ELSEVIER

Ecological Modelling 132 (2000) 151–166

ECOLOGICAL MODELLING

www.elsevier.com/locate/ecolmodel

# Determining temporal pattern of community dynamics by using unsupervised learning algorithms

Tae-Soo Chon [a,*], Young-Seuk Park [a], June Ho Park [b]

[a] *Department of Biology, Pusan National University, Pusan 609-735, South Korea*
[b] *Department of Electrical Engineering, Pusan National University, Pusan 609-735, South Korea*

**Abstract**

Analysis of patterns of temporal variation in community dynamics was conducted by combining two unsupervised artificial neural networks, the Adaptive Resonance Theory (ART) and the Kohonen network. The field data used as input for training represented monthly changes in density and species richness in selected taxa of benthic macroinvertebrates collected in the Suyong River in Korea from September 1993 to October 1994. The sampled data for each month was initially trained by ART, the weights of which preserved conformational characteristics among communities during the process of the training. Subsequently these weights were rearranged sequentially from 2 to 5 months, and were provided as input to the Kohonen network to reveal temporal variations in communities. The network was then able to extract the features of community dynamics in a reduced dimension covering the specified input period.


*Keywords:* Community dynamics; Pattern analysis; Temporal variation; Artificial neural network; Adaptive Resonance Theory; Kohonen network; Unsupervised learning algorithm; Benthic macroinvertebrate community

## 1. Introduction

Data for community dynamics are difficult to analyze as they describe many species varying in different locations and times in a non-linear fashion. Although there are numerous accounts and reviews on various topics in communities (Strong et al., 1984; Mallin and Paerl, 1994; McPeek and Miller, 1996) and on the classification of communities through multivariate analyses in ecology (Legendre and Legendre, 1987; Ludwig and Reynolds, 1988), few studies have been conducted on patterning community dynamics per se. Historically communities were classified in static terms, not in dynamic terms. Although the data might have been collected sequentially over varying times, temporal interpretation was usually made on the results after conventional multivariate analyses had been conducted in static terms, mainly based on ordination technique and correspondence analysis (Bunn et al., 1986; Quinn et al., 1991; Stromberg and Griffin, 1996).

In describing dynamics in succession quantitatively, Tanner et al. (1994) used sensitivity analysis based on transition matrix. However this

* Corresponding author. Tel.: +82-51-5102261; fax: +82-51-5812962.

0304-3800/00/$ - see front matter 
PII: S0304-3800(00)00312-4

study, based on matrix theory, was more oriented to examining the mechanism of community dynamics in relation to stability problem and estimating how rapidly community equilibrium is regained after a disturbance. Law and Morton (1996) also worked on assembly of communities by utilising the concept of Lotka-Volterra dynamics and average Lyapunov function so that 'permanence' of assembly of communities could be investigated. This study also mainly focused on stability of communities to see if a certain combination of species could coexist as time proceeds. Legendre et al. (1985) and Legendre (1987) reviewed methods for classifying communities in temporal domain including ordination and segmentation techniques in a multivariate data series, and proposed a method for chronological clustering to represent the succession of species within a community.

Development of methods to describe patterns in community dynamics, however, is currently a topic of considerable interest in ecosystem management. It is necessary to characterise the sequence of community dynamics if the target ecosystem needs to be assessed. This requires methods that characterise how a community varies in space and over time simultaneously, caused either by natural agents or anthropogenic disturbances. Artificial neural networks have been widely applied in interpreting complex and nonlinear phenomena in machine intelligence in engineering (Lippmann, 1987; Wasserman, 1989; Hecht-Nielsen, 1990; Zurada, 1992). The networks are problem oriented and flexible, and analytical solutions are not necessarily required before running them. The capability of learning in artificial neural networks has been recently applied in natural sciences in the field of chemistry, for example, regarding clustering high dimensional data on molecule structures (Ferrán and Ferrara, 1992; Melssen et al., 1993; Reibnegger et al., 1993).

In ecology, artificial neural networks have been applied to classifying groups (Chon et al., 1996; Levine et al., 1996), or patterning relationships between variables (Huntingford and Cox, 1996; Lek et al., 1996). Recently Chon et al. (1996) utilised the Kohonen network to classify community data. Artificial neural networks were also implemented in estimating time development of populations (Elizondo et al., 1994; Boudjema and Chau, 1996; Recknagel et al., 1997; Stankovski et al., 1998). Grassland community changes were predicted by Tan and Smeins (1996). They were mostly based on the backpropagation algorithm for patterning the relational effects with environmental factors.

One advantage of artificial neural network is its flexibility in combining with other neural networks, such as multi-layer perceptron (Rumelhart et al., 1986; Lippmann, 1987) and counterpropagation (Hecht-Nielsen, 1987, 1990), which enhanced flexibility of perceiving non-linearity of input data. In this study we attempted to use the combined networks for patterning community dynamics. However, the above-mentioned networks were not effective for this purpose, as the former requires training under supervised learning (Lippmann, 1987) while the latter is mainly utilised in learning relational characteristics between the variable sets (Hecht-Nielsen, 1990). Our objective is to see what patterns would form on the time basis if various community data were given at different sites and at different sampling times. We describe a method to decipher temporal patterns, without previous knowledge of patterns of community changes. This requires unsupervised learning on sequential input data. By elaborating the flexibility and the capability of information extraction by the artificial neural networks we now attempt to combine two unsupervised networks to pattern temporal variation in community dynamics in this study.

## 2. Methods

### 2.1. Network process

The schematic diagram of applying the combined unsupervised networks, the Adaptive Resonance Theory (ART; Carpenter and Grossberg, 1987) and the Kohonen network (Kohonen, 1989), in determining temporal pattern is shown in Fig. 1. Initially all the community data for

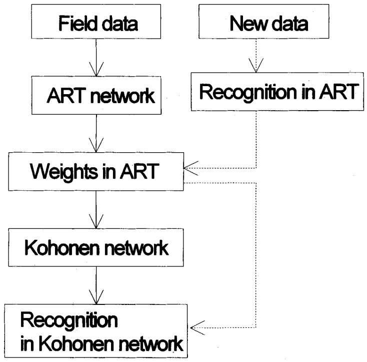


Fig. 1. A flowchart for training and recognition of temporal variations in community dynamics with two unsupervised learning algorithms (artificial neural network), the Adaptive Resonance Theory (ART) and the Kohonen network.

one-time sampling from the field were trained by ART (Figs. 1 and 2). A modified algorithm in ART (Pao, 1989) was used in this study. Bottom up weights, $b_{ji}(0)$, between output node $j$ and input node $i$ were initialised with some small numbers. After the input $x_i$, density and species richness in selected taxa, is given to the network, distance, $d_j(t)$, for each output node, $j$, is calculated as follows:

$$d_j(t) = \sqrt{\sum_{i=0}^{n-1} [b_{ji}(t) - x_i]^2}$$

where $n$ is the number of input nodes. The distance, $d_j(t)$, measures the degree of similarity between weights and input data, and is used as a criterion for grouping inputs through the training process.

As each new input enters the network the distance is calculated and the output node $j$ which has minimum distance is selected as $j*$. If $d_{j*}(t)$ is

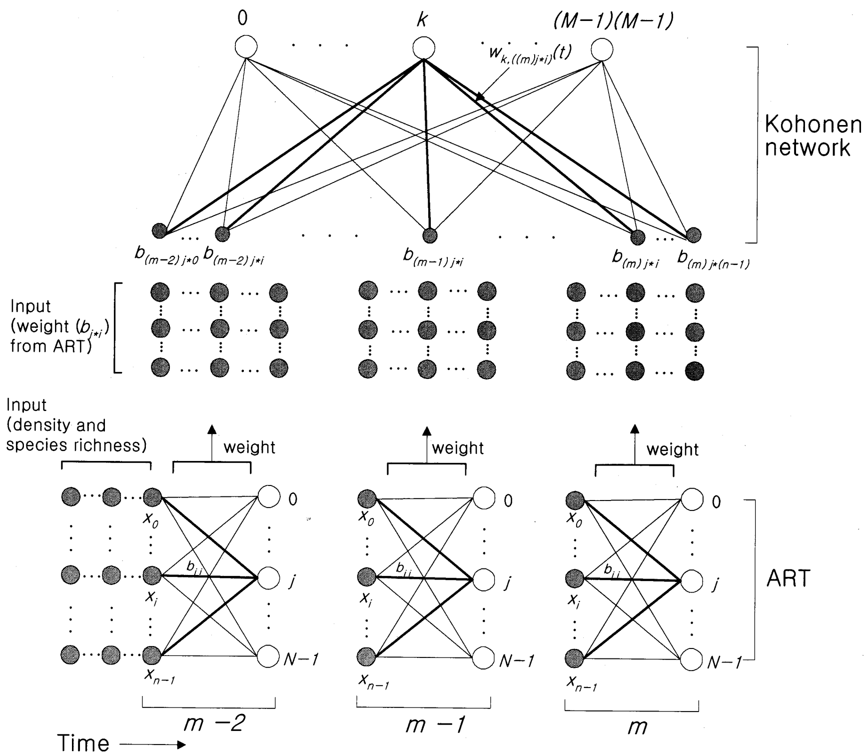


Fig. 2. A schematic diagram representing the algorithm for the combined network of ART (16 × 11) and Kohonen ((3 × 16) × (9 × 9)) for training temporal variations in community dynamics ($b_{ji}$, bottom up weights between output node $j$ and input node $i$ in ART; $b_{j*i}$, converged weight of ART which is used as input data in Kohonen network; $m$, sampling month in the sequential period; $M$; order of output node (9) for the Kohonen network; $n$, number of input nodes (16) for ART; $N$, number of output nodes (11) for ART; $x_i$, input data at the node $i$ in ART; $w_{k,(mj*i)}(t)$; weight in Kohonen network, see text for explanation in detail).

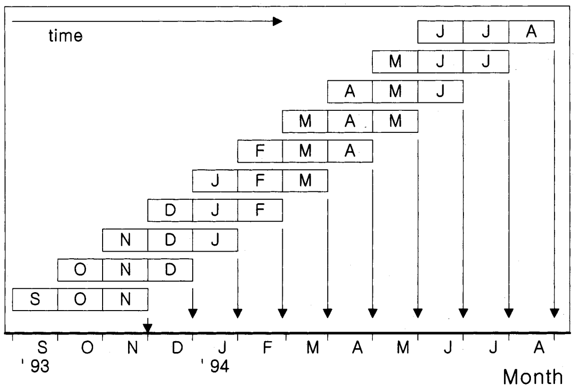


Fig. 3. An example of input of weights when 3-month changes in communities were to be trained with the Kohonen network. At each target time weights produced in ART in 2 previous months were appended to those for the target month for training.

smaller than $\rho$, which is a threshold parameter in determining vigilance, the input is assigned to output node $j*$. $\rho$ was empirically determined in this study as 0.61 based on efficiency of grouping of input data. The weight for the node $j*$, $b_{j*i}(t)$, then, is updated as follows:

$$b_{j*i}(t+1) = \frac{c}{c+1} b_{j*i}(t) + \frac{1}{c+1} x_i$$

where $c$ is the number of sample units classified to node $j*$.

If $d_{j*}(t)$ is larger than $\rho$ the input is assigned to new output. This means that the entering input forms to a new pattern, not belonging to one of the previously existing patterns. Then its weight $b_{j*i}(t)$ is newly assigned as follows:

$$b_{j*i}(t+1) = x_i.$$

These weights produced by ART preserve the conformational characteristic of input data for grouping, and through them the associations among the communities are projected into the space defined in ART (Zurada, 1992). Since the 1-month sampling data were characterised by weights, it was supposed that, if the weights for previous months were appended to the target month, they would represent temporal changes in the community during the specified period. Based on this reason the weights trained for 1 month in ART were combined sequentially and used in the Kohonen network as input. If community changes in 3 months were to be analyzed by the Kohonen network, weights obtained by ART for 3 months were aligned sequentially (Fig. 3). If March was the target month for training, for example, weights for the 2 previous months were appended in front of March (i.e. January–February–March). In total $T \times n$ weights were used for input where $T$ and $n$, respectively, represented the number of months and variables for training as input for ART. This is equivalent to creating a window of a specified input period (3 months in this case) and scanning all the target sampling data through the survey period.

The Kohonen network is good for self-organization and mapping the data feature in a reduced dimension. In the Kohonen network an array of $T \times n$ artificial neurons, the number of weights produced in ART after the sequential arrangement, was used for the input layer. For output, $M^2$ neurons were used. The number of output neurons could be empirically determined based on the efficiency of convergence and capability of discrimination of patterned neurons. In this case a two-dimensional square array of order nine was used. The weights in the Kohonen network were represented as $w_{k,(mj*i)}(t)$. Since $j*$ was determined as a winner in ART and all winner neurons were selected for input to the Kohonen network, $j*$ was set to a constant for training. Among designating digits for input node, sampling month in the sequential period, $m$, and input node for ART, $i$, were varied in this case. Similar to ART, the weights were randomly assigned in small values initially for convergence. When the input vector was put through the network, the summed distance, $d'_k(t)$, for each output neuron, $k$, between input vector and weights was given as:

$$d'_k(t) = \sum_{m=0}^{T-1} \sum_{i=0}^{n-1} [b_{j*i}(t) - w_{k,(mj*i)}(t)]^2.$$

The neuron, the weight vector of which has the shortest distance to the input vector, was chosen to be the winning neuron. The winning neuron (and possibly its neighboring neurons) was allowed to learn by changing the weights in the manner to further reduce the distance between the weight and the input vector as shown below:

$$w_{k,(mj*i)}(t+1) = w_{k,(mj*i)}(t) + \eta(t)[b_{mj*i} - w_{k,(mj*i)}(t)]Z_{mj*i}\ .$$

$Z_{(mj*i)}$ is assigned 1 for the winning and its neighbor neurons while it is assigned 0 for the rest of the neurons. This allows adjusting weight for the winning and neighbor neurons through the training process. $\eta(t)$ denotes the fractional increment of the correction and, initially chosen to be close to 0.4, was decreased gradually in an arbitrary scale to a small value 0.1 as iterative calculation at $t$ proceeded to convergence in this study. For details of the Kohonen network, including the determination of neighbor neurons and fractional increment of the correction, see Lippmann (1987), Kohonen (1989), Hecht-Nielsen (1990), Zurada (1992) and Chon et al. (1996). After training was complete, new data for community changes collected for a specified period were provided to the trained network for recognition (Fig. 1). New data were given to ART and weights were subsequently updated. The updated weights were then arranged sequentially for a given period and were provided to the trained Kohonen network for recognition.

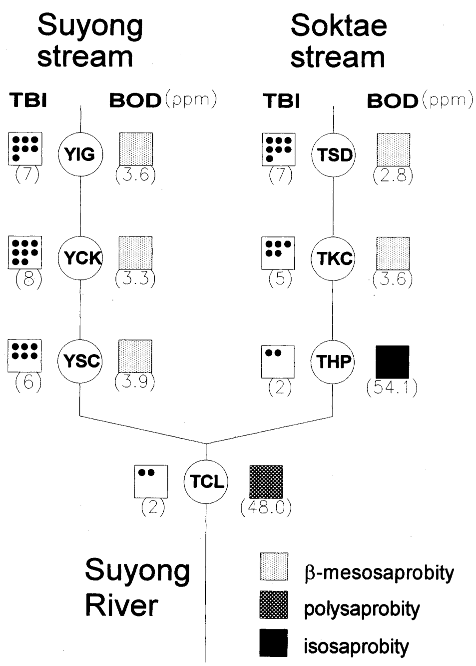


Fig. 4. Saprobity, BOD (in ppm) and TBI (Trent Biotic Index) in averages at the sample site for benthic macroinvertebrates in the Suyong River in Korea from September 1993 to August 1994 (modified from Kwon and Chon (1993) and Kang et al. (1995)). The alpha codes, three characters in the figure, designate the names of sample sites: TCL, Chungli; THP, Hapansong; TKC, Kochon; TSD, Sadeungkol; YCK, Changki; YIG, Imgog; YSC, Shinchon.

## 2.2. Field data

Benthic macroinvertebrate communities were collected monthly in the Suyong and Soktae streams of the Suyong River from September 1993 to August 1994. Saprobity was measured at each location (BOD in ppm), and a wide range of organic pollution occurred over a period at the study sites from oligosaprobity to polysaprobity (Kwon and Chon, 1993; Kang et al., 1995). Community status indicating water quality, measured by the Trent Biotic Index (TBI; Woodiwiss, 1964), correspondingly varied over a wide range (Fig. 4). Since the number of species (132) was too large to train, species were grouped into seven selected taxa: Chironomidae, Diptera (except Chironomidae), Trichoptera, Ephemeroptera, miscellaneous Insecta, Oligochaeta, and miscellaneous Macroinvertebrates. The selected taxa represented the ecological status of water quality in the studied sites of the Suyong River (Kang et al., 1995). Densities (number of individuals per m$^2$) and species richness (number of species) in each taxa, as well as the total density and the total richness, were given as input for training with ART. The total number of input node for ART was 16. Monthly changes in densities and species richness at the sampled sites are shown in Figs. 5 and 6. Ecological information for benthic macroinvertebrates in the Suyong River was reported in Kwon and Chon (1993) and Kang et al. (1995). When samples were missing, mean values between the immediately preceding and immediately following sample were used for representing the input data for the target time. A period length of 5 months was used as input.

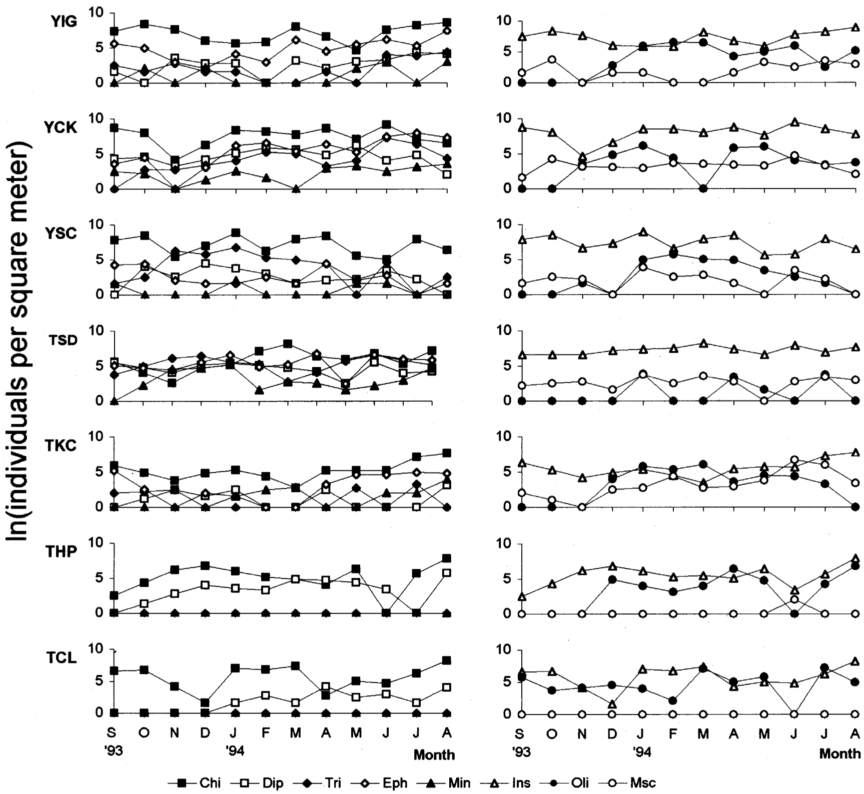


Fig. 5. Monthly changes in densities (number of individuals per m$^2$) in selected taxa of benthic macroinvertebrate communities collected at the sample sites in the Suyong River from September 1993 to August 1994 (Chi, Chironomidae; Dip, Diptera except Chironomidae; Eph, Ephemeroptera; Ins, total Insecta; Min, miscellaneous Insecta; Msc, miscellaneous Macroinvertebrates; Oli, Oligochaeta; Tri, Trichoptera) (modified from Kang et al. (1995)). The alpha codes designating sample sites are explained in the caption of Fig. 4.

## 3. Results and discussion

### 3.1. Classification of one-time samples

Fig. 7a is an example of classification of communities after training by ART. The number of trained communities were 84, while the numbers of input and output nodes for ART were 16 and 11, respectively. In overall terms, classification by the network reflected environmental impacts. Sample communities collected from polluted sites at TCL and THP (Fig. 4) were grouped closely, and were separated from the other less-polluted sites. The community data collected from TKC, which were in a medium range between clean and polluted status, appeared diverse, and were divided into many small groups, separately (e.g. TKC4-2 and TKC4-3 of neuron 7 in Fig. 7a), or in groups with the other similar sites (e.g. TKC3-9 and TKC4-8 of neuron 0 in Fig. 7a). Although BOD levels were similar between TKC and other less-polluted sites including YIG, YCK and TSD, the biological indicator TBI appeared differently among sampled sites, five on TKC while seven and eight on less-polluted sites (Fig. 4). This indicated that the community mapping by the network reflected TBI more effectively than BOD in this case.

Among less-polluted sample sites, communities were generally grouped according to topographical conditions. Sample sites from the same stream had a higher tendency of grouping. Communities collected from YIG, YCK and YSC from the Suyong stream (Fig. 4) showed generally higher chance of grouping. Sample communities from TSD, a less-polluted site in the Soktae stream, were generally separated from the other communities in less-polluted sites in the Suyong stream.

To see if the weights produced by ART correctly extracted the data features and carry the information through the subsequent process, training of the Kohonen network ($9 \times 9$ neurons) was conducted directly on weights produced from ART using sample data from a single month (Fig. 7b). The mapping by the Kohonen network correspondingly reflected the characteristics observed at the classification results from ART, while ART had used the raw data as input (Fig. 7a): The polluted sites (THP and TCL) were closely located and separated from the less-polluted sites. Communities collected from the medium pollution in TKC formed small groups widely dispersed on the map, while the less-polluted sample sites were generally divided based on topographical conditions. This confirmed that the features of the input data were accordingly conveyed to the Kohonen training through the weights produced in ART.

### 3.2. Patterns of monthly changes in communities

Fig. 8 shows results after the Kohonen training of temporal variations from 2 to 5 months. In general the trained results reflected environmental impacts and showed the general characteristics as observed in the results from 1-month samples (Fig. 7a, b): grouping was mainly based on pollution level and topographical condition.

In the 2-month sequences (Fig. 8a), several

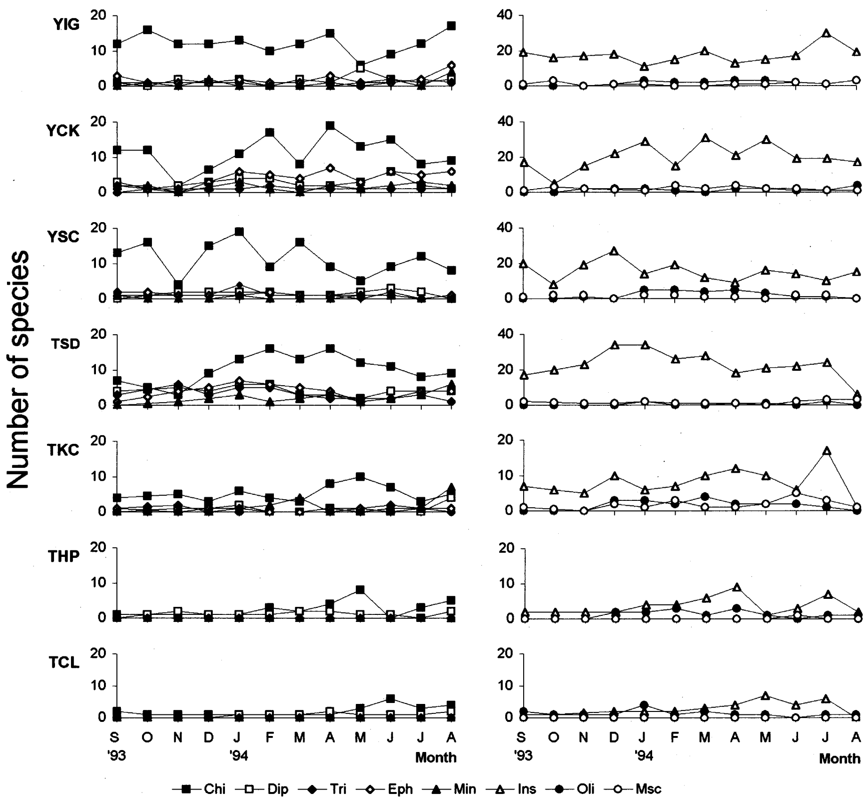


Fig. 6. Monthly changes in number of species in selected taxa of benthic macroinvertebrate communities collected at the sample sites in the Suyong River from September 1993 to August 1994 (modified from Kang et al. (1995)). The selected taxa names are explained in the caption of Fig. 5 and the alpha codes designating sample sites are explained in the caption of Fig. 4.

a)

| Neuron | Number | Sample community grouped by the network |
|---|---|---|
| 0 | 14 | YIG3-9 YIG3-10 YIG3-11 YIG4-7 YCK3-9 YCK3-10 YCK4-3 YSC3-9 YSC3-10 YSC3-12 YSC4-7 TSD3-9 TKC3-9 TKC4-8 |
| 1 | 12 | YIG3-12 YIG4-1 YIG4-2 YIG4-3 YIG4-4 YIG4-5 YCK3-12 YSC4-2 YSC4-4 TKC4-1 TKC4-4 TKC4-5 |
| 2 | 12 | YIG4-6 YIG4-8 YCK4-1 YCK4-2 YCK4-4 YCK4-5 YCK4-6 YCK4-7 YCK4-8 TSD4-1 TSD4-4 TSD4-7 |
| 3 | 6 | YCK3-11 YSC3-11 YSC4-6 YSC4-8 TSD4-5 TKC3-10 |
| 4 | 2 | YSC4-1 YSC4-3 |
| 5 | 3 | TSD3-10 TSD3-11 TSD3-12 |
| 6 | 4 | TSD4-2 TSD4-3 TSD4-6 TSD4-8 |
| 7 | 2 | TKC4-2 TKC4-3 |
| 8 | 2 | TKC4-6 TKC4-7 |
| 9 | 8 | TKC3-11 THP3-9 THP3-10 THP3-11 THP4-6 TCL3-11 TCL3-12 TCL4-6 |
| 10 | 19 | YSC4-5 TKC3-12 THP3-12 THP4-1 THP4-2 THP4-3 THP4-4 THP4-5 THP4-7 THP4-8 TCL3-9 TCL3-10 TCL4-1 TCL4-2 TCL4-3 TCL4-4 TCL4-5 TCL4-7 TCL4-8 |

b)

| | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 |
|---|---|---|---|---|---|---|---|---|---|
| 0 | | | | | | THP3-9 TKC3-10<br>TKC3-11 THP3-11<br>TCL3-11 TCL3-12<br>THP4-6 TCL4-6 | | | |
| 1 | | | | | | | | | |
| 2 | | | | | | | | | |
| 3 | TCL3-9 TCL3-10<br>TKC3-12 THP3-12<br>THP4-1 TCL4-1<br>THP4-2 TCL4-2<br>THP4-3 TCL4-3<br>THP4-4 TCL4-4<br>YSC4-5 THP4-5<br>TCL4-5 THP4-7<br>TCL4-7 THP4-8<br>TCL4-8 | | | | | | | YIG3-12 YCK3-12<br>YIG4-1 TKC4-1<br>YIG4-2 YSC4-2<br>YIG4-3 YIG4-4<br>YSC4-4 TKC4-4<br>YIG4-5 TKC4-5 | |
| 4 | | | | TKC3-10 YCK3-11<br>YSC3-11 TSD4-5<br>YSC4-6 YSC4-8 | | | | | |
| 5 | | | | | | | | | |
| 6 | | | | | | | | | |
| 7 | TSD3-10 TSD3-11<br>TSD3-12 | TSD4-2 TSD4-3<br>TSD4-6 TSD4-8 | YCK4-1 TSD4-1<br>YCK4-2 YCK4-4<br>TSD4-4 YCK4-5<br>YIG4-6 YCK4-6<br>YCK4-7 TSD4-7<br>YIG4-8 YCK4-8 | YIG3-9 YCK3-9<br>YSC3-9 TSD3-9<br>TKC3-9 YIG3-10<br>YCK3-10 YSC3-10<br>THP3-10 YIG3-11<br>YSC3-12 YCK4-3<br>YIG4-7 YSC4-7<br>TKC4-8 | TKC4-2 TKC4-3 | | | | |
| 8 | | | | | YSC4-1 YSC4-3 | | TKC4-6 TKC4-7 | | |

Fig. 7. Patterns of benthic macroinvertebrates in 1-month samples, collected in the Suyong River as structured by ART (a), and the Kohonen network (b). In (a), sample communities associated with specified neurons through ART training are listed in groups. In (b), the neurons were arranged in two dimensions, and sample communities patterned by the Kohonen network to a specified neuron ($i$, $j$ ) are grouped together in the associated table position ($i$, $j$ ). The alpha codes for the names of sample communities are explained in the caption of Fig. 4. The first numerical digit appearing after the alpha codes represents the year of collection (i.e. 3 for 1993 and 4 for 1994) while the second numerical digit following the dash stands for the month of collection (e.g. 1 for January, 2 for February, etc.).

large groups appeared. A large number of samples from polluted sites of TCL and THP were grouped together (group A) at (6; $x$ axis, 0; $y$ axis) in Fig. 8a. This was also the case in the 1-month mapping (Fig. 7b). Many sample communities from YCK and TSD formed another large group (group B) at (3, 7) of the map. Communities collected at YIG in the early part of 1994 also formed a group (group C) at (6, 3). Communities from TKC were widely spread on the map with small groups, similar to Fig. 7b. In contrast to 1-month sampling, however, a slight difference was observed in the 2-month map. Sample data from TSD were absorbed into group B in the 2-month map (Fig. 8a). Although separately located in the 1-month map (Fig. 7b), most of them were also close to a large group mainly consisting of YCK data, which is in fact group B. Sample data from YSC, which were mostly located close to, or inside group B in Fig. 7b, tended to drop

a)

| | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 |
|---|---|---|---|---|---|---|---|---|---|
| 0 | | | THP3-10 THP3-11 TCL3-12 | | YSC4-2 YSC4-4 | TKC4-4 | TCL3-10 THP4-1 THP4-2 THP4-3 THP4-4 THP4-5 THP4-6 TCL4-2 TCL4-3 TCL4-4 TCL4-5 TCL4-6 | | |
| 1 | | | | | | | TCL3-11 TCL4-6 THP4-6 | | |
| 2 | TSD3-11 TSD3-12 | | | | TKC3-11 YCK3-12 | | TKC4-1 | | TKC3-12 TKC4-1 THP3-12 THP4-7 TCL4-7 |
| 3 | | | | | | | YIG4-1 YIG4-2 YIG4-3 YIG4-4 YIG4-5 TKC4-5 | | |
| 4 | TKC4-7 TKC4-8 | TKC4-6 | TKC4-2 | | TKC4-3 | | | | |
| 5 | | | | | | | YSC4-5 | | |
| 6 | | | | | | | | | |
| 7 | YSC4-1 | | YSC4-6 YSC4-6 | YIG3-10 YIG3-11 YIG4-6 YIG4-7 YCK3-10 YCK4-1 YCK4-4 YCK4-5 YCK4-6 YCK4-7 YCK4-8 YSC3-10 YSC4-3 TSD3-10 TSD4-1 TSD4-2 TSD4-3 TSD4-4 TSD4-6 TSD4-7 TSD4-8 | | TKC3-10 YCK3-11 YCK4-2 YSC3-11 YSC4-8 | | | |
| 8 | YIG3-12 | | YSC3-12 | YCK4-3 TSD4-5 YIG4-7 | | | | | |

Fig. 8. Mapping of benthic macroinvertebrate communities collected in the Suyong River when the temporal variations were trained by the Kohonen network from the period of 2–5 months. (a) 2 months; (b) 3 months; (c) 4 months; and (d) 5 months. The alpha codes and numerical digits designating the name of sample units are explained in the captions of Figs. 4 and 7.

b)

| | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 |
|---|---|---|---|---|---|---|---|---|---|
| 0 | THP3-11 TKC4-8 | | | YSC4-3 YSC4-5 | YCK3-11 YSC3-11 | YIG3-12 YCK3-12 YSC4-2 YSC4-4 | TKC3-11 | | |
| 1 | YIG3-11 YIG4-6 YIG4-7 YIG4-8 YCK4-1 YCK4-2 YCK4-3 YCK4-4 YCK4-5 YCK4-6 YCK4-7 YCK4-8 YSC4-1 YSC4-8 TSD3-11 TSD3-12 TSD4-1 TSD4-2 TSD4-3 TSD4-4 TSD4-5 TSD4-6 TSD4-7 TSD4-8 | YSC4-7 | | | | | YSC3-12 | YIG4-1 | YIG4-2 YIG4-3 YIG4-4 YIG4-5 YSC4-6 |
| 2 | TKC4-6 TKC4-7 | | | | THP4-2 THP4-3 THP4-4 THP4-5 TCL4-3 TCL4-4 TCL4-5 | | | | TKC4-2 |
| 3 | TKC4-1 | | THP4-1 THP4-8 TCL4-2 TCL4-8 | | | | THP3-12 TCL4-1 | | TKC4-5 |
| 4 | | | | | | | | | TKC4-4 |
| 5 | | | | TKC4-3 | | | | | |
| 6 | | | | | | | | TCL3-11 TCL4-6 THP4-6 | |
| 7 | THP4-7 TCL4-7 | | | TCL3-12 | | | | | |
| 8 | | | | | | | | | TKC3-12 |

Fig. 8. (*Continued*)

off at group B in the 2-month map. Similar to the situation of TSD, however, sample communities did not move far way from group B. This suggested in overall terms, although there were slight changes between 1- and 2-month maps, that over time the general characteristics of communities did not change on a larger scale.

In the maps describing community pattern of periods longer than 2 months, the characteristics shown in the 2-month map were generally preserved (Fig. 8b, c, d). Most sample communities in groups B and C were consistently found inside the groups as the input period increased. Sample communities from TKC and YSC also showed a similar tendency as seen in the 2-month map (Fig. 8a). In group A, however, the size of samples was gradually decreased. In the 3-month map, for example, THP4-1, THP4-6, TCL4-2 and TCL4-6 were separated from group A, while in groups B and C only slight changes were observed. In group B all sample communities in the 2-month map were preserved and two communities, YCK4-2 and YCK4-3, joined additionally, while one sample community, YIG4-1, moved from group C. As the sampling period increased, communities that belonged to group A were further disbanded to produce the many small isolated groups appearing on the last 5-month map (Fig. 8d). This indicated that variations in community dynamics increased as the observation time increased at polluted sites. This might be related to population dynamics of dominant species at polluted sites, such as *Limnodrillus hoffmeisteri* in Oligochaeta and *Chironomus flaviplumus* in Diptera, and would be important ecological assessment indicators of water quality in the stream ecosystem.

As mentioned before, groups B and C generally maintained their sizes as the observation interval increased. In 5-month mapping (Fig. 8d) group B was divided into two neurons at (0, 8) and (1, 8). However, since the neurons are located next to

c)

| | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 |
|---|---|---|---|---|---|---|---|---|---|
| 0 | | THP4-3 THP4-4 THP4-5 TCL4-4 TCL4-5 | | THP4-7 TCL4-7 | | | | | TKC4-5 |
| 1 | | | | | | | | | |
| 2 | TKC4-2 | | | | THP4-6 TCL4-6 | | | | THP4-8 TCL4-8 |
| 3 | | | YSC4-7 | | | | TCL4-1 | | |
| 4 | THP4-2 TCL4-3 | | | | THP4-1 TCL4-4 | | | | TCL3-12 |
| 5 | | | | | | | TKC3-12 | | |
| 6 | | | | | | | | | YSC3-12 |
| 7 | | | | | | YIG4-2 YIG4-3 TIG4-4 YIG4-5 | YCK3-12 YIG4-1 | YSC4-2 YSC4-4 YSC4-6 | TKC4-8 |
| 8 | TKC4-1 | | TKC4-3 | | TKC4-6 TKC4-7 | TKC4-4 | THP3-12 | YIG4-6 YIG4-7 YIG4-8 YCK4-1 YCK4-2 YCK4-3 YCK4-4 YCK4-5 YCK4-6 YCK4-7 YCK4-8 TSD3-12 TSD4-1 TSD4-2 TSD4-3 TSD4-4 TSD4-5 TSD4-6 TSD4-7 TSD4-8 | YIG3-12 |

Fig. 8. (*Continued*)

d)

| | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 |
|---|---|---|---|---|---|---|---|---|---|
| 0 | TKC4-5 | | | TKC4-2 | | THP4-2 TCL4-3 | | TKC4-3 | |
| 1 | | | | | | | | | |
| 2 | THP4-4 THP4-5 TCL4-5 | | | | | TCL4-1 | | THP4-1 | |
| 3 | | | | TCL4-2 | | | | | |
| 4 | YIG4-1 YIG4-2 YIG4-3 YIG4-4 YIG4-5 | | | | | THP4-3 THP4-4 | | TKC4-6 | TKC-7 |
| 5 | | | | THP4-6 TCL4-6 | | | | | |
| 6 | YSC4-7 | YSC4-8 | | | | TKC4-1 | | | TKC4-4 |
| 7 | YSC4-5 YSC4-6 | TKC4-8 | | THP4-8 TCL4-8 | | | | | |
| 8 | YIG4-6 YIG4-7 YIG4-8 YCK4-1 YCK4-2 YCK4-3 YCK4-4 YSC4-1 YSC4-2 YSC4-3 YSC4-4 | TSD4-1 TSD4-2 TSD4-3 TSD4-4 TSD4-5 TSD4-6 TSD4-7 TSD4-8 YCK4-5 YCK4-6 YCK4-7 YCK4-8 | | | | | | | |

Fig. 8. (*Continued*)

each other, the patterns would be generally similar. Communities from the medium polluted site, TKC, preserved the same characteristics in the map for a short period (Fig. 8a, b, c), forming many small groups, separated widely on the map. Some communities from YSC, mostly those collected from February to August in 1994, were generally located in separate neurons close to group B. These observations indicated that in overall terms, community dynamics could be effectively characterised by the use of combined artificial neural networks.

### 3.3. Recognition

Once training of community changes was complete, recognition for new input data by the network was possible. Benthic communities collected at the sites from Suyong stream (YIG, YCK, and YSC) from September to November in 1994 were used for recognition. Initially, the data were given to ART and weights were updated (Fig. 1). The updated weights were then arranged sequentially for a given period (for 2 or 3 months in this case), and were provided to the trained Kohonen net-

## Two-month

| | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 |
|---|---|---|---|---|---|---|---|---|---|
| 0 | | | | | | | | | |
| 1 | | | | | | | | | |
| 2 | | | | | | | | | |
| 3 | | | | | | | | | |
| 4 | | | | | | | | | |
| 5 | | | | | | | | | |
| 6 | | | | | | | | | |
| 7 | | YSC4-10 | | YIG4-9 YIG4-10 YIG4-11 YCK4-9 YCK4-10 YCK4-11 YSC4-9 YSC4-11 | | | | | |
| 8 | | | | | | | | | |

## Three-month

| | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 |
|---|---|---|---|---|---|---|---|---|---|
| 0 | YCK4-11 | | | | | | | | |
| 1 | YIG4-9 YIG4-10 YIG4-11 YCK4-9 YCK4-10 YSC4-10 | | | | | | YSC4-9 | YSC4-11 | |
| 2 | | | | | | | | | |
| 3 | | | | | | | | | |
| 4 | | | | | | | | | |
| 5 | | | | | | | | | |
| 6 | | | | | | | | | |
| 7 | | | | | | | | | |
| 8 | | | | | | | | | |

Fig. 9. Recognition of newly collected benthic macroinvertebrate communities to the trained Kohonen network in the period from 2 to 3 months. The alpha codes and numerical digits designating the names of sample units are explained in the captions of Figs. 4 and 7.

work for recognition (Fig. 9). Generally most of input data were recognised as belonging to group B (see location at (3, 7) in Fig. 8a for the 2-month map and location at (0, 1) in Fig. 8b for the 3-month map), which was the main community group formed from the Suyong stream. A few communities from YSC were recognised separately from group B. Recognised neurons in Fig. 9, however, were mostly equal to or close to patterned neurons for YSC communities of input training (Fig. 8a, b). The results revealed an expected pattern, based on the field experience.

### 3.4. *Efficiency of pattern detection*

This study demonstrates the method's feasibility in feature extraction of structure or patterns from dynamic data, using two unsupervised networks in combination. This encourages us to suggest that neural networks can be used for comprehensive understanding of data features in community dynamics in the spatio-temporal domain. This should assist in the development of strategic tools for the management of the ecosystem. The mappings do not provide exact answers; things like unexpected random variation in data or possible error accumulations in the highly iterative calculation process may bring biases into results and these should be guarded against. But we believe that in many situations this method could enhance one's comprehensive understanding of community data and other unexpected aspects of the dynamic behavior of communities may be revealed. A comprehensive understanding of ecosystem dynamics is essential in assessing ecological impacts and in providing continuing ecosystem management.

The algorithm in ART is generally feasible for classification of data (Carpenter and Grossberg, 1987). We tried to train raw input-data from 1-month samples, both with the ART and with the Kohonen network as a preliminary test. ART was better at discovering patterns than the Kohonen network. The Kohonen network was able to decipher patterns in the community data (Chon et al., 1996). In Chon et al. (1996), however, the data were based on densities of species that had high noise levels, i.e. many species with low or zero density. In this study we summed the species density into selected taxa, and this may have contributed to stabilising the process. With lower degree of noises, ART was more feasible in classification. This is why we used ART for classification of 1-month data. If a more powerful computer is available and species data are to be trained, the Kohonen network may be applied for initial 1-month training, and more detail information at the species level could be revealed. However further investigation may be required to handle noises in such large scale data.

The Kohonen network projects the data feature on a map (Kohonen, 1989; Hecht-Nielsen, 1990). As shown in Fig. 7a, ART neurons were not structured spatially, while contrasting this, in the Kohonen network it is possible to project the patterning neurons into the map in spatial dimensions (conveniently two or three). The output results are more comprehensible in characterising the conformation of neurons, as can be seen in Fig. 7b. This is the reason we chose to use the Kohonen network after training with ART. In total, these processes of combined training by the artificial neural networks serve as an alternative to non-linear filtering in ecological data; eliminating the input noise and carrying information to the output nodes in the network.

Chon et al. (1996) mentioned that the problem of objectivity exists with neural networks, since the network is based on random effects and iterative calculations. In fact each configuration after convergence may be different, depending upon the initial training. In this study, however, since the training was repeated at different sequential periods, general grouping characteristics can be observed with consistency. Although locations of the groups in the map can be different in each configuration, large groups A, B and C, and other small groups representing TKC, YSC and TSD, would consistently appear. This should increase confidence in the objectivity of the Kohonen network in community study.

Although the method looks effective, some points needed to be explored further in determining patterns of temporal variation in communities. In the field of electronics and computer sciences attention has been given to dynamic neural networks for patterning spatio-temporal data including recurrent networks (Kung, 1993; Giles et al., 1994; Haykin, 1994). Pattern analysis in electrical engineering, however, is different from interpreting community data. In engineering, pattern analysis is usually based on the concept of recurrence, with a very close time difference between samples such as machine movement operation, dynamic imaging, etc. Also pattern determination is usually conducted under supervised learning. In this study the sample points are relatively far apart in time, and recurrence is not of major interest. Identification of cycles, which could be investigated effectively with methods in time series such as periodograms or correlograms, is also not directly relevant to our study. Instead, changing patterns in community development are the target for pattern analysis. Also, unsupervised learning is required, which generally produces more complex problems in the analysis of time-space data series. However, since there have been many examples and analytical studies in spatio-temporal patterning in electronics engineering, further studies are needed before broader application of these methods to ecological data. Also the concept of recurrence deserves attention since periodicity could be a main theme in investigation of fluctuation in communities over longer periods.

By self-organising the temporal community patterns, the practice of pattern analysis on a time series could provide basic data for further analytical work such as determining stability in community assembly (Tanner et al. 1994; Law and Morton 1996). Wray and Green (1994) reported that artificial neural networks could be utilised for investigating parameters in non-linear dynamics. They showed that an artificial neural network, a multi-layer perceptron, could be used to find Volterra series, a well-known method of describing non-linear dynamics. The network could effectively estimate Volterra kernel in terms of its internal parameters. Further study is required to describe analytically the results of non-linear dynamics as summarised by the unsupervised networks utilised in this study.

Legendre et al. (1985) compared methods in classifying succession of species in a community. However their comparison was generally empirical, based on feasibility in practical use. It is difficult to compare on a temporal basis the efficiency of pattern analysis objectively, for example, in terms of recognition rate. This is understandable if we consider that unsupervised spatio-temporal pattern recognition is difficult due to the complexity and size of the problem, and it is not easy to quantitatively measure how correctly and effectively the learning method perceives dynamic behavior of input data. Along with cumulating empirical data and field experience, further analytical studies are needed to develop methodologies for comparison of efficiency between different classification methods, including artificial neural networks and other statistical methods in classifying temporal variation of communities (Legendre et al., 1985).

## Acknowledgements

This research was supported by the Non Directed Research Fund, Korea Research Foundation in 1993 and 1996. We thank Dr Jim Colbert for his helpful comments and correction of the manuscript.

## References

Bunn, S.E., Edward, D.H., Loneragan, N.R., 1986. Spatial and temporal variation in the macroinvertebrate fauna of streams of the northern jarrah forest, Western Australia: community structure. Freshwater Biol. 16, 67–91.

Boudjema, G., Chau, N.P., 1996. Revealing dynamics of ecological systems from neural recordings. Ecol. Model. 91, 15–23.

Carpenter, G.A., Grossberg, S., 1987. ART2: self-organization of stable category recognition codes for analog input patterns. Appl. Optics 26 (3), 4919–4930.

Chon, T.-S., Park, Y.S., Moon, K.H., Cha, E.Y., 1996. Patternizing communities by using an artificial neural network. Ecol. Model. 90, 69–78.

Elizondo, D.A., McClendon, R.W., Hoongenboom, G., 1994. Neural network models for predicting flowering and physiological maturity of soybean. Trans. ASAE 37 (3), 981–988.

Ferrãn, E.A., Ferrara, P., 1992. Clustering proteins into families using artificial neural networks. CABIOS 8, 39–44.

Giles, C.L., Kuhn, G.M., Williams, R.J., 1994. Dynamic recurrent neural networks: theory and applications. IEEE Trans. Neural Networks 5, 153–156.

Haykin, S., 1994. Neural Networks. Macmillian, New York.

Hecht-Nielsen, R., 1987. Counterpropagation networks. In: Proceedings of the International Conference on Neural Networks, vol. II. IEEE Press, New York, pp. 19–32.
Hecht-Nielsen, R., 1990. Neurocomputing. Addison-Wesley, New York 433 pp.
Huntingford, C., Cox, P.M., 1996. Use of statistical and neural network techniques to detect how stomatal conductance responds to changes in the local environment. Ecol. Model. 97, 217–246.
Kang, D.H., Chon, T.S., Park, Y.S., 1995. Monthly changes in benthic macroinvertebrate communities in different saprobities in the Suyong and Soktae streams of the Suyong River. Korean J. Ecol. 18, 157–177.
Kohonen, T., 1989. Self-Organization and Associative Memory. Springer, Berlin 312 pp.
Kung, S.Y., 1993. Digital Neural Networks. Prentice Hall, Englewood Cliffs, NJ, p. 444.
Kwon, T.-S., Chon, T.-S., 1993. Ecological studies on benthic macroinvertebrates in the Suyong River. III. Water quality estimations using chemical and biological indices. Korean J. Limnol. 26, 105–128.
Law, R., Morton, R.D., 1996. Permanence and the assembly of ecological communities. Ecology 77 (3), 762–775.
Legendre, P., 1987. Constrained clustering. In: Legendre, P., Legendre, L. (Eds.), Developments in Numerical Ecology. Springer, Berlin, pp. 289–307.
Legendre, P., Legendre, L., 1987. Developments in Numerical Ecology. Springer, Berlin 585 pp.
Legendre, P., Dallot, S., Legendre, L., 1985. Succession of species within a community: chronological clustering, with applications to marine and freshwater zooplankton. Am. Nat. 125, 257–288.
Lek, S., Delacoste, M., Baran, P., Dimopoulos, I., Lauga, J., Aulagnier, S., 1996. Application of neural networks to modelling non-linear relationships in ecology. Ecol. Model. 90, 39–52.
Levine, E.R., Kimes, D.S., Sigillito, V.G., 1996. Classifying soil structure using neural networks. Ecol. Model. 92, 101–108.
Lippmann, R.P., 1987. An introduction to computing with neural nets. IEEE Acoustics Speech Signal Processing Magazine April, 4–22.
Ludwig, J.A., Reynolds, J.F., 1988. Statistical Ecology: A Primer of Methods and Computing. Wiley, New York 337 pp.
Mallin, M.A., Paerl, H.W., 1994. Planktonic trophic transfer in an estuary: seasonal, diel, and community structure effects. Ecology 75 (8), 2168–2184.
McPeek, M.A., Miller, T.E., 1996. Special feature: evolutionary biology and community ecology. Ecology 55 (5), 1319–1378.
Melssen, W.J., Smits, J.R.M., Rolf, G.H., Kateman, G., 1993. Two-dimensional mapping of IR spectra using a parallel implemented self-organising feature map. Chemometr. Intell. Lab. Syst. 18, 195–204.
Pao, Y.-H., 1989. Adaptive Pattern Recognition and Neural Networks. Addison-Wesley, New York 309 pp.
Quinn, M.A., Halbert, S.E., Williams, L. III, 1991. Spatial and temporal changes in aphid (Homoptera: Aphididae) species assemblages collected with suction traps in Idaho. J. Econ. Entomol. 84, 1710–1716.
Recknagel, F., French, M., Harkonen, Yabunaka, K.I., 1997. Artificial neural network approach for modelling and prediction of algal blooms. Ecol. Model. 96, 11–28.
Reibnegger, G., Weiss, G., Wachter, H., 1993. Self-organizing neural networks as a means of cluster analysis in clinical chemistry. Eur. J. Clin. Chem. Clin. Biochem. 31, 311–316.
Rumelhart, D.E., Hinton, G.E., Williams, R.J., 1986. Learning internal representations by error propagation. In: Rumelhart, D.E., McClelland, J.L. (Eds.), Foundations. In: Parallel Distributed Processing: Explorations in the Microstructure of Cognition, vol. 1. MIT Press, Cambridge, MA, pp. 318–362.
Stankovski, V., Debeljak, M., Bratko, I., Adamic, M., 1998. Modelling the population dynamics of red deer (*Cervus elaphus* L.) with regard to forest development. Ecol. Model. 108 (1–3), 145–153.
Stromberg, M.R., Griffin, J.R., 1996. Long-term patterns in coastal California grasslands in relation to cultivation, gophers, and grazing. Ecol. Applications 6, 1189–1211.
Strong, D.R. Jr, Simberloff, D., Abele, L.G., Thistle, A.B. (Eds.), 1984. Ecological Communities: Conceptual Issues and the Evidence. Princeton University Press, Princeton, NJ, p. 614.
Tan, S.S., Smeins, F.E., 1996. Predicting grassland community changes with an artificial neural network model. Ecol. Model. 84, 91–97.
Tanner, J.E., Hughes, T.P., Connell, J.H., 1994. Species coexistence, keystone species, and succession: a sensitivity analysis. Ecology 75 (8), 2204–2219.
Wasserman, P.D., 1989. Neural Computing: Theory and Practice. Van Nostrand Reinhold, New York 230 pp.
Woodiwiss, F.S., 1964. The biological systems of stream classification used by the Trent River Board. Chem. Ind., 443–447.
Wray, J., Green, G.G.R., 1994. Calculation of the Volterra kernels of non-linear dynamic systems using an artificial neural network. Biol. Cybern. 71, 187–195.
Zurada, J.M., 1992. Introduction to Artificial Neural Systems. West, New York 683 pp.