Modern Management based on Big Data III
A.J. Tallón-Ballesteros (Ed.)

doi:10.3233/FAIA220123

Restoration Effect Evaluation of Artificial Reefs Based on Bayesian Networks

Xin LV[a], Yifei ZHANG[a1], Xin FANG[a,b1] , Zonghao HOU[a]
[a]*Second Institute of Oceanography, Ministry of Nature Resources, Hangzhou 310012, China*
[b]*School of Geography and Ocean Science, Nanjing University, Nanjing 210093, China*

Abstract. To meet the increasingly urgent need for ecological restoration effect evaluation, the ecosystem health status of reefs was described, and an ecosystem health evaluation model of artificial reefs (ARs) based on Bayesian networks (BNs) was established. By comparing the probability of the ecological health status between restored areas and control areas, we assessed the AR ecological restoration effect of the Qinhuangdao project in 2012. The results show that this project had a remediating effect in May and September, with clear repair effects on the water environment, sediment environment, and fishery resources. The sensitivity of each indicator was calculated, and organic carbon and benthos biomass was identified as the most sensitive factors. This study will provide the basis for the further development of restoration measures.

Keywords. Artificial reefs (ARs), ecological restoration, Bayesian networks (BNs), effect assessment, ecosystem health

1. Introduction

Significant degradation of marine ecosystems has occurred owing to human disturbances [1, 2]. Artificial reefs (ARs) both repair the water environment and provide habitats for fish, which is conducive to fishery proliferation [3, 4, 5], and helps to counteract the degradation of the ecosystem. A lot of ecological restoration work has been conducted on ARs. Therefore, it is crucial to evaluate the ecological restoration effect of ARs objectively.

The ecological restoration evaluation of ARs has been mainly carried out using two methods: the comparison of individual indicators and the calculation of a comprehensive index. The former has been conducted using various aspects, such as fishery resources [6], the benthic community level [7], the water quality environment, sediment status, and marine biological resources [8] compared the biomass and species diversity of fishery resources in before and after reef construction. Single indicator comparison is simple, but requires accurate and complete data [9]. Moreover, Zhao et al. and Tong et al. comprehensively evaluated the ecosystem health of restoration area and control area by calculating the ecosystem health index (EHI) [10, 11]. Fu et al. assessed the ecological

[1] Corresponding Author, Yifei Zhang, Second Institute of Oceanography, Ministry of Nature Resources, Hangzhou 310012, China; E-mail: zhang1213@163.com.
Xin Fang, Second Institute of Oceanography, Ministry of Nature Resources, Hangzhou 310012, China; E-mail: fangx@sio.org.com.

restoration effect by calculating the weighted average of the changing rate of each indicator [12]. Wang et al. used the exergy value to reflect the community structure, stability, and organization degree, and thus to evaluate the health status of animal communities in the control and restoration area. The comprehensive evaluation index provides a direction for artificial reef ecological restoration [13]. Furthermore, the concept of ecosystem health contains kinds of possibilities, using probability to evaluate ecosystem health uncertainty can help managers understand the mechanism behind the model and make more scientific decisions.

The Bayesian network (BN) can conduct probability reasoning from a parent node to a child node and vice versa. At the same time, it is especially well suited to handle missing data, making it suitable for risk assessment, diagnosis, and decision analysis [14, 15]. Based on the logical characteristics of the BN, we evaluated the ecological restoration effect of ARs in Qinhuangdao, China. Therefore, we regarded the ecosystem health status as a variable with its own probability distribution, and demonstrated the uncertainty of ecosystem health status in the study area in the form of a probability so as to evaluate the restoration effect of ARs.

The objectives of present research taking the Qinhuangdao AR restoration project as a case study, are: (1) to propose an index integrating water quality indicators, sediment environment indicators, biotic indicators and fishery environment; (2) to build a BN model based on the index for assessing the healthy state of the Qinhuangdao AR ecosystem; (3) to compare the healthy state of the control area and restoration area thus evaluating the restoration effect of ARs

2. Study Area and Data Collection

2.1. Study Area

The Qinhuangdao coastal areas are important fish spawning sites and feeding grounds in China. In recent years, they have suffered from problems such as the extinction of marine organisms, reduction of biodiversity, and declining catches due to overfishing, reclamation, and water pollution [16]. To restore the deteriorating coastal environment of Qinhuangdao, the relevant departments conducted a restoration project in 2010, including the creation of an AR restoration area (Figure 1).

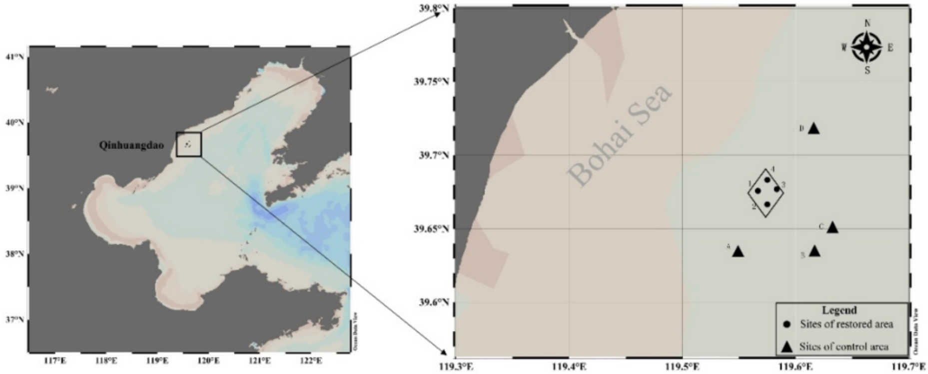

Figure 1. Study area.

2.2. Data Collection

This study involved data including biochemical oxygen demand, organic carbon, the phytoplankton diversity index, the zooplankton diversity index, and the species number of swimming animals and biomass, which were all collected from Xu et al. [17]. These data were measured on two survey voyages in May and September 2012, respectively. The distribution of survey stations is shown in Figure 1. Sites 1 to 4 are the AR restoration areas, while A to D is the control site outside the reefs.

3. Methods

3.1. Theoretical Background and Applications

BNs, also known as Bayesian belief networks, are composed of directed acyclic graphs and node probability tables (NPTs), which are used to describe the dependence between variables and the quantitative relationship between variables, respectively [18].

The theoretical basis of BNs is the Bayesian formula, shown in Equation (1). BN is a representation of the decomposition of joint probability distribution. The probability distribution of the target node is expressed by a joint distribution containing n variables, as shown in Equations (2) and (3):

$$P(\pi(X_i) \mid X_i) = \frac{P(\pi(X_i))P(X_i \mid \pi(X_i))}{\sum_{i=1}^{n} P(\pi(X_i))P(X_i \mid \pi(X_i))} \tag{1}$$

$$P(X_i) = \sum_{i=1}^{n} P(\pi(X_i))P(X_i \mid \pi(X_i)) \tag{2}$$

$$p(X_1, \ldots, X_n) = \prod_{i=1}^{n} P(X_i \mid \pi(X_i)) \tag{3}$$

where $P(\pi(X_i) \mid X_i)$ is the posterior probability, $P(\pi(X_i))$ is prior probability, $P(X_i)$ is the edge distribution of the root node, $P(X_i \mid \pi(X_i))$ is the conditional probability, and $p(X_1, \ldots, X_n)$ is the joint probability. The probability of the parent node is corrected while the observation value is updating, and the observation results can be input at any node.

BNs have the advantage of combining expert knowledge with objective data. Forio et al. used this characteristic to predict the water quality of typical multi-functional tropical watersheds [19]. Havron et al. and Mo used BNs' expression in the form of probability [20, 21]. The former simulated the probability of several macrobenthos' occurrence in different geographical locations to analyze the habitat suitability, and the latter predicted ecological vulnerability. The results of the abovementioned studies indicate that BNs can be used to analyze the uncertainty of the ecological environment.

3.2. Research Framework

A systematic framework is proposed in six steps to describe how BNs can be applied in assessing the ecological restoration effect:

Step 1. Confirm the evaluation system. Based on the evaluation requirements and the availability of data, we took ecosystem health as the target node took fishery resources, water environment, sediment environment, and biological environment (B1 to B4) as the criterion layer nodes, and selected several indicators as the parent node to construct a hierarchical model.

Step 2. Build the BNs' structure. Using some structural forms to define the graphic part of BNs can accelerate the development process of BNs and improve their quality [22]. The definitional/synthetic idiom is the most common form, and the characteristic of this kind of structure defines the synthetic node according to the parent node. In this paper, the intermediate node was defined by classifying the indicators, and then the target node was defined according to the concept of the intermediate node so as to construct AR ecological health evaluation system.

Step 3. Calculate the NPTs. Determining the BN parameters involves constructing NPTs, which can be constructed by manually filling out the probability table and inputting the eigenvalues of the measured data. We selected the latter, which mainly includes the following points:

① Data standardization

The data set was standardized according to Equation (4), and each index was dimensionless.

$$x^* = \frac{x - x_{\min}}{x_{\max} - x_{\min}} \tag{4}$$

where x is measured data; x^* is standardized data; and $x_{\min}$ and $x_{\max}$ are the minimum and maximum value in a dataset, respectively.

② Ranked node

The prior probability of root nodes in the model needs to be given when using BNs for prediction. Hierarchical nodes make it easier to build and edit BNs than others [23]. Five evaluation grades were set on root nodes. Considering that the standardized data fell in the interval of [0,1], we divided this interval into five grades as 'VL/L/M/H/VH,' with 0.2 as the interval, and preliminarily judged the evaluation grade of each data point. The data were divided into four categories according to seasons and regions. Averaging all site values for each evaluation index by classification, we determined the comprehensive grade of the root node as the prior knowledge according to the interval of the average value.

When using hierarchical nodes to describe variables, it is necessary to ensure that the node state is consistent with the good or bad trend of the evaluation target; that is, the higher the reverse index value in this study, the worse the corresponding node state [22].

③ Produce NPTs

The truncated normal distribution (TNormal) was introduced to define the statistical function in the interval [0,1]. The distribution is a finite endpoint, all nodes are limited in [0,1], and the expression is $TN\text{ormal}(\mu,\sigma^2,0,1)$. NPTs could be obtained by inputting the mean expression and variance of the relationship between nodes. Considering the different comparison objects, the corresponding standardized station data were selected to calculate the mean and variance.

Step 4. Allot the weight of evaluation indicator. The mean of the superior node could be linearly added according to the mean of the inferior node multiplied by the weight; the variance of superior nodes could be represented by the reciprocal of the sum of the weights of inferior nodes.

The weight value of intermediate node B was determined by equal weight, and the weight value of root node C was calculated by principal component analysis in SPSS25.0 software (IBM).

Step 5. Deduce the probability distribution of target node. In order to facilitate the comparison of results, the intermediate node B was divided into five evaluation grades, and the target node A was divided into seven evaluation grades. According to the distribution of probability in different levels, the levels of target nodes and intermediate nodes were determined.

Step 6. Identify sensitive factors. Sensitivity can reflect the average influence of the change of variable state on the state of the target node [24]. Through sensitivity calculation and analysis, the key factors affecting the health of AR ecosystem could be identified, and the key repairing indexes in ecological restoration projects could be clarified. The calculation of sensitivity is as follows:

$$I_a(C_k)=\frac{1}{P_k}\sum_{c_k}^{P_k}\frac{|P(C_k=c_k\mid A=a)-P(C_k=c_k)|}{P(C_k=c_k)} \tag{5}$$

where c_k is the state of root node C_k, P_k is the number of states, and a is the state of the target node A.

4. Model Construction

4.1. Index System for Assessing AR Ecosystem Health

The BN model established in this study contains three hierarchical structures (Figure 2). The root node is each specific index. The intermediate node is obtained by classifying and defining the meaning of the root node, and the target node is the AR ecological health status (A).

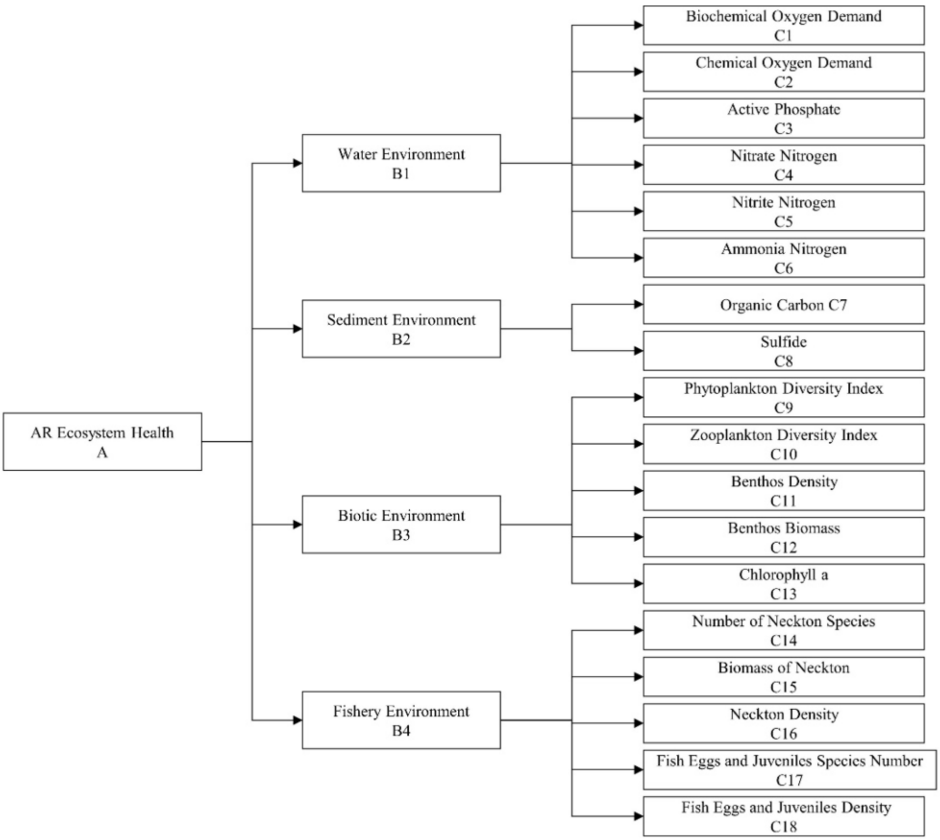

Figure 2. Health evaluation index system of the AR ecosystem in Qinhuangdao city.

4.2. Construction of BNs Structure

According to the hierarchical structure between nodes shown in Figure 2, the BN topology for evaluating the health of AR ecosystem is constructed (Figure 3) (take the situation of the restored area in May as an example).

4.3. Data Processing

The mean, variance, and weight of each index are determined to generate NPTs of variables, and the measured data are classified as the prior knowledge of the input model. The mean and variance of standardized station data are shown in Table 1. C17 and C18 (number and density of eggs, larvae, and juveniles) data in September were 0, so we manually set the probability of 'very low' in the node probability Table to 1, and the probability of other grades was 0.

The interval [0,1] was divided into five grades with intervals of 0.2. The positive index (C9–C18) was divided into grades from 'very low' to 'very high' and the reverse index (C1–C8) was the opposite.

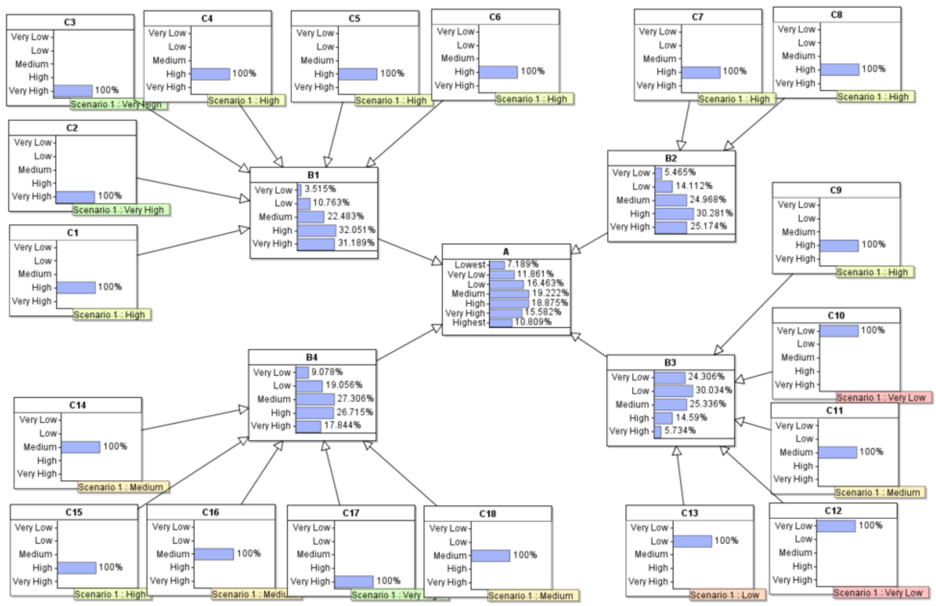

Figure 3. BN model built in AgenaRisk for assessing the AR ecosystem health.

Table 1. Processed data and divided grades.

Indicators	Weight (%)	May				September			
		Mean	Variance	Grade of restored area	Grade of control area	Mean	Variance	Grade of restored area	Grade of control area
C1	27	0.400	0.115	H	M	0.417	0.097	M	L
C2	15	0.334	0.102	VH	M	0.426	0.100	VH	H
C3	27	0.375	0.172	VH	L	0.330	0.083	H	M
C4	13	0.448	0.101	H	L	0.640	0.073	L	VL
C5	9	0.438	0.105	H	L	0.309	0.123	VH	L
C6	9	0.392	0.157	H	M	0.323	0.087	M	L
C7	51	0.388	0.124	H	H	0.215	0.040	H	M
C8	49	0.511	0.088	H	L	0.397	0.077	H	L
C9	16	0.408	0.091	H	L	0.597	0.101	VH	H
C10	17	0.433	0.136	VL	M	0.610	0.115	VH	VH
C11	17	0.342	0.110	M	VL	0.364	0.090	M	VL
C12	26	0.208	0.094	VL	VL	0.189	0.099	L	VL
C13	24	0.250	0.096	L	VL	0.478	0.113	L	VL
C14	24	0.542	0.109	M	M	0.500	0.083	M	L
C15	14	0.488	0.126	H	VL	0.499	0.152	VH	L
C16	20	0.304	0.091	M	VL	0.234	0.103	M	VL
C17	15	0.500	0.250	VH	VL	-	-	VL	VL
C18	27	0.340	0.084	M	VL	-	-	VL	VL

VL: very low; L: low; M: medium; H: high; VH: very high.

5. Results and Discussion

Using BNs, the mean (or the function used to calculate the mean) and variance were input into the AgenaRisk software (Agena) to generate the prior probability distribution, and then the level of the root node was input as the observation data to complete the probability calculation. By comparing the probability of each AR ecosystem health status of the restored area and control area in Qinhuangdao, we found that that the ecological status of the restored area was better than that of the control area after restoration, and the ARs had some restoration effect on ecology.

5.1. Assessment of the AR Ecological Restoration Effect

Taking the data of the restored area in May, for example, the probability distribution of each node in the restored area was obtained by inputting the pre-calculated root node level as the observation value into the software AgenaRisk (Figure 3). According to the maximum probability principle, the ecosystem health grade of the restored area in May was ‘middle’; likewise, the grade of node B1 was ‘high’.

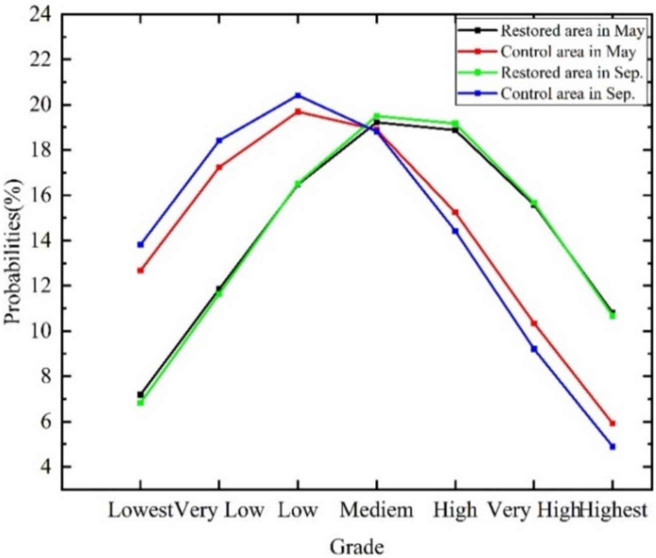

Figure 4. Probability distribution of ecosystem health at all levels.

The probability distributions of the target nodes in the restored and control area in May and September were counted and are shown in Figure 4, and the corresponding criterion level results are shown in Table 2. In May, the ecosystem health rating of the restored area was ‘medium’, and that of the control area was ‘low’; in September, the ecosystem health rating of the restored area was ‘medium,’ and that of the control area was ‘low’. The overall ecosystem health status of the restored area was better than that of the control area. Therefore, we argue that the AR in May and September had a restoration effect on the ecological environment of Qinhuangdao. Moreover, the water environment, sediment environment, and fishery resources were repaired.

The results of this study are somewhat different from those of Xu. The remediation effect of sediments in May and September was not obvious, but the results of this study show that there was some remediation effect. The biotic environment remediation effect was obvious in May, while our results show that both the restored and control area had the ‘low’ grade, revealing that the repairing effect was not obvious. Xu assigned value to each indicators through differential significance analysis and calculated the comprehensive index value to assess the restoration effect, which was subjective.

Different assignment gradients also affect the final results. This paper defines variables as hierarchical nodes, and the difference in the hierarchical led to the change of the evaluation level, which led to a difference in the evaluation results. The results of this study are consistent with those of Xu's study on water environment and fishery resources. Xu's results show that the repairing effect in May and September was obvious and slight, respectively, which is consistent with the results of this paper.

Table 2. Results of ecosystem health assessment in the criterion layer.

Grade	May		September	
	Grade of restored area	Grade of control area	Grade of restored area	Grade of control area
Water environment (B1)	High	Medium	High	Low
Sediment environment (B2)	High	Medium	High	Medium
Biotic environment (B3)	Low	Low	Medium	Low
Fishery resources (B4)	Medium	Very Low	Low	Very Low

5.2. Identification of Sensitive Factors

Based on the model constructed, it was assumed that the ecosystem health is in a good state, and we set the observation value of the target node as 'Highest', adversely reasoning the probability distribution of each root node according to Equation (1). Moreover, the sensitivity of each root node was calculated according to Equation (5) (Figure 5). It can be seen that the indicators C1, C7, C12, and C18 in B1, B2, B3, and B4, respectively, had a high level of sensitivity. Similarly, we set the observation value of the target node as 'High', and the result shows that C1 (biochemical oxygen demand), C7 (organic carbon), C12 (benthos biomass), and C18 (density of fish eggs and larvae) had the highest influence. Similarly, analysis of the sensitive node in the case of September showed that C3 (active phosphate), C7 (organic carbon), C12 (benthos biomass), and C16 (swimming animal density) were the more sensitive factors.

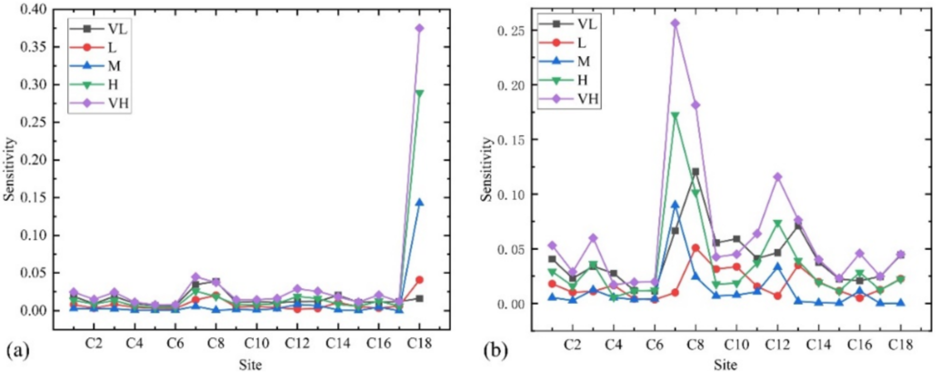

Figure 5. BN model built in AgenaRisk for assessing AR ecosystem health.

When conducting surveys, the value of sensitive factors should be carefully measured to improve the accuracy of evaluation results; in future restoration work, the content of labile phosphate in the water environment should be strictly controlled, an appropriate-biochemical environmental-condition is required, and the biomass of benthic animals and the biological density of swimming animals should be increased, which will all help improve the AR ecological restoration effect.

6. Conclusion

Procedures for assessing AR ecological restoration effect were proposed in the present study. As a case study of the Qinhuangdao restoration project, a set of indicators involving the water environment, sediment environment, biological environment, and fishery resources was presented and applied to the Qinhuangdao ARs ecological restoration effect assessment. The assessment results show that the restoration project had a comprehensive repair effect on the ecology. Sensitivity analysis revealed that organic carbon and benthos biomass were important variables determining ecological health. The results indicate that BN can intuitively express the relationship between variables, and it is feasible to evaluate the effect of AR restoration based on the BN. Moreover, the proposed method can be used in other types of comprehensive ecological restoration evaluation.

References

[1] Jouffray JB, Blasiak R, Norström AV, Österblom H, Nyström M. The blue acceleration: The trajectory of human expansion into the ocean. One Earth. 2020;2(1):43-54.
[2] Li YQ and Tang XX. Marine restoration ecology. Qingdao: China Ocean University Press;2016.153 p.
[3] Lima JS, Zalmon IR, Love M. Overview and trends of ecological and socioeconomic research on artificial reefs. Marine Environmental Research. 2019;145:81-96.
[4] Pitcher TJ, Buchary EA, Hutton T. Forecasting the benefits of no-take human-made reefs using spatial ecosystem simulation. ICES Journal of Marine Science. 2002;59(suppl):S17-S26.
[5] Zhang R, Liu H, Zhang Q, Zhang H, Zhao J. Trophic interactions of reef-associated predatory fishes (Hexagrammos otakii and Sebastes schlegelii) in natural and artificial reefs along the coast of North Yellow Sea, China. Science of The Total Environment. 2021;791:148250.
[6] Lima JS, Sanchez-Jerez P, dos Santos LN, Zalmon IR. Could artificial reefs increase access to estuarine fishery resources? Insights from a long-term assessment. Estuarine, Coastal and Shelf Science. 2020;242:106858.
[7] Hammond M, Bond T, Prince J, Hovey RK, McLean DL. An assessment of change to fish and benthic communities following installation of an artificial reef. Regional Studies in Marine Science. 2020;39:101408.
[8] Han YY, Yin R, Sun GQ, Yu, QH, Liu, YF, He ZW and Gong CG. Effect evaluation of artificial reef construction in Beidaihe National Marine Ranch Demonstration Zone. Hebei Fisheries. 2020(04):24-7+57.
[9] Liu SB, Wang ZH, Lin LW, Yang PH and Zhang S Y. Preliminary evaluation on effects of artificial reef project in Shengsi. Journal of Shanghai Ocean University. 2007;16(3):297-302.
[10] Tong F, Zhang XM, Wu ZX and Zhang PD. The ecosystem health assessment of artificial reef areas in Lidao Island, Rongcheng. Periodical of Ocean University of China. 2014;44(04):29-36.
[11] Zhao RR, Gong PH, Zhang Y, Yuan W, Guan CT, Shi MM, Xie ZH and Li J. Ecosystem health assessment of artificial reef area in Long Island. Progress in Fishery Sciences. 2019;40(06):9-17.
[12] Fu XM, Tang JY, Wu WQ and Zhang S. Evaluation of ecological restoration performance in Haizhou Bay, Lianyungang. Journal of Dalian Fisheries University. 2017;32(01):93-8.
[13] Wang Y, Hou CQ, Wang XY, Zhang BL, Chen W and Xu XF. Preliminary evaluation on proliferation effects of artificial reefs in Live Oyster Reef Area, Tianjin, China. Transactions of Oceanology and Limnology. 2018(02):137-45.
[14] Yao XW, Xu KL, Yan F and Tang GC. Research on the risk evaluation model for composite techniques and its field application. Journal of Northeastern University (Natural Science). 2015;36(5):737-42.
[15] Chen J, Wu YW, Li HZ, Qian XM and Yuan MQ. Risk analysis of burning and explosion of gas pipeline network based on dynamic Bayesian network. Transactions of Beijing Institute of Technology. 2021;41(7):696-705.
[16] Li XW, Zhao JM, Liu H, Zhang H and Hou XY. Status, Problems and optimized management of spawning, feeding, overwintering grounds and migration route of marine fishery resources in Bohai Sea and Yellow Sea. Transactions of Oceanology and Limnology. 2018(05):147-57.

[17] Xu YF, Li YQ, Zhang HP, Gao WB, Wang S Z, Hu B C and Liu S J. evaluation index system and case application of ecological restoration effect in Artificial Fish (Algae) Reef Area. Hebei Fisheries. 2017;(01):42-50.

[18] Pearl J. Bayesian Networks: A model of self-activated memory for evidential reasoning. Proc of Cognitive Science Society (CSS-7). Pearl_CSS851985.

[19] Forio MAE, Landuyt D, Bennetsen E, Lock K, Nguyen THT, Ambarita MND, et al. Bayesian belief network models to analyse and predict ecological water quality in rivers. Ecological Modelling. 2015;312:222-38.

[20] Havron A, Goldfinger C, Henkel S, Marcot BG, Romsos C, Gilbane L. Mapping marine habitat suitability and uncertainty of Bayesian networks: a case study using Pacific benthic macrofauna. Ecosphere. 2017;8(7):e01859.

[21] Mo DY. A Bayesian network model and its application on ecological environment vulnerability assessment. Beijing: The University of Chinese Academy of Sciences. 2017.

[22] Fenton NE, Neil M. Risk assessment and decision analysis with Bayesian network. New York: CRC Press; 2012.p.314-68

[23] Fenton, N. E., Neil, M. Ranked nodes: A simple and effective way to model qualitative judgements in large-scale Bayesian networks. London; 2005.

[24] Wu XG, Li TJ, Lin JY, Ma J, Zhang LM and Liu WL. safety risk assessment of metro shield tunnel-induced adjacent bridge damage based on rough set and Bayesian network. Journal of Civil Engineering and Management. 2016;33(03):9-15+29.