs that can be used to explore different possible causal relationships likely to modulate coral bleaching for the Bahian reefs (Figs. 4–7). These fully directed BNs were obtained by adding expert knowledge gathered from the literature to the skeleton produced by the BNPC (Fig. 3) as follows: 1) that ENSO proxies MEI and MEI_7 do not have parent nodes due to its remote origin, i.e., no other node on the BN can influence them; 2) that bleaching intensity is an obligatory child node, meaning that it cannot influence other nodes but can be influenced by anything in the BN; 3) that PC1MaxSST can be a direct or indirect cause of SSTAc5d and MaxSST due to their related origin; and 4) that |W|, (U) and (V) can be direct or indirect cause of SSTAc5d. Once conditional independencies were established, the specialized BNs were constructed based on the directed arcs and nodes expressing different likely relationships between variables. These specialized BNs have the purpose of learning the most likely causal relationships affecting coral bleaching and thus test for the best candidates of BNs based on prediction efficiency.

Specialized BNs were: 1) Thermal BN, emphasising the thermal environment (Fig. 4); 2) Thermal-Atmospheric BN, integrating thermal and atmospheric (wind and rain precipitation) variables (Fig. 5); 3) Thermal-Eolic Resultant BN (Fig. 6); and 4) Thermal-Eolic Meridional BN (Fig. 7). The Thermal-Eolic Resultant and Thermal-Eolic Meridional BNs integrate thermal variables with the resultant wind and its meridional component, respectively. The class intervals obtained (using the BNPC software, see Table 3) for all variables were used with the EM algorithm of Netica to calculate the CPTs.

To test for prediction efficiency, all BN parent nodes for bleaching were set to their observed values relative to each of the 16 bleaching occurrences (Table 1), and the resulting probabilities recorded for each bleaching node stated as strong or weak intensity. Predictive rates of bleaching intensity were computed as: 1) “high probability”, when prediction was 100% correct; 2) “likely state”, when it is the highest probability among all states; and 3) “unlikely state”, when bleaching state does not display the highest probability. The “assigned prediction” or total predictive rate value is the sum of correct predictions for all cases pointed as “high probability” and “likely state”. A rate of 14 correct predictions out of 16 (88%) was achieved for the total of strong and weak bleaching cases in all 4 specialized BNs. It is important to highlight that the five-day accumulated SST (represented by the proxy SSTAc5d) was the single effective parent node for bleaching in all BNs (Figs. 4–7). Thus the Markov blanket (Pearl, 1988) for the bleaching node is formed by a single variable, namely SSTAc5d, i.e. when it is known, the information on ancestor nodes is not propagated to the bleaching node. To understand how the other variables may affect the bleaching state we tested the specialized BNs by allowing evidences to propagate without using information on SSTAc5d (see Table 4).

The Thermal-Eolic Resultant BN (Fig. 6) presented the best results in terms of propagation of evidences across the network maintaining a total predictive rate of 88% and the prediction power (probability values) whether SSTAc5d is known or not (Table 4). The Thermal (Fig. 4) and Thermal-Atmospheric (Fig. 5) BNs provided a total predictive rate of 69%, with loss of prediction power. When SSTAc5d is known, all weak cases are correctly predicted with 6 (out of 10) of them pointed with 100% chance of

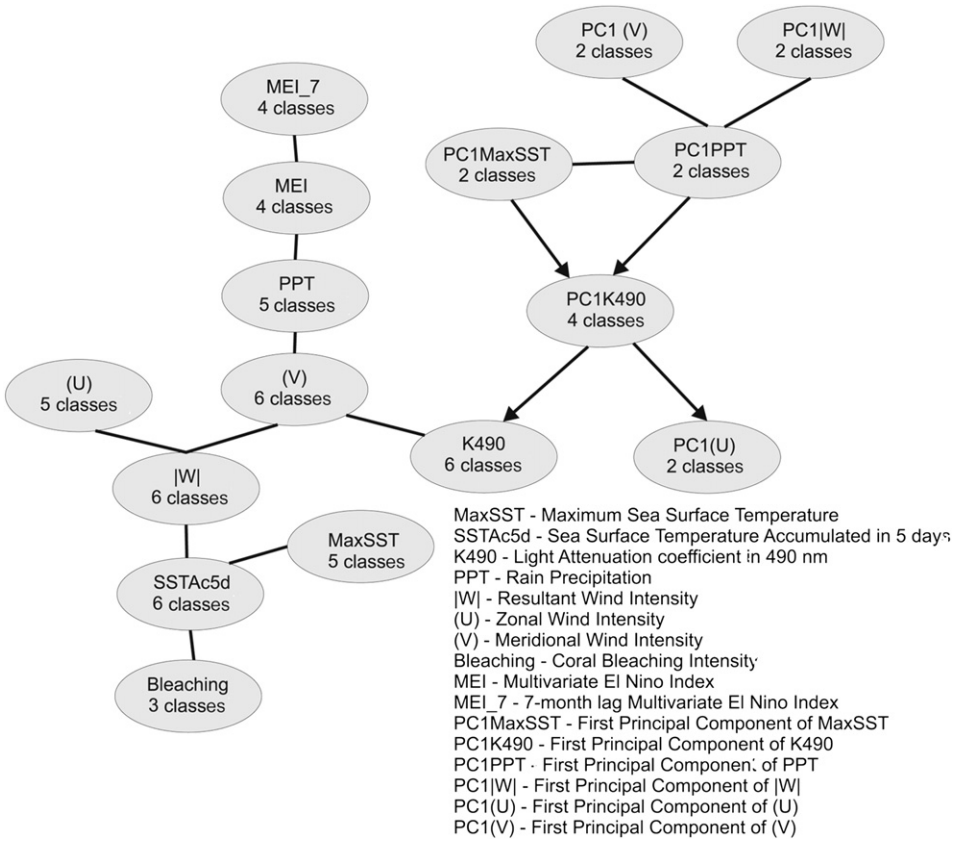

Fig. 3. First approach on the Bayesian network to estimate coral bleaching at Bahia based on all available variables with no expert knowledge added. Local environmental variables are more related to bleaching than El Niño remote forcing (MEI) and the characteristic variability patterns (PC1s).

occurrence. Once this variable is treated as not known, only 3 cases are given as certain (100% chance), 6 are now pointed as most likely and 1 was wrongly predicted (prediction rate of 69%, Table 4). The Thermal-Eolic Meridional BN (Fig. 7) also performed well when SSTAc5d node is known, with a prediction rate of 88%, but it falls to 63% when SSTAc5d is not known. Despite the fact that |W| and PPT nodes had little or no influence at all on the bleaching node in the Thermal-Atmospheric BN, these have been maintained in the graph (Fig. 5) to highlight that the conditional independence test yielded no direct relationship between PPT and bleaching intensity.

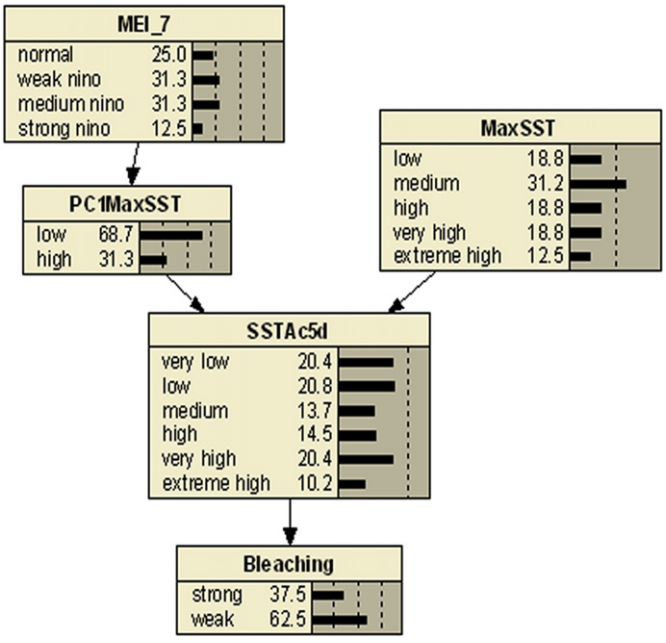

Fig. 4. Thermal Bayesian network. The nodes on the BN show the probabilities for each of their states when all nodes condition is unknown.

It is also useful to evaluate the level of regret associated with wrong predictions of the specialized BNs when the most likely state is used, for instance, in the calculation of predictive rates. The logarithmic loss is 0.2387 (from 0 to infinity, with 0 indicating the best performance), the Brier score is 0.1667 (between 0 and 2, with 0 being best) and the spherical payoff is 0.9045 (between zero and 1, with 1 being best) for all BNs, indicating that they do not differ in their abilities to predict the different bleaching states.

The 'leave-one-out' validation applied to the Thermal-Eolic Resultant BN when SSTAc5d is known reached a prediction rate of 63% (4 strong bleaching and 6 weak bleaching cases correctly predicted with 100% probability), and if information on ancestor nodes was allowed to propagate (MaxSST and |W|), i.e. SSTAc5d was treated as unknown, this rate decreased to 56% of correct predictions. Here, 5 weak cases were correctly predicted with 100% and 4 were correctly pointed as most likely. In both cases some of the bleaching prognostics were 50% weak and 50% strong when it was actually a weak bleaching event. If one considers that a 50% chance of bleaching (as opposed to, e.g., 5%) should be considered as a serious threat, then prediction rates would rise to 88% and 63%, whether SSTAc5d was used or not, respectively.

Results of the sensitivity test (Table 5) of the bleaching node to other nodes demonstrated that SSTAc5d exerts the strongest

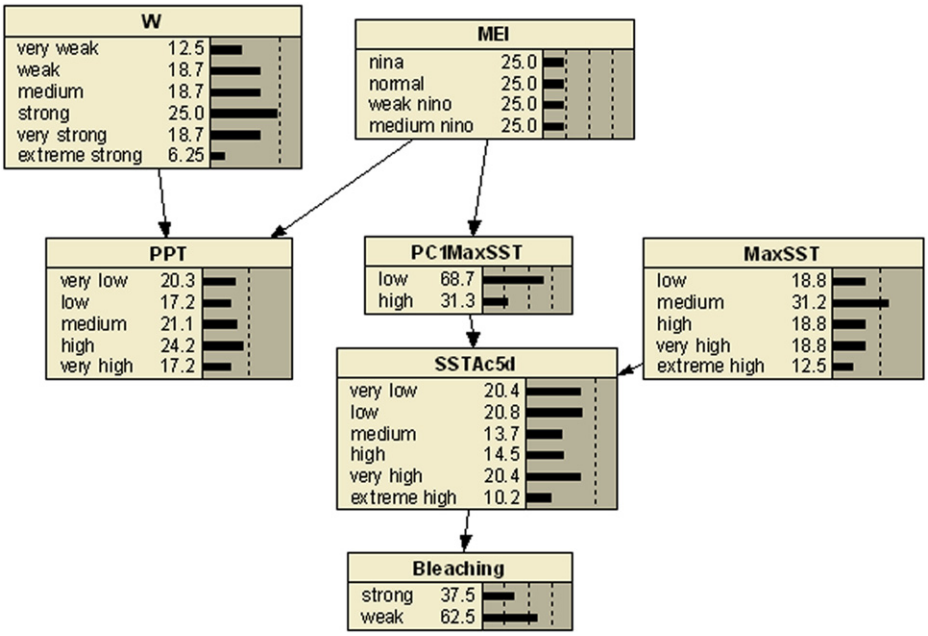

Fig. 5. Thermal-atmospheric Bayesian network. The nodes on the BN show the probabilities (percentage values and horizontal bars) for each of their states when all nodes condition is unknown.

influence on all BNs, followed by MaxSST and |W| for strong and weak bleaching, respectively (Fig. 8). It is also worth noting that wind conditions during bleaching events have a stronger influence than typical summertime conditions for maximum SST and K_{490} (compare the percent of entropy reduction in Table 5). Bleaching seems relatively less sensitive to MEI, although its influence on bleaching probability can be stronger in weak bleaching events. The influence of MEI in the outcome of a bleaching state is stronger when the influence of wind conditions was stronger (see T-E |W| and T-E (V) BN in Table 5). Strong and weak bleaching states are more sensitive to MaxSST variations, while variations in |W| had more influence on the weak bleaching state (Fig. 8b). Together these results suggest that thermal variables are the most important factors influencing the intensity of bleaching for this reef system.

4. Discussion

Environmental conditions are key in determining bleaching variability within and between reef sites, in particular geographic location, circulation pattern of surface waters and the presence of atmospheric systems (Maina et al., 2008; McClanahan et al., 2002, 2007). Therefore an obvious step to predict bleaching of Bahia's reefs is to examine the interplay of these variables at a regional scale. A useful starting point is examining ocean and atmospheric variability during summer. Both north–south and onshore–offshore patterns are evident in the PC1 maps (Fig. 2) and these are important to help interpret the results of the dependency analysis. The low spatial variability in the PC1 maps for K_{490}, PPT and wind data (Fig. 2), and the generally low number of classes obtained for the PC loadings (mostly two classes, see Fig. 3), suggest

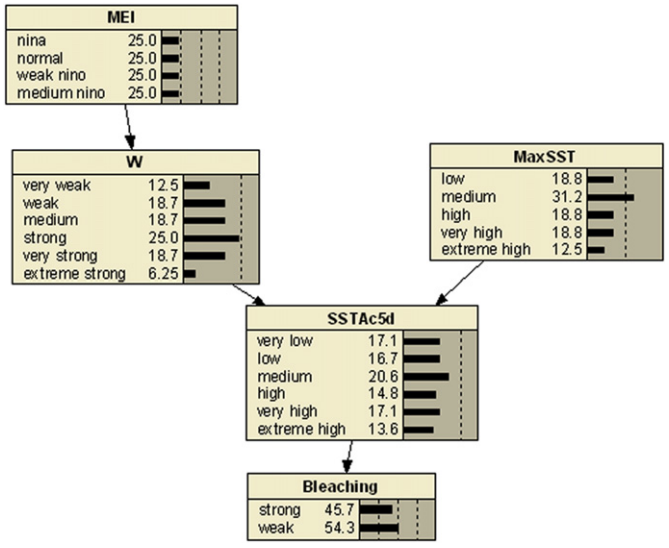

Fig. 6. Thermal-Eolic resultant Bayesian network. The nodes on the BN show the probabilities (percentage values and horizontal bars) for each of their states when all nodes condition is unknown.

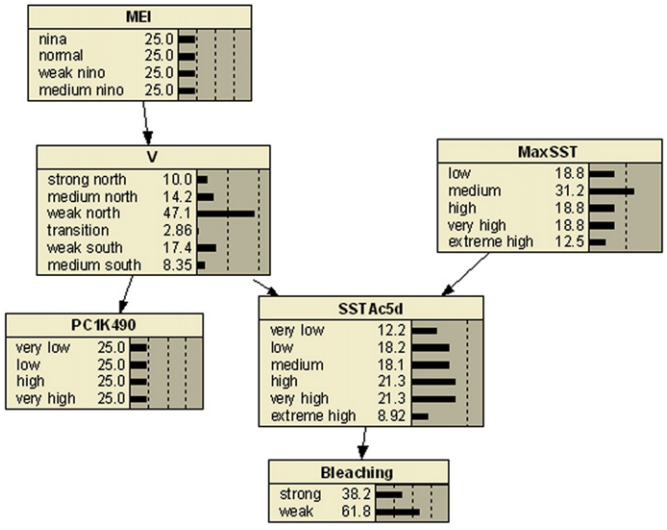

Fig. 7. Thermal-Eolic meridional Bayesian network. The nodes on the BN show the probabilities (percentage values and horizontal bars) for each of their states when all nodes condition is unknown.

Table 3
Based on frequency, the variables are discretized into class intervals later used to calculate the conditional probabilities. The nodes presented in one or more Bayesian Networks are: bleaching intensity; summer characteristic amplitude variation of maximum sea surface temperature (PC1MaxSST); maximum sea surface temperature (MaxSST); sea surface temperature accumulated in 5 days; resultant wind (|W|); meridional wind (V); summer characteristic amplitude variation of light attenuation coefficient (PC1 K_{490}); rain precipitation (PPT); El Niño index without and with a 7 month lag (MEI and MEI_7, respectively).

Bleaching (%)		*PC1MaxSST*	
Weak	$(0 < x \leq 20)$	Low	$(x \leq 0.0095)$
Strong	$(x > 20)$	High	$(x > 0.0095)$
MaxSST (°C)		*SSTAc5d (°C)*	
Low	$(x < 28.3)$	Very low	$(x < 141.2)$
Medium	$(28.3 \leq x < 28.5)$	Low	$(141.2 \leq x < 141.7)$
High	$(28.5 \leq x < 28.8)$	Medium	$(141.7 \leq x < 142.3)$
Very high	$(28.8 \leq x < 29.2)$	High	$(142.3 \leq x < 143.8)$
Extreme high	$(x \geq 29.2)$	Very high	$(143.8 \leq x < 145.9)$
		Extreme high	$(x \geq 145.9)$
|W| (m s^{-1})			
Very weak	$(x < 3.08)$	*PPT (mm d^{-1})*	
Weak	$(3.08 \leq x < 4.06)$	Very low	$(x < 1.68)$
Medium	$(4.06 \leq x < 5.38)$	Low	$(1.68 \leq x < 2.47)$
Strong	$(5.38 \leq x < 5.83)$	Medium	$(2.47 \leq x < 5.52)$
Very strong	$(5.83 \leq x < 6.87)$	High	$(5.52 \leq x < 6.46)$
Extreme strong	$(x \geq 6.87)$	Very high	$(x \geq 6.46)$
(V) (m s^{-1})			
Strong North	$(x < -2.32)$	*MEI*	
Medium North	$(-2.32 \leq x < -1.65)$	Niña	$(x \leq -0.082)$
Weak North	$(-1.65 \leq x < -0.9)$	Normal	$(-0.082 < x \leq 0.261)$
Transition	$(-0.9 \leq x < 0.27)$	Weak Niño	$(0.261 < x \leq 0.684)$
Weak South	$(0.27 \leq x < 1.38)$	Medium Niño	$(x > 0.684)$
Medium South	$(x \geq 1.38)$		
PC1 K_{490}		*MEI_7*	
Very low	$(x > -0.0032)$	Normal	$(x < 0.289)$
Low	$(-0.0061 < x \leq -0.0032)$	Weak Niño	$(0.289 \leq x < 0.596)$
High	$(-0.0096 < x \leq -0.0061)$	Medium Niño	$(0.596 \leq x < 0.964)$
Very high	$(x \leq -0.0096)$	Strong Niño	$(x \leq 0.964)$

that at present they can only have a limited contribution to the analysis. The conditional dependence tests indicate that the structural relations between the characteristic environments (PC1 maps) and the environmental conditions during bleaching events are conditionally independent given the knowledge of K_{490} (Fig. 3). This indicates that specific local conditions tend to be more informative (and may be more important) than the knowledge on the long term summertime environmental variability for predicting bleaching intensity. It is surprising that K_{490} does not seem to be an important variable, given that light availability has been cited as an important moderating factor towards bleaching extent/severity (e.g. Baker et al., 2008; Suggett and Smith, 2011). It is worth mentioning that Bahia's reefs are clearly subjected to wide temporal and spatial variations in K_{490} (Suggett et al., 2012) and corals have developed adaptations to this condition.

Reduced spatial resolution of PPT (2.5°) and wind (1.9°) data might also explain the segregation of their proxies (PCs) into a cut-set (or subset) given K_{490} (Fig. 3) during the conditional test. With this in mind, we can assume that the robustness of this method minimizes the risks of drawing conclusions from poor data and perhaps highlights what information should be considered most relevant to ultimately refining the predictive power. It is also worth noting that (U) and (V), being orthogonal components of |W|, are parent nodes of this variable, which in turn has an influence on SSTAc5d. The influence diagram of Fig. 3 also connected the temperature proxies to the bleaching node, confirming the empirical knowledge of their causal relations (Berkelmans et al., 2004; Brown, 1997; Fitt et al., 2001; Glynn, 1996).

Here it should be stressed that no previous knowledge (no node ordering) was used to run the conditional dependence test, yet the dependency relations that were established are in good agreement with what is perceived as a plausible natural relationship. Examples for this are the independence of PC1PPT and PC1MaxSST given knowledge of PC1 K_{490}, since rainfall and SST are known to influence productivity and water transparency. Our observations indeed suggest that the southern reef areas, located on the Royal Charlotte and Abrolhos banks, exhibit higher K_{490} values, and hence lower water transparency than northern reef areas. These patterns are consistent with field observations reported by Leão et al. (2003). In spite of these differences, K_{490} is high throughout for coral reef systems and thus corals here must be well adapted to a high turbidity environments where light availability is highly variable (Leão et al., 2003). Such adaptation could provide enhanced tolerance to anomalous (light) conditions and thus further confound K_{490} (and/or the product of K_{490} and incident light intensity) as a predictor of bleaching intensity.

Studies are increasingly observing that the susceptibility to bleaching (induced mortality) declines with SST variability (e.g. Oliver and Palumbi, 2011); however, such a relationship is clearly not universal (e.g. Maina et al., 2008 for Western Indian Ocean corals). Two possible explanations are give: (1), corals living in areas with high SST variability are probably more adapted to

Table 4
Predictive rates of bleaching states when sea surface temperature accumulated in 5 days (SSTAc5d) is treated as unknown. Values represent "high probability", when prediction was 100% correct; "likely" state, when it is the highest probability among states; and "unlikely state", when bleaching state does not display the highest probability. The "assigned prediction" value is the sum of correct predictions for all cases considered as "high probability" and "likely state". The thermal-Eolic resultant BN produces the same predictive rates whether SSTAc5d is known or unknown.

BN	Thermal			Thermal-atmospheric		
	Strong	Weak	All cases	Strong	Weak	All cases
High probability	2/6 (33%)	3/10 (30%)	5/16 (31%)	2/6 (33%)	3/10 (30%)	5/16 (31%)
Likely state	0	6/10 (60%)	6/16 (38%)	0	6/10 (60%)	6/16 (38%)
Unlikely state	4/6 (67%)	1/10 (10%)	5/16 (31%)	4/6 (67%)	1/10 (10%)	5/16 (31%)
		Assigned prediction	11/16 (69%)		Assigned prediction	11/16 (69%)
	Thermal-Eolic resultant			Thermal-Eolic meridional		
	Strong	Weak	All cases	Strong	Weak	All cases
High probability	4/6 (67%)	6/10 (60%)	10/16 (63%)	0	6/10 (60%)	6/16 (38%)
Likely state	0	4/10 (40%)	4/16 (25%)	0	4/10 (40%)	4/16 (25%)
Unlikely state	2/6 (33%)	0	2/16 (13%)	6/6 (100%)	0	6/16 (38%)
		Assigned prediction	14/16 (88%)		Assigned prediction	10/16 (63%)

Table 5
Sensitivity test of bleaching in relation to the other nodes of thermal Bayesian Network (T BN), thermal-atmospheric BN (T-A BN), thermal-Eolic resultant BN (T-E |W| BN) and thermal-Eolic meridional BN (T-E (V) BN) based on entropy reduction (expressed as percentage) and variance of belief which measures the effect of one node on another.

	Node	PC1 K_{490}	PC1MaxSST	MEI	MEI_7	V	W	PPT	MaxSST	SSTAc5d	Bleaching
T BN	Distance from bleaching	n.a.	2	n.a.	3	n.a.	n.a.	n.a.	2	1	0
	Entropy reduction	n.a.	0.30	n.a.	0.04	n.a.	n.a.	n.a.	12.3	60.8	100
	Variance of belief	n.a.	0. 0009350	n.a.	0.0001190	n.a.	n.a.	n.a.	0.0366040	0.1437532	0.2343213
T-A BN	Distance from bleaching	n.a.	2	3	n.a.	n.a.	5	4	2	1	0
	Entropy reduction	n.a.	0.299	0.0164	n.a.	n.a.	0	0.0031	12.3	60.8	100
	Variance of belief	n.a.	0.0009350	0.0000510	n.a.	n.a.	0	0.0000103	0.0366041	0.1437532	0.2343213
T-E |W| BN	Distance from bleaching	n.a.	n.a.	3	n.a.	n.a.	2	n.a.	2	1	0
	Entropy reduction	n.a.	n.a.	0.764	n.a.	n.a.	6.24	n.a.	3.77	68.4	100
	Variance of belief	n.a.	n.a.	0.0026023	n.a.	n.a.	0.0209846	n.a.	0.0127326	0.1720137	0.2481160
T-E (V) BN	Distance from bleaching	3	n.a.	3	n.a.	2	n.a.	n.a.	2	1	0
	Entropy reduction	0	n.a.	13.2	n.a.	13.3	n.a.	n.a.	19.7	67.9	100
	Variance of belief	0	n.a.	0.0499750	n.a.	0.0400779	n.a.	n.a.	0.0471650	0.1617051	0.2361670

extreme conditions; and (2) shallow water circulation is mainly controlled by wind forced Ekman transport and imposes distinct regulation of SST and thus coral bleaching. Strong winds can enhance vertical mixing leading to water cooling and cause onshore advection of warm surface waters increasing the persistence of warm SST. Weak winds can reduce water mixing and contribute to warm SST conditions, leading to thermal stress (Gleason and Wellington, 1993). Therefore, the complex patterns arising from these processes may develop within days or weeks and the long-term proxies may be only able to capture part of the inherent driving signal. That said, the dependency analysis algorithm considered the MEI and MEI_7 as influencing PPT, despite their remote origin. This agrees with findings that establish ENSO events induced major remote forcing of interannual climate variations for many parts of South America. In particular, drier than normal conditions over northeast Brazil and along the Bahia coast can be directly related to El Niño events (Kayano et al., 2009, and references therein).

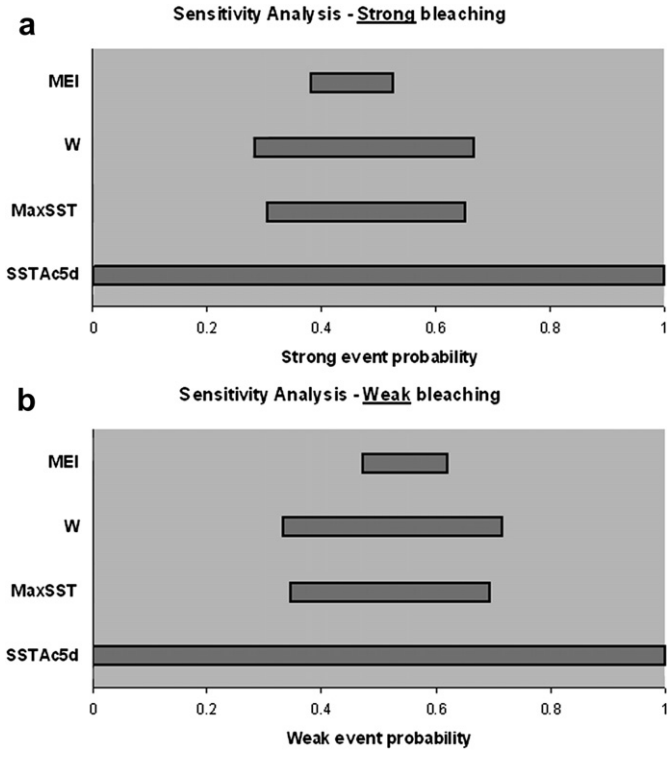

Fig. 8. Sensitivity analysis for a) strong and b) weak bleaching to individual alterations at thermal-Eolic resultant BN nodes.

When previous expert knowledge (known relationships) is added, the robust relations learnt from the dependency analysis algorithm produced specialized BNs to investigate how different variables influence the likelihood of bleaching. The "belief" that the thermal environment is the primary cause of coral bleaching has been captured in all four specialized BNs; i.e. a pattern that conforms the well established paradigm that elevated SSTs primarily drive coral bleaching (Berkelmans et al., 2004; Glynn, 1993). It is important to note that PC1MaxSST and MaxSST are conditionally independent given SSTAc5d. Thus historical SST patterns and episodic SST extremes, although directly linked to heat stress, play important roles in the persistence of stressful conditions leading to coral bleaching. The SSTAc5d proxy was devised based on the three days accumulated maximum SST (Max3day) proposed by Berkelmans et al. (2004), which correlates with coral bleaching in the GBR. Wooldridge and Done (2004) also used the Max3day in their BNs, which directly influenced coral bleaching and mortality in the GBR during the 2002 bleaching event. The reefs of Bahia would therefore seem to respond to accumulated thermal stress in a similar manner as with other major reef complexes.

Specific wind conditions influenced coral bleaching, as expressed by the connections of the resultant wind (Thermal-Eolic Resultant, Fig. 6) and its meridional (V) component (Thermal-Eolic Meridional, Fig. 7) with the SSTAc5d node. More importantly than the predictive capability is the fact that a direct probabilistic relation between wind intensity and the persistence of high SST has been found. A total of nine, out of 16 bleaching cases, occurred during high wind intensities ($|W| > 5$ m s^{-1}), from which six cases were recorded as weak bleaching (Table 1). Setting |W| with probability of 100% for extreme strong winds, the Thermal-Eolic Resultant BN will predict a strong/weak bleaching ratio of 42.4/57.6%, respectively. In contrast, setting |W| to 100% very weak winds will predict a bleaching ratio of 36.1/63.9%, respectively. Inspection of raw data indicates that the magnitudes of |W| are similar to (U), and magnitudes for (V) are much lower, but may also contribute to the bleaching process. Setting (V) to 100% strong northerly wind, the Thermal-Eolic Meridional BN predicts 18.2/81.8% of strong/weak bleaching ratio. On the other hand, setting (V) to 100% weak northerly wind will predict a 58.3/41.7% ratio of strong/weak bleaching ratio. During summer, there is a predominance of easterly and northeasterly winds (Rao et al., 1993) that contributes to the advection of warm water from the equatorial Atlantic. Reducing the meridional component of the northerly wind may thus have contributed to increase the

residence time of warm waters in the Bahia coast. With the present data set, it is not possible to establish a deterministic relationship between high wind velocity and weak bleaching but the BNs seem to be capable of incorporating these uncertainties in a coherent fashion.

All of the specialized BNs incorporate the external forcing represented by the MEI and MEI_7 in a way that deserves some attention. Apparently, there seems to have occurred some confusion when the Thermal BN (Fig. 4) connected the MEI_7 to PC1MaxSST as parent node, and the Thermal-Atmospheric BN connected MEI to the same child node. Maximum correlations between ENSO and the tropical Atlantic SST occur with a lag of up to seven months (Enfield and Mayer, 1997; Klein et al., 1999; Lanzante, 1996). The Thermal BN (Fig. 4) captured this relation but this was not the case for the Thermal-Atmospheric BN (Fig. 5). The reason for that may lie in the fact that the former BN included nodes for |W| and PPT, for which maximum correlations are achieved with zero lag. This mechanism includes the upper-tropospheric Rossby-wave train that extends from the equatorial eastern Pacific to the northern tropical Atlantic and the east–west displacement of the Walker circulation during El Niño years (Hastenrath, 1976, 2006; Kayano et al., 1988). Such ability to learn this relationship has been confirmed in the Thermal-Eolic Resultant and the Thermal-Eolic Meridional BNs that incorporated ENSO information via MEI without time lag.

Finally, the cross validation procedure demonstrated that the overall bleaching prediction accuracy is maintained by enabling evidences to propagate from the ancestor nodes (MaxSST and W) to the bleaching intensity node. This propagation was even more encouraging considering that the sensitivity analysis showed that knowledge of SSTAc5d tends to have a large impact on the probability of bleaching. Such behaviour might have been influenced by the small data sample used to build the BNs, thus, showing the ability of BNs to help generalize from small data samples. Future studies that have a large number of samples will enable further exploration of this issue.

In summary, our new insights into likely (environmental) drivers of coral bleaching intensity for Bahia's reefs from the BN approach presented here appear to be in strict agreement with observations in many other reef provinces (Fitt et al., 2001; Maynard et al., 2008; Wooldridge and Done, 2004); i.e. changes in thermal environment are key. The use of BNs to learn robust relationships from a limited data set contrasts with predictive ecological models that are generally 'data-hungry'. However, for operational prediction purposes the BN model has to be fed with new environmental data on a regular basis, whether originated from *in situ* sensors or from remote sensing. Data availability for our current BN is obviously limited in extent (hence the choice of a BN over a GIS model) and additional future (targeted) data collection will enable a means to independently test the current BN. It will also allow an iterative refinement of our approach to improve predictive capacity further; notably care must be taken in an objective and consistent means to classify bleaching (see Chen and Pollino, 2012; Suggett and Smith, 2011) and thus pay close attention to weak and or null bleaching events during anomalous conditions. Analysis of historical bleaching records, e.g. from coral cores, may provide an additional means to further refine the BNs ahead of any (unpredictable) anomalous events. Interestingly, despite the observation that SST is a key driver of bleaching events here and thus consistent with observations for other reefs systems, it is clear that current predictive tools (e.g. NOAA's coral watch SST proxies) are not applicable to Bahia's reef system and therefore refining the BN is seen as a priority for effective regional management of Brazil's reefs.

Advancing the BN for Bahia's reef system remains a long term priority. Current operational coral bleaching prediction tools provided by the Coral Reef Watch program of the National Oceanic and Atmospheric Administration (NOAA) (http://coralreefwatch.noaa.gov/satellite/), are based exclusively on SST stress. None of the bleaching events used in the present study was reported by the Coral Reef Watch site as submitted to a thermal stress above the 1 °C bleaching threshold. Some events (April/1998, April/2002, March/2003 and March–April/2005) were indicated as hotspots but below this bleaching threshold. This may be the result of the way stressful conditions are estimated using the hottest climatological monthly SST. It is, however, beyond the scope of the present paper to compare and discuss specific bleaching prediction tools. The main point here is to highlight the complexities involved in the bleaching scenarios that have been identified for the Bahia state reefs and to offer a prediction tool that incorporates evidences drawn from a larger set of environmental information.

Acknowledgements

We would like to acknowledge M. Travassos, R.B. Souza, L.P. Pezzi, M.A.L. Caetano and M.A Soppa for their collaboration with data and lively discussions, and two anonymous reviewers for their helpful comments that greatly improved the paper. L.A. Krug was supported by a CAPES fellowship (Brazil). Z.M.A.N. Leão and R.K.P. Kikuchi benefit from CNPq fellowships. Collaboration between D.J. Suggett and R.K.P. Kikuchi was supported through an EU Marie Curie networking grant "SymbioCoRe" (PIRSES-GA-2011-295191). This is a contribution of the Pro-Abrolhos project funded by CNPq (420219/2005-6).

References

Aguilera, P.A., Fernández, A., Fernández, R., Rumí, R., Salmerón, A., 2011. Bayesian networks in environmental modelling. Environ. Modell. Softw. 26, 1376–1388.
Alameddine, I., Cha, Y., Reckhow, K.H., 2011. An evaluation of automated structure learning with Bayesian networks: an application to estuarine chlorophyll dynamics. Environ. Modell. Softw. 26, 163–172.
Baker, A.C., Glynn, P.W., Riegl, B., 2008. Climate change and coral reef bleaching: an ecological assessment of long-term impacts, recovery trends and future outlook. Estuar. Coast. Shelf Sci. 80, 435–471.
Berkelmans, R., De'ath, G., Kininmonth, S., Skirving, W.J., 2004. A comparison of the 1998 and 2002 coral bleaching events on the Great Barrier Reef: spatial correlation, patterns, and predictions. Coral Reefs 23, 74–83.
Bickel, J.E., 2007. Some comparisons among quadratic, spherical, and logarithmic scoring rules. Decis. Anal. 2, 49–65.
Brown, B.E., 1997. Coral bleaching: causes and consequences. Coral Reefs 16, 129–138.
Carpenter, K.E., Abrar, M., Aeby, G., Aronson, R.B., Banks, S., Bruckner, A., Chiriboga, A., Cortés, J., Delbeek, J.C., DeVantier, L., Edgar, G.J., Edwards, A.J., Fenner, D., Guzmán, H.M., Hoeksema, B.W., Hodgson, G., Johan, O., Licuanan, W.Y., Livingstone, S.R., Lovell, E.R., Moore, J.A., Obura, D.O., Ochavillo, D., Polidoro, B.A., Precht, W.F., Quibilan, M.C., Reboton, C., Richards, Z.T., Rogers, A.D., Sanciangco, J., Sheppard, A., Sheppard, C., Smith, J., Stuart, S., Turak, E., Veron, J.E.N., Wallace, C., Weil, E., Wood, E., 2008. One-third of reef-building corals face elevated extinction risk from climate change and local impacts. Science 321, 560–563.
Castro, C.B., Pires, D.O., 1999. A bleaching event on a Brazilian coral reef. Rev. Bras. Oceanogr. 47, 87–90.
Charniak, E., 1991. Bayesian networks without tears. Artif. Intell. Mag. 12, 50–63.
Chen, S.H., Pollino, C.A., 2012. Good practice in Bayesian network modelling. Environ. Modell. Softw. 37, 134–145.
Cheng, J., Griener, R., Kelly, J., Bell, D., Liu, W., 2002. Learning Bayesian networks from data: an information-theory based approach. Artif. Intell. 137, 43–90.
Chow, C.K., Liu, C.N., 1968. Approximating discrete probability distributions with dependence trees. IEEE Trans. Inform. Theory 14, 462–467.
Donner, S.D., Knutson, T.R., Oppenheimer, M., 2007. Model-based assessment of the role of human-induced climate change in the 2005 Caribbean coral bleaching event. Proc. Natl. Acad. Sci. U. S. A. 104, 5483–5488.
Eastman, R.J., 1992. Time series map analysis using standardized principal components. ASPRS/ACSM/RT 92 Technical Papers. Glob. Change Educ. 1, 195–204.
Enfield, D.B., Mayer, D.A., 1997. Tropical Atlantic SST and its relation to El Niño-Southern Oscillation. J. Geophys. Res. 102, 929–945.
Fitt, W.K., Brown, B.E., Warner, M.E., Dunne, R.P., 2001. Coral bleaching: interpretation of thermal tolerance limits and thermal thresholds in tropical corals. Coral Reefs 20, 51–65.

Ginsburg, R.N., Kramer, P.A., Lang, J.C., Sale, P., 1998. AGRRA, Atlantic and Gulf Rapid Reef Assessment. Available online at: http://agrra.org (accessed 25.01.12).
Gleason, D.F., Wellington, G.M., 1993. Ultraviolet radiation and coral bleaching. Nature 365, 836–838.
Glynn, P.W., 1993. Coral reef bleaching. Coral Reefs 12, 1–17.
Glynn, P.W., 1996. Coral reef bleaching: facts, hypotheses and implications. Glob. Change Biol. 2, 495–509.
Goreau, T., McClanahan, T.R., 2000. Conservation of coral reefs after the 1998 global bleaching event. Conserv. Biol. 14, 5–15.
Gupta, M.R., Chen, Y., 2010. Theory and use of the EM algorithm. Found. Trends Signal Process. 4, 223–296.
Hastenrath, S., 1976. Variations in low-latitude circulation and extreme climatic events in the tropical Americas. J. Atmos. Sci. 33, 202–215.
Hastenrath, S., 2006. Circulation and teleconnection mechanisms of northeast Brazil droughts. Progr. Oceanogr. 70, 407–415.
Kahn, C.E., Roberts, L.M., Shaffer, K.A., Haddawy, P., 1997. Construction of a Bayesian network for mammographic diagnosis of breast cancer. Comput. Biol. Med. 27, 19–29.
Kayano, M.T., Rao, B., Moura, A.D., 1988. Tropical circulations and the associated rainfall anomalies during two contrasting years. J. Climatol. 8, 477–488.
Kayano, M.T., Oliveira, C.P., Andreoli, R.V., 2009. Interannual relations between South American rainfall and tropical sea surface temperature anomalies before and after 1976. Int. J. Climatol. 29, 1439–1448.
Kelmo, F., Attrill, M.J., Jones, M.B., 2003. Effects of the 1997–1998 El Niño on the cnidarian community of a high turbidity coral reef system (Northern Bahia, Brazil). Coral Reefs 22, 542–550.
Kilpatrick, K.A., Podestá, G.P., Evans, R., 2001. Overview of the NOAA/NASA advanced very high resolution radiometer Pathfinder algorithm for sea surface temperature and associated match up database. J. Geophys. Res. 16, 9179–9197.
Klein, S.A., Soden, B.J., Lau, N., 1999. Remote sea surface temperature variations during ENSO: evidence for a tropical atmospheric bridge. J. Clim. 12, 917–932.
Kragt, M.E., Newham, L.T.H., Bennett, J., Jakeman, A.J., 2011. An integrated approach to linking economic valuation and catchment modelling. Environ. Modell. Softw. 26, 92–102.
Krug, L.A., Gherardi, D.F.M., Stech, J.L., Leão, Z.M.A.N., Kikuchi, R.K.P., 2012. Characterization of coral bleaching environments and their variation along the Bahia state coast, Brazil. Int. J. Remote Sens. 33, 4059–4074.
Kuikka, S., Hilden, M., Gislason, H., Hansson, S., Sparholt, H., Varis, O., 1999. Modeling environmentally driven uncertainties in Baltic cod (*Gadus morhua*) management by Bayesian influence diagrams. Can. J. Fish. Aquat. Sci. 56, 629–641.
Lanzante, J.R., 1996. Lag relationships involving tropical sea surface temperatures. J. Clim. 9, 2568–2578.
Lauritzen, S.L., Spiegelhalter, D.J., 1988. Local computations with probabilities on graphical structures and their application to expert systems. J. Roy. Stat. Soc. B 50, 157–224.
Leão, Z.M.A.N., Kikuchi, R.K.P., Testa, V., 2003. Corals and coral reefs of Brazil. In: Cortez, J. (Ed.), Latin American Coral Reefs. Elselvier, New York, pp. 9–52.
Leão, Z.M.A.N., Kikuchi, R.K.P., Oliveira, M.D.M., 2008. Branqueamento de corais nos recifes da Bahia e sua relação com eventos de anomalias térmicas nas águas superficiais do oceano. Biota Neotropica 8, 69–82.
Little, L.R., Kuikka, S., Punt, A.E., Pantus, F., Davies, C.R., Mapstone, B.D., 2004. Information flow among fishing vessels modelled using a Bayesian network. Environ. Modell. Softw. 19, 27–34.
Maina, J., Venus, V., McClanahan, T.R., Ateweberhan, M., 2008. Modelling susceptibility of coral reefs to environmental stress using remote sensing data and GIS models. Ecol. Model. 212, 180–199.
Maskery, S.M., Hu, H., Hooke, J., Shriver, C.D., Liebman, M.N., 2008. A Bayesian derived network of breast pathology co-occurrence. J. Biomed. Inform. 41, 242–250.
Maynard, J.A., Turner, P.J., Anthony, K.R.N., Baird, A.H., Berkelmans, R., Eakin, C.M., Johnson, J., Marshall, P.A., Packer, G.R., Rea, A., Willis, B.L., 2008. ReefTemp: an interactive monitoring system for coral bleaching using high-resolution SST and improved stress predictors. Geophys. Res. Lett. 35, L05603.
McClanahan, T.R., Polunin, N., Done, T., 2002. Ecological states and the resilience of coral reefs. Conserv. Ecol. 6, 2–18.
McClanahan, T.R., Ateweberhan, M., Muhando, C.A., Maina, J., Mohammed, M.S., 2007. Effects of climate and seawater temperature variation on coral bleaching and mortality. Ecol. Monogr. 77, 503–525.
Monahan, A.H., Fyfe, J.C., Ambaum, M.H.P., Stephenson, D.B., North, G.R., 2009. Empirical orthogonal functions: the medium is the message. J. Clim. 22, 6501–6514.
Mueller, J.L., 2000. SeaWiFS algorithm for the diffuse attenuation coefficient, K(490), using water-leaving radiances at 490 and 555 nm. In: Hooker, S.B., Firestone, E.R. (Eds.), SeaWiFS Postlaunch Technical Report Series. SeaWiFS Postlaunch Calibration and Validation Analyses, Part 3, vol. 11. NASA Goddard Space Flight Center, Maryland, pp. 24–27.
Mumby, P.J., Crisholm, J.R.M., Edwards, A.J., Andrefouet, S., Jaubert, J., 2001. Cloudy weather may have saved Society Island reef corals during the 1998 ENSO event. Mar. Ecol. Prog. Ser. 222, 209–216.
Neapolitan, R.E., 2004. Learning Bayesian Networks. Prentice Hall, Upper Saddle River.
Needham, C.J., Bradford, J.R., Bulpitt, A.J., Westhead, D.R., 2007. A primer on learning in Bayesian networks for computational biology. PLoS Comput. Biol. 3, 129.
Norsys Software Corporation, 2006. Sensitivity to Findings. Netica Software Documentation. www.norsys.com.
Oliver, T., Palumbi, S.R., 2011. Do fluctuating temperature environments elevate coral thermal tolerance? Coral Reefs 30, 429–440.
Pearl, J., 1988. Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference. Morgan Kaufmann, San Mateo.
Pearl, J., 2000. Causality: Models, Reasoning, and Inference. Cambridge University Press, Cambridge.
Pérez-Miñana, E., Krause, P.J., Thornton, J., 2012. Bayesian networks for the management of greenhouse gas emissions in the British agricultural sector. Environ. Modell. Softw. 35, 132–148.
Rao, V.B., Lima, M.C., Franchito, S.H., 1993. Seasonal and interannual variations of rainfall over eastern Northeast Brazil. J. Clim. 6, 1754–1763.
Santos, E.B., Hruschka Jr., E.R., Hruschka, E.R., Ebecken, N.F.F., August 2011. Bayesian network classifiers: beyond classification accuracy. Intell. Data Anal. 15 (3), 279–298.
Sheppard, C.R.C., 2003. Predicted recurrences of mass coral mortality in the Indian Ocean. Nature 425, 294–297.
Spalding, M.D., Ravilious, C., Green, E.P., 2001. World Atlas of Coral Reefs. University of California Press, Berkeley.
Spirtes, P., Glymour, C., Scheines, R., 2000. Causation, Prediction and Search, second ed. MIT Press, New York.
Suggett, D.J., Smith, D.J., 2011. Interpreting the sign of coral bleaching: friend versus foe. Glob. Change Biol. 17, 45–55.
Suggett, D.J., Kikuchi, R., Oliveira, M.D.M., Spano, S., Carvalho, R., Smith, D.J., 2012. Photobiology of corals from Brazil's near-shore marginal reefs of Abrolhos. Marine Biol. 159, 1461–1473.
Ticehurst, J.L., Curtis, A., Merritt, W.S., 2011. Using Bayesian networks to complement conventional analyses to explore landholder management of native vegetation. Environ. Modell. Softw. 26, 52–65.
Uusitalo, L., 2007. Advantages and challenges of Bayesian networks in environmental modelling. Ecol. Model. 203, 312–318.
Varis, O., Kuikka, S., 1997. Joint use of multiple environmental assessment models by a Bayesian meta-model: the Baltic salmon case. Ecol. Model. 102, 341–351.
Wilks, D.S., 2006. Statistical Methods in the Atmospheric Sciences. Academic Press, London.
Wolter, K., Timlin, M.S., 1998. Measuring the strength of ENSO – how does 1997/98 rank? Weather 53, 315–324.
Wooldridge, S., Done, T., 2004. Learning to predict large-scale coral bleaching from past events: a Bayesian approach using remotely sensed data, in-situ data, and environmental proxies. Coral Reefs 23, 96–108.
Yang, Y., Webb, G.I., 2003. On why discretization works for naive-Bayes classifiers. In: Proceedings of the 16th Australian Joint Conference on Artificial Intelligence, pp. 440–452.
Yang, Y., Webb, G.I., 2009. Discretization for naive-Bayes learning: managing discretization bias and variance. Mach. Learn. 74, 39–74.