World Academy of Science, Engineering and Technology
International Journal of Marine and Environmental Sciences
Vol:13, No:4, 2019

Application of a Generalized Additive Model to Reveal the Relations between the Density of Zooplankton with Other Variables in the West Daya Bay, China

Weiwen Li, Hao Huang, Chengmao You, Jianji Liao, Lei Wang, Lina An

***Abstract*—**Zooplankton are a central issue in the ecology which makes a great contribution to maintaining the balance of an ecosystem. It is critical in promoting the material cycle and energy flow within the ecosystems. A generalized additive model (GAM) was applied to analyze the relationships between the density (individuals per m³) of zooplankton and other variables in West Daya Bay. All data used in this analysis (the survey month, survey station (longitude and latitude), the depth of the water column, the superficial concentration of chlorophyll a, the benthonic concentration of chlorophyll a, the number of zooplankton species and the number of zooplankton species) were collected through monthly scientific surveys during January to December 2016. GLM model (generalized linear model) was used to choose the significant variables' impact on the density of zooplankton, and the GAM was employed to analyze the relationship between the density of zooplankton and the significant variables. The results showed that the density of zooplankton increased with an increase of the benthonic concentration of chlorophyll a, but decreased with a decrease in the depth of the water column. Both high numbers of zooplankton species and the overall total number of zooplankton individuals led to a higher density of zooplankton.

***Keywords*—**Density, generalized linear model, generalized additive model, the West Daya Bay, zooplankton.

I. Introduction

ZOOPLANKTON include a group of small animals which weekly swims, floats or drifts in water [1]. It plays an important role in the ecosystem which acts as a primary and secondary link in the food chain. Zooplankton provides food for most juveniles and small fishes [2], which shifts the primary products into a higher tropic level. Understanding the variables that control the abundance of zooplankton inhabiting the local bay is very important for assessing the fish community shift. This topic has been discussed for a long time in recent decades, and all the methods employed in the analysis were linear regressions, i.e. prey-driven control of predator assemblages of the zooplankton abundance were proved by Pintar and Resetarits [3], zooplankton growth and biomass was linked to organic matter concentrations, while weakly correlated with chlorophyll and biogenic silica in Rose Sea [4].

Linear regression such as GLM proved to be a powerful tool for linear analysis [5], and it gave us an insight into the significant relationship factors among the mass unclear factors. The factors were significantly related to the independent variables, but always had complicated non-linear relationships [6]. As we know, the marine ecosystem is complex with its species and abundance [7], and the knowledge of the physical and biological processes' impact on the abundance of zooplankton is critical to know from the ecosystem perspective. Therefore, a method which gives us a comprehensive way to deal with this problem is needed. GAM is a model which provides an objective way for the analysis among several factors related to the independent variable [7]. And it has been widely used in inhabited preference studies [8], and ecology studies [9].

The objective of this study is to obtain the density of zooplankton related to other variables including the survey months, survey station (longitude and latitude), the depth of the water column, the superficial concentration of chlorophyll a, the benthonic concentration of chlorophyll a, the number of zooplankton species and the number of zooplankton species collected from January to December 2016. The prediction of this relationship would be useful to assess the fish community and implementation of sustainable exploitation schemes of fish stocks in the West Daya Bay.

II. Materials and Methods

A. Survey Description

The geographical scientific survey area of the West Daya Bay is depicted in Fig .1. Totally 15 survey stations distributed in the West Daya Bay were exploited in this study. All the data applied in this study were from scientific survey, including the survey month, survey station (longitude and latitude), the depth of water column, the superficial concentration of chlorophyll a, the benthonic concentration of chlorophyll a, the number of zooplankton species and the abundance of total zooplankton.

Weiwen Li, Jianji Liao, Lei Wang, and Lina An are with the Laboratory of Marine Biology and Ecology, Third Institute of Oceanography, State Oceanic Administration, Xiamen, Fujian 361005, China.

Chengmao You is with the Operations Department, Daya Bay Nuclear Power Operations and Management Co., LTD., Shenzhen, Guangdong, 518124, China.

Hao Huang is with the Laboratory of Marine Biology and Ecology, Third Institute of Oceanography, State Oceanic Administration, Xiamen, China (e-mail: huanghao@tio.org.cn).

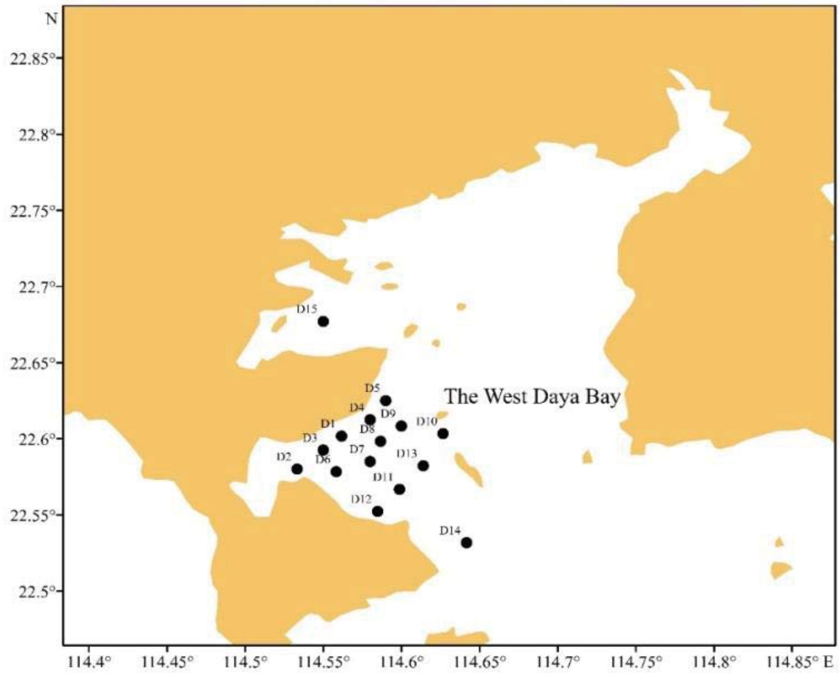

Fig. 1 Survey distribution in the West Daya Bay

III. Data Analysis

A. Chlorophyll a Analysis

Sampling and determination procedures were carried out in accordance with the marine survey code [10]. The chlorophyll a was determined by extraction fluorimetric method [11].

B. Zooplankton Analysis

Zooplankton samples were collected from January to December 2016 within the West Daya Bay. The zooplanktons were collected by the vertical nets (145 cm long in sock, 50 cm in mouth diameter, and 0.50 mm in mesh) from bottom to the surface. All the collected samples by the vertical nets were fixed with the 5% neutral formaldehyde solution until taken back to the laboratory for further analysis. Stereo microscope and optical microscope were used to identify the species, and the Counting Chamber for zooplankton was used to count the number of zooplanktons, then the normal density of zooplankton was calculated as ind/m^3.

C. GLM Analysis

GLM performs a response variable to a set of predicted variables through a link function [12]. A linear model which showed the response variable and the explanatory variables as:

$$\mathrm{f}(\mu_i) = x_i^T \times \beta,$$

where f represents the link function, μ_i is the predicted response, x_i^T is a transposed vector of explanatory variables, and β is a vector of parameter.

D. GAM Analysis

GAM is a non-parametric and semi-parametric regression method for exploring the response variable related to predicted variables, it is an extension of GLM [13]. There is no need to define the prior assumption on the functions which links the predicted variables and response variable. The influence factors upon the abundance of zooplankton are poorly known in the West Daya Bay. So, in this study, we choose the GAM to analysis the abundance of zooplankton and other significant related factors. So the GAM can be expressed as [14]:

$$\mathrm{g(y)} = \sum_{i=1}^{n} f_i\,(x_i) + \varepsilon\,,$$

where g is the link function, ε is random error term, f_i is a function for the $x_i$independent variable, while x_i is a vector of independent variables, $n$is number of vectors of independent variables. GAMs were fitted using the R statistical programming environment with package *mgcv* [15].

IV. Results

A. Independent Variables Selection

For the GLM, it requires the response variable display with normal or near normal distribution. We transferred the density of the zooplankton into log(log(log(density))) and the response variable log(log(log(density))) displays near normal distribution, as shown in the normal Q-Q plot (Fig. 2).

In this study, 9 independent variables including the survey month, survey station number, longitude, latitude, the depth of water column, the superficial concentration of chlorophyll a, the benthonic concentration of chlorophyll a, the number of zooplankton species and the total zooplankton were recorded which are all employed in the GLM for variables selection. The number of zooplankton species, total zooplankton and the survey month significance relates to the density of zooplankton, and the benthonic concentration of chlorophyll a and the depth of water column positive effect on the density of zooplankton (as shown in Table I).

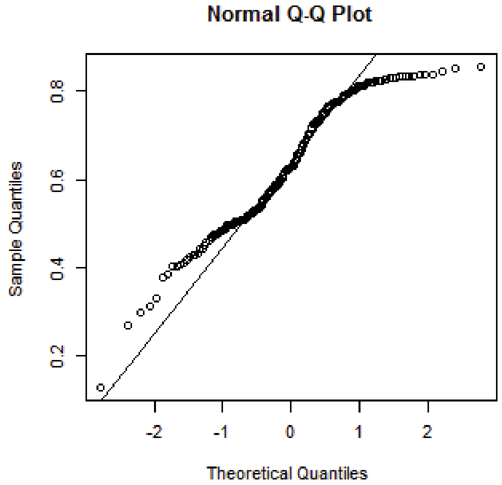

Fig. 2 Quantile-quantile plots of GLM

TABLE I
TEST OF THE SIGNIFICANT VARIABLES IN THE GLM FITTED TO THE LOG(LOG(LOG(DENSITY)))

	Standard Error	t Value	Pr(>|t|)	Note
Longitude	2.78E+10	0.002	0.99844	
Latitude	1.61E+11	-0.002	0.99844	
Superficial concentration of chlorophyll a	2.78E-04	-0.151	0.880264	
Benthonic concentration of chlorophyll a	9.27E-04	2.059	0.041275	*
The depth of water column	6.17E-03	-2.114	0.036211	*
Number of zooplankton species	1.22E-03	4.467	1.56E-05	***
Total zooplankton	6.05E-07	6.449	1.50E-09	***
as.factor(station)	1.55E+09	-0.002	0.99844	
as.factor(month)	0.002623	-0.67	<2e-16	***

B. GAM Analysis

After all the variables were selected by GLM, the number of zooplankton species, total zooplankton, the benthonic concentration of chlorophyll a, the depth of water column and the survey month were applied in the GAM. It can be expressed as:

$$\text{Gam}\ (\log(\log(\log(DEN))))\sim s(BCC)+s(DEP)+s(NUM)+s(SPE)+ \text{as.factor}(MON)+\varepsilon,$$

where DEN is the density of zooplankton, BCC is the benthonic concentration of chlorophyll a, DEP is the depth of water column, NUM is the total number of zooplankton capture, SEP is the number of zooplankton species capture and the MON is the survey month.

The solid lines represent the mean relative effect of the benthonic concentration of chlorophyll a and the depth of water column upon the density of zooplankton, and the dashed lines represents the margin of the error or uncertainty (Figs. 3 and 4). These uncertainties increase when there is little available data. So, the density of zooplankton increases with the benthonic concentration of chlorophyll a, but, when the concentration was larger than 20 mg/m^3 we obtained a larger uncertainty (Fig. 3). In general, the density of zooplankton decreased with the depth of water column (Fig. 4).

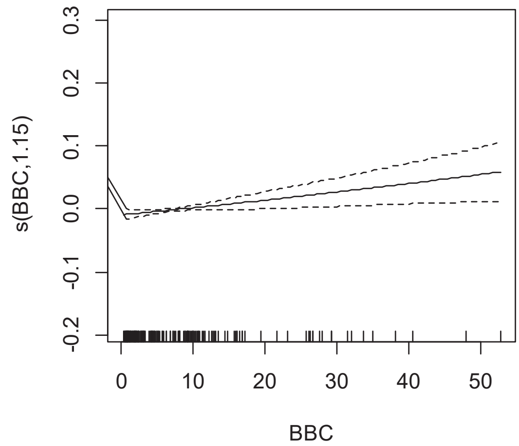

Fig. 3 Effects of the benthonic concentration of chlorophyll a on the density of zooplankton in the West Daya Bay

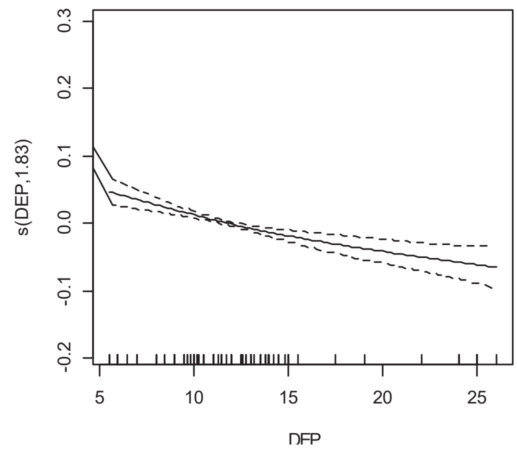

Fig. 4 Effects of the depth of water column on the density of zooplankton in the West Daya Bay

The density of zooplankton increased with the species (Fig. 5). When the species were more than 15, the density of zooplankton increased with species (Fig. 5). The density of zooplankton nearly linear relates to the number of total zooplankton captured, and roughly increased when the number total zooplankton was larger than 10000 individuals (Fig. 6).

From January to April, the density of zooplankton increased rapidly, then a sharp decrease from April to May (Fig. 7). During this time, the density of zooplankton among the 15 sampling locations varied significantly between each other. From June to December, the density of zooplankton was much lesser than that of January to May, but the density of zooplankton among the 15 sampling locations were much closed (Fig. 7).

C. Comparison of Normal and Standard Density of Zooplankton

From February to April, the standard density of zooplankton was larger than the normal density of zooplankton (Fig. 8). For most months, the normal density of zooplankton was larger than the standard density of zooplankton. But, normal density of the zooplankton and standard density plankton were much closed. Both from the normal density of zooplankton and standard density of zooplankton, April was the highest density among the survey year while October was the lowest (Fig. 8).

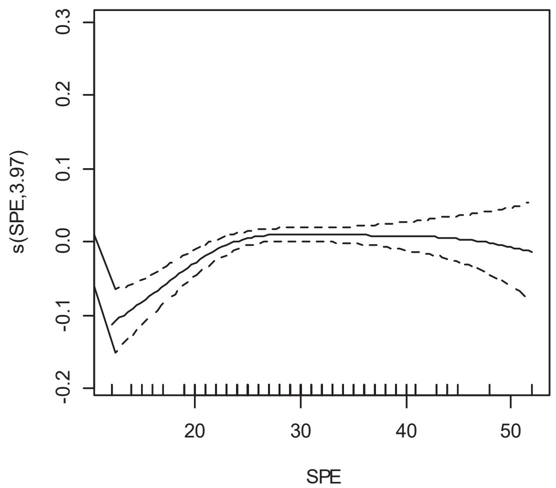

Fig. 5 Effects of the number of zooplankton species captured on the density of zooplankton in the West Daya Bay

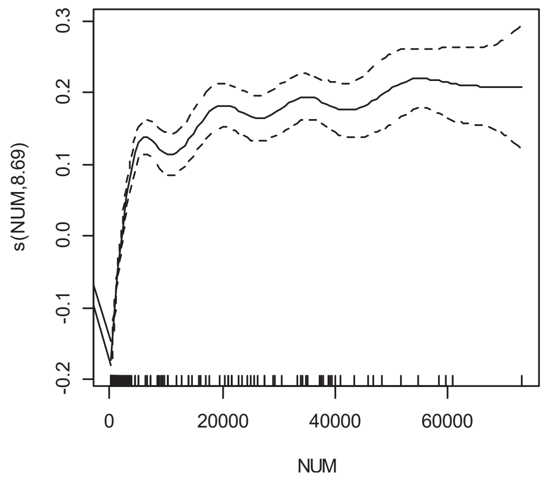

Fig. 6 Effects of the total number of zooplankton capture on the density of zooplankton in the West Daya Bay

V. Discussion

Understanding factors that control the density of zooplankton in its habitats are a key issue in the local ecology in the West Daya Bay. Changes in the abundance of zooplankton in the West Daya Bay have a direct effect on the primary productivity and fish stocks local habitat [16] which in turn is beneficial for assessing fish stock activities and managements in this area.

A. The Importance of Factors Selection

Due to lack of knowledge on the mechanism of factors driving the density of zooplankton, we always consider more factors involved in our surveys. Many large-scale factors involve trying to build interpretable models linking a large set of potential variables to a response variable in a linear or nonlinear function [17]. Not all the factors included in the models are effective, and more ineffective factors brings a low exploitation rate to the model, what is more, it gives a large uncertainty for the results.

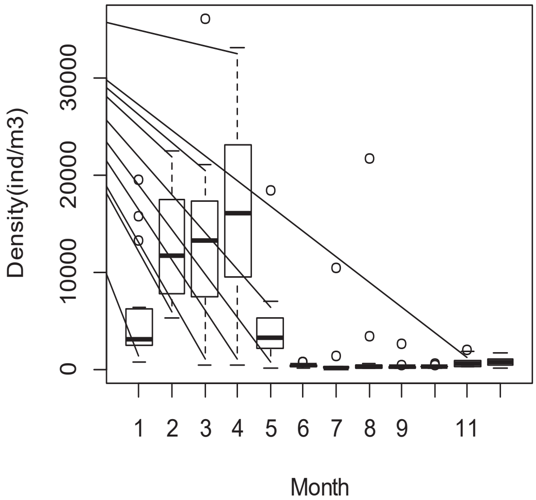

Fig. 7 Effects of the survey month on the density of zooplankton in the West Daya Bay

In order to deal with this problem, more and more methods were employed to select useful factors. Such as Akaike information criterion (AIC) [18], Bayesian information criterions (BIC) [19], Principal Component Analysis (PCA) [20], model-based clustering [21], and empirical likelihood method [22] have been successful applied into the variables selection. And currently, variable selection with p-values proved to be an easy way to identify the significance factors [23]. So, in this study, we used the GLM with significance p-values to choose variables for correlation analysis. As the results showed that the number of zooplankton species, total zooplankton, the benthonic concentration of chlorophyll a, the depth of water column and the survey month were significant relate to the density of zooplankton in the West Daya Bay.

B. The Effect of Factors

The density of zooplanktons distribution was associated with environmental factors [24]. The number of zooplankton species, total zooplankton, the benthonic concentration of chlorophyll a, the depth of water column and the survey month were significant factors associated with the density of zooplankton distribution selected by the GLM. GAM was employed to analyze the relationship between the density of zooplankton and these significant factors.

The density of zooplankton significance effectuated by the benthonic concentration of chlorophyll a (Table I), and a linear relationship was determined (Fig. 3), but, superficial concentration of chlorophyll a was not effective related to the density of zooplankton. A bottom-up regulation by the benthonic concentration of chlorophyll a was determined by the density of zooplankton in the West Daya Bay, a similar conclusion was found by the Voutilainen et al 2016 [24] when determining the effect of plankton on zooplankton. A higher number of zooplankton species and total zooplankton roughly contributed to a larger density of zooplankton, which was determined by the linear relationship (Figs. 5 and 6).

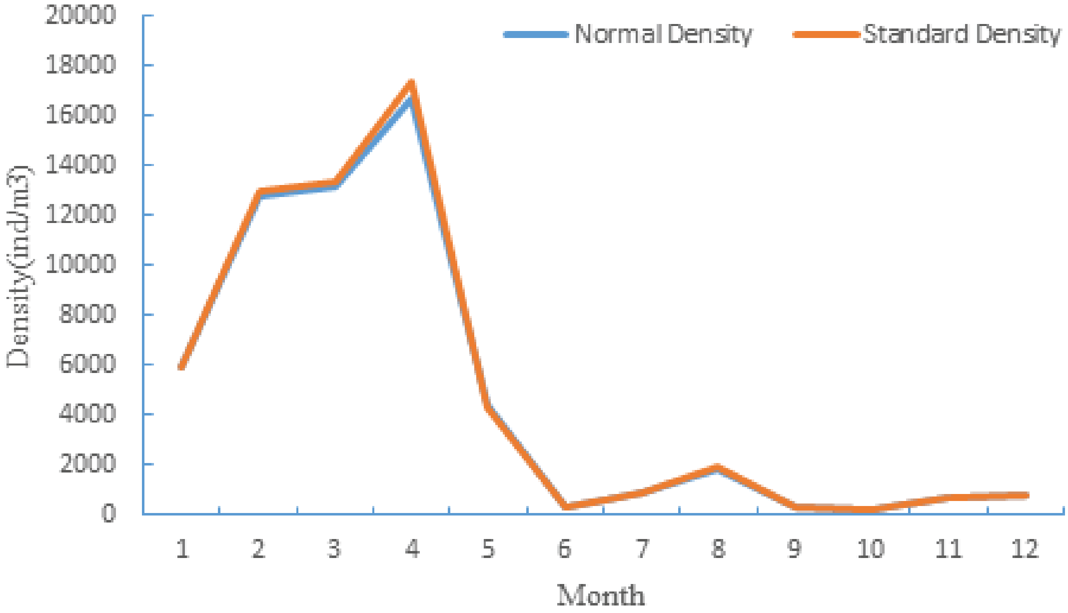

Fig. 8 Comparing standardized density and nominal density of zooplankton in different month

Light is the basic element for the chlorophyll a, and it attenuates in the water column [25], decreasing the concentration of chlorophyll a. In addition, the deeper the water column, the less concentration of oxygen it contains. Zooplankton are aerobic biological species, and its density decline with the decrease in oxygen concentration (Fig. 4).

Marine plankton density can be affected by the increase in temperature [26] and is the key food resources for the zooplankton. In general, highest density zooplankton are recorded in summer while the lowest is in winter [27], [28]. But this balance has been broken down by the human activities, such as seasonal closure. From May to September, four-month seasonal closure have been carried out in 2016 for fish stock recovering. Increasing the fish stock consumes much more zooplankton, and the density of zooplankton decreases during this time, thus the reason why the density of zooplankton seasonal distribution in this study is different from other marine studies (Fig. 7).

C. The Importance of Standardized Density of Zooplankton

Normal density of zooplankton always influences the biomass calculation, metered volume etc. conducted by different technicians. It is hard to determine the zooplankton stocks that provide a bias information on the zooplankton abundance [29], [30]. GAM has been useful tool for quantifying these biases, and predicting more reliable results. As the results produced by the GAM, shows that the normal density and standard density of the zooplankton are much closer to each other indicating that the technology applied with the concentration of chlorophyll a and the zooplankton is very skillful. The seasonal shift of the density of zooplankton provides a better understanding capacity for supporting the fish stock in the West Daya Bay.

ACKNOWLEDGEMENT

This work described in this paper was supported by the National Key Research and Development Program of China (No. 2018YFC0507205), the 973project (No. 2015CB452902) and the Marine Public Welfare Project of China (No. 201505009), the Natural Sciences Foundation of China (No. 40706042) and the scientific research foundation of Third Institute of Oceanography, State Oceanic Administration(No. HE 170704-17(1)).We would also like to thank Dr Jiaqi Wang from Shanghai Ocean University for mapping the sample distribution and the Dr Kindong Richard form Shanghai Ocean University for editing the language. The authors report no conflicts of interest. The authors alone are responsible for the content and writing of the article.

REFERENCES

[1] Vinobaba P. (2012). Impact of water quality on species composition and seasonal fluctuation of planktons of Batticaloa lagoon, Sri Lanka. Journal of Ecosystem &Ecography, 02(4)doi:10.4172/2157-7625.1000117.
[2] Yongo E, Outa N. (2017).Spatial distribution and abundance of zooplankton communities in Lake Victoria, Kenya. International Journal of Fisheries and Aquatic Research, 2(1):33-35.
[3] Pintar M R,Resetarits W J. (2017). Prey-driven control of predator assemblages: zooplankton abundance drives aquatic beetle colonization. Ecology (2017). doi:10.1002/ecy.1914
[4] Smith W. O, Delizo L M, Herbolsheimer C, et al. (2017). Distribution and abundance of mesozooplankton in the Ross Sea, Antarctica. Polar Biology: Doi: 10.1007/s00300-017-2149-5
[5] Suárez E, Pérez C, M Rivera, R, et al. (2017). 9. Generalized Linear Models. Applications of Regression Models in Epidemiology. John Wiley & Sons, Inc.
[6] Dreano D, Tsiaras K, Triantafyllou G, et al. (2017). Efficient ensemble forecasting of marine ecology with clustered 1d models and statistical

lateral exchange: application to the red sea. Ocean Dynamics, 67(7), 935-947.
[7] Solanki H U, Bhatpuria D, Chauhan P. (2015). Applications of generalized additive model (gam) to satellite-derived variables and fishery data for prediction of fishery resources distributions in the Arabian Sea. Geocarto International, 32(1), 1-29.
[8] Chen X J, Tian S Q, Chen Y, Liu B L. （2010）. A modeling approach to identify optimal habitat and suitable fishing grounds forneon flying squid (Ommastrephesbartramii) in the Northwest Pacific Ocean. Fishery Bulletin. 108(1):1–14
[9] Saad S A, Wade C M. (2017). Seasonal and Spatial Variations of Saltmarsh Benthic Foraminiferal Communities from North Norfolk, England. Microbial Ecology, 73(3):539-555.
[10] China Standard Committee. 2008. China state standard GB/T 12763.6-2007. The specification for oceanographic survey, the sixth part: marine biological research. Beijing: China Standard Press. 159 pp.
[11] Ostrowska M, Matorin D N, Ficek D. (2000). Variability of the specific fluorescence of chlorophyll in the ocean. Part 2. Fluorometric method of chlorophyll a determination. Oceanologia, 42(2):221-229.
[12] McCullagh P, Nelder JA. (1989). Generalized Linear Models, 2nd ed. Chapman and Hall, London
[13] Wood SN. (2004). Stable and efficient multiple smoothing parameter estimation for generalized additive models. Journal of the American Statistical Association, 99(467):673-686.
[14] Hastie T, Tibshirani R. (1990). Generalized Additive Models.London:Chapman& Hall
[15] Wood SN. (2011). Fast stable restricted maximum likelihood and marginal likelihood estimation of semiparametric generalized linear models. Journal of the Royal Statistical Society: Series B (Statistical Methodology), 73(1): 3-36.
[16] Wu J, Yan B, Feng Z, et al. (2011). Zooplankton ecology near the Tianwan Nuclear Power Station. Acta Ecologica Sinica, 31(22):6902-6911.
[17] Candes E, Fan Y, Janson L, et al. (2017). Panning for Gold: Model-free Knockoffs for High-dimensional Controlled Variable Selection. arXiv: 1610.02351v3
[18] Yamashita T, Yamashita K, Kamimura R. (2007). A stepwise aic method for variable selection in linear regression. Communications in Statistics - Theory and Methods, 36(13), 2395-2403.
[19] Tokuda T, Van Mechelen I, Claeskens G, et al. (2012). Bic selection of the number of classes in latent class models with background variables. International Statistical Institute/international Association for Statistical Computing.
[20] Nizami N, Prasad N. (2017). Factor Analysis and PCA Analysis on Decent Work Indicators. Decent Work: Concept, Theory and Measurement. Springer Singapore.
[21] Raftery, A. E, Dean, N. (2017). Variable selection for model-based clustering. Journal of the American Statistical Association, 101(473), 168-178.
[22] Variyath A M, Chen J, Abraham B. (2010). Empirical likelihood based variable selection. Journal of Statistical Planning & Inference, 140(4), 971-981.
[23] Mosomtai G, Evander M, Sandström P, et al. (2016). Association of ecological factors with rift valley fever occurrence and mapping of risk zones in kenya. International Journal of Infectious Diseases Ijid Official Publication of the International Society for Infectious Diseases, 46(4), 49-55.
[24] Voutilainen A, Jurvelius J, Lilja J, et al. (2016). Associating spatial patterns of zooplankton abundance with water temperature, depth, planktivorous fish and chlorophyll. Boreal Environment Research, 21(1-2), 101-114.
[25] Loiselle S A, Azza N, Cózar A, et al. (2008). Variability in factors causing light attenuation in Lake Victoria. Freshwater Biology, 53(3):535–545.
[26] Moreau S, Mostajir B, Almandoz G O, et al. (2017). Effects of enhanced temperature and ultraviolet b radiation on a natural plankton community of the beagle channel (southern argentina): a mesocosm study. European Journal of Immunology, 10(8), 577-82.
[27] Fazeli N, Savari A, Nabavi S M B, et al. (2013). Seasonal variation of zooplankton abundance, composition and biomass in the chabahar bay, oman sea. International Journal of Aquatic Biology, 1(6), 1411-1419.
[28] Golmarvi D, Kapourchali M F, Moradi A M, et al. (2017). Influence of physico-chemical factors, zooplankton species biodiversity and seasonal abundance in anzali international wetland, iran. Open Journal of Marine Science, 07(1), 91-99.
[29] Harley S J, Myers R A, Dunn A. (2001). Is catch-per-unit-effort proportional to abundance?. Canadian Journal of Fisheries and Aquatic Sciences, 58(9): 1760-1772.
[30] Maunder M N, Punt A E. (2004). Standardizing catch and effort data: a review of recent approaches. Fisheries research, 70(2): 141-159.