SHORT COMMUNICATION

Statistical assessment of trophic conditions: squared Euclidean distance approach

Monte Kietpawpan[1], Parichart Visuthismajarn[2], and Chatchai Ratanachai[3]

Abstract

Kietpawpan, M.[1], Visuthismajarn, P.[2], and Ratanachai, C.[3]

Statistical assessment of trophic conditions: squared Euclidean distance approach

Songklanakarin J. Sci. Technol., 2003, 25(3) : 359-365

The classification of trophic conditions of water bodies may often face contradictory cases where a given lake is classified into a trophic category from a trophic variable, whereas it is classified into another trophic category from other trophic variables. To solve this problem, this paper proposes a new methodology based on the concepts of squared Euclidean distance and the boundary values recommended by the OECD (Organization for Economic Cooperation and Development). This methodology requires that a trophic variable data set of a water body under consideration and such boundary values be compared by a measure of similarity computed by using basic statistical techniques to determine the trophic condition of a given water body. The methodology has been tested by applying it to two sample data sets: the Pattani Dam Reservoir and the North Adriatic Sea data sets, which were taken from Kietpawpan (2002) and Zurlini (1996), respectively. The squared Euclidean distance analysis were then applied to the above data sets in order to classify trophic conditions, based on four trophic variables comprising total nitrogen, total phosphorus, chlorophyll-*a*, and Secchi depth. Our results show that the squared Euclidean distance analysis is a useful methodology for preliminarily classifying trophic conditions and solving contradictory classifications, which often arise when applying the present OECD methodology alone.

Key words : eutrophication, OECD, squared Euclidean, trophic state classification

[1]M.S. (Environmental Management), [2]Dr.phil. (Biogeography), [3]Ph.D. (Environmental Engineering), Faculty of Environmental Management, Prince of Songkla University, Hat Yai, Songkhla 90112 Thailand.
Corresponding e-mail: kietpawpan@yahoo.com
Received, 19 August 2002 Accepted, 10 March 2003

บทคัดย่อ

มนตรี เกียรติเผ่าพันธ์ ปาริชาติ วิสุทธิสมาจาร และ ฉัตรไชย รัตนไชย

การประเมินสภาพโทรฟิกของแหล่งน้ำทางสถิติ: วิธีสแควร์ยูคลิเดียน

ว. สงขลานครินทร์ วทท. 2546 25(3) : 359-365

การประเมินสภาพโทรฟิก (trophic) ของแหล่งน้ำโดยใช้วิธีการแบบดั้งเดิม เช่น การเปรียบเทียบค่าเฉลี่ยระหว่างค่าตัวแปรชี้วัดสภาพโทรฟิกของแหล่งน้ำที่ศึกษากับค่าอ้างอิงที่กำหนดขึ้นโดยองค์กรเพื่อการร่วมมือและพัฒนาทางด้านเศรษฐกิจ (OECD) อาจได้ผลการประเมินที่มีความขัดแย้งระหว่างตัวแปรชี้วัดสภาพโทรฟิก เช่น แหล่งน้ำนั้นถูกจัดเป็นโทรฟิกระดับหนึ่งเมื่ออิงตัวแปรชนิดหนึ่ง และจัดเป็นโทรฟิกระดับอื่นเมื่ออิงตัวแปรชนิดอื่น บทความนี้เสนอวิธีแก้ไขปัญหาข้างต้น โดยประยุกต์ใช้การเปรียบเทียบความแตกต่างแบบหลายตัวแปรระหว่างชุดข้อมูลตัวอย่างกับชุดข้อมูลอ้างอิงด้วยวิธีสแควร์ยูคลิเดียน (squared Euclidean) ผู้วิจัยได้ทดสอบวิธีนี้กับข้อมูลตัวแปรชี้วัดสภาพโทรฟิก ได้แก่ ค่าคลอโรฟิลล์-เอ ความโปร่งใส ไนโตรเจนทั้งหมด และฟอสฟอรัสทั้งหมด ของอ่างเก็บน้ำเขื่อนปัตตานี (Kietpawpan, 2002) และของ North Adriatic Sea (Zurlini, 1996) กับชุดข้อมูลอ้างอิงซึ่งเป็นเกณฑ์สำหรับการประเมินสภาพโทรฟิกของ OECD ผลการทดสอบแสดงให้เห็นถึงประโยชน์ของวิธีสแควร์ยูคลิเดียนสำหรับการประเมินสภาพโทรฟิกของแหล่งน้ำ และการแก้ไขปัญหาความไม่ชัดเจนของผลการประเมินสภาพโทรฟิกที่เกิดจากการประเมินโดยวิธีดั้งเดิม

คณะการจัดการสิ่งแวดล้อม มหาวิทยาลัยสงขลานครินทร์ อำเภอหาดใหญ่ จังหวัดสงขลา 90112

Eutrophication is the process of change in lake types from oligotrophic lake, mesotrophic lake, to eutrophic lakes (Okada, 1989). These trophic categories are conventionally identified using the boundary values for trophic categories proposed by the Organization of Economic and Cooperation Development (OECD, 1982) as reference criteria. The adoption of these proposed values, however, may often face contradictory cases where a given lake is classified into eutrophic lake from phosphorus concentration, whereas it is classified into oligotrophic lake from nitrogen concentration (Okada, 1989). This problem can occur even when taking different variables simultaneously into account such as when applying the Trophic State Indices (TSI) of Carlson (1977). This contradictory classification, therefore, could discourage the use of these systems (Zurlini, 1996).

Recently, Zurlini (1996) formulated a new trophic classification procedure, based on the combined probability method developed by Fisher (1954), to extend the applicability of the OECD system to multi-parametric sample space. Although this procedure provides a solution to the above problem, its main disadvantage is its difficult application by managers and technical personnel with only limited statistical training.

Here we focus on a new simple approach for initially characterizing trophic conditions of water bodies and solving the problem of contradictory classification. This approach was based on the concept of squared Euclidean distance described by Subhash (1996) and on the reference values for trophic state classification recommended by the OECD (OECD, 1982). We tested this approach by applying it to two practical cases: the Pattani Dam Reservoir (Kietpawpan, 2002) and the North Adriatic Sea (Zurlini, 1996) in order to show its usefulness.

Methodology

Data sources

Data on trophic variables comprising total nitrogen (TN), total phosphorus (TP), chlorophyll-*a*, and Secchi depth, were taken from sampling campaigns carried out in the Pattani Dam Reservoir, from May 2000 to May 2001. Complete methodology on water sampling and analysis are presented in Kietpawpan (2002). Another set of data on trophic variables were taken from the monitoring program carried out in the North Adriatic Sea from 20 August to 26 September 1992 (Zurlini, 1996). An exhaustive description of the physical and biological environment of this area can be found in Vollenweider *et al*. (1992). Des-criptive statistical summarizes of the trophic variables considered are given in Table 1 and 2.

Data on the boundaries for trophic classification recommended by the Organization for Economic Cooperation and Development (OECD, 1982) were taken as reference criteria for eutrophic level assessment of the sample cases (Table 3).

Normality testing was undertaken in both graphical (Histogram) and statistical (Shapiro-Wilk test) methods, with an aid of SPSS for Windows (SPSS, Inc., Chicago, IL). Right-skewed data were normalized by square root transformation as recommended by Mateu (1997). Accordingly, geometric mean values were utilized for trophic classification based on the concept of squared Euclidean distance.

Table 1. Descriptive statistics of trophic variables of the Pattani Dam Reservoir data set (Kietpawpan, 2002)

Variable	n	Min	Max	Mean	S.D.
Chlorophyll-*a* (µg/L)[a]	117	0.2	62.2	12.7	15.1
Secchi depth (cm)[b]	117	20	150	68.59	0.4
Total nitrogen (µg/L)	117	32	175	107	35
Total phosphorus (µg/L)	17	2.9	34.6	16.9	5.4

[a] mean after transformation was 3.1
[b] mean after transformation was 7.9

Table 2. Descriptive statistics of trophic variables of the North Adriatic data set (Zurlini, 1996)

Variable	n	Min	Max	Mean	S.D.
Chlorophyll-*a* (µg/L)	413	0.03	35.8	1.47	-
Secchi depth (cm)	415	120	1520	710	-
Total nitrogen (µg/L)	250	90	2140	475.4	-
Total phosphorus (µg/L)	250	2.0	300	11.5	-

Table 3. Trophic boundaries (Mean) for lakes from OECD (1982)

Variable	Oligotrophy	Mesotrophy	Eutrophy
Chlorophyll-*a* (µg/L)[a]	1.7	4.7	14.3
Secchi depth (cm)[b]	990	420	245
Total nitrogen (µg/L)	661	753	1875
Total phosphorus (µg/L)	8.0	26.7	54.4

[a] means after transformation was 1.3, 2.2, and 4.1
[b] mean after transformation was 32.4, 21.3, and 16.1

2. The Squared Euclidean Distance Approach

The concept of distance is simple. Distance is a measure of how far apart two data sets are (Subhash, 1996). Distance measure is smallest for cases that are most similar. A commonly used distance measure is the squared Euclidean distance, which is the sum of the squared differences over all of the variables.

Here we apply the squared Euclidean distance to the above data sets to determine the trophic condition of the water bodies. The squared Euclidean distance is the sum of the squared differences over all of the standardized trophic variables. The distance is smallest for the trophic category to which a water body condition is most likely to belong.

To compute the squared Euclidean distance for each of the three trophic conditions (i.e. oligotrophy, mesotrophy, and eutrophy), we used Equation 1 with values of x and j obtained from Table 1 or 2, and 3.

$$D_j^2 = \sum_{i=1}^{4}(x_i - j_i)^2 \qquad (1)$$

where,

D_j^2 = squared Euclidean distance between a sample data set of a given water body and a reference data set of the trophic condition *j*

x_i = standardized trophic variable *i* of a given water body

j_i = standardized trophic variable *i* of the reference trophic condition *j*

i = trophic variables (i.e. TN, TP, Chlorophyll-*a*, and Secchi Depth)

Σ = to sum together

The standardized form (*Z*-score) indicated in Table 4 was used to eliminate the effect on computed distances from the difference in measurement units and magnitudes among trophic variables. The *Z*-scores for trophic variable *i* of trophic condition *j* were computed, using SPSS for Windows (SPSS, Inc., Chicago, IL), based on the mean and standard deviations for all four cases (i.e. oligotrophic, mesotrophic, eutrophic, and a sample water body).

The application of the OECD and the TSI approaches are also presented to provide examples of the contradictory classification. Further, a comparison between the trophic category of the North Adriatic Sea data set classified by the combined probability approach (Zurlini, 1996) and by our approach was made to show the agreement in the classification results.

Results

1. OECD method

The contradictory classifications were found when traditionally comparing the OECD's reference values to the values of trophic variables of Pattani Dam Reservoir. That is, the reservoir water was considered eutrophic by chlorophyll-*a* and Secchi depth, whereas it was considered oligotrophic by total nitrogen and mean total phosphorus.

Table 4. *Z*-scores for trophic variables indicated in Table 1 and 3

Category \ Variable	Chl-a	Sechi	TN	TP
Oligotrophy	-1.14484	1.26473	-0.25367	-0.92008
Mesotrophy	-0.39549	0.18276	-0.12953	0.00995
Eutrophy	1.18647	-0.32410	-1.38438	1.38758
PDR*	0.35386	-1.12339	-1.00118	-0.47745

*** Pattani Dam Reservoir**

2. Trophic state indices method

Based on the application of TSI by Kietpawpan (2002), the different trophic variables of Pattani Dam Reservoir did also not agree with the same trophic category, facing a problem of contradictory classifications. The reservoir water was classified into eutrophic category from Secchi depth, whereas it was classified into mesotrophic category from chlorophyll-*a* and phosphorus concentrations.

3. Squared euclidean distance method

The screening for normality showed that data on all trophic variables, except chlorophyll-*a* and Secchi depth, were normally distributed. The positive skewness of chlorophyl-*a* (Coef. Skew. = 1.727) and of Sechhi depth (Coef. Skew. = 0.489) were normalized by square root transformation (Mateu, 1997). Accordingly, the geometric mean values of 3.1 for chlorophyll-*a* and of 0.79 for Sechhi depth were used in analyzing the squared Euclidean distance. For sensible comparison, the OECD's reference values of chlorophyll-*a* and Secchi depth were also changed by squared root transformation. The transformed values of chlorophyll-*a* and Secchi depth (as indicated beneath Table 1 and 3) were then employed in analyzing the squared Euclidean distance.

From Equation 1 and data on Table 1 and 3, the squared Euclidean distance (D_j^2) of the Pattani Dam Reservoir data set, based on the standardized trophic variables for oligotrophic category (Table 4), was $[-1.14484-0.35386]^2+[1.26473-(-1.12339)]^2+[-0.25367-(-1.00118)]^2+[-0.92008-(-0.47745)]^2$, or 8.704. Computed accordingly, the D_j^2 for mesotrophic and eutrophic categories were, respectively, 3.265 and 10.501. It is evident from this outcome that the distance measure for the Pattani Dam Reservoir data set was smallest for the mesotro-phic category (3.265), which indicates that the reservoir water body was most likely to be a mesotrophic water body.

4. Comparison of sample classification procedures

A comparison of the two classification procedures: the combined probability (Zurilni, 1996) and the squared Euclidean distance, was performed based on the data set presented in Table 5. The North Adriatic Sea data set was chosen as a practical case for this comparison because the contradictory classification has been found and successfully solved by another complex statistical method. According to Zurlini (1996), the water body of the North Adriatic Sea is classified as oligotrophic (*P*=0.657), based on the combined probabilities for oligotrophic, mesotrophic, and eutrophic categories of 0.657, 0.066, and 0.000, respectively.

To the same data set, the squared Euclidean distance approach was applied in order to test if these different classification approaches may provide a same result. As computed by Equation 1 with the values obtained from Table 4, the distance measure of the North Adriatic Sea data set for oligotrophic, mesotrophic, and eutrophic categories were 0.845, 1.781, and 15.564, respectively. It appears that the distance measure was smallest for the oligotrophic category (0.845), which indicates

Table 5. *Z*-scores for trophic variables indicated in Table 2 and 3

Variable / Category	Chl-a	Sechi	TN	TP
Oligotrophy	-0.63819	1.21652	-0.44235	-0.81195
Mesotrophy	-0.13993	-0.52245	-0.29706	0.07338
Eutrophy	1.45452	-1.05635	1.47485	1.38482
NDS*	-0.67639	0.36229	-0.73545	-0.64625

*** North Adriatic Sea**

that the North Adriatic Sea water body was most likely to be oligotrophic.

Discussion

The results outlined above represent an applied basic statistical methodology for trophic state classification. It appears that the application of OECD reference values alone was found difficult classification because of disagreement among trophic variables. These finding support a conclusion that trophic variables compared with the trophic boundary values can yield contradictory classification (Okada, 1989; Giovanardi and Tromellini, 1992; Zurlini, 1996). The results demonstrate that the application of the squared Euclidean distance to the reference values of the OECD can effectively solve the problem of contradictory classification in a simple manner. It requires just small data sets and simple calculation using basic statistical techniques.

In the case of Pattani Dam Reservoir data set, however, uncertainty arising from the uncoupling role played by Secchi depth can provide an overestimate of trophic level membership. Kietpawpan (2002), for instance, classified the water body of Pattani Dam Reservoir into a eutrophic water body based on Carlson's Trophic State Index when taking low observed Secchi depth into account alone. In fact, such low Secchi depth results from a high loading of suspended sediments from the Pattani River. The observed low Secchi depth, therefore, implicates high turbidity rather than high trophic condition. On the contrary, the squared Euclidean distance approach presented here takes all trophic variables into account simultaneously. This yields a more reliable classification result, compared to an indices-based approach, such as the Carlson's Trophic State Index (Carlson, 1977).

Compared to the combined probability methods, a more complex trophic classification procedure proposed by Zurlini (1996), the squared Euclidean approach provides a relatively reliable classification of trophic group membership, as shown by comparison results of the different classification procedures applied to the North Adriatic Sea data set.

Conclusion

One of commonly used methods for classifying the trophic condition of a water body is the OECD methodology. Most of difficulties encountered in applying the OECD methodology might have arisen because of the contradictory response of trophic variables (Zurlini, 1996). The trophic classification approach developed here is based on the application of OECD classification system and the concept of the squared Euclidean distance. This simple approach can solve doubtful membership of samples and provide the same classification results as a more complex approach can do (e.g. Zurlini, 1996). Nevertheless, it must be stressed that the reliability of the squared Euclidean approach presented here should be treated with some caution until the misclassification percentage of this approach are determined by comparing its classification results, based on more sample data sets, with results of other reliable approaches.

Acknowledgment

This work was supported in part by fund from the Graduate School, Prince of Songkla University. We thank the three anonymous reviewers for their useful comments. This paper could not have been written without the contribution of time, knowledge, and keen vision of our previous benefactors and benefactresses.

References

Carlson, R.E. 1977. A Trophic State Index for Lakes, Limnology and Oceanography. 22(2): 361-369.

Fisher, R.A. 1954. Statistical Methods for Research Workers, Oliver and Boyd, Edinburgh.

Giovanardi, F. and Tromellini, E. 1992. Statistical assessment of trophic conditions. Application of the OECD methodology to the marine environment, Sci. Total Environ., Suppl. 1992: 211-233.

Kietpawpan, M. 2002. Ecological Risk Associated with Eutrophication and Lead Contamination in Pattani Dam Reservoir, M.S. thesis, Prince of Songkla University.

Mateu, J. 1997. Methods of assessing and achieving normality applied to environmental data, Environmental Management, 21(5): 767-777.

OECD 1982. Eutrophication of Waters, Monitoring, Assessment and Control, Organization for Economic Cooperation and Development. Paris.

Okada, M. 1989. No. 8 Eutrophication, Department of Chemical Engineering, Faculty of Engineering, Tokyo University of Agriculture & Technology, Tokyo.

Subhash, S. 1996. Applied Multivariate Techniques, John Wiley & Son, Inc, New York.

Vollenweider, R.A., Rinaldi, A. and Montanari, G. 1992. Eutrophication, structure and dynamics of marine coastal system: results of ten-year monitoring along the Emilia Romagna coast (Northwest Adriatic Sea). Sci. Total Environ., Suppl. 1992: 63-106.

Zurlini, G. 1996. Multiparametric classification of trophic conditions. The OECD methodology extended: combined probabilities and uncertainties—application to the North Adriatic Sea, Sci. Total Environ., 182: 169-185.