Determining Dark Diversity of Different Faunal Groups in Indian Estuarine Ecosystem: A New Approach with Computational Biodiversity

Moumita Ghosh[1], Anirban Roy[2], and Kartick Chandra Mondal[1(✉)]

[1] Jadavpur University, Kolkata, India
[2] West Bengal Biodiversity Board, Kolkata, India

Abstract. Computational Biodiversity can broadly be understood as the effort of computational approaches for exploring, interpreting, and analyzing biodiversity data. An enormous load of growing biodiversity data needs algorithmic care for accurate data management, and therefore the term computational biodiversity comes. Instead of relying purely on presence data, the probabilistic forecast of member distribution including the regions of not occurrence can neutralize biodiversity loss by restoring potential ecosystems. This paper is aiming at revealing the perspective of computational biodiversity as a counteract for biodiversity loss by correlating the concept of dark diversity. The computation of the dark diversity is accompanied by a data mining algorithm for establishing rules with more nobility to manage the depletion of biodiversity. We generate a dataset for the Indian estuarine ecosystem and show the use of our approach by ending up with rules worthwhile for the ecologists. These would step up reinforcing biological diversity via introducing or rehabilitating specific faunal groups to an estuary under survey.

Keywords: Ecosystem · Computational biodiversity · Dark diversity · Data mining · Indian estuarine ecosystem

1 Introduction

Background Study. Biodiversity loss is a global threat [4,21] which implies local loss or reduction of members in an individual habitat. To prevent this loss, endless studies and experiments on behalf of ecological researchers and conservationists are going on and different paths are followed. Ecological theory has plenty of research articles on measuring biological diversity [14,17] within species, between species as well as between ecosystems. Hypothesis establishment for species richness, abundance, and modeling distribution patterns are attempted numerous times over a range of datasets [2,20]. The influence of regional species richness on local species, co-existence statistics, diversification factors, etc. are

attributed by multiple authors [8,9]. Species biodiversity monitoring [7,22] has been studying for assessing the changes that occur over time and implementing suitable measures for different management strategies. Biodiversity hotspot conservation [18,23] is treated as the most effective way for compensation towards the loss. All of these research articles are directed at quantifying the gathered biodiversity information of the recorded individuals, or we can say all of these are dealing with the presence data of the individuals.

Motivation Towards Dark Diversity. The concern for the missing part of biodiversity data raises a few more questions like whether a particular site is capable of providing shelter for more species or is it possible to expand diversity as well as the richness of species. But absent data could not be estimated merely from survey data. It indicates all the absent members from the total species pool that occur in a particular geographical region having similar environmental conditions [25]. All these absent species are considered to have the potentiality for being present and expand diversity. These locally absent species from the regional species pool are termed as dark diversity [6,12]. Quantifying dark diversity is proved to be [11] an additional information source for restoration ecology.

Contribution. In this paper we have dealt with a dataset of Indian Estuarine Biodiversity consisting of faunal data [5]. It contains species count for each faunal group at different estuaries. It can be noted that multiple estuaries have not even reported any member count from the respective faunal groups. Our objective in this study is to employ the computation of dark diversity upon the dataset before processing rule mining tasks so that all possible likelihood of faunal presence could be studied. These rules would generate the potential habitat for the faunal groups. In this regard, we would adapt a rule mining task [16] which is solely based upon biclustering [13] and association rule mining [24].

Different Parts of the Paper. The whole paper is subdivided into multiple sections. Section 2 illustrates the dataset we have used and the step-by-step explanation of our followed procedure. Our observation is portrayed in Sect. 3 in detail. Section 4 concludes the paper.

2   Data and Methodology

Indian estuarine data of faunal groups is outlined on page 2, Fig. 1 of the book: "Indian Estuarine Biodiversity" [5]. This data is manually curated and detailed in Table 1 in our paper. It is referring to the number of species counted for each faunal group and stored estuary-wise.
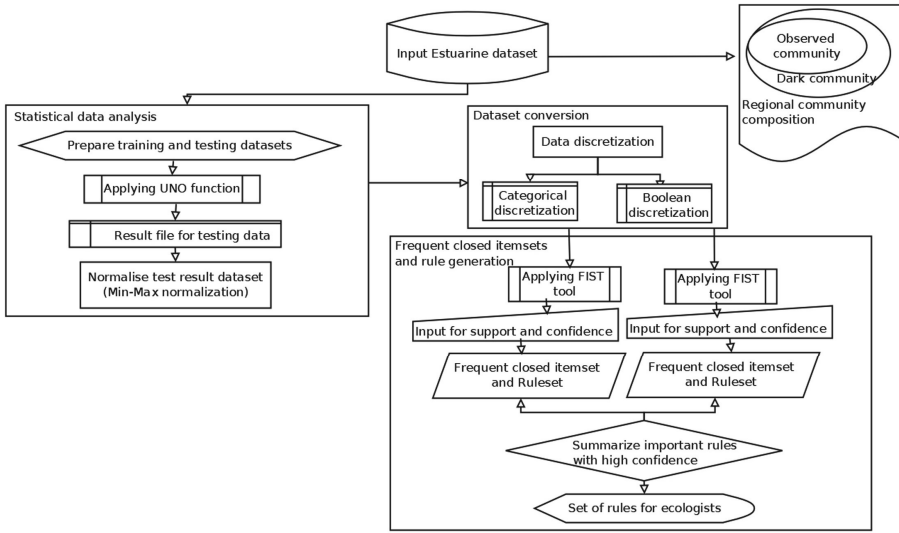
Fig. 1. Flowchart for the presented approach

In this paper, the role of dark diversity may be thought of as the task to forecast the occurrence of a faunal group based on the composition pattern of the community. The composition pattern relates to the different faunal groups in a community. It may appear that some specific groups are not detected at all in some particular regions. Using the calibration dataset, the researchers have shown a unique approach (UNO) [3] for predicting probable occurrences in those regions. The most widely used Beals function [1] is already compared to the UNO approach and UNO proves its better efficiency. Therefore, we follow this approach for predicting the dark diversity of our dataset. UNO function uses correspondence analysis for the ordination of species samples. In this regard, ordination [19] is the task of discovering a gradient for the presence/absence of samples along the geographical region. Primarily, it follows an exploratory analysis technique to reveal the distribution pattern of species. Correspondence analysis in the ordination method uses the weighted average of the data where the species occur and follows one scaling technique among some given choices [10]. UNO function provides 3 different choices and based on applicability, users specify anyone from those. The whole procedure is shown in the flowchart below in Fig. 1.

Table 1. Faunal data at Indian Estuaries:- *E1:Hooghly-Matla, E2: Subarnarekha, E3: Baitarani-Brahmani, E4: Mahanadi, E5: Rushikulya, E6: Bahuda, E7: Vamsadhara, E8: Nagavali, E9: Godavari, E10: Krishna, E11: Penner, E12: Ennore, E13: Adyar, E14: Veller, E15: Cauveri, E16: Cochin, E17: Zuari, E18: Mandovi, E19: Tapi, E20: Narmada*

Taxonomic group	Major estuaries of India																			
	E1	E2	E3	E4	E5	E6	E7	E8	E9	E10	E11	E12	E13	E14	E15	E16	E17	E18	E19	E20
Faunal groups																				
Protozoa					25	26				20	3			23					4	3
Foraminifera				5					47	11						73	14			
Porifera	1							2								2	2			
Cnidaria	24	12		11	20	5		13	3	10						34	3	3		
Ctenophora	1			1	2					1	1				1					
Rotofera	5								14		2	16		13						
Nematoda	2			11													20			
Acanthocephala									1											
Sipuncula	1	1																		
Mollusca	83	49	19	152	47		28	43	73	103	82	10		11	51	26	40	41	30	32
Annelida	91	37	11	34	19		13	4	70	45				24	48	47	70	3		
Arthropoda	377	53	88	45	159	99	24	17	88	118	125	56	58	35	55	167	72	21	25	60
Bryozoa	4																			
Brachiopoda	1	2							1											
Chaetognatha	4			3	6	2			1	2	3					4	6	6		3
Echinodermata	22	6	1					7	2	1										
Hemichordata	1																			
Urochordata				3	6	4				3						1				
Pisces	314	146	157	177	45	91	64	71	307	268	63	17	135	82	135	126	73	44	64	49
Amphibia	13	3	14						4		3					27				
Reptilia	57	5	45	6	2	1	1	1	1	10		1				4	7	7		
Aves	156	108	269	46	1		52	75	17							45	43	150	23	23
Mammalia	41	2	27	4					1	11		8			2	5				
Total	1198	424	631	478	352	228	130	188	652	630	320	118	193	188	291	560	336	291	146	170

2.1 Applying UNO Function

We use 75% of the dataset as training data. Then we predict the probable occurrences of not-found classes at multiple estuaries for the remaining 25% of the dataset. UNO method uses any of the 3 different methods (minobs(), minpred(), binminpred()) to calculate threshold value for predicting abundances. After having tested with all the 3 different methods, it has been found that minobs() and minpreds() are suitable for our case. minpred() uses the smallest predicted value for cases where a class observed and minobs() uses the smallest positive value observed for a class. Data is first binarized in the case of binminpred(). Results obtained from the UNO function will be preprocessed.

It has been observed that minpred() and minobs() generate quite similar result and they scaled the data in proportion. To justify this observation, a graphical view is generated which shows the rescaled result of the first four classes in Fig. 2. For each pair of values, the first one is from the observation of minpred().
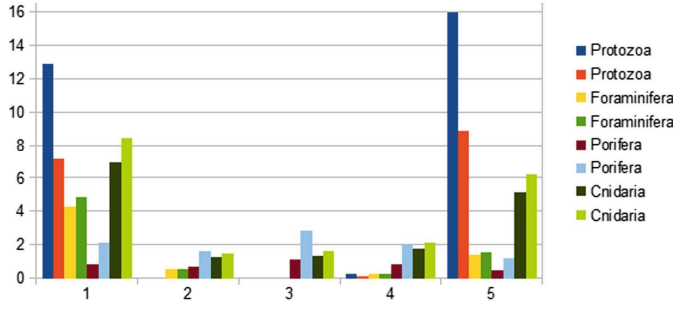
Fig. 2. Result comparison for minpred()(first one from each pair) and minobs() for the first 4 classes: X Axis shows 5 estuaries; Y axis shows probability of occurrence

From the achieved resultset, multiple sites with the potential presence of currently absent members could be found. Comparing with the initial dataset, one can identify the probable occurrence at different locations. All absent sites are not likely for each group under consideration. Like, in the case of Protozoa, it can be seen from the original dataset (Table 1) that this class is not reported from estuaries E1, E2, E3, and E4. But the application of the UNO function identifies E1 as most likely to be expected and E4 as likely to be expected (labeled as discretization referred to in Table 2). Considering Cnidaria, their original occurrence data and the data after computing dark diversity are displayed in Fig. 3, where E3 is turned up to have the probability of appearance.

2.2   Normalization and Binarization of the Result Dataset

We adopt the dataset obtained by applying minobs() function. Here, we feel the necessity of bringing down all the attributes on the same scale fluctuating between 0 and 1. Thus min-max normalization is performed on the result of UNO Function. We discretize all the obtained values and apply textual leveling to make the dataset comprehensible. The obtained result is presented in Table 2.
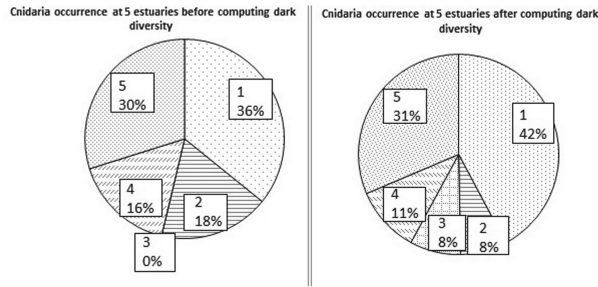

Fig. 3. Result comparison between before and after computing dark diversity

Table 2. Normalized and discretized dataset; discretization follows: 1 = Maximum frequency (MP); $1 > X > 0.60$ = Most likely to occur (ML); $0.60 > X > 0.30$ = Less likely to occur (LL); $0.30 > x > 0.01$ = Least likely to occur (LTL); 0 = Not probable (NP)

Class	Min-max normalization					Categorical discretization					Binarization				
	E1	E2	E3	E4	E5	E1	E2	E3	E4	E5	E1	E2	E3	E4	E5
Protozoa	0.81	0	0	0.02	1	ML	NP	NP	LTL	MP	1	0	0	1	1
Foraminifera	1	0.12	0	0.06	0.33	MP	LTL	NP	LTL	LL	1	1	0	1	1
Porifera	0.59	0.29	1	0.53	0	LL	LTL	MP	LL	NP	1	1	1	1	0
Cnidaria	1	0	0.03	0.1	0.69	MP	NP	LTL	LTL	ML	1	0	1	1	1
Ctenophora	0.98	0	0.02	0	1	ML	NP	LTL	NP	MP	1	0	1	0	1
Rotofera	1	0.02	0.27	0	0.07	MP	LTL	LTL	NP	LTL	1	1	1	0	1
Nematoda	0.14	0.11	0	0.46	1	LTL	LTL	NP	LL	MP	1	1	0	1	1
Acanthocephala	1	0.7	0.31	1	0	MP	ML	LL	MP	NP	1	1	1	1	0
Sipuncula	1	0.43	0.95	0.25	0	MP	LL	ML	LTL	NP	1	1	1	1	0
Mollusca	0.81	0.42	0	1	0.34	ML	LL	NP	MP	LL	1	1	0	1	1
Annelida	1	0.33	0	0.8	0.17	MP	LL	NP	ML	LTL	1	1	0	1	1
Arthopoda	1	0	0.17	0.01	0.41	MP	NP	LTL	LTL	LL	1	0	1	0	1
Bryozoa	1	0.24	0.74	0	0.16	MP	LTL	ML	NP	LTL	1	1	1	0	1
Brachiopoda	1	0.53	0.81	0.53	0	MP	LL	ML	LL	NP	1	1	1	1	0
Chaetognatha	0.96	0	0.32	0.1	1	ML	NP	LTL	LTL	MP	1	0	1	1	1
Echinodermata	1	0.33	0.74	0.19	0	MP	LL	ML	LTL	NP	1	1	1	1	0
Hemichordata	1	0.24	0.74	0	0.16	MP	LTL	ML	NP	LTL	1	1	1	0	1
Urochordata	0.46	0	0	0.19	1	LL	NP	NP	LTL	MP	1	0	0	1	1
Class Pisces	1	0.36	0.32	0.53	0	MP	LL	LL	LL	NP	1	1	1	1	0
Class Amphibia	1	0.15	0.7	0	0.06	MP	LTL	ML	NP	LTL	1	1	1	0	1
Class Reptilia	0.92	0.28	1	0	0.03	ML	LTL	MP	NP	LTL	1	1	1	0	1
Class Aves	0.76	0.37	1	0.18	0	ML	LL	MP	LTL	NP	1	1	1	1	0
Class Mammalia	1	0.21	0.85	0	0.07	MP	LTL	ML	NP	LTL	1	1	1	0	1

This table is also showing the binarized data based on the discretization (1 is replacing all other categorical values except NP, while NP is replaced with 0).

2.3   Applying Data Mining Methodology

A combined approach of biclustering and association rule mining [15,16] is applied to both the categorical discretized data and corresponding binarized data (Table 2). The result is interpreted in detail in the following section.

3   Result Analysis

In this section, we will assess the consequence of measuring dark diversity before the rule mining task. Here we justify the usage of the dark diversity function as it replaces most of the null entries, followed by extracting additional information.

To accomplish this, we attempt to mine twice; on the binarized dataset obtained corresponding to the initial table and the datasets after applying the
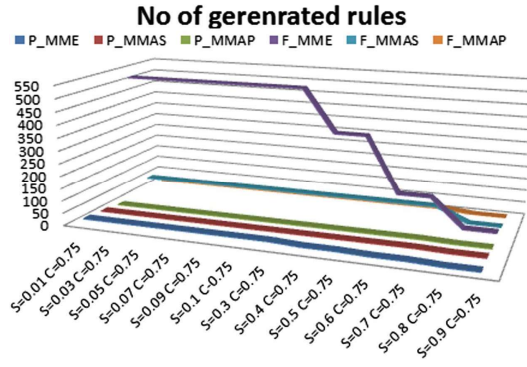
Fig. 4. Study on the no of rules generated in previous case and following case of incorporating dark diversity function; x axis: Threshold of support and confidence ; y axis: no of rules

dark diversity function (Table 2). We carry out a comparison of the data generated in both cases. It is found that they differ significantly. Figure 4 is determining the difference in the number of rules and confirms that larger rules are brought about where the dark diversity function is used. Here, we look at the scenario for the 5 estuaries that are taken as testing data. In Fig. 4, P_MME, P_MMAS, and P_MMAP denote the preceding scenario, i.e. the rules generated without using the dark diversity function. F_MME, F_MMAS, and F_MMAP denote the rules generated by exploiting the dark diversity function. Here, MME corresponds to exact rules with support value 1. MMAS and MMAP both conform to approximate rules where support value is always greater than zero, asserting the possibility of appearing.

Table 3. FCI obtained from the original and derived dataset

ClosedSet	Support	Object list
Result obtained from original binarized dataset before applying dark diversity function		
FCI1: [E1 = 1]	19	3, 4, 5, 6, 7, 9, 10, 11, 12, 13, 14, 15, 16, 17, 19, 20, 21, 22, 23
Result obtained from derived dataset after applying dark diversity function		
FCI2: [E1 = MP]	14	2, 4, 6, 8, 9, 11, 12, 13, 14, 16, 17, 19, 20, 23
FCI3: [E1 = ML]	6	1, 5, 10, 15, 21, 22
FCI4: [E1 = LL]	2	3, 18

Knowledge retrieved by employing computational biodiversity: specifically frequent closed itemsets (FCI) and approximate rules: FCIs are able to find a large set of items that are taking place more times than user-specified support values in a dataset. Hence, FCIs are capable to derive useful knowledge. Also, exact rules are derived from the facts that are demonstrated in the dataset, whereas approximate rules are dealing with the probability of occurring. Below, in this

section, we are moving to discuss the information that we may derive from the FCIs and rules.

– Table 3 is showing the resultset consisting of FCIs, before and after applying the dark diversity function. It is explicit from the available dataset that 19 members are identified to have habitat at Hooghly-Matla (E1) estuary. But the adaptation of the dark diversity function mentions 22 total members that may occur. Thus, information related to 6 more faunal data is obtained having the occurrence data categorized as maximum frequency, most likely to occur, and least likely to occur. The illustration is given below,
   • Performing FCI2 - FCI1: 2, 8; It indicates that Foraminifera and Acanthocephala are the maximum probable species missing at E1.
   • Performing FCI3 - FCI1: 1; It shows that Protozoa is the most likely faunal group that should exist at E1.
   • Performing FCI4 - FCI1: 3; It suggests that Porifera may have the chance of occurrence, but it is less likely to occur.
   The diversity of Foraminifera is inadequate. It is less available as its existence is reported merely from 5 estuaries (Rushikulya, Krishna, Penner, Cochin, and Mandovi). Being highly sensitive to physicochemical characteristics, this group can be chosen as an indicator of oceanic and climatic information. On that account, the resulting rule serves as noteworthy information from the ecological perspective. Acanthocephala is reported from only Godavari estuary. As Hooghly-Malta is a highly diverse estuary, it may be the apparent site next to Godavari estuary for this parasite. For a similar reason, Protozoa and Porifera can also be expected at estuary Hooghly-Malta.
– Findings based on presence-only data are noted in Table 4, where the antecedent part is the same for all the cases. Comparing the consequent and object list, we can draw the following conclusions:

Table 4. Rules (Homogeneous antecedent) from binarized derived dataset

Rule	Antecedent	Consequent	Sup.	Conf.	Estuary list
R1	[S6 = 1, S13 = 1, S17 = 1, S20 = 1, S21 = 1, S23 = 1]	[S5 = 1, S12 = 1, S4 = 1, S15 = 1]	3	0.75	[1, 3, 5]
R2	[S6 = 1, S13 = 1, S17 = 1, S20 = 1, S21 = 1, S23 = 1]	[S2 = 1, S7 = 1, S10 = 1, S11 = 1]	3	0.75	[1, 2, 5]
R3	[S6 = 1, S13 = 1, S17 = 1, S20 = 1, S21 = 1, S23 = 1]	[S3 = 1, S8 = 1, S9 = 1, S14 = 1, S16 = 1, S19 = 1, S22 = 1]	3	0.75	[1, 2, 3]

   • Estuary list (R1) - Estuary list (R2) = 3; also Antecedent R1 is the same as the Antecedent of R2. It implies that estuary 3 has the probability of finding all the faunal groups numbered S2, S7, S10, and S11 with a confidence level of 75% those are in the consequent part of R2.

- Oppositely, Estuary list (R2) - Estuary list (R1) = 2, and the antecedents of R1 and R2 are matched. So, like before, it can be estimated that estuary 2 has the probability of sustaining all the faunal groups present in the consequent part of R1 i.e. the groups numbered S5, S12, S4, S15 with a confidence level of 75%.
  - Similarly, from rule R1 and R3, it can be stated that estuary 5 has the likelihood of occurrence for S3, S8, S9, S14, S16, S19, S22 with a confidence level of 75%.
  - Rule R2 and R3 could derive that the estuary 3 has the chance to exhibit the groups numbered S3, S8, S9, S14, S16, S19, S22 with a confidence level of 75%.
- Table 5 is displaying two rules where the object list is the same for all the cases. Comparing the antecedent and the consequent, we can claim that the antecedent and consequent parts are highly associated as they complement each other. These faunal groups are building a closed set with a 75% confidence level.
- An identical example is shown in Table 6. The estuary list is the same for all. It can be observed that antecedent and consequent together are forming a closed group. The occurrence of any class is linked to all others belonging to the same group with 75% confidence.

  It is known that ([5]) E1 (Hooghly-Matla) is a highly diverse estuary. Again, the estuaries E2 (Subarnarekha) and E3 (Baitarani-Brahmani); E4 (Mahanadi) and E5 (Rushikulya) are situated side by side and those are closely related regions. Therefore, they are expected to have similar physico-chemical properties. So, closely associated member lists are found which are justified in Tables 5, 6.

Table 5. Rules (similar estuary list) from binarized derived dataset

Rule	Antecedent	Consequent	Sup.	Conf.	Estuary list
R1	[S6 = 1, S13 = 1, S17 = 1, S20 = 1, S21 = 1, S23 = 1]	[S3 = 1, S8 = 1, S9 = 1, S14 = 1, S16 = 1, S19 = 1, S22 = 1]	3	0.75	[1, 2, 3]
R2	[S3 = 1, S8 = 1, S9 = 1, S14 = 1, S16 = 1, S19 = 1, S22 = 1]	[S6 = 1, S13 = 1, S17 = 1, S20 = 1, S21 = 1, S23 = 1]	3	0.75	[1, 2, 3]

Therefore, through this study, we have shown how data mining knowledge discovery can be associated with ecological research and discover new diversity patterns with proper justification. We have shown that with the help of the algorithmic solution, we can estimate the likelihood of being present in a locally absent faunal group and verified our findings from the standpoint of ecology. These data would certainly assist ecologists in practicing ecological restoration through habitat and range improvements for species under study. Contrarily it could also be stated that it nourishes the sustainability of an ecosystem under surveillance via assimilating new species.

Table 6. Rules from binarized derived dataset with similar object lists

Rule	Antecedent	Consequent	Sup.	Conf.	Estuary list
R1	[S4 = 1]	[S1 = 1, S18 = 1, S2 = 1, S7 = 1, S10 = 1, S11 = 1, S15 = 1]	3	0.75	[1, 4, 5]
R2	[S15 = 1]	[S1 = 1, S18 = 1, S2 = 1, S4 = 1, S7 = 1, S10 = 1, S11 = 1]	3	0.75	[1, 4, 5]
R3	[S2 = 1]	[S1 = 1, S18 = 1, S4 = 1, S7 = 1, S10 = 1, S11 = 1, S15 = 1]	3	0.75	[1, 4, 5]
R4	[S7 = 1]	[S1 = 1, S18 = 1, S2 = 1, S4 = 1, S10 = 1, S11 = 1, S15 = 1]	3	0.75	[1, 4, 5]
R5	[S10 = 1]	[S1 = 1, S18 = 1, S2 = 1, S4 = 1, S7 = 1, S11 = 1, S15 = 1]	3	0.75	[1, 4, 5]
R6	[S11 = 1]	[S1 = 1, S18 = 1, S2 = 1, S4 = 1, S7 = 1, S10 = 1, S15 = 1]	3	0.75	[1, 4, 5]

4　Conclusions

This study incorporates data mining along with statistics and directs us towards a competent solution for biodiversity restoration specific to a particular region of study. This paper has introduced the proposition of applying the dark diversity function to the presence-absence dataset before the process of rule mining. The reason behind this is to gain information related to the absent part of the occurrence data. The usefulness of deploying the dark diversity function is illustrated by visualizing the number of rules generated with and without applying the dark diversity function. It is understood that the underlying reason behind the greater number of rules is more non-zero values in the dataset. Our study helps in proper management in a survey or resurvey aiming at finding new sites for probable habitat for a particular faunal group. The generated results can suggest the likelihood of occurrence for specific faunal groups to a degraded estuary for the member introduction or restoration purpose.

In the future, we would like to concentrate on developing a more efficient algorithmic solution that needs hardly any human intervention in analyzing the result set. In addition to this, we would like to estimate the dark diversity on the littoral forest of Sundarban mangrove as declining mangrove cover is threatening for human existence. Thus, this kind of approach would be helpful for safeguarding the precious ecosystem.

References

1. Beals, E.W.: Bray-curtis ordination: an effective strategy for analysis of multivariate ecological data. In: Advances in Ecological Research, vol. 14, pp. 1–55. Elsevier (1984)

2. Booth, T.H., Nix, H.A., Busby, J.R., Hutchinson, M.F.: BIOCLIM: the first species distribution modelling package, its early applications and relevance to most current maxent studies. Divers. Distrib. **20**(1), 1–9 (2014)
3. Brown, J.J., et al.: A novel method to predict dark diversity using unconstrained ordination analysis. J. Veg. Sci. **30**, 610–619 (2019)
4. Cardinale, B.J., et al.: Biodiversity loss and its impact on humanity. Nature **486**(7401), 59 (2012)
5. Chandra, K., Raghunathan, C., Dash, S.: Current status of Estuarine Biodiversity in India: 1-575. the Director, Zool. Surv. India, Kolkata (2018)
6. De Bello, F., et al.: Measuring size and composition of species pools: a comparison of dark diversity estimates. Ecol. Evol. **6**(12), 4088–4101 (2016)
7. Dube, R.J.: Stand Structure and Diversity Analysis of Timber Tree Species at Lupane State University Paddock, Matabeleland North, Zimbabwe. Ph.D. thesis, Lupane State University (2018)
8. Harrison, S., Cornell, H.: Toward a better understanding of the regional causes of local community richness. Ecol. Lett. **11**(9), 969–979 (2008)
9. Kamruzzaman, M., Minhaj-Uj-Siraj, M., Ahmed, S., Osawa, A.: Regeneration status of mangrove species under mature stands in the oligohaline zone of the Sundarbans, Bangladesh. Reg. Stud. Mar. Sci **16**, 15–20 (2017)
10. Legendre, P., Gallagher, E.D.: Ecologically meaningful transformations for ordination of species data. Oecologia **129**(2), 271–280 (2001)
11. Lewis, R.J., et al.: Applying the dark diversity concept to nature conservation. Conserv. Biol. **31**(1), 40–47 (2017)
12. Lewis, R.J., Szava-Kovats, R., Pärtel, M.: Estimating dark diversity and species pools: an empirical assessment of two methods. Methods Ecol. Evol. **7**(1), 104–113 (2016)
13. Madeira, S.C., Oliveira, A.L.: Biclustering algorithms for biological data analysis: a survey. IEEE/ACM Trans. Comput. Biol. Bioinform. (TCBB) **1**(1), 24–45 (2004)
14. Magurran, A.E.: Biological diversity. Curr. Biol. **15**(4), R116–R118 (2005)
15. Mondal, K.C.: Algorithms for Data Mining and Bio-informatics. Ph.D. thesis, Université Nice Sophia Antipolis (2013)
16. Mondal, K.C., Pasquier, N., Mukhopadhyay, A., Maulik, U., Bandhopadyay, S.: A new approach for association rule mining and bi-clustering using formal concept analysis. In: International Workshop on Machine Learning and Data Mining in Pattern Recognition, pp. 86–101. Springer (2012)
17. Mondal, K.C., Pasquier, N.: Galois closure based association rule mining from biological data. In: Biological Knowledge Discovery Handbook, pp. 761–802 (2013)
18. Myers, N., Mittermeier, R.A., Mittermeier, C.G., Da Fonseca, G.A.B., Kent, J.: Biodiversity hotspots for conservation priorities. Nature **403**(6772), 853 (2000)
19. Okland, R.H.: Are ordination and constrained ordination alternative or complementary strategies in general ecological studies? J. Veg. Sci. **7**(2), 289–292 (1996)
20. Reiss, H., Cunze, S., König, K., Neumann, H., Kröncke, I.: Species distribution modelling of marine benthos: a north sea case study. Mar. Ecol. Prog. Ser. **442**, 71–86 (2011)
21. Roe, D., Seddon, N., Elliott, J.: Biodiversity loss is a development issue: a rapid review of evidence. Issue Pap. Int. Inst. Environ. Dev. (IIED) Lond. **798**, 678–683 (2019)
22. Sreelekshmi, S., Nandan, S.B., Kaimal, S.V., Radhakrishnan, C.K., Suresh, V.R.: Mangrove species diversity, stand structure and zonation pattern in relation to environmental factors–a case study at Sundarban delta, east coast of India. Reg. Stud. Mar. Sci. **35**, 101111 (2020)

23. Zachos, F.E., Habel, J.C.: Biodiversity hotspots: distribution and protection of conservation priority areas. Springer Science & Business Media (2011)
24. Zhang, C., Zhang, S.: Association Rule Mining: Models and Algorithms. Springer-Verlag (2002)
25. Zobel, M., Otto, R., Laanisto, L., Naranjo-Cigala, A., Pärtel, M., Fernández-Palacios, J.: The formation of species pools: historical habitat abundance affects current local diversity. Glob. Ecol. Biogeogr. **20**(2), 251–259 (2011)