Environmental decision-making using Bayesian networks: creating an environmental report card

Sandra Johnson[a]**, Murray Logan**[b]**, David Fox**[c,d]**, John Kirkwood**[e]**,
Uthpala Pinto**[e]** and Kerrie Mengersen**[a*†]

Environmental report cards are popular mechanisms for summarising the overall status of an environmental system of interest. This paper describes the development of such a report card in the context of a study for Gladstone Harbour in Queensland, Australia. The harbour is within the World Heritage-protected Great Barrier Reef and is the location of major industrial development, hence the interest in developing a way of reporting its health in a statistically valid, transparent and sustainable manner. A Bayesian network (BN) approach was used because of its ability to aggregate and integrate different sources of information, provide probabilistic estimates of interest and update these estimates in a natural manner as new information becomes available.

BN modelling is an iterative process, and in the context of environmental reporting, this is appealing as model development can be initiated while quantitative knowledge is still under development, and subsequently refined as more knowledge becomes available. Moreover, the BN model helps build the maturity of the quantitative information needed and helps target investment in monitoring and/or process modelling activities to inform the approach taken. The model is able to incorporate spatial and temporal information and may be structured in such a way that new indicators of relevance to the underlying environmental gradient being monitored may replace less informative indicators or be added to the model with minimal effort.

The model described here focuses on the environmental component, but has the capacity to also incorporate social, cultural and economic components of the Gladstone Harbour Report Card. Copyright © 2016 John Wiley & Sons, Ltd.

Keywords: environmental report card; Bayesian network; healthy harbour

1. Introduction

Report cards are widely used for reporting on issues such as the state of emergency care [1], food and nutrition of children [2], and are rapidly gaining popularity for timely and consistent dissemination of aquatic health information and for effective communication to the public and stakeholders. Thus, such a report card was requested by a consortium of industry and government bodies under the auspices of the Gladstone Healthy Harbour Partnership (GHHP). GHHP commissioned a review to identify the key characteristics of a successful report card with respect to its ability to capture and communicate the health of the harbour to stakeholders and the public [3]. That review considered 14 aquatic health report cards and concluded that a successful report card has clearly defined goals, engages its stakeholders, is flexible in its implementation, is able to communicate effectively and is based on rigorous science [3, 4].

Environmental report card indicators are selected to adequately reflect the current environmental health of an area and should be guided by the simple, measurable, achievable, realistic and time-limited (SMART) principle [5]. They are then recorded and visually represented in a report card to communicate this information to a wide and diverse audience [6]. Some popular statistical approaches for calculating environmental report card indices are principal component analysis, cluster analysis, analysis of variance and factor analysis [7, 8]. A variety of software is used to create these report cards including statistical packages, database packages, mapping and modelling software, custom built software and specialist desktop publishing software [3].

[a]*Queensland University of Technology, Brisbane, Queensland, Australia*
[b]*Australian Institute of Marine Science, Townsville, Queensland, Australia*
[c]*Environmetrics Pty Ltd, Melbourne, Victoria, Australia*
[d]*University of Melbourne, Melbourne, Victoria, Australia*
[e]*Gladstone Healthy Harbour Partnership, Gladstone, Queensland, Australia*
*Correspondence to: Kerrie Mengersen, Queensland University of Technology, Brisbane, Queensland, Australia.
[†]*E-mail: k.mengersen@qut.edu.au*

The use of Bayesian networks (BN) for environmental and resource management purposes is widely acknowledged and is popular for combining the available quantitative and qualitative information on the subject matter of interest [9]. The ability to visually represent the various elements of the model, to perform diagnostic and predictive modelling, include uncertainty and assist in decision-making has added to its wide appeal [10]. More recently this modelling framework has been used to develop a regional industrial harbour report card, which integrates social, cultural and economic indicators [6], and to develop a sustainability scorecard for the dairy industry in Australia, combining the social, economic and environmental sustainability perspectives [11, 12].

A BN modelling approach was therefore recommended for Gladstone Harbour Study for several reasons: Firstly, it provided a visual representation of the structure of the report card and the way in which the values reported in the report card were generated. Secondly, it enabled data from different sources to be combined in a simple and robust manner. Third, it incorporated the various sources of uncertainty in the estimates in a statistically valid manner. Fourthly, it allows estimates to be obtained at whole-of-harbour and zonal levels in a consistent manner. Lastly, it provided a model that could be used interactively, allowing for evaluation of 'what-if' scenarios by changing input values and observing the consequent changes taking into account the other factors in the model, and allowing for continual updating of information in a transparent manner.

The model is structured in a modular fashion, with the resultant probabilities represented as scores for each zone and for the harbour as a whole. For each zone, the available data are used to inform measures (e.g. turbidity), which are aggregated into indicators (e.g. physico-chemical), which are further aggregated into indicator groups (e.g. water quality) and then aggregated into subcomponents (e.g. water and sediment quality), and finally, aggregated into the environmental component report card grade.

The Gladstone Harbour Report Card was released as a pilot in 2014, and the first full release was in early 2016. The pilot report card showed the potential of the BN approach to report on the health of Gladstone Harbour. This paper describes a potential extension of the model used for the pilot study, using example data from the 2013–2014.

2. Methods

The Gladstone Harbour Report Card was envisaged by the GHHP as a summary of four key components of overall health: environmental, social, cultural and economic. Each of these components was to be composed of a set of subcomponents. For example, the environmental component, which is the focus for this paper, was defined as comprising four subcomponents: water and sediment quality, habitats, connectivity, and fish and crabs. Each of these in turn comprised a set of indicators that could be measured by available data, as described earlier. The report card was mandated to provide grades A (very good) to E (very poor) for each zone in the Harbour and for the Harbour as a whole. For water and sediment quality, these grades were to be based on data that were provided by the Port Curtis Integrated Monitoring Program (PCIMP).

This modular design of the report card lent itself naturally to the choice of a BN as an underpinning model. A BN hierarchically aggregates available data into measures, which are themselves aggregated into indicators, then indicator groups, then subcomponents and finally components. This model has a number of appealing features. It is a structured, logical design that can be collapsed and expanded to various broad or detailed levels, depending on the audience and motivation. Moreover, the design provides flexibility, so that entire sub-models may be replaced with alternative sub-models, or measures may be replaced with alternative measures, facilitating adaptive management processes.

A wide range of water and sediment quality measures was available from the PCIMP. The GHHP Independent Science Panel recommended that measures be selected according to five criteria: conceptual relevance (measure reflects ecosystem health), feasibility (measure is measurable), response variability (spatial and temporal signal, extent and variability understood), ability to detect change (along a defined gradient) and interpretation/utility (indicator values and bin ranges understood). The final decision also needed to take into account factors such as the scientific merit of including or excluding measures, and the suitability of the monitoring in capturing the state of the measure and the reliability and accuracy of the data. For example, some measures may be more sensitive to adverse changes in the health of the indicator and would therefore be preferable in providing an 'early warning system' that there are potential problems at a monitoring site, or more widely in a particular zone, or more generally in the harbour. The measure may also exhibit naturally fluctuating behaviour, so that the variation observed in the monitoring data may not accurately reflect an improvement or degradation of the measure, and consequently, the indicator. Finally, each measure was required to have a baseline or reference condition that could be used for comparative evaluation of obtained data.

Following a review of methods for developing baseline or reference conditions for indicators, attention was restricted to those measures for which Australian and New Zealand Environment and Conservation Council (ANZECC) or Queensland Department of Environment and Heritage Protection (DEHP) guidelines were available. The DEHP derived guidelines had the advantage of being specific for each zone within Gladstone Harbour, but they are based on a relatively limited period of historical data. These local guidelines were used for physicochemical parameters and nutrient levels and were calculated

as the mean levels of each measure over the period for which data were available [13]. In contrast, the ANZECC guidelines were used for metal concentrations and were based on the toxicity of those metals to aquatic organisms [14–16].

Five data grading bins were used, from A (very good) to E (very poor). The quantitative thresholds for these bins were determined using the PCIMP data and the corresponding guideline. The annual average value for each measure within each zone was estimated from four quarterly samples from each site within that zone.

An index was calculated for each of these averages and expressed relative to each guideline value. Fox [17] recommended that an index calculation needs to be targeted, relevant, fit-for-purpose, scientifically credible, statistically defensible and address key management issues. A scaled modified amplitude method was selected for the water and sediment quality indexes and capped at 10-fold change, as follows:

$$e_i = \begin{cases} log_2 \left(\frac{x_i}{benchmark_i} \right)^{-1} & if \ x_i > benchmark \ means \ failing \ guideline \ value \\ log_2 \left(\frac{x_i}{benchmark_i} \right)^{1} & if \ x_i < benchmark \ means \ failing \ guideline \ value \end{cases}$$

$$E_i = \begin{cases} log_2 (10) & if \ e_i > \ log_2(10) \\ -log_2 (10) & if \ e_i < \ -log_2(10) \\ e_i & else \end{cases}$$

$$For \ E_i \neq \ log_2 (10) :$$

$$Index_i = \ 1 + \frac{1}{\left(log_2 (10) \right) * (E_i - log_2(10))}$$

Using the scaled modified amplitude method, the deviation of the raw data mean from the guideline value was calculated and then scaled to lie in the range -1 to $+1$. Letting B denote the guideline and x denote the mean, the scaled value was taken to be $y = log_2(B/x)$. The value of y was then assigned to a grading bin using the thresholds A (0.6, 1.0), B (0.2, 0.6), C (−0.2, +0.2), D (−0.6, −0.2) and E (−1.0, −0.6). Special consideration was given to situations where the guideline was 'close to' a measure's detection limit, as defined by an expert panel, whereby probabilities within this range were apportioned amongst the graded data bins. The probability in each bin was then estimated as the proportion of the monitoring sites that were assigned that bin. When considering the grading bins using the whole of harbour view, the A to E grading was interpreted as the probability that the indicator has been assigned to a particular bin. Furthermore, a modal view was created whereby the actual grading bin assigned to the indicator was the bin with the highest probability. Zones, which did not have any data for a measure, were excluded from the calculations.

Letting a_i, b_i, c_i, d_i, e_i denote, respectively, the probabilities of a node state of A, B, C, D and E for the ith node in the BN, a utility function was used to calculate a score by multiplying the midpoint of each of five equal data bin intervals over [0, 1] with the corresponding probability of being in that data bin:

$$s_i = \left[(0.9 * a_i) + (0.7 * b_i) + (0.5 * c_i) + (0.3 * d_i) + (0.1 * e_i) \right] * 100$$

These scores are shown as 'expected utility' in the BN model. Each score can then be mapped to a report card grade that communicates the perceived health of the measure, indicator, indicator group, subcomponent or component. This grade may be for a particular zone, or by default, aggregated across all zones in the harbour.

In the construction of the BN, all child node conditional probability tables (CPTs) contain a matrix that aggregates the grading bins (A–E) of its parent nodes. In the absence of information to the contrary, the CPT assigns an equal weight to each grade in the combination, so that the resulting distribution across the grades of the child node equates to the unweighted average of the parent nodes. Note that the CPT matrix rapidly increases in size. The number of entries in the CPT is 5^n, where n is the numbr of parent nodes. The summary CPT nodes were generated programmatically and added to the model so it is able to aggregate measures into indicators, indicators into indicator groups, indicator groups into subcomponents and subcomponents into the environmental component at run-time.

The steps of a BN approach to the development of a report card are summarised in Figure 1.

In addition to the aggregated view across all zones, the report card was also required to provide a zonal perspective of the health of the harbour. A BN submodel was created to facilitate this display. The zone nodes depicted the composition of the whole of harbour perspective for each of the indicator groups: water quality, sediment quality and seagrass. Zones for which there were no monitoring data were set to zero, and in the absence of other guidance, the remaining zones all contributed equally to the overall scores and grades. The BN was structured so that the whole of harbour view provided a baseline for zone selection. Therefore, selecting a specific zone replaced the whole of harbour view of the report card model with the information for that zone, and switching the zone off, reverted back to the whole of harbour view.
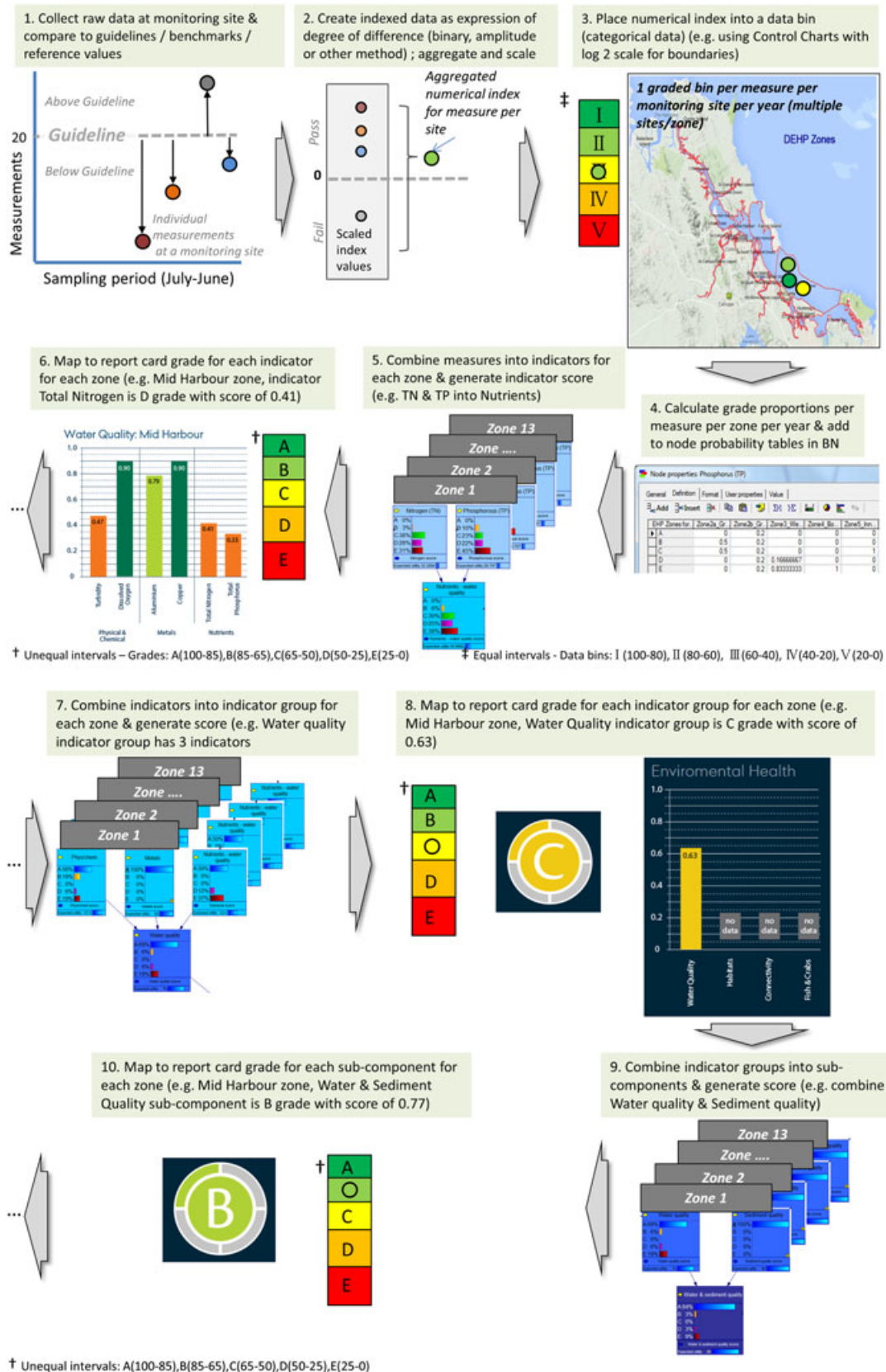
Figure 1. Overview of the steps employed in construction of the Bayesian network (BN) for the Gladstone Healthy Harbour Partnership (GHHP) Report Card. Note that the grades shown in this figure are for illustration purposes only.
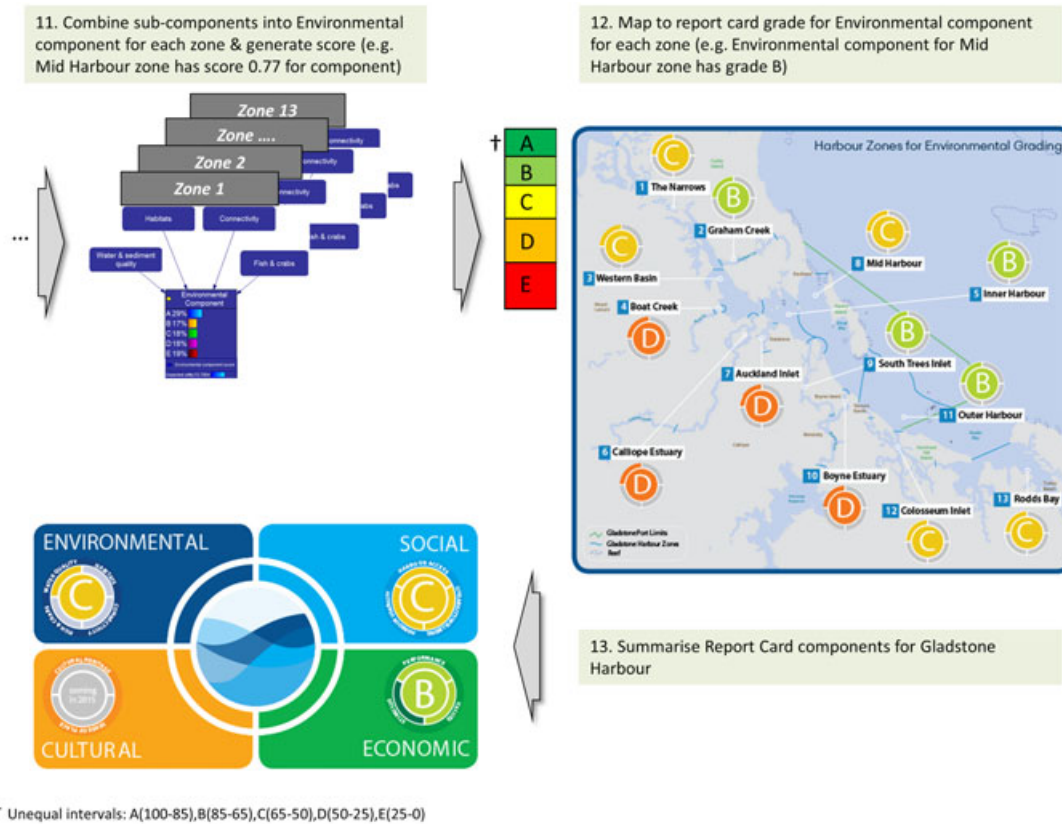
Figure 1. Continued.

Table I. Example raw monitoring data.

Site	Raw data, date 1	Raw data, date 2	Raw data, date 3	Average (3 data points)	Modified amplitude (scaled)	Graded bin
S1	140	130	200	157	−0.21	IV
S2	150	130	120	133	0.02	III
S3	220	150	185	185	−0.45	IV
S4	200	150	110	153	−0.18	III
S5	160	180	100	147	−0.12	III
S6	120	150	130	133	0.02	III
S7	150	160	160	157	−0.21	IV

The initial BN model was developed using GeNIe Modeler and the SMILE Engine software [18]. The model was then run in R [19] to calculate the indices, quantify the BN model and to generate the report card grades.

2.1. Worked example

The BN approach is illustrated through a worked example that focuses on total nitrogen (TN) in the Gladstone Mid Harbour zone. This measure feeds into the nutrients indicator for water quality. Note that the model and results described here are intended only for illustration and bear no relationship to the Gladstone Harbour report cards produced by GHHP. Reportable values may be found in the information published by GHHP (http://rc.ghhp.org.au/report-cards).

Step 1 *Collect monitoring data at the sites in the selected zone.* Table I shows an illustration of raw monitoring data from seven monitoring sites in one zone for three sampling dates.

Step 2 *Calculate averages for each monitoring site in the zone.* This is indicated in the 'average' column in Table I. Here, deviations from the guideline value for TN were quantified using the modified amplitude indexation technique. If required, an index can optionally be scaled to a range defined by a certain fold (e.g. twofold) difference from guidelines. For the water and sediment quality indexes, an example default fold could be 2, meaning that indexes would be rescaled
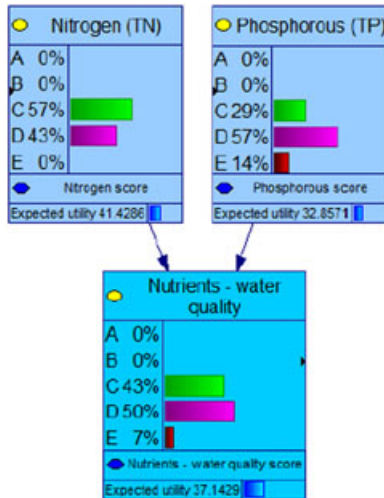
Figure 2. Example of combining measures into indicators (step 5).

to a maximum of $\log_2(2)$ and a minimum of $-\log_2(2)$. For this example, the average would be compared with 135, which is the guideline value for TN for this zone. Therefore, the modified amplitude index value for site S1 would be log2 $(135/156.7) = -0.22$

Step 3 *Assign to data bins*. Using equal boundaries (I: +1.0 to +0.6, II:+0.6 to +0.2, III: +0.2 to +0.2, IV: +0.2 to +0.6, V: +0.6 to +1.0), TN for site S1 would be assigned to data bin 'IV'.

Step 4 *Complete the node probability table in the BN model*. This shows the proportion of sites falling into each of the data bins (I to V), for each zone of a measure. Continuing with the example in Table I, the proportions for grades were 0 for data bins I, II and V; 3/7 = 0.43 for bin IV, and 4/7 = 0.57 for bin III. The 'expected utility' node then calculates a score from the probabilities of the node, using the midpoint of each data bin. For this example, the score is $[(0.9 * 0) + (0.7 * 0) + (0.5 * 0.57) + (0.3 * 0.43) + (0.1 * 0)] * 100 = 41$. Based on the grade intervals that were used in the 2014 pilot report card, A (100–85), B (85–65), C (65–50), D (50–25), E (25–0), this value means that a report card grade of D would be assigned to TN in this zone.

Step 5 *Combine measures into indicators*. Only two measures, TN and total phosphorous (TP), are included in this example to quantify the nutrients – water quality indicator (Figure 2). The resulting proportions for the data grading bins are an unweighted average of the two measures. A similar procedure can be followed to translate the proportion of grading bins for the indicator into a score.

Step 6 *Map the score to a report card grade*. The score for the water quality nutrients indicator in this zone is 37 (Figure 2), so nutrients would have a grade of D in the report card.

Step 7 *Combine indicators into indicator groups for each zone and generate score*. In this example, the nutrients indicator is aggregated with two other indicator nodes: metals and physico-chemical to calculate a score for the water quality indicator group (Figure 3). The expected utility node generates a score for the indicator group, which is 63 for this zone.

Step 8 *Map to report card*. The score of 63 would be assigned a grade of C.

Step 9 *Combine indicator groups into subcomponents and generate score*. The process described earlier could be repeated for each indicator group to create subcomponents.

Step 10 *Map to report card grade for each subcomponent for each zone*. The earlier process could be repeated for all the other Gladstone Harbour zones.

Step 11 *Combine subcomponents into environmental component for each zone and generate score*. The overall grades and scores for all the zones could be combined into an overall harbour grade and score. In this example, all zones were determined to contribute equally to the overall grade and score, but differential weights could be used.

Step 12 *Map to report card grade for environmental component in each zone*. The derived scores can be mapped as described in Figure 1.

Step 13 *Summarise report card components for Gladstone Harbour*. The case study information can be integrated into the overall report card.
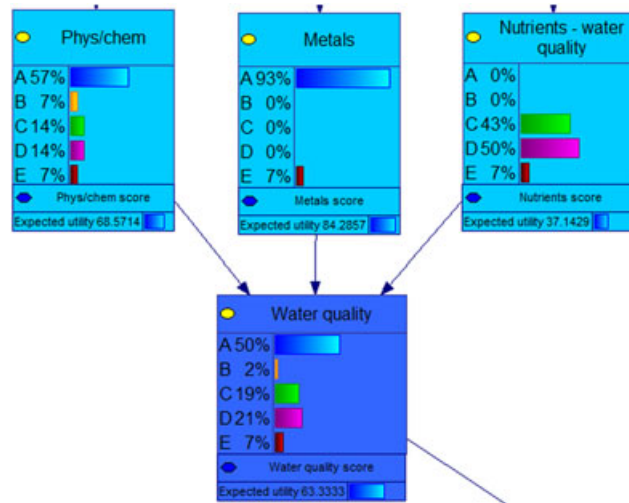
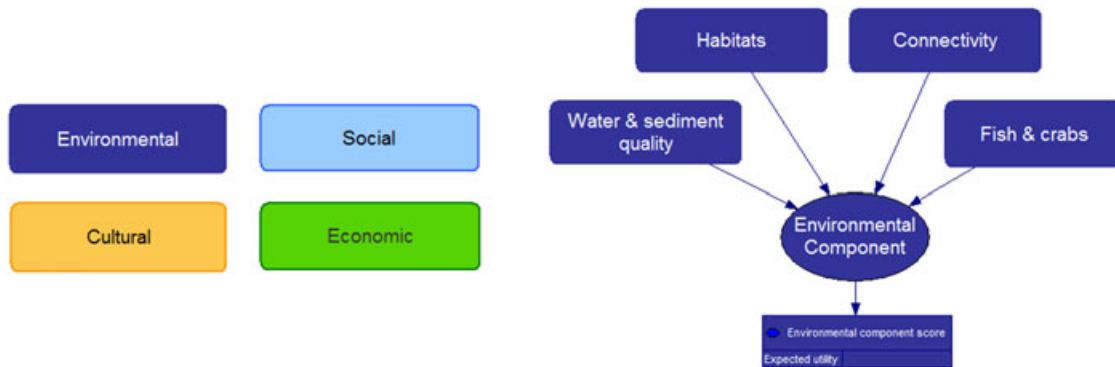**Figure 3.** Example of combining indicators into an indicator group (step 7).

Figure 4. High level view (left) and environmental view (right) of the GHHP Report Card.

3. Results

3.1. Construction of the Bayesian network

The 13 spatially defined zones in the study area are depicted in Figure 1, step 13.

Two levels of the report card model are shown in Figure 4: the high level consists of four key components of harbour health (environmental, social, cultural and economic), and the next level down shows four subcomponents of the environmental health component (water and sediment quality, habitats, connectivity and fish and crabs). The pilot report card included only the water quality indicator group, but the BN framework enables a straightforward way to integrate additional components and subcomponents as illustrated in Figure 4.

The model structure of the water and sediment quality subcomponent of the environmental component consists of two indicator groups, water quality and sediment quality, which have three and one indicators, respectively (Figure 5). The hierarchical nature of the model structure is further illustrated with each of the indicators having an associated BN sub-model containing the measures used to quantify it. For example, Figure 6 shows one node the nutrients – water quality indicator as a BN sub-model with six measures:chlorophyll a, phosphorous (TP - total phosphorous), orthophosphate (FRP - filterable reactive phosphorus), nitrogen (TN - total nitrogen) and nitrogen oxides (NOx) (note that only TN and TP were included in the pilot report card.)

The BN sub-models each generate a probability distribution across the data grading bins, which is representative of the proportion of the monitoring sites having been assigned to each of them. For example, an indicator may have 70% assigned to the bin indicating 'very good', 10% to a bin indicating 'good and 20% to a third bin across the monitoring sites in a particular zone. This can be interpreted that for this indicator, 70% of the monitoring sites in that zone were considered to be very good, 10% were considered good and 20% of the sites were less healthy.
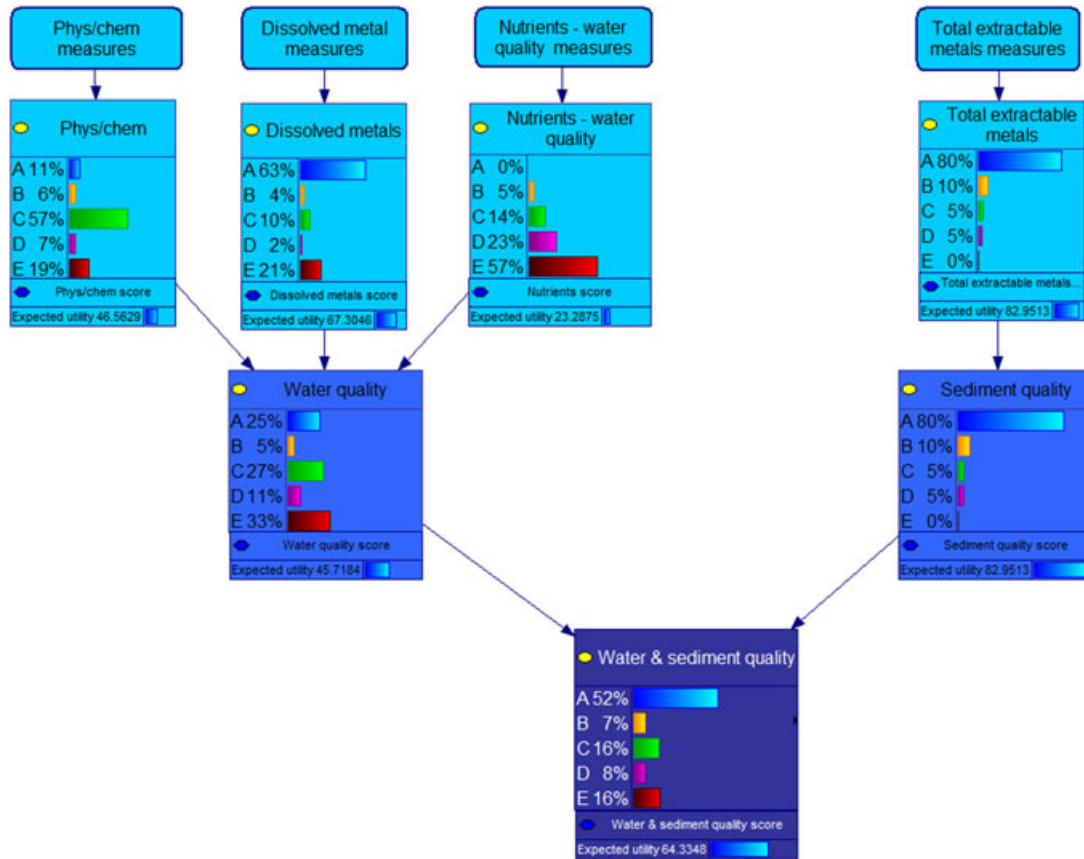
Figure 5. Example BN for a water and sediment quality subcomponent, combining two indicator groups, water quality and sediment quality, and their respective indicators.
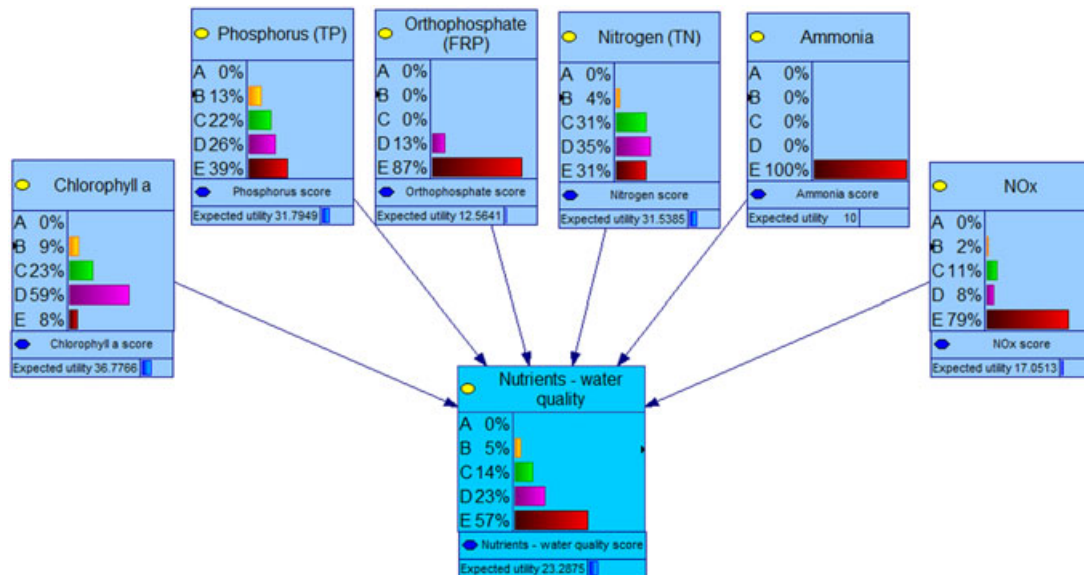

Figure 6. Example Bayesian Q3 network (BN) for nutrients – water quality indicator aggregating six nutrient measures.

3.2. Results for worked example

Using the data analyses based on supplied guidelines and objectives, potential measure scores, data bins and report card grades can be calculated for the whole of Gladstone Harbour and for zones within the harbour. The report card structure illustrated earlier can generate scores and probabilities across data grading bins for each of the zones so that any areas

of interest or concern may be explored in more detail. A BN subnetwork model facilitates a zonal view as well as the aggregation across all zones for the whole of harbour perspective (Figure 7).

Figure 8 shows how the data grading bins and scores for the environmental component can be aggregated across all the zones of Gladstone Harbour. In this example, the score for the environmental component is 58, which would yield a report card grade of C for the harbour as a whole.

In this example, the distribution of the proportions in the various data bins of the water and sediment quality subcomponent was skewed towards the healthy range, and with a score of 64 was assigned a report card grade of C (Figure 5). Moreover, the next layer of detail showed that the sediment quality indicator group, consisting of only one indicator, total extractable metals, was even more heavily skewed towards the very good range, and the score of 83 maps to a report card grade of B. On the other hand, the water quality indicator group had probability mass in roughly equal proportions in the data bins that are meeting and failing guideline values and had a score of 46, which would map to a report card grade of D using the method described earlier.
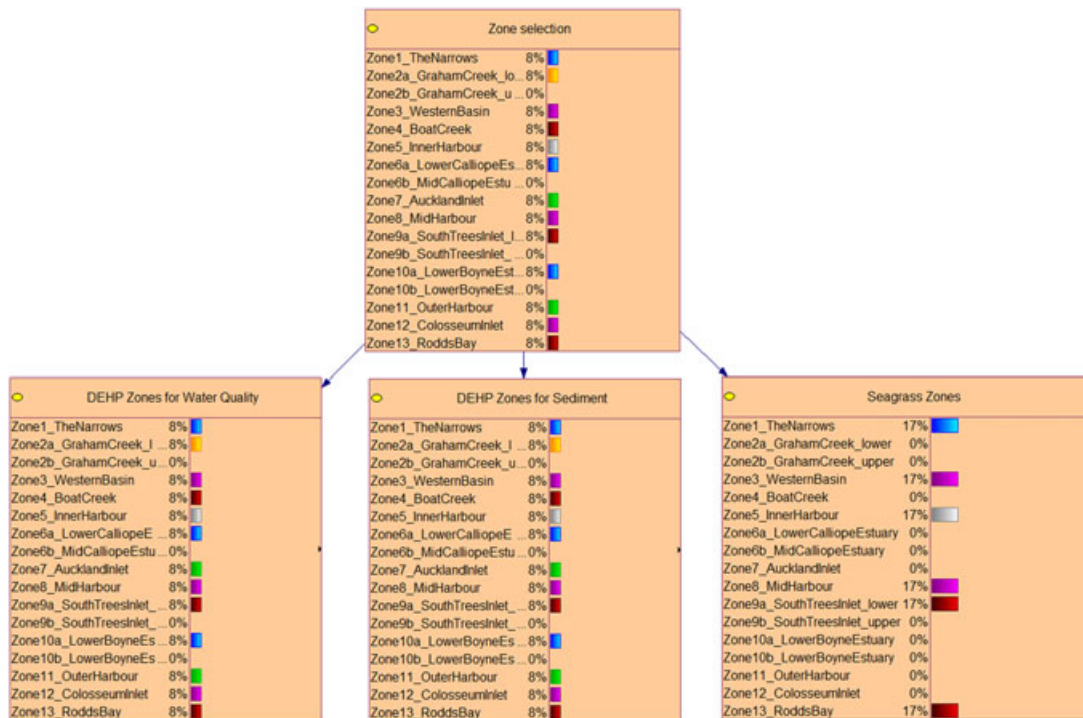

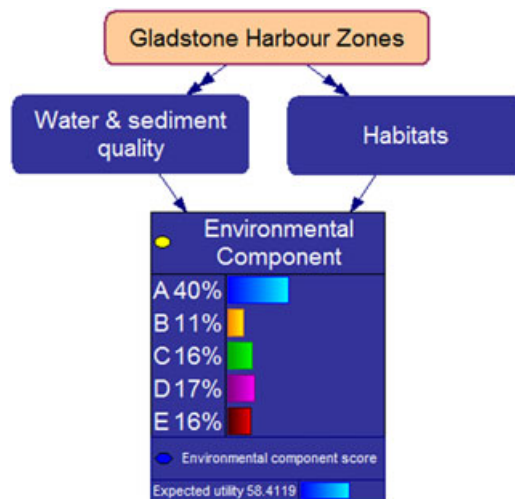**Figure 7.** Illustration of a zone selection sub-model for Gladstone Harbour.

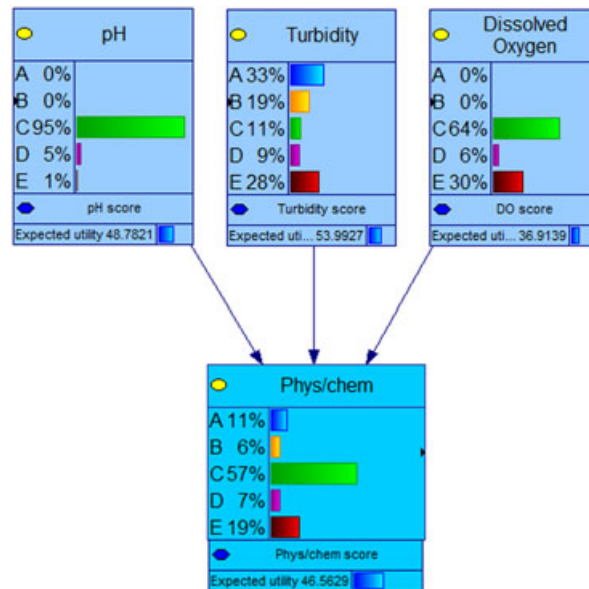**Figure 8.** Example aggregation of results for the whole of Gladstone Harbour.

Figure 9. Example results for the physico-chemical water quality indicator.
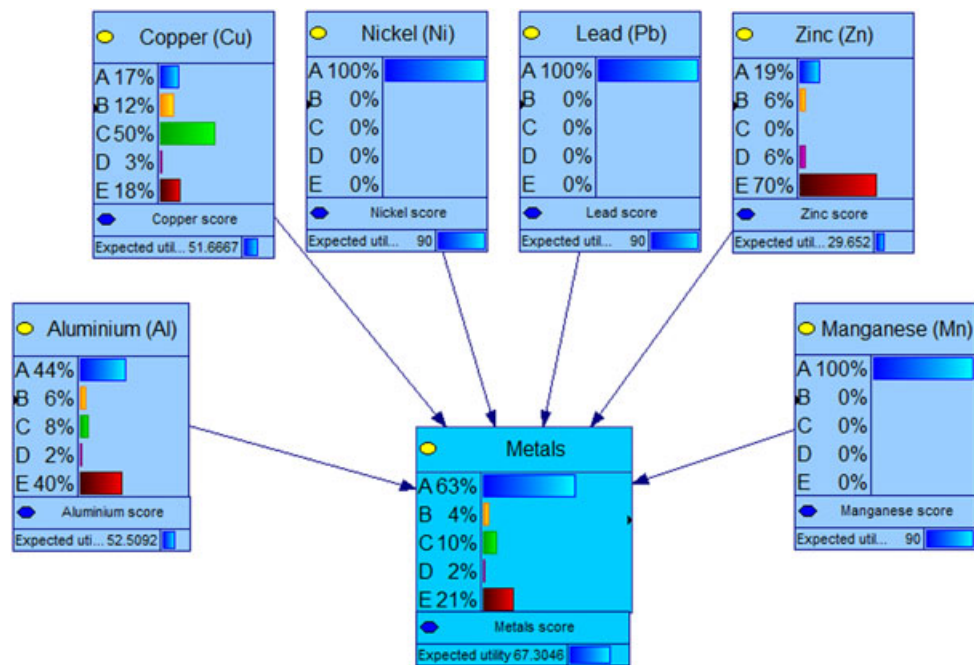

Figure 10. Example grading bins and scores for the metals water quality indicator.

An example of data grading bins and scores for a physico-chemical water quality indicator (node *Phys/chem*) for the whole of harbour is shown in Figure 9. The three inputs into this indicator are pH, turbidity and dissolved oxygen (DO). Here, the modal grading bin generally meets guideline values (data bin III has 57%), but it is important to note that 26% of the indicator fails acceptable levels (data bins IV 7% and V 19%). Furthermore, the probability distributions across the data bins of the three measures differ quite markedly; notably turbidity tells a different story to DO. The shape of the distribution for pH and DO is broadly similar. However, DO has a substantial proportion failing the guidelines, 36%, compared with just 6% for pH. In this case, the score attached to the physico-chemical water quality indicator is 46.6, which would map to a report card grade of D. The measure indexes aggregated in this indicator were computed using the scaled modified amplitude indexation method.

In a similar manner to that described earlier, the water quality metals indicator can be obtained by aggregating measures such as aluminium, copper, nickel, lead, zinc and manganese (Figure 10). The water quality nutrients indicator can also
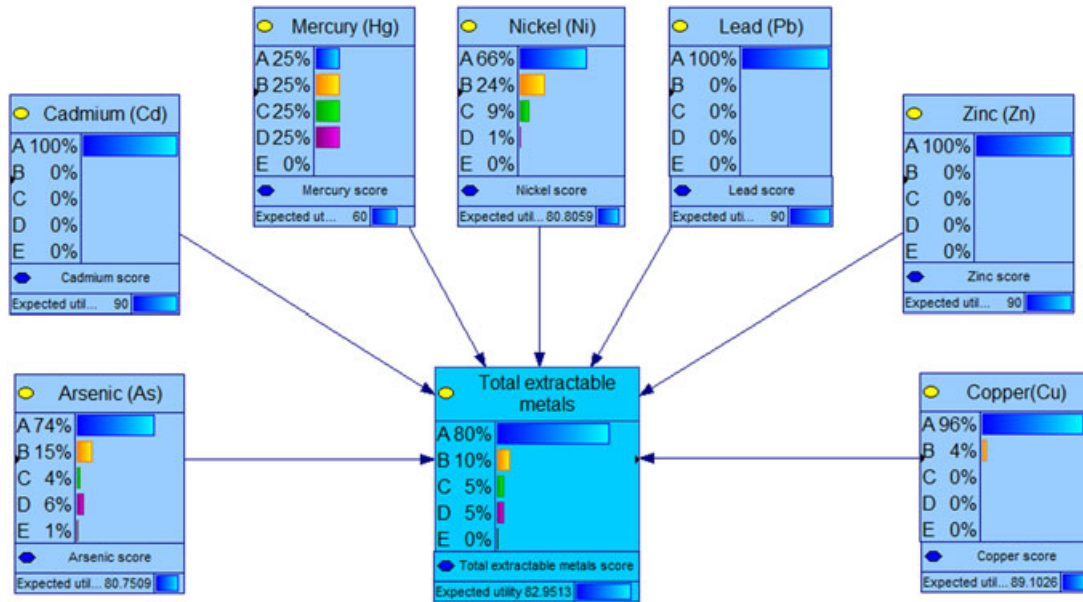
Figure 11. Example results for the sediment quality total extractable metals indicator.

Table II. Report card structure based on output of BN. Additional rows can be added for the overall Harbour or other zones, and additional columns for other components, subcomponents and indicator groups.

Grades and scores

Grade	A	B	C	D	E
Score	100 - 85	85 - 65	65 - 50	50 - 25	25 - 0

	Component		Sub-Component		Indicator Group			
	Environmental		Water & Sediment Quality		Water Quality		Sediment Quality	
Zone	Grade	Score	Grade	Score	Grade	Score	Grade	Score
Zone xxx	C	58.4	C	59.4	D	45.7	B	83.0
Zone yyy	D	48.0	B	66.9	D	38.8	B	80.0
Etc								

combine measures such as chlorophyll a, phosphorus, nitrogen, orthophosphate, NOx and ammonia (Figure 6). For the sediment quality indicator total extractable metals, seven candidate measures could be included arsenic, cadmium, mercury, nickel, lead, zinc and copper (Figure 11). With the exception of mercury, this indicator has a fairly consistent modal assessment across the measures, with six of the seven measures having a high probability of being in the 'very good' bin (66% to 100%). Mercury, the exception, is an interesting case because its guideline value is below the detection limit. Rules for dealing with special cases such as this were developed by GHHP; this resulted in 25% being assigned to the top four data bins. Thus, the report card grade for the extractable metals indicator would be B (score 83).

The final grades and scores based on the earlier example can be represented as a report card, as illustrated in Table II.

4. Discussion

As illustrated in this paper, in addition to providing an aggregated summary, a report card can be valuable in reflecting inconsistencies with the health of the harbour. The divergent scores may be further investigated by considering, for each measure, the characteristics of the raw monitoring data and the suitability of the guideline values. For example, unusual weather events may affect certain measures more than others, or a monitoring device may have been malfunctioning. Moreover, the transparency of the approach used is further illustrated by the worked example. In this instance, an average of the raw data would have been a suitable alternative. However, in other cases such a simple average would not be suitable, and an indexation method such as the modified amplitude method is preferable. The use of distance-to-guidelines (amplitude) over binary compliance for individual measures is recommended by Fox [17].

Developing environmental indicators is a complex process, and the set of indicators deemed most suitable for an environmental report card may change over time as new research comes to light, more sophisticated monitoring methods are employed, more sensitive monitoring equipment become available and economically viable to use or as other scientific considerations that directly affect the purpose or scope of the report card evolve. Furthermore, it would be prudent to investigate the value that each measure contributes to the associated indicator. In other words, what is the optimal number of measures and which measures are best to communicate the health of that indicator? If two measures are very highly correlated, it would likely be of minimal or no value to include both. However, these correlations must be assessed in the context of the scientific knowledge about the ability of each measure to reflect the health of the indicator under all pertinent external conditions. For example, in most instances, the two measures may appear to be perfectly correlated, but under some extreme or abnormal weather conditions, this strong correlation may no longer hold and information from both measures is needed to understand the health of the indicator.

A feature of the BN modelling approach is that it can incorporate expert information and update prior probabilities as more data become available. For example, in the context of the Gladstone Harbour Report Card, the importance attributed to certain zones in influencing the actual or perceived health of Gladstone Harbour may be modified in accordance with expert judgement to afford more or less credibility or importance to certain zones within the harbour. The weighting of the DEHP zones for each of the indicator groups (Figure 7) could then be updated to reflect this altered composite view of the health of Gladstone Harbour and the model rerun to generate the new distributions for all the measures, indicators, indicator groups, subcomponents and components and the associated scores and report card grades. Moreover, the utility of equating the proportion of monitoring sites in each zone to the probability that the zone is in the said class, assumes that each monitoring site contributes equally to the perceived health of that zone. This assumption needs to be further discussed in light of scientific and geographical knowledge of the harbour and the environmental indicators, to recommend alternative weighting strategies or conversely to endorse the current approach if no scientifically rigorous and defensible alternatives are found.

The BN model presents a visual image of the measures and indicators of a report card, and the expected utility, which denotes the final classification of each report card indicator. As can be seen in Figure 5 such a classification may fall into a class that has a very low probability itself. In this instance, the transparency of a BN modelling approach enables the end-user to assess whether this is desirable behaviour or not, and this apparent anomaly could be reported and discussed.

In addition to the environmental focus of this paper, it is also important to consider an economic perspective, because Gladstone is a busy and important harbour in Australia, supporting several industries and importing and exporting many commodities. The port is one of the world's top five coal export ports [20, 21]. The recommended modelling approach for the social, cultural and economic components in the Gladstone Harbour Report Card was also a BN framework [6, 21]. Similar to the hierarchical structure of the environmental component of the Gladstone Harbour Report Card, the economic component consisted of three subcomponents: direct economic footprint, economic stimulus and recreational value. To quantify these subcomponents of the economic health of the harbour, three indicators were aggregated for direct economic footprint (tourism, commercial fishing and shipping activity); two indicators were aggregated for economic stimulus (unemployment and socio-economic status) and three for recreational value (recreational fishing, beach recreation and land-based recreation). This feature of seamless integration of quite disparate information in a BN via modular subnetworks enables entire sub-models to be added or replaced with alternative sub-models, or measures to be replaced with alternative measures. This facilitates adaptive management processes and provides the flexibility when combining various types of complex information for reporting purposes [3].

Applied Stochastic
Models in Business
and Industry

References

1. Hirshon JM, Alson RL, Blunk D, Brosnan DP, Epstein SK, Gardner AF, Lum DL, Moskovitz JB, Richardson LD, Stankus JL, Kivela PD, Wilkerson D, Price C, Bromley M, Calaway N, Geist M, Gore L, Singh C, Wheeler G. America's emergency care environment a state-by-state report card: 2014 edition. *Annals of Emergency Medicine* 2014; **63**:100–243.
2. Olstad DL, Raine KD, Nykiforuk CIJ. Development of a report card on healthy food environments and nutrition for children in Canada. *Preventive Medicine* 2014; **69**:287–295.
3. Connolly R, Bunn S, Campbell M, Escher B, Hunter J, Maxwell P, Page T, Richmond S, Rissik D, Roiko A, Smart J, Teasdale P. *Review of the Use of Report Cards for Monitoring Ecosystem and Waterway Health*. Gladstone Healthy Harbour Partnership: Queensland, 2013.
4. McIntosh EJ, Poiner IR, ISP Members. Gladstone Harbour Report Card Framework recommendation – March 2014, 2014.
5. Dauvin JC, Fisson C, Garnier J, Lafite R, Ruellet T, Billen G, Deloffre J, Verney R. A report card and quality indicators for the Seine estuary: from scientific approach to operational tool. *Marine Pollution Bulletin* 2008; **57**:187–201.
6. Pascoe S, Tobin R, Windle J, Cannard T, Marshall N, Kabir Z, Flint N. Developing a social, cultural and economic report card for a Regional Industrial Harbour. *PLoS ONE* 2016; **11**:e0148271, 1–19.
7. Whittaker G, Lautenbach S, Volk M. What is a good index? Problems with statistically based indicators and the Malmquist index as alternative. *2012 International Congress on Environmental Modelling and Software Managing Resources of a Limited Planet*, 2012;925–932.
8. Fox DR. Probability weighted indices for improved ecosystem report card scoring. *Environmetrics* 2014; **25**:351–360.
9. Barton DN, Kuikka S, Varis O, Uusitalo L, Henriksen HJ, Borsuk M, de la Hera A, Farmani R, Johnson S, Linnell JDC. Bayesian networks in environmental and resource management. *Integrated Environmental Assessment and Management* 2012; **8**:418–429.
10. Johnson S, Fielding F, Hamilton G, Mengersen K. An integrated Bayesian network approach to *Lyngbya majuscula* bloom initiation. *Marine Environmental Research* 2010; **69**:27–37.
11. Buys L, Mengersen K, Johnson S, van Buuren N, Chauvin A. Creating a sustainability scorecard as a predictive tool for measuring the complex social, economic and environmental impacts of industries, a case study: assessing the viability and sustainability of the dairy industry. *Journal of Environmental Management* 2014a; **133**:184–192.
12. Buys L, Mengersen K, Johnson S, van Buuren N, Miller E. *A Triple Bottom Line Planning Tool for Measuring Sustainability: A Systems Approach to Sustainability Using the Australian Dairy Industry as a Case Study*. Chartridge Books Oxford: Oxford, UK, 2014b.
13. DEHP. *Curtis Island, Calliope River and Boyne River Basins Environmental Values and Water Quality Objectives: Basins 131, 132 and 133 including all waters of Gladstone Harbour, the Narrows, Curtis Island, Calliope and Boyne River Basins and adjacent coastal waters*. Department of Environment and Heritage Protection: Brisbane, 2014.
14. ANZECC/ARMCANZ. Australian and New Zealand guidelines for fresh and marine water quality, 2000. Volume 1: The guidelines. Australia and New Zealand Environment and Conservation Council and Agriculture and Resource Management Council of Australia and New Zealand, National Water Quality Management Strategy Paper no. 4, Canberra.
15. ANZECC. Water quality guidelines for the protection of aquatic ecosystems. Toxicants manganese – marine water, 2014,1–11. Australian and New Zealand Environment and Conservation Council, Canberra.
16. Golding LA, Angel BM, Batley GE, Apte SC, Krassoi R, Doyle CJ. Derivation of a water quality guideline for aluminium in marine waters: derivation of a marine water quality guideline for aluminium. *Environmental Toxicology and Chemistry* 2015; **34**:141–151.
17. Fox DR. Statistical issues associated with the development of an ecosystem report card, 2013.
18. BayesFusion L. GeNIe Modeler & SMILE Engine, 2015. http://www.bayesfusion.com [accessed 23 May 2014].
19. R Core Team. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria, 2014. http://www.R-project.org/ [accessed 11 July 2014].
20. Greer L, Kabir Z. Guidance for the selection of social, cultural and economic indicators for the development of the GHHP Report Card. In *Report to the gladstone healthy harbour partnership*. School of Human Health and Social Science: CQUniversity Australia, Rockhampton, 2013; 1–50.
21. Pascoe S, Tobin R, Windle J, Cannard T, Marshall N, Kabir Z, Flint N. Piloting of social, cultural and economic indicators for the Gladstone Healthy Harbour Partnership Report Card, 2014. CSIRO Final report for the Gladstone Healthy Harbours Partnership.