ECOLOGICAL MODELLING 203 (2007) 26–33

available at www.sciencedirect.com

ScienceDirect

journal homepage: www.elsevier.com/locate/ecolmodel

Community patterns of benthic macroinvertebrates collected on the national scale in Korea

Young-Seuk Park[a], *Mi-Young Song*[b], *Young-Cheol Park*[c], *Kyung-Hee Oh*[d], *Eungchun Cho*[e], *Tae-Soo Chon*[b,*]

[a] *Department of Biology and The Korea Institute of Ornithology, Kyung Hee University, Dongdaemun-gu, Seoul 130-701, Republic of Korea*
[b] *Division of Biological Sciences, Pusan National University, Busan 609-735, Republic of Korea*
[c] *Project Management Unit, UNDP/GEF Korea Wetlands Project, National Institute of Environmental Research, Seo-gu, Incheon 404-170, Republic of Korea*
[d] *Biological Resources Division, National Institute of Environmental Research, Seo-gu, Incheon 404-170, Republic of Korea*
[e] *Division of Mathematics and Sciences, Kentucky State University, Frankfort, KY 40601, USA*

ARTICLE INFO

Article history:
Received 4 January 2005
Received in revised form
16 April 2006
Accepted 23 April 2006
Published on line 2 February 2007

Keywords:
Benthic community
Species richness
Abundance
Community classification
Regional classification

ABSTRACT

From 1997 to 2002, benthic macroinvertebrates were collected with a Surber sampler from the 1970 sampling sites from clean to intermediately polluted areas in streams in Korea. The Self-Organizing Map (SOM) was used for patterning and visualizing community patterns of macroinvertebrates collected on the national scale. The SOM accordingly classified the samples to eight groups of communities, and the groups firstly corresponded to the geographical regions of the sampling sites, distinguishing the samples from mountain streams with high species richness from the samples from lowland areas with low species richness. Relations between species richness and abundance showed three patterns according to the SOM groups: (1) high species richness and high abundance in the clean area, (2) intermediate species richness and high abundance in intermediately polluted area, and (3) low species richness and low abundance in severely damaged area. Biological indices (Shannon diversity index, evenness, EPT richness, EPT abundance, and BMWP score) were also accordingly differentiated among the different clusters. The clustered communities were correspondingly related to the patterns of species richness and abundance, and the trained SOM appears to be efficient in defining the patterns of ecoregions.

© 2006 Elsevier B.V. All rights reserved.

1. Introduction

The community composition depends on stability of their habitats which provide the resources for development of the residing populations (Cummins, 1979; Ward and Stanford, 1979; Malmqvist and Otto, 1987). The structures of community assemblages are potentially determined by various environmental factors acting on the habitats in different spatial and time scales (Stevenson, 1997; Snyder et al., 2002). Regarding that different species show different spatial–temporal dynamics in responding to the impact of environments, understanding community abundance patterns is a fundamental step to achieve the sustainable management of aquatic ecosystems.

With the advantages of taxonomic diversity, sedentariness in behavior and long life cycles, benthic macroinvertebrates respond to environmental disturbances in an integrated and continuous manner. Consequently, macroinvertebrates have

* *Corresponding author*. Tel.: +82 51 510 2261; fax: +82 51 512 2262.
E-mail address: tschon@pusan.ac.kr (T.-S. Chon).
0304-3800/$ – see front matter © 2006 Elsevier B.V. All rights reserved.
doi:10.1016/j.ecolmodel.2006.04.032

been widely used for assessing ecological water quality in aquatic ecosystems. There have been a numerous accounts of benthic macroinvertebrates used for indicators of the short- and long-term environmental changes in running waters (Hellawell, 1978; Lenat, 1988; Smith et al., 1999; Hawkins et al., 2000).

In natural conditions, species richness (i.e. the number of species occurring in a given area) has been commonly used as an integrative descriptor of the community (Lenat, 1988), since species richness is influenced by a large number of environmental factors in a continuous period with regards to environmental stability (Cummins, 1979; Ward and Stanford, 1979), ecosystem productivity (Lavandier and Décamps, 1984) and heterogeneity (Malmqvist and Otto, 1987), and other biological factors (MacArthur, 1965; Feminella and Resh, 1990). The interactions of the environmental factors can consequently determine the gradients in species richness in streams (Vannote et al., 1980; Minshall et al., 1985). Additionally, species richness is useful for revealing the impact of environmental factors in disturbed conditions. The species richness of aquatic invertebrates has been influenced by natural and/or anthropogenic disturbances (Rosenberg and Resh, 1993), which may lead to spatial discontinuities of predictable gradients (Ward and Stanford, 1979, 1983) and loss of taxa (Brittain and Saltveit, 1989).

Along with analysis in species richness, investigation of species abundance patterns has been regarded as an important topic in elucidating patterns of communities responding to the disturbances. Preston's canonical log-normal distribution has been the most widely accepted formalization of the relative commonness and rarity of species (Preston, 1962; Brown, 1981). Regarding that the species are often vulnerable to various environmental disturbances, the existence of rare species is a key issue in community ecology in relation to risk assessment. This type of complex relationships in 'community changes and environment disturbances' would be accordingly addressed by studying community compositions in relation to abundance patterns per taxa, i.e. relations between species richness and abundance.

Considering that the relationships between species and abundance are closely associated with overall ecological conditions, and communities are influenced by the given environmental factors limited by the geographic regions, the species–abundance patterns would be reflected by ecoregions. The effect of ecoregions would appear especially on a large scale. Many studies have been carried out to classify ecoregions based on the distribution patterns of aquatic organisms on regional or national scale, showing the importance of geographic differences in biotic and abiotic characteristics of streams (Tison et al., 2005).

Similarly, overall variation in community composition would be reflected on a map characterizing the community patters. When processing the large-scale data, the first step is usually to derive a classification of sites or systems by analysing the degree of association among community data. In this study, we apply the technique in ecological informatics, Self-Organizing Map, to mining the large-scale community data and to further relating the community patterns to variation caused by geographic distribution and different degrees of disturbances. We aim (1) to classify benthic macroinvertebrate communities on the national scale in Korea, (2) to accordingly reveal the relationships between community grouping and environmental factors, and (3) to elucidate association of community clusters with the "species–abundance" patterns.

2. Materials and methods

2.1. *Ecological data*

Benthic macroinvertebrates were collected in relatively clean to intermediately polluted areas in South Korea, being initiated by the National Natural Environment Monitoring Project, Ministry of Environment of Korea. Benthic macroinvertebrates were sampled from the 1970 sampling sites in total in the major rivers in Korea from 1997 to 2002. The samples were collected mostly at the 1st–3rd stream order using a Surber sampler (30 cm × 30 cm). The number of collected individuals was estimated in the unit area (1 m^2). The general techniques for sampling benthic macroinvertebrates and measuring environmental variables have been conducted according to the National Natural Environment Monitoring protocol (Ministry of Environment, 1997). We used four variables to describe environmental conditions for the sampling site: altitude, depth, current velocity and conductivity.

In the dataset, 687 species was recorded in total. Diptera (including Chironomidae), Ephemeroptera, Trichoptera, Oligochaeta and Gastropod were mostly dominant. In addition to species richness and abundance, community and biological indices were estimated to evaluate ecological status of the sampling sites: Shannon diversity and evenness indices, and Biological Monitoring Working Party (BMWP) score (National Water Council, 1981; Walley and Hawkes, 1997). We also checked abundance of three major insect orders, i.e. Ephemeroptera, Plecoptera, and Trichoptera (EPT), which are commonly identified at the species level in freshwater studies. Subsequently, the EPT richness (i.e. the number of species occurring in a given area) and abundance were recorded at each sampling site.

The differences of variables between the different groups identified through the models were evaluated by the Unequal N HSD multiple comparison test. Statistical analyses were conducted with the software package, Statistica (StatSoft, 2004). Abundance of each species was proportionally normalized between 0 and 1 (in the range of the minimum and maximum values) as input data for modelling.

2.2. *Modelling procedure*

First we extracted information on the relationships between species richness and abundance using scatter plots and log-normal distribution plots. Then we utilized the Self-Organizing Map (SOM) (Kohonen, 1982) for patterning macroinvertebrate communities. The SOM approximates the probability density function of the input data, and has been frequently used for clustering, visualization, and abstraction of ecological data (i.e. the idea of which is to show the dataset in another, more usable, representation form) (Kohonen, 2001; Park et al., 2003).

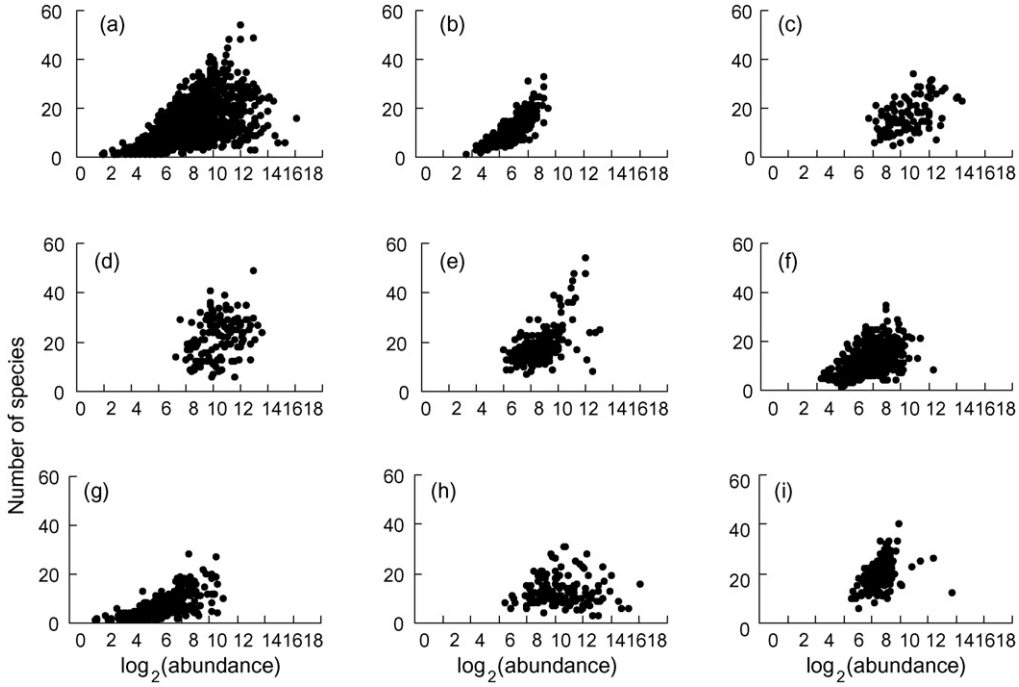

Fig. 1 – Relations between number of individuals in log scale and species richness in the 1970 samples used in this study. Each point indicates each sampling site. (a) All samples, (b–i) the samples separately grouped in clusters 1, 2, 3,. . ., 8, respectively.

The SOM consists of two layers: input and output layers connected by connection intensities (weights). Input layer gets information from data matrix, while output layer visualizes the computational results. When an input vector $\boldsymbol{x}$ (abundance of species) is sent through the network, each neuron k of the network computes the distance between the weight vector $\boldsymbol{w}$ and the input vector $\boldsymbol{x}$. The output layer consists of D output neurons, which are usually arranged into a two-dimensional grid for better visualization. There are no strict rules regarding the choice of the number of output neurons. To choose a suitable map size, we trained the SOM with different map sizes. Based on the recommendations of field ecologists in analyzing the data and the results from different map sizes, we chose 300 ($D = 20 \times 15$) neurons as the number of output neurons in this study. Each output neuron is a computational unit in the learning process. The best arrangement for the output layer is a hexagonal lattice, as it does not favor horizontal and vertical directions as much as a rectangular array (Kohonen, 2001).

Among all the output neurons, the best matching unit (BMU) with minimum distance between the weight and input vectors is the winner. For the BMU and its neighborhood neurons, the weight vectors $\boldsymbol{w}$ are updated using the SOM learning rule. A detailed description of the SOM algorithm could be provided by Kohonen (2001) for computation and by Park et al. (2003) for ecological application.

The learning process of the SOM was carried out using the SOM Toolbox (Alhoniemi et al., 2000) developed by the Laboratory of Information and Computer Science in the Helsinki University of Technology (http://www.cis.hut.fi/projects/somtoolbox/) in Matlab environments (The Mathworks, 2001), and we adopted the initialization and training methods suggested by the authors of the SOM Toolbox that allow the algorithm to be optimized (Vesanto et al., 1999).

3. Results

3.1. *Relations between species richness and abundance*

Species richness in the samples was in the broad range from 1 to 54. It showed the maximum value around the abundance of 12 of log 2 (individuals) (Fig. 1a). While species richness showed the widest range in the middle point of abundance level, the range of species richness became narrower at either lower or higher values of abundance. Fig. 2 shows the log-normal model of community structure by grouping species into abundance categories (octaves, i.e. power of 2). Although the distribution approached to a normal distribution, the distribution was not statistically significant for the normal distribution ($p < 0.01$ for the Lilliefors normality test).

3.2. *Patterning benthic communities*

Benthic macroinvertebrate communities were patterned according to the similarity of community compositions through training with the SOM. The classified samples were accordingly visualized on the map (Fig. 3). The node includes another hexagon within itself in order to indicate the size of the patterned samples (Fig. 3a). As the number of patterned samples increase in a SOM output unit, the size of hexagon correspondingly increases. Simultaneously, the gray level becomes darker with the increasing number of the

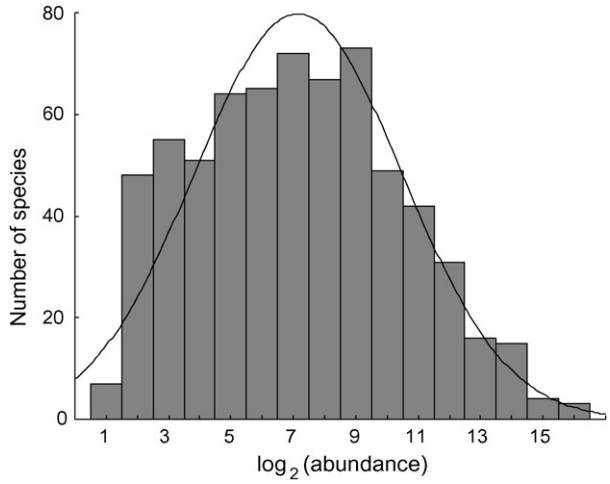

Fig. 2 – Log-normal model of community structure by grouping species into abundance categories (octaves, i.e. power of 2).

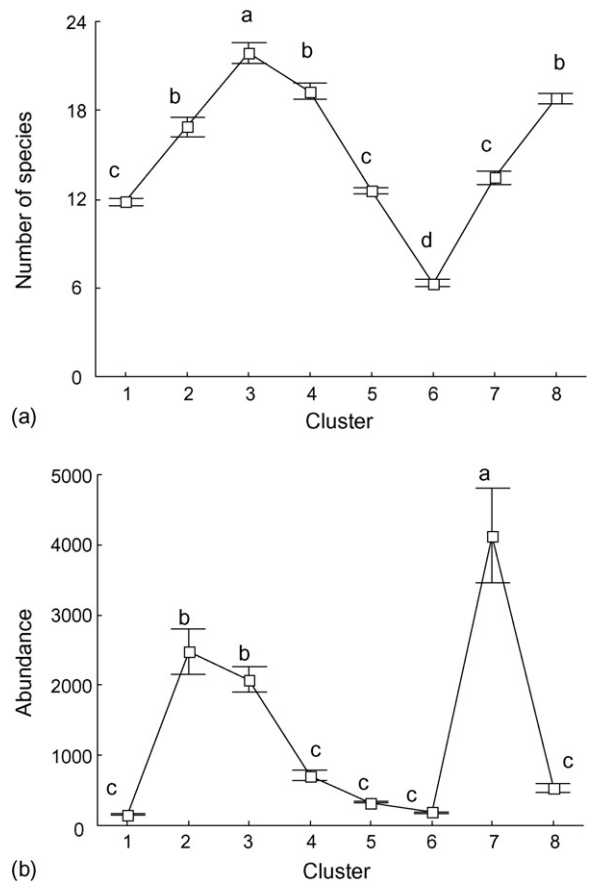

Fig. 4 – Community characterization in different clusters according to the SOM (Fig. 3a). (a) Species richness (number of species), (b) abundance (different alphabets indicate significant differences between the clusters based on the Unequal N HSD multiple comparison test ($p = 0.05$). Error bars indicate mean and standard error of each variable).

patterned samples in each output unit. The number of the patterned samples in the different output units ranged from 0 to 37. Overall, the SOM units were classified to two main clusters based on the dendrogram of the cluster analysis with the Ward's linkage method (Fig. 3b). The clusters were further divided to eight subclusters at different levels of the Euclidean distance. The clusters accordingly reflected geographical differences among the sampling sites.

Fig. 3c shows the representative location for each cluster in the geographical map of Korea. Cluster 1 was characterized by the samples collected mainly from the four regions Goyang, Paju, Gongju, Imsil and Jinhae, whereas cluster 2 was from Uljin in the southern area of Korea. However, cluster 6 did not show any specific geographic areas, being scattered widely over Korea. Therefore, cluster 6 was not indicated in the map in Fig. 3c.

According to the clusters shown in Fig. 3a, the patterns of species richness and abundance were correspondingly differentiated in different clusters (Fig. 1b–i). Species richness and abundance were statistically significant in different clusters (Fig. 4). Abundance was extremely low in cluster 1, while species richness in this cluster was in the intermediate range (Fig. 1b). The samples in cluster 1 were mainly collected at the pool zones with higher values of depth and lower values

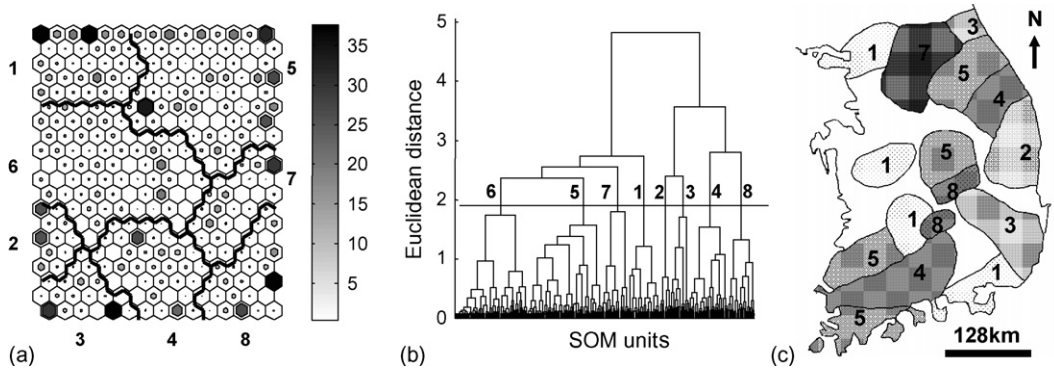

Fig. 3 – Classification of the samples according to the trained SOM. (a) The SOM units were grouped to eight clusters, (b) the dendrogram according to the Ward linkage method based on Euclidean distance, (c) geographical location of the sampling sites matching to clusters according to the SOM (Fig. 3a). Cluster 6 is not indicated on the geographical map because it did not show any specific geographic area.

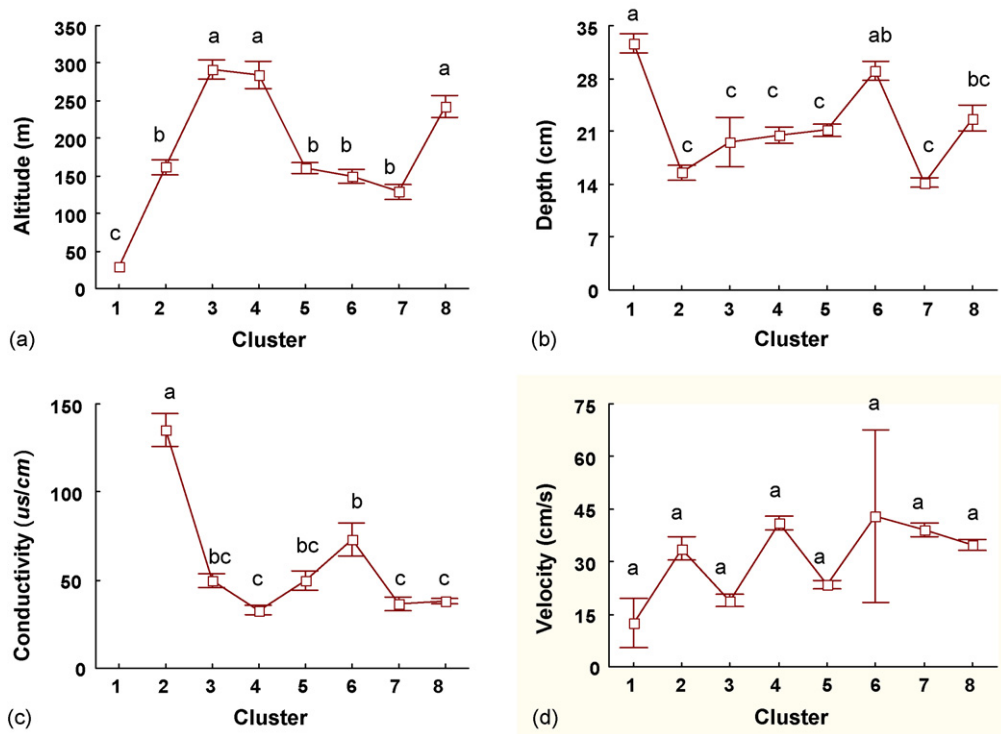

Fig. 5 – Environmental variables in different clusters according to the SOM (Fig. 3a). (a) Altitude, (b) depth, (c) conductivity, (d) velocity (different alphabets indicate significant differences between the clusters based on the Unequal N HSD multiple comparison test ($p = 0.05$). Error bars indicate mean and standard error of each variable. Conductivity was not available at the samples in cluster 1).

of velocity at the low altitude areas (Fig. 5). In contrast, the samples in cluster 2 (Fig. 1c) displayed the high level of abundance and the intermediate level of species richness, and were mainly collected in the southeastern areas in Korea. The sample sites in these areas were characterized with high electric conductivity (Fig. 5). This was due to the wide distribution of the limestone in this area.

Overall, the clusters projected on the trained SOM were in accordance with the different geographical areas in Korea. While species richness was in the highest range in cluster 3 in the mountainous area (Fig. 1d), the highest range of abundance was observed in cluster 7 in the urbanized area (Fig. 1h). The samples in cluster 7 were mainly collected from the slightly polluted streams in urbanizing area in the suburbs of Seoul. It was notable that the bottom areas of the SOM map showed high values of both species richness and abundance, including cluster 3 (Fig. 1d), cluster 4 (Fig. 1e), and cluster 8 (Fig. 1i) (Fig. 4). This was clearly distinguished at the level of the main clusters obtained by the Ward linkage method (Fig. 3b). The samples of clusters 3, 4 and 8 were collected mostly from the mountain streams showing relatively high altitude (Fig. 5a). Generally, clusters showing high species richness displayed high abundance in this study.

The sampling sites of clusters 4 and 5 were closely located each other in the geographical map (Fig. 3c). However, their community compositions were greatly different, separately belonging to the two main groups of the dendrogram (Fig. 3b). While the samples in cluster 4 (Fig. 1e) showed higher species richness in a broad range, the samples in cluster 5 (Fig. 1f) were in the lower ranges for both species richness and abundance. It was notable that cluster 8 showed unique patterns with high species richness and low abundance (Fig. 1i). The samples in cluster 8 were collected mainly in the Jiri Mountain (Fig. 3c), indicating clean water with low conductivity (Fig. 5). Cluster 7 (Fig. 1h) was characteristic with regards that communities belonging to this cluster showed the negative relationships with communities grouped in cluster 8 in between species richness and abundance.

Community patterns in cluster 6 were characteristic with the lowest range in both species richness and abundance (Fig. 1g). The samples in cluster 6 were identified as the communities collected from the widely distributed disturbed areas in the country due to the construction and rehabilitation programs carried out in the streams during the survey period.

We evaluated each cluster with biological indices BMWP score, EPT richness, EPT abundance, Shannon diversity, and evenness (Fig. 6). In general, biological indices matched with the gradients of environmental disturbances shown by the trained SOM (Fig. 3a). EPT richness ranged from 0 to 38, showing lower values in clusters 1 and 6, and higher values in clusters 3, 4 and 8 (Fig. 6a). Clusters 2, 5, and 7 showed intermediate values. EPT richness accordingly revealed the degree of disturbances according to the clustered communities. EPT abundance ranged from 0 to 10,756, showing low values in clusters 1 and 6 and high values in clusters 2, 3 and 7 (Fig. 6b). The differences in EPT abundance among clusters were in gen-

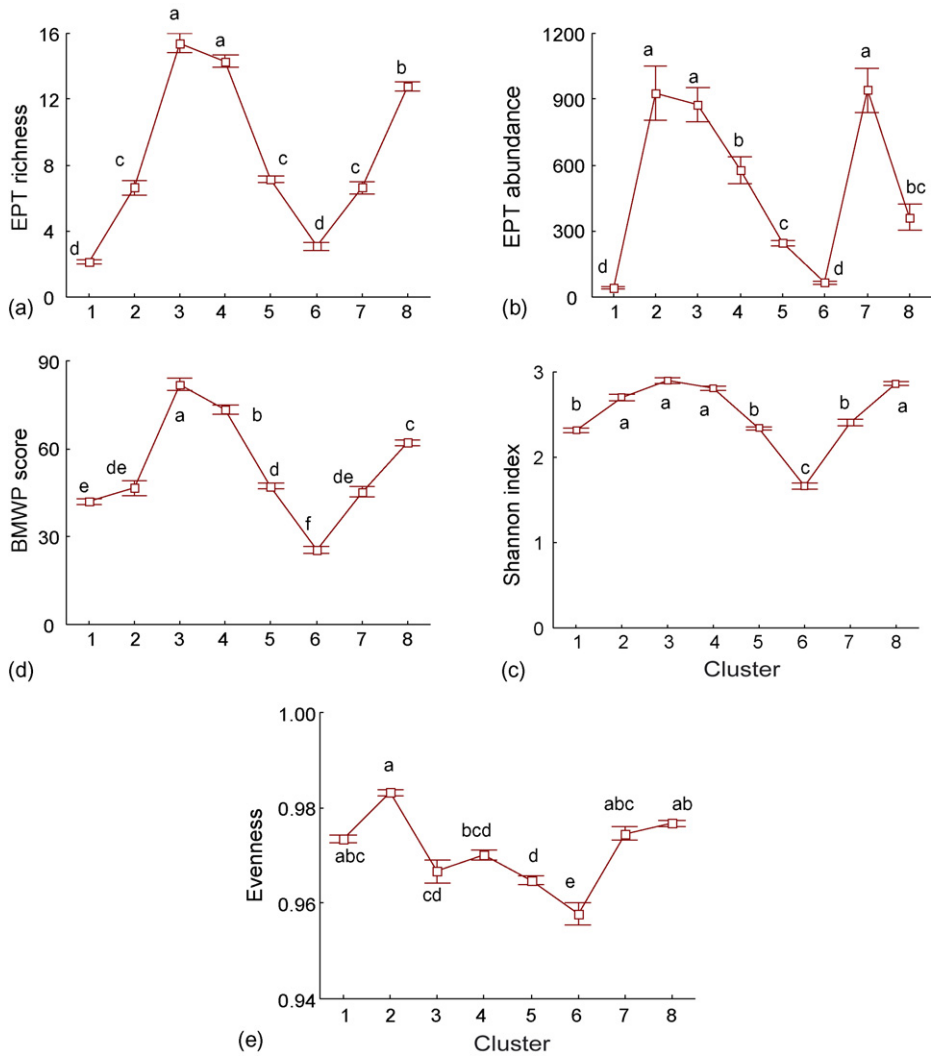

Fig. 6 – Variation in biological indices in different clusters according to the SOM (Fig. 3a). (a) EPT richness, (b) EPT abundance, (c) Shannon diversity index, (d) Biological Monitoring Working Party (BMWP) score, (e) evenness (different alphabets indicate significant differences between the clusters based on the Unequal N HSD multiple comparison test ($p = 0.05$). Error bars indicate mean and standard error of each variable).

eral similar to the patterns of overall abundance, although EPT abundance was in the highest range in clusters 2 and 7, while overall abundance was in the highest range only in cluster 7 (Fig. 4b). BMWP score was between 0 and 146, showing the highest value in cluster 3 and the lowest value in cluster 6 (Fig. 6c).

Community parameters also accordingly revealed the group characters of the clusters based on the trained SOM (Fig. 3a). As expected, Shannon diversity index was in the highest range in cluster 3, and was in the lowest range in cluster 6 (Fig. 6d). Evenness appeared to be lower in cluster 6 along with low values of species richness and abundance, while the

Table 1 – Correlation coefficients among community parameters and biological indices used in the datasets

	Species richness	Abundance	EPT richness	EPT abundance	BMWP score	Shannon index
Abundance	0.192					
EPT richness	0.878	0.141				
EPT abundance	0.414	0.504	0.407			
BMWP score	0.897	0.129	0.849	0.347		
Shannon index	0.891	0.173	0.722	0.337	0.841	
Evenness	0.186	0.122	0.108	0.1491	0.162	0.844

Table 2 – Community parameters, indicator species and environmental descriptions in different clusters

Cluster	Species richness	Abundance	Environments	Indicator species
1	Intermediate	Low	Lowland, pool zone	*Physa acuta*, *Pantala flavescens*
2	Intermediate	High	High conductivity	*Baetis nla*, *Ephemera strigata*
3	High	High	High altitude	*Ecdyonurus kibunensis*, *Epeorus latifolium*
4	High	Intermediate	High altitude	*Ecdyonurus kibunensis*, *Epeorus curvatulus*
5	Intermediate	Low	Intermediate altitude	*Uracanthella rufa*, *Ecdyonurus levis*
6	Intermediate	Low	Intermediate conductivity	*Baetis ursinus*, *Simulium uchidai*
7	Intermediate	High	Low depth	*Limnodrilus socialis*, *Uracanthella rufa*
8	High	Low	High altitude	*Semisulcospira libertine*, *Epeorus latifolium*

index showed the highest value in cluster 8 with high species richness and low abundance (Fig. 6e).

The patterns on the SOM well reflected the degree of association among water quality indices. Correlation coefficient was 0.878 between overall species richness and EPT richness were higher, 0.897 between overall species richness and BMWP score, and was 0.849 between EPT richness and BMWP score (Table 1). Shannon diversity index showed the higher correlation coefficients with species richness, EPT richness, and evenness.

Each cluster was also characteristically associated with the indicator species in the corresponding clusters. In Table 2, indicator species (or taxa) characterizing each cluster were presented along with summary of species richness and abundance.

4. Discussion and conclusion

In this study, we classified benthic macroinvertebrate communities collected at the relatively less polluted areas (e.g., mountains, suburbs, etc.) for the purpose of conservation and diversity recording. The SOM accordingly classified the sample sites to eight groups based on species compositions. The classification matched with geographical distribution of the sampling sites, and showed that spatial variation was the main factor for characterizing benthic macroinvertebrate communities collected in Korea on a large scale (Fig. 3). Geographic location was effectively identified with the clusters according to the trained SOM.

Moreover, we demonstrated that the geographically identified communities matched with the species–abundance patterns (Fig. 1b–i): lowland area (cluster 1, Fig. 1b), the limestone area (cluster 2, Fig. 1c), mountain area (clusters 3 (Fig. 1d), 4 (Fig. 1e), and 8 (Fig. 1i)), and slightly organic polluted area (cluster Fig. 1h). The SOM demonstrated feasibility in mapping of the clusters regarding provision of information on geographic distribution and species–abundance patterns at the same time.

Regional or national surveys of the stream ecosystems provide large volumes of the site-specific data, which may carry some valuable information to derive certain spatial patterns of biological communities on different scales (i.e. from a local to a regional area). It has been a key issue in defining community patterns on a large scale for sustainable ecosystem management. There have been many attempts to classify streams based on the distribution patterns of aquatic organisms (Huet, 1954; Holmes et al., 1998; Wright et al., 1998, 2000; Cowx et al., 2004). Wasson et al. (2002) defined hydro-ecoregions in surface water bodies in France based on geology, relief, and climate, and validated the classification with benthic macroinvertebrate fauna in streams. Tison et al. (2005) also validated the hydro-ecoregions in France with diatoms in streams.

Through the learning process of the SOM, we demonstrated that the characteristics of the samples on a large scale were distinctively identified in the clusters. The samples from mountain streams were grouped in cluster 8, showing high species richness and low abundance (Fig. 4). Samples from the disturbed areas due to the rehabilitation and construction on streams were grouped together in cluster 6, showing extremely low species richness and low abundance. This is commonly observed in the physically disturbed areas with poor condition of colonization by aquatic organisms.

Efficiency of mapping was further demonstrated in the clusters located closely. The sampling sites of clusters 4 and 5, for instance, were close each other in the geographical map (Fig. 3c). However, their community compositions were greatly different: while the samples in cluster 4 (Fig. 1e) showed higher species richness in a broad range, the samples in cluster 5 (Fig. 1f) were in the lower range for both species richness and abundance. This feasibility in identifying critical characteristics by the SOM could be efficiently used for characterizing typology of the river systems.

However, to clearly define typology based on the benthic macroinvertebrates, we need the sufficient number of reference sites along with records of physical and chemical environmental factors and sufficient geographical information. Further studies are required to define typologies in Korea on the national scale.

In conclusion, the SOM was efficient in patterning and visualizing community characteristics and accordingly classified the samples corresponding to the geographical regions and to the degree of disturbances. The communities clustered by the trained network were also characterized with the species-abundant patterns. The SOM was efficient in linking various information on community composition, geographic distribution, and 'species–abundance' patterns in an integrative manner. The SOM could serve as an efficient ecological map for specifying ecoregions and for providing comprehensive view on ecological status of the communities sampled on a large scale.

Acknowledgments

This work was supported by Korea Research Foundation and The Korean Federation of Science and Technology Societies

Grant funded by Korea Government (MOEHRD, Basic Research Promotion Fund).

REFERENCES

Alhoniemi, E., Himberg, J., Parhankangas, J., Vesanto, J., 2000. SOM Toolbox, http://www.cis.hut.fi/projects/somtoolbox.

Brittain, J.E., Saltveit, S.J., 1989. A review of the effect of river regulation on mayflies (Ephemeroptera). Reg. Riv.: Res. Manage. 3, 191–204.

Brown, J.H., 1981. Two decades of homage to Santa Rosalia: toward a general theory of diversity. Am. Zool. 21, 877–888.

Cowx, I.G., Noble, R.A., Nunn, A.D., Harvey, J.P., 2004. Flow and level criteria for coarse fish and conservation species. Environment Agency R&D Report, W6-096. Bristol.

Cummins, K.W., 1979. The natural stream ecosystem. In: Ward, J.V., Stanford, J.A. (Eds.), The Ecology of Regulated Streams. Plenum Press, New York, pp. 7–24.

Feminella, J.W., Resh, V.H., 1990. Hydrologic influences, disturbance, and intraspecific competition in a stream caddisfly population. Ecology 71, 2083–2094.

Hawkins, C.P., Norris, R.H., Gerritsen, J., Hughes, R.M., Jackson, S.K., Johnson, R.K., Stevenson, R.J., 2000. Evaluation of the use of landscape classifications for the prediction of freshwater biota: synthesis and recommendations. J. N. Am. Benthol. Soc. 19, 541–556.

Hellawell, J.M., 1978. Biological Surveillance of Rivers. Water Research Center, Stevenage Laboratory, England.

Holmes, N.T.H., Boon, P.J., Rowell, T.A., 1998. A revised classification system for British rivers based on their aquatic plant communities. Aquat. Conserv. -Mar. Freshwater Ecosyst. 8 (4), 555–578.

Huet, M., 1954. Biologie, profils en long et en travers des eaux courantes. Bull. Fran. Pisci. 175, 41–53.

Kohonen, T., 1982. Self-organized formation of topologically correct feature maps. Biol. Cybern. 43, 59–69.

Kohonen, T., 2001. Self-Organizing Maps. Springer, Berlin.

Lavandier, P., Décamps, H., 1984. Estaragne. In: Whitton, B.A. (Ed.), Ecology of European Rivers. Blackwell, London, pp. 237–264.

Lenat, D.R., 1988. Water quality assessment of streams using a qualitative collection method for benthic macroinvertebrates. J. N. Am. Benth. Soc. 7, 222–233.

MacArthur, R.H., 1965. Patterns of species diversity. Biol. Rev. 40, 510–533.

Malmqvist, B., Otto, C., 1987. The influence of substrate stability on the composition of stream benthos: an experimental study. Oikos 48, 33–38.

Ministry of Environment, 1997. The National Natural Environment Monitoring Protocol. Ministry of Environment, Seoul.

Minshall, G.W., Petersen, R.C., Nimz, C.F., 1985. Species richness in streams of different size from the same drainage basin. Am. Nat. 125, 16–38.

National Water Council, 1981. River Quality: The 1980 Survey and Future Outlook. National Water Council, London.

Park, Y.-S., Céréghino, R., Compin, A., Lek, S., 2003. Applications of artificial neural networks for patterning and predicting aquatic insect species richness in running waters. Ecol. Model. 160, 265–280.

Preston, F.W., 1962. The canonical distribution of commonness and rarity. Part I. Ecology 43, 185–215.

Rosenberg, D.M., Resh, V.H. (Eds.), 1993. Freshwater Biomonitoring and Benthic Macroinvertebrates. Chapman & Hall, London.

Smith, M.J., Kay, W.R., Edward, D.H.D., Papas, P.J., Richardson, K.St.J., Simpson, J.C., Pinder, A.M., Cale, D.J., Horwitz, P.H.J., Davis, J.A., Yung, F.H., Norris, R.H., Halse, S.A., 1999. AusRivAS: using macroinvertebrates to assess ecological condition of rivers in Western Australia. Freshwater Biol. 41, 269–282.

Snyder, E.B., Robinson, C.T., Minshall, G.W., Rushforth, S.R., 2002. Regional patterns in periphyton accrual and diatom assemblages structure in a heterogeneous nutrient landscape. Can. J. Fish. Aquat. Sci. 59, 564–577.

StatSoft, Inc., 2004. STATISTICA (data analysis software system), version 7, www.statsoft.com.

Stevenson, R.J., 1997. Scale-dependent determinants and consequences of benthic algal heterogeneity. J. N. Am. Benth. Soc. 16, 248–262.

The Mathworks Inc., 2001. MATLAB Version 6.1. The Mathworks, Inc., Massachusetts.

Tison, J., Park, Y.-S., Coste, M., Wasson, J.G., Ector, L., Rimet, F., Delmas, F., 2005. Typology of diatom communities and the influence of hydro-ecoregions: a study at the French hydrosystem scale. Water Res. 39, 3177–3188.

Vannote, R.L., Minshall, G.W., Cummins, K.W., Sedell, J.R., Cushing, C.E., 1980. The river continuum concept. Can. J. Fish. Aquat. Sci. 37, 130–137.

Vesanto, J., Himberg, J., Alhoniemi, E., Parhankangas, J., 1999. Self-organizing map in Matlab: the SOM Toolbox. In: Proceedings of the MATLAB Digital Signal Processing Conference, Espoo, Finland, pp. 35–40.

Walley, W.J., Hawkes, H.A., 1997. A computer-based development of the Biological Monitoring Working Party score system incorporating abundance rating, biotope type and indicator value. Water Res. 31 (2), 201–210.

Ward, J.V., Stanford, J.A., 1979. Ecological factors controlling stream zoobenthos with emphasis on thermal modification of regulated streams. In: Ward, J.V., Stanford, J.A. (Eds.), The Ecology of Regulated Streams. Plenum Press, New York, pp. 35–55.

Ward, J.V., Stanford, J.A., 1983. The intermediate disturbance hypothesis: an explanation for biotic diversity patterns in lotic systems. In: Fontaine, T.D., Bartell, S.M. (Eds.), Dynamics of Lotic Ecosystems. Ann Arbor Sciences, Ann Arbor, Michigan, pp. 347–356.

Wasson, J.G., Chandesris, A., Pella, H., Blanc, L., 2002. Typology and reference conditions for surface water bodies in France—the hydro-ecoregion approach. In: Proceeding of Typology and Ecological Classification of Lakes and Rivers, Finnish Environment Institute (SYKE), Helsinki, Finland, October 25–26, 2002 (published in TemaNord).

Wright, J.F., Sutcliffe, D.W., Furse, M.T. (Eds.), 2000. Assessing the Biological Quality of Fresh Waters: RIVPACS and Other Techniques. Freshwater Biological Association, Ambleside.

Wright, J.F., Furse, M.T., Moss, D., 1998. River classification using invertebrates: RIVPACS applications. Aquat. Conserv. -Mar. Freshwater Ecosyst. 8 (4), 617–631.