SELF ORGANISATION FEATURE MAPS FOR EUTROPHICATION ANALYSIS OF TAIHU LAKE

Q. CHEN [1] and A. E MYNETT [2]

[1] Professor, RCEES Chinese Academy of Sciences, 18 Shuangqing Road, Haidian District, Beijing, 100085, China
(Tel: +86-10-6284-9326, Fax: +86-10-6284-9326, e-mail: qchen@mail.rcees.ac.cn)
[2] Professor, WL | Delft Hydraulics, Rotterdamsweg 185, P.O. Box 177, 2600 MH Delft, The Netherlands
(Tel: +31-15-285-8571, Fax: +31-15-285-8582, e-mail: arthur.mynett@wldelft.nl)

Recently artificial neural networks (ANN) is increasingly used in modelling of hydraulics, hydrological and ecological process (Hall and Minns, 1993; van den Boogaard, 1998). Most of the applications of ANN in aquatic systems focus on matching the relations of input-output by using Multi Layer Perceptron (MLP) networks or Radial Basic Functions (RBF) networks. An essential step of such applications is the construction of networks' input-output pattern (or networks structure), that even determines the success or failure of the practical application. Therefore physical knowledge of the specific case is used as much as possible in data pre-processing. In order to identify the structure or assess the distribution of the data patterns, techniques for multi-variable analysis are needed.

Commonly used multi-variable analysis methods include Principal Component Analysis, Empirical Orthogonal Functions, auto-correlation, cross-correlation, and clustering, possibly in combination with statistical techniques (van den Boogaard, 1998). Recent work shown that self organizing feature maps (SOFM), being a type of ANN, has been successfully applied to multi-variable data analysis, and is most suitable for examination and interpretation of data sets where little knowledge is available about the inner structure. Comparing with traditional multi-variable data analysis techniques, SOFM has an important and attractive points that a multi-dimensional input ensemble is mapped into a one or two dimensional space, which is much easier to be displayed and visualised, and more convenient to be interpreted. Unlike MLP and RBF networks, a SOFM networks has no output layer, but map the input signals into a certain feature. Moreover, SOFM networks implements unsupervised learning paradigm. This paper is to apply SOFM method to analyze aquatic data from Taihu Lake, China.

Taihu Lake, the third largest freshwater lake in China, is located in the most developed Yangtze delta. It has strategic significance since the Taihu Basin is the most populated and developed area, which provides more than 10% of the national GDP. However, in association with the rapid development of agriculture, urbanization and industries, the water quality in Taihu basin deteriorate very quickly. According to water quality assessment, almost 80% of the tributaries are polluted. In some local areas of the lake eutrophication is very severe, especially in spring and summer when red tide happens locally, which cause fish mortality, bad smell and pipe block. Water pollution badly hampers the economic development and affects amenity seriously. In order to get an image about the spatio-temporal patterns of the eutrophication in the lake, some analyses are conducted, and the results are presented below.

For cluster 1 that has 535 cases (about 78%), all of the considered parameters have low value, which means the water quality in the lake is generally good. According to the

monitoring data, this class represents the water in the centre and east Taihu Lake.

For cluster 2 that contains 7% cases, both NH_4 and Chla have high value although TIP is not high, which means that NH_4 has great effect on chlorophyll concentration. This is identical with the correlation coefficient. Cluster 2 is the water located in the north part, which is also called Meiliang Lake, where the water is stagnant. In spring and summer, the prevailing wind is from the sea, which blows a lot of Chla to Meiliang Lake, and then the nutrients are accumulated here year after years, which causes vicious circle.

For cluster 3, which has 15% cases, TIP has high values, while both TIN and NH_4 are low, and correspondingly the Chla is not high, which means phosphorus is not limiting factor.

From cluster 2 and cluster 3, it can be concluded that phosphorus is enough in both cases, while nitrogen is limiting factor.

Although cluster 2 has only 42 cases, the impact on aquatic is quite great, so eutrophication in Taihu Lake is mostly contributed cluster 2, that is NH_4.

Clustering with respect to temporal is also conducted in this case study. Take April, June, August and October as variable, and then cluster the distribution of TIP, NH4 and Chla respectively. The detailed result are not presented here since space limit, and the conclusion is: high concentration of TIP and NH_4 usually happen in April, and in October in some local area, while the high concentration of Chla take place in Jun and August. One reason is that strong wind in winter makes it easy for sediment resuspension and nutrient release, which leads to high TIP and NH_4 in spring and then high Chla in summer when light penetration becomes enough. Although in some local places the TIP and NH_4 are high in October, while the Chla is still low, because at this moment the irradiance becomes the predominant factor.

The case study show that the water quality in Taihu Lake is generally good expect for some local areas, such as Meiliang Lake, West Taihu Lake, which are severely eutrophicated. Eutrophication happens mainly in June and August when the nutrient and irradiance are both enough. The productivity of the eutrophicated areas of the lake is limited by NH_4, while the phosphorus is enough for phytoplankton growth.

The advantage of SOFM is to map multiple-dimensional system into two or three dimensional maps while preserve the inherent features, which is much easier to be analysed by powerful visualisation tools. The traditional methods for water quality assessment are single parameter evaluation or integral index assessment, or fuzzy logic sometimes. Comparing with those techniques, SOFM consider different importance by assigning weights, moreover, it can explore some inherent relations among those multiple variables. Further study on this work will analyse the sources of nutrients and determine the main contributor.

REFERENCE

Ali, Md.A., 1997. Application of Self Organisation Feature Maps for analysis of ydrological and Ecological Data Sets. *M. Sc. Thesis* H.H. 315, IHE, Delft, The Netherlands

Boogaard, H.F.P. van den, Mynett, A.E., Ali, Md.A., 1998. Self organising maps for the analysis of hydrological and ecological data sets. *Hydroinformatics '98* ISBN 90 5410 9831

Hall, M.J., Minns, A.W., 1993. Rainfall-runoff modelling as a problem in artificial intelligence: experience with a neural network. *Proc. 4th Nat. Hydrol. Symp.* Cardiff, British Hydrological Society, London

Kohonen, T., 1982. Self Organized formation of topological correct feature maps. *Iological Cybernetics 43.*