

# Implementation of Self-Organizing Maps (SOM) to analyses of environmental parameters and phytoplankton biomass in a macrotidal estuary and artificial lake

ROKSANA JAHAN<sup>1</sup>, HYU CHANG CHOI<sup>2</sup>, YOUNG SEUK PARK<sup>3</sup>, YOUNG CHEOL PARK<sup>4</sup>, JI HO SEO<sup>1</sup>  
AND JOONG KI CHOI<sup>1</sup>

<sup>1</sup>Department of Oceanography, INHA University, Nam-gu, Incheon 402-751, Korea, <sup>2</sup>Korea Hydro & Nuclear Power Co., Ltd, Gangnam-gu, Seoul 135-791, Korea, <sup>3</sup>Department of Biology, Kyung Hee University, Dongdaemun-gu, Seoul 130-701, Korea, <sup>4</sup>Ecocean Co., Ltd, E Dong, 101 BL-3L, 661-16, Gojan-dong, Namdong-gu, Incheon 405-819, Korea

*Self-Organizing Maps (SOM) have been used for patterning and visualizing ten environmental parameters and phytoplankton biomass in a macrotidal (>10 m) Gyeonggi Bay and artificial Shihwa Lake during 1986–2004. SOM segregated study areas into four groups and ten subgroups. Two strikingly alternative states are frequently observed: the first is a diverse non-eutrophic state designated by three groups (SOM 1–3), and the second is a eutrophic state (SOM 4: Shihwa Lake and Upper Gyeonggi Bay; summer season) characterized by enhanced nutrients (3 mg l<sup>-1</sup> dissolved inorganic nitrogen, 0.1 mg l<sup>-1</sup> PO<sub>4</sub>) that act as a signal and response to that signal as algal blooms (24 µg chlorophyll-a l<sup>-1</sup>). Bloom potential in response to nitrification is affiliated with high temperature (r = 0.26), low salinity (r = -0.40) and suspended solids (r = -0.27). Moreover, strong stratification in the Shihwa Lake has accelerated harmful algal blooms and hypoxia. The non-eutrophic states (SOM 1–3) are characterized by macro-tidal estuaries exhibiting a tolerance to pollution with nitrogen-containing nutrients and retarding any tendency toward stratification. SOM 1 (winter) is more distinct from SOM 4 due to higher suspended solids (>50 mg l<sup>-1</sup>) caused by resuspension that induces light limitation and low chlorophyll-a (<5 µg l<sup>-1</sup>). In addition, eutrophication-induced shifts in phytoplankton communities are noticed during all the seasons in Gyeonggi Bay. Overall, SOM showed high performance for visualization and abstraction of ecological data and could serve as an efficient ecological map that can specify blooming regions and provide a comprehensive view on the eutrophication process in a macrotidal estuary.*

**Keywords:** Self-Organizing Maps, Gyeonggi Bay, Shihwa Lake, summer blooms, eutrophication

Submitted 15 October 2011; accepted 10 April 2012; first published online 13 August 2012

## INTRODUCTION

Plankton models are developed to understand how global change, the result of natural and anthropogenically induced climate change, will impact on the structure and function of the planktonic marine ecosystem. Marine planktonic modelling remains enigmatic because it is generated from the non-linear combination of biotic (physiological responses and predator–prey interactions) and abiotic (temperature, pH, light, nutrient supply, contaminant exposure, etc.) ecosystem forcing and has various plankton functional types including phytoplankton, for example, and its diverse groups such as diatoms, coccolithophorids, nitrogen fixers, picoplankton, phytoflagellates and dinoflagellates (Totterdell *et al.*, 1993) with their different emergent properties. Anderson (2005) recently identified a number of problems with the plankton functional types modelling including poorly understood

ecology, the difficulty of aggregating diversity within functional groups into meaningful state variables, the sensitivity of outputs to parameters choice and the representation of external physical and chemical environment. So, we need to understand the ecology of plankton well enough to do the model and also need to formulate and test a greater variety of models; perhaps it is time to think beyond the traditional planktonic ecosystem model and look to a different model (Franks, 2009).

Traditional conventional multivariate methods (i.e. cluster analysis (CA), discriminate analysis (DA), principal component analysis (PCA), factor analysis (FA), absolute principal component score multiple linear regression (APCS-MLR), factor analysis-multiple regression (FA-MR), etc.) are somewhat limiting for revealing the non-linear and complex dynamic nature that is frequently associated with analysing and synthesizing ecological data because they generally apply for linear data and are less flexible for data handling (e.g. noise and uncertainty) (Chon, 2011; Su *et al.*, 2011). In recent years, artificial neural network (ANN) techniques have become popular in ecological modelling by virtue of their powerful performance. There are two types of ANNs

**Corresponding author:**  
J.K. Choi  
Email: jkchoi@inha.ac.kr

according to the learning algorithms: supervised ANNs are used for data estimation (e.g. prediction and environmental community causality relationships) based on *a priori* knowledge and unsupervised ANNs are used when deriving information from data (e.g. ordination and classification) without previous knowledge (Kohonen, 1982). The Self-Organizing Map (SOM), based on an unsupervised neural network (Kohonen, 1982, 2001), appears to be an effective method for feature extraction and classification. It maps high-dimensional input data onto a low dimensional (usually 2-d) space while preserving the topological relationships between the input data. As a pattern recognition and classification tool, the SOM finds widespread use across a number of disciplines (Kaski *et al.*, 1998; Oja *et al.*, 2002). The SOM has also been applied in oceanography by Ainsworth (1999) and Ainsworth & Jones (1999) for chlorophyll estimates from satellite data, by Silulwane *et al.* (2001) and Richardson *et al.* (2002) to identify ocean chlorophyll profiles, by Hardman-Mountford *et al.* (2003) to relate satellite altimeter data with the recruitment of the Namibian sardine, by Ultsch & Roske (2002) to predict sea level, and by Richardson *et al.* (2003) and Risien *et al.* (2004) to extract sea surface temperature (SST) and wind patterns from satellite data. Nevertheless, for oceanographers unfamiliar with neural network techniques, the SOM remains a 'black box' with associated scepticism. In this paper, the SOM is applied for visualization and abstraction of the complexity of environmental-phytoplankton relationships in the macrotidal Gyeonggi Bay, Korea.

The entire Gyeonggi Bay is a useful site for comparative estuarine science because it comprises two connected, but distinct subsystems: Gyeonggi Bay (GB) and Shihwa Lake (SL). GB is a shallow macrotidal and well-mixed estuary limiting the accumulation of organic matter; but unpredictable inputs from Han River discharge can maintain nutrient availability. SL, on the other hand, is the artificial saltwater lake constructed from 1986 to 1994 and suffered from severe eutrophication, anoxia as well as environmental disaster (Han & Park, 1999). By the late 1980s to mid-1990s, the entire GB is heavily impacted by eutrophication, caused by nutrient input from the densely populated and industrialized catchment area, resulting in an increase in phytoplankton biomass and primary production including alternations to species distributions, composition and phenology (annual bloom dynamics) (Park & Park, 2000; Yang *et al.*, 2008). However, in the past (1980s), GB's phytoplankton seasonality followed a single spring diatom bloom that was triggered by increasing daily irradiance and atmospheric heat input that stratifies the water column after winter mixing brings nutrients to the surface (Choi & Shim, 1986c). In contrast, recently (2000s), GB's waters present diverse seasonal patterns with large variability from diatom (siliceous) blooming during winter to non-diatom blooming during summer related to complex interactions among physical, chemical and biological processes (Yang *et al.*, 2008). In general, phytoplankton seasonality at the estuaries is driven by more than a few climatic factors (Cloern & Jassby, 2008). This is a fundamental ecological distinction from the open marine and terrestrial biomes. It confirms Longhurst's (1995) insightful conclusion about the unpredictability of oceanographic processes along the margins of the oceans, where it is exceedingly difficult to generalize the processes which determine seasonality of plankton production. Hence, coastal ecosystem models are tools that offer an explicit framework for integration of the knowledge

gained as well as detailed investigation of the underlying dynamics and the reason for it, into a management approach.

The purpose of this study is to apply SOM as a modelling approach to pattern, classification, clustering and visualization of ten main environmental parameters (temperature, salinity, pH, DO, SS, COD, NO<sub>3</sub>, NO<sub>2</sub>, NH<sub>4</sub> and PO<sub>4</sub>) and phytoplankton biomass (chlorophyll-*a*) on the GB and SL during 1986–2004. We also discuss the underlying mechanisms of phytoplankton blooms and the eutrophication impacts on phytoplankton community structure based on present and past surveys.

## MATERIALS AND METHODS

### Study area

#### GYEONGGI BAY

The GB (Figure 1) has a number of features that typify shallow and coastal plain estuaries, including: (1) morphology characterized by a broad shallow channel of 10–20 m depth flanked by tidal flats >3 km width (Choi & Shim, 1986a); (2) macrotidal (tidal amplitude >10 m) bay specialized by semi-diurnal strong tidal currents (1.2–2.3 and 0.9–1.9 ms<sup>-1</sup> during spring and neap tides, respectively) and strong winter monsoon (3.77 ms<sup>-1</sup>) sweeps over the bay to introduce vertical mixing causing the resuspension of the bottom sediment (KMA, 2010); (3) wet summer season represented by huge Han River discharge (55 × 10<sup>6</sup> m<sup>3</sup>d<sup>-1</sup>) induces higher compensation depth that is inversely related with turbidity, and light favourable for phytoplankton blooms (Park *et al.*, 2000); alternatively (4) huge turbidity by winter mixing resuspended sediment particles and upward flux of nutrients in spite of low river discharge responsible for tychopelagic plankton (Choi & Shim, 1986c). Despite huge Han River flow, symptoms of stratification in lower GB are quite absent due to a well-mixed water body, except the Han River estuarine region during summer where vertical gradients of salinity stratification (salinity difference >5 psu) have often been observed (NFRDI, 2008; Park *et al.*, 2000). No hypoxia has been reported.

#### SHIHWA LAKE

The reclamation of an intertidal flat created Shihwa Lake (SL) in Gyeonggi province on the western coast of Korea during December 1986 to January 1994. The lake is enclosed by 12.7 km of sea dike and has a surface area of 42.3 km<sup>2</sup>, water storage capacity of 332 million tons with a management water level at -1.0 m, a maximum depth of 18 m, and a total seawater flux of 380 million tons per year (MOMAF, 2006). The artificial saline lake was expected to transform into a freshwater one to be used for irrigation purposes and, however, the drainage structure of the lake does not allow the entrapment of Yellow Sea water to be fully replaced by freshwater from its hinterland, which led to the project failure. Furthermore, the severe deterioration of lake-water quality in the mid-1990s prompted evaluations of environment impact (Park *et al.*, 2003a; Yoo *et al.*, 2009) and eutrophication, consequently, progressed rapidly brought about by the untreated sewage and wastewater flowing in from the Shihwa adjacent area (Kim *et al.*, 2004). The main tributaries of the lake consist of nine streams: four waterways traversing the industrial area, i.e. the Okgu, Gunga, Jeongwang, and

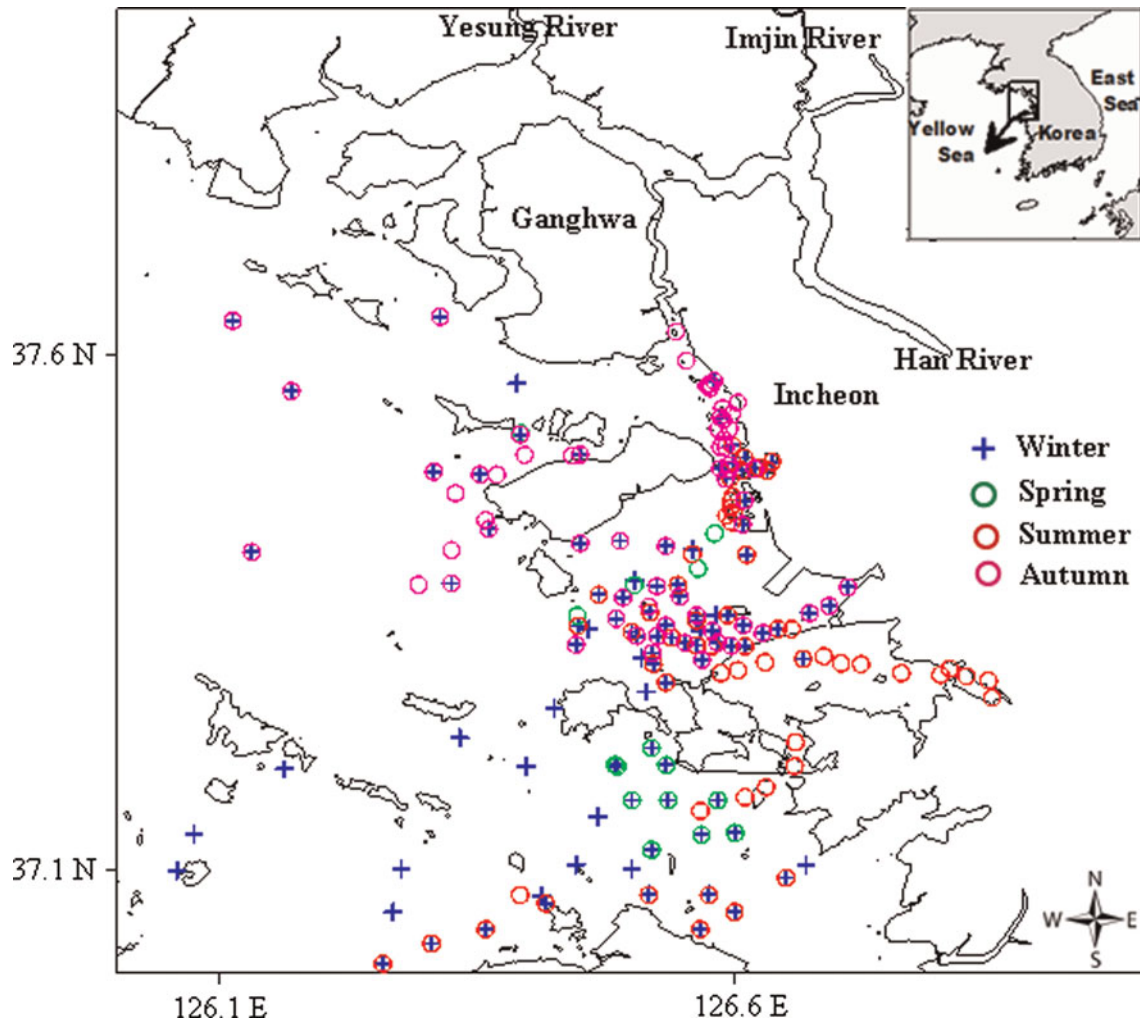


Fig. 1. Map showing the study area and geographical distribution of study stations. Dotted circle represents 'Upper Gyeonggi Bay' regions.

Siheung streams, and the Singil, Ansan, Banweol, Dongwha, and Samwha streams, with the last three passing through the Shihwa constructed wetland (Oh *et al.*, 2010).

### Field data

Our database was built up based on the environmental data collected from environment research reports which were released by several institutions (Korea Ocean Research and Development Institute (KORDI), Ministry of Construction and Transportation (MOCT), Ministry of Land Transport and Maritime Affairs (MLTM), Centre for Coastal Environments of Yellow Sea (CCEYS), Korea Water Resources Corporation (K-water), Korea Electric Power Corporation (KEPCO), Incheon Free Economic Zone (IFEZ), Ocean Science and Technology Institute Inha University (OSTI), Sudokwon Landfill Site Management Corp (SLC) and Korea Aggregates Association Incheon Branch (KAA)) over 19 years (1986–2004). These surveys were conducted at surface waters in the study area (Figure 1). We would consider particularly ten physico-chemical parameters (temperature, salinity, pH, dissolved oxygen, suspended sediment, chemical oxygen demand, ammonium, nitrate, nitrite and phosphate) and chlorophyll concentrations. Temperature and salinity were measured

using CTD, STD or T-S bridge. Dissolved oxygen (DO) concentration was measured using a DO meter (YSI), CTD or Winkler method and pH was measured using CTD or pH meter. Suspended sediment (SS) and chemical oxygen demand (COD) were determined by gravimetric analysis using the glass fibre filter and dichromate reflux methods, respectively. Nutrient concentration and chlorophyll-*a* were determined using the methods of Parsons *et al.* (1984). For clustering and organizing of the study area using SOM, we used about 800 study stations, which included 10 parameters from 1986–2004. Note that, during 1986–1994, our data sets contain some missing values and the SOM is a good method to recover them. The idea is as simple as to use the centre of each subclass to estimate the missing values of a value of a given observation. The virtue of the SOM regarding this problem is twofold: first, it is a non-parametric regression procedure that does not suppose any underlying models of the data set; and secondly it uses the information from similar observations to refine the positions of subclasses' centres and hence gives better estimation (Latif & Mercier, 2010). However, we used continuous ten years (1995–2004) data for temperature, salinity and chlorophyll-*a*, for the analysis on the long-term variation of phytoplankton dynamics which ignored the spatial variability at surface water of the entire bay.

## Self-Organizing Map (SOM)

In order to extract the structure of the high-dimensional data formed by the sample units (site-year), a method based on ANNs using an unsupervised algorithm called SOM (Kohonen, 2001) has been used. SOMs are different from other artificial neural networks in the sense that they use a neighbourhood function to preserve the topological properties of the input space (Kohonen, 2001). This makes SOMs useful for visualizing low-dimensional views of high-dimensional data akin to multidimensional scaling.

The SOM consists of two layers: input and output layers connected by connection intensities (weights). The input layer gets information from data matrix, while the output layer visualizes the computational results. When an input vector  $\mathbf{x}$  is sent through the network, each neuron  $k$  of the network computes the distance between the weight vector  $\mathbf{w}$  and the input vector  $\mathbf{x}$ . 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 (Park *et al.*, 2007). In this study, we used 10 environmental parameters as input units and 200 ( $20 \times 10$ ) neurons as the number of output neurons for a hexagonal lattice. The optimum map size was chosen based on minimum values of quantization and topographic errors (Kiviluoto, 1996; Kohonen, 2001), and ecological knowledge about the study area. SOM can be interpreted as a non-linear projection of the high-dimensional input data onto an output array of units. The best arrangement for the output layer is a hexagonal lattice, as it does not favour horizontal and vertical directions as much as rectangular arrays (Kohonen, 2001). Among all the  $D$  output neurons, the best matching unit (BMU) with minimum distance between the weight and input vectors is the winner. For the BMU and its neighbourhood neurons, the weight vectors  $\mathbf{w}$  are updated using the SOM learning rule. As a result, the network is trained to classify the input vectors according to the weight vectors that are closest to them.

A detailed description of the SOM algorithm has already been given by Kohonen (2001) for the theory and Park *et al.* (2003b, 2004) 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).

To test for the difference of environmental parameters including chlorophyll-*a*, one-way analysis of variance (ANOVA) was applied and Tukey's *post-hoc* test for multiple comparisons among means to detect differences using SPSS for Windows version 12.0.1 (SPSS Inc, Chicago, III) was followed. Differences are significant at 95% ( $P < 0.05$ ). In order to quantitatively analyse and confirm the relationships between chlorophyll-*a* and environmental parameters in each group, the Pearson's correlation analysis was applied.

## RESULTS

Figure 2A, B illustrates the temporal variation of temperature and salinity during 1995–2004. Water temperature ranged

from 0.0 to 30°C (mean 15.6°C) with lower values during winter and maximum values in summer (Figure 2A). Salinity variation was a reverse trend of temperature. Salinity ranged from 7.1 to 33.2 psu (mean 27.7 psu) with lower values during the summer wet season and a higher value in winter (Figure 2B). The lower values of salinity ( $< 10$  psu) were particularly recorded in August 1995, July 1996 and July 1997 at the SL region because of untreated sewage and waste water inputs from the watersheds and limited physical mixing during the stratified periods (1994–1999). The characteristics of the lake water quality slightly changes after 1999, the period when seawater dilution was allowed to improve the deteriorated water quality. Moreover, the salinity of the stations located at the upper bay and at near the tributaries was lower than those of middle and lower bays due to the freshwater input from rivers (spatial data not shown).

In this study, chlorophyll-*a* concentration ranged from 0.7 to 210.7 (mean 18.3)  $\mu\text{g l}^{-1}$  (Figure 2C). High chlorophyll-*a* concentration ( $> 10 \mu\text{g l}^{-1}$ ) was recorded during the summer season (i.e. July 1996, August 1996, August 1998) in the upper GB, Incheon harbour and vicinity of SL (spatial data not shown). The phytoplankton biomass increased about twofold from the mid-1990s to the mid-2000s and phytoplankton blooms were often detected in all areas of the bay through the seasons. In addition, winter blooms ( $> 110 \mu\text{g l}^{-1}$  in December of 1997) were recorded because of the increase in abundance of *Thalassiosira nordenskioldii* Cleve. The peak of

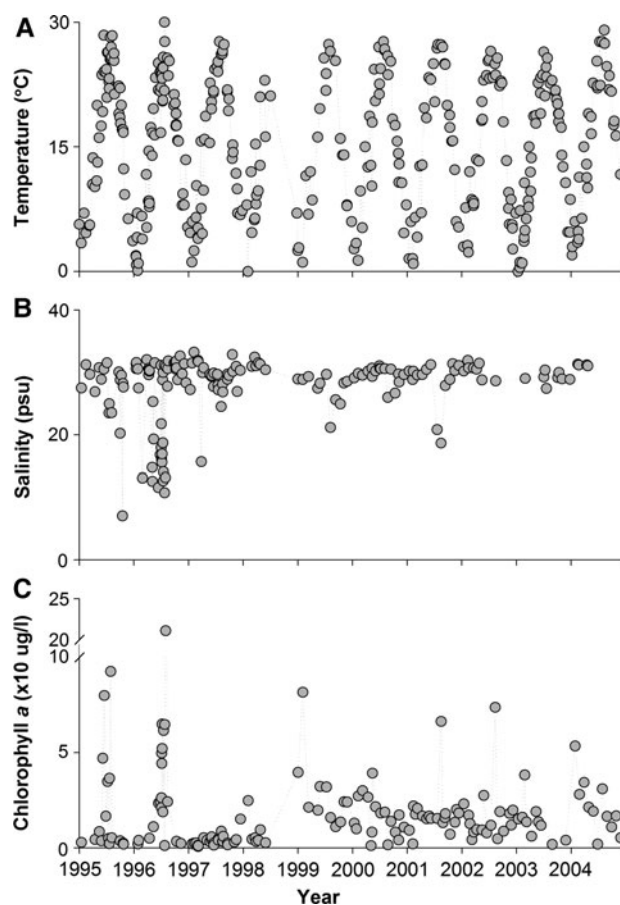


Fig. 2. The temporal variation of temperature (A), salinity (B) and chlorophyll-*a* concentration (C) during the last decade (1995–2004) in Gyeonggi Bay.

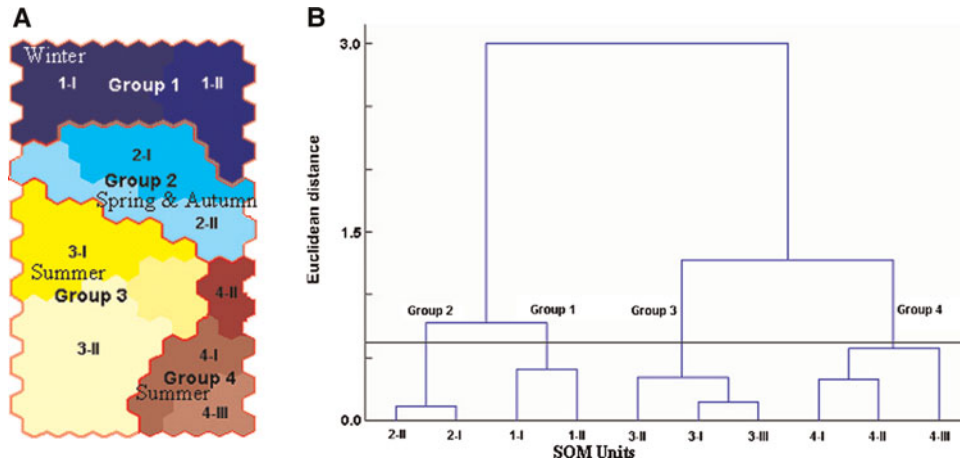


Fig. 3. Classification of study stations on the Self-Organizing Maps (SOM) map trained with environmental parameters and subgroup in each large group (A), and a dendrogram of hierarchical cluster analysis using Ward linkage method with Euclidean distance showing relations among groups defined in the SOM map (B).

phytoplankton biomass also frequently occurred during spring and autumn over the past decade.

### Model result

After learning the process of the SOM with environmental parameters, the study stations were divided into four large groups with ten different subgroups at different linkage distances according to the hierarchical cluster analysis with Ward's linkage method (Figure 3A, B). The study stations,

as results of SOM and cluster analysis, first were divided according to seasonal characteristics. The group 1, which was located at the upper part on the map, was characterized as winter, whereas groups 3 and 4 located at the lower part were characterized as summer. The group 2 showed the intermediate characteristics between group 1 and groups 3 and 4, characterized as spring and autumn. Each group was subdivided into 2 and 3 sub-groups according to the characteristics of environmental parameters. Figure 4 A–D shows the representative location for each SOM group in the geographical

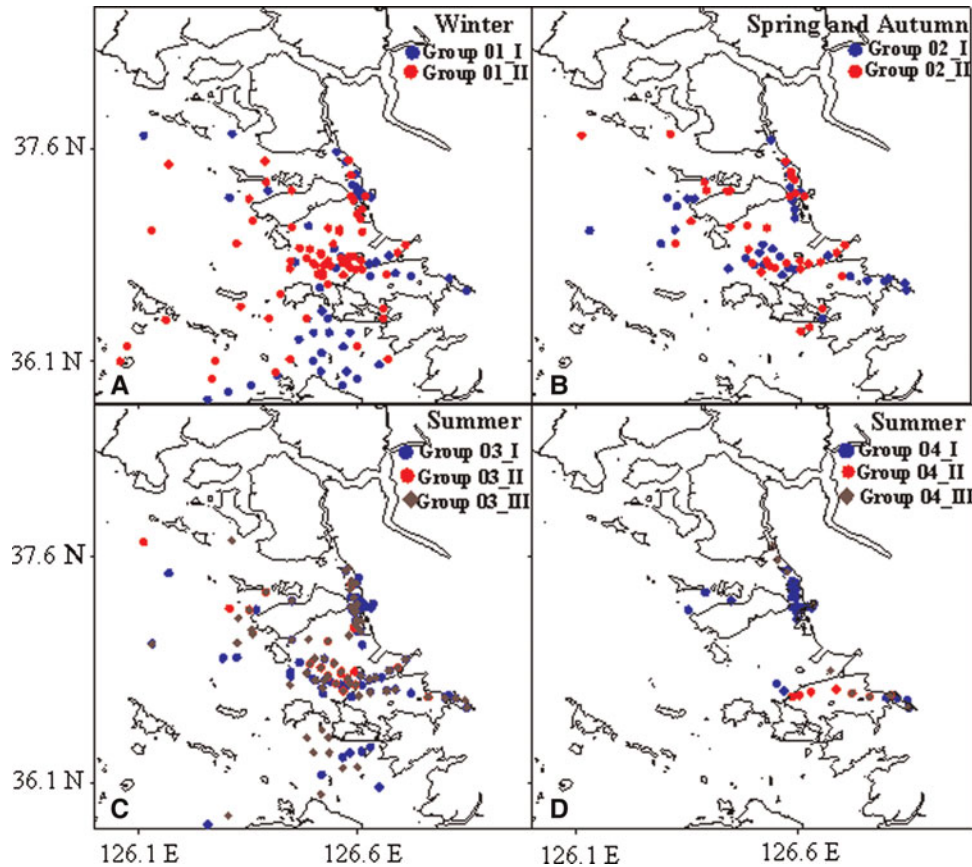


Fig. 4. Geographical distribution of study stations based on the trained Self-Organizing Maps in each group and subgroup (A, Group 1; B, Group 2; C, Group 3; D, Group 4).

map of the GB and SL. SOM group 1 (winter) encompassed sites from the enormous GB and SL regions with 2 sub-groups based on site-specific environmental values. All sites from group 2 (spring and autumn) belonged to the upper-middle GB and SL stations, and were partitioned into 2 sub-groups. Finally, group 3 (summer) also corresponded to large GB and SL regions, whereas group 4 (summer) was strictly located at upper GB, inside and outside of SL with 3 sub-groups. It is remarkable that most of the sub-group sites during summer were still overlapped by others due to site-specific environmental values and coastal hydrological processes.

Estimated values of each parameter in the trained SOM map are visualized in grey scale in Figure 5. Dark areas on the map represent high values of each parameter while light areas represent low values. Each environmental parameter shows significantly different distribution patterns among different groups on the SOM map (one-way ANOVA,  $P < 0.05$ ) (Figure 6). Temperature was low in group 1 (winter) and high in groups 3 and 4 (summer). In contrast, the salinity showed a different pattern, with high in group 1 (winter: 34) and low in group 4 (summer: 18). Group 2 (spring and autumn) and group 3 (summer) showed somewhat similar salinity gradient. In case of pH, group 2 showed slightly higher concentrations than other groups. Dissolved oxygen concentration and suspended sediment concentration were the highest in group 1 (winter), while chemical oxygen demand and nutrient concentrations were the highest in group 4 (summer). Nutrient and COD profiles divided into two parts among the four SOM groups. The SOM group 4 was characterized by the highest nutrients and COD, whereas the remaining SOM groups (1–3) were designated by lower nutrients and COD.

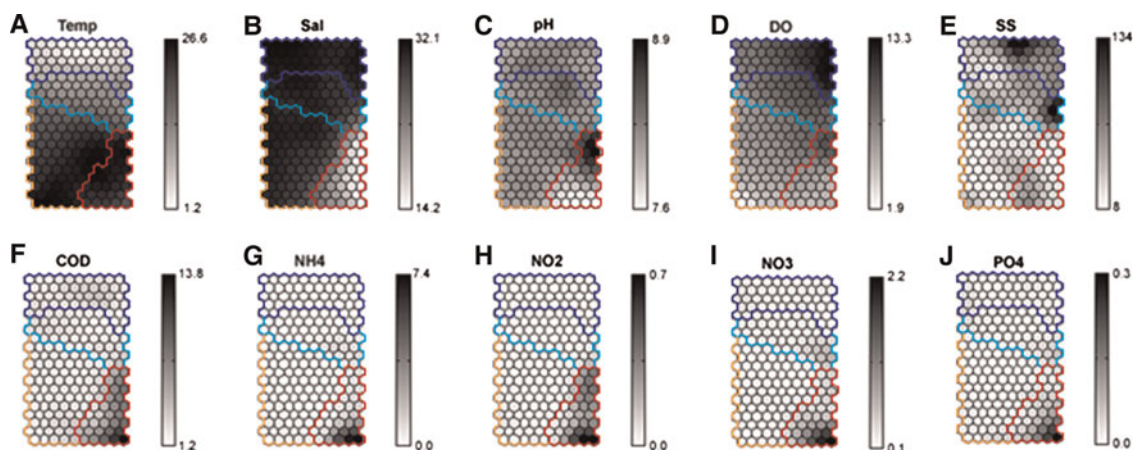
The phytoplankton biomass also showed significant difference (one-way ANOVA,  $P < 0.05$ ) among groups (Figure 7). The highest chlorophyll-*a* concentration was detected in group 4 (summer). These stations belonging in group 4 were considerably influenced by massive nutrient inputs in summer. The chlorophyll-*a* concentrations in group 2 (spring and autumn seasons) were a little higher than in other groups (groups 1 and 3), though there is no significant difference among these three SOM groups.

Correlations between phytoplankton biomass of each group and environmental parameters were examined by Pearson's correlations analysis (Table 1). In SOM groups 1 and 2, phytoplankton biomass was significantly correlated with salinity ( $r = -0.27$ ) and temperature ( $r = -0.22$ ), respectively. In SOM groups 3 and 4, chlorophyll-*a* concentration was positively correlated with temperature and DO ( $r = 0.24$ ,  $r = 0.26$ , respectively), and negatively correlated with salinity ( $r = -0.28$ ,  $r = -0.40$ , respectively). This weak relationship between phytoplankton, and temperature and salinity, reveals it to have a pronounced eurythermal and euryhaline nature. Phytoplankton biomass showed insignificant relationships with nutrients in SOM groups 3 and 4 (summer). In addition, phytoplankton biomass was negatively correlated with suspended solids ( $r = -0.27$ ) in SOM group 4.

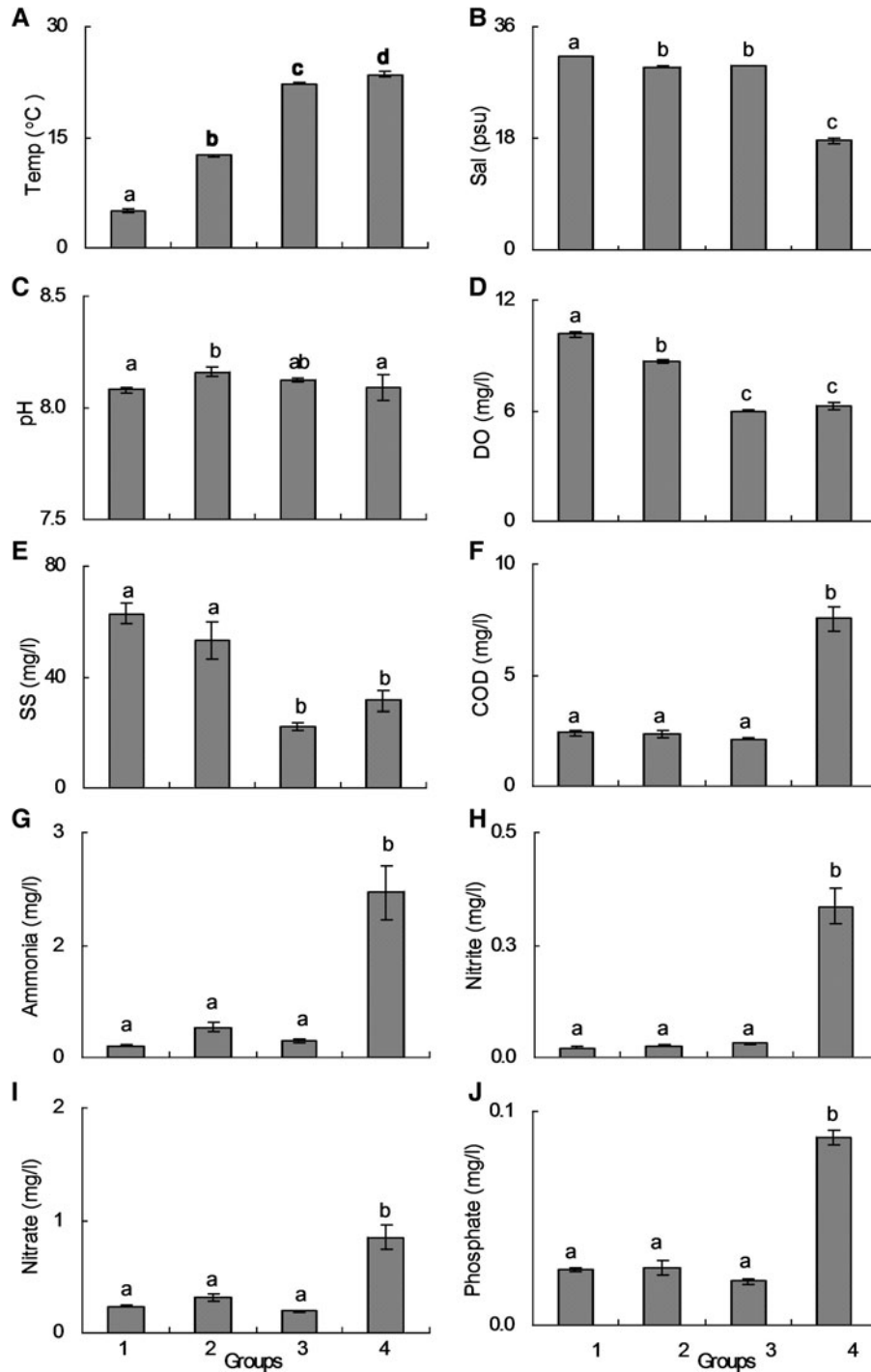
## DISCUSSION

### The SOM techniques

The non-linearity and complexity of variables involved in water quality have led many researchers to use the ANN model to simulate these variables due to the ability of such models to handle complex, non-linear relationships (Richardson *et al.*, 2002, 2003; Park *et al.*, 2003b, 2004; Song *et al.*, 2007). In this particular study, SOMs demonstrated their effectiveness for assessing four seasonal patterns from multidimensional environmental time-series data during 1986–2004 (Figure 3). Park *et al.* (2004) compared the SOM and PCA, and found that SOM grouping was more relevant to ecology, revealing different effects of pollution states, and impact of spatial and temporal variations in environment. For example, the SOM, by explaining total variance in the data, was able to describe more directly the discriminatory power of input variables in mapping, while PCA explained less than 30% of the total variance in the data (Park *et al.*, 2004). One of the most significant characteristics and contributions of using SOM is that the results obtained by SOM can be visualized from its topology map (Figure 5). Once the SOM has converged, it stores the most relevant information about the process in its topology map and allows all such information to be displayed, something that is not possible with the standard



**Fig. 5.** Visualization of environmental parameters calculated in the trained Self-Organizing Maps (SOM) in grey scale. The values were calculated during the learning process (A, temperature; B, salinity; C, pH; D, dissolved oxygen; E, suspended sediment; F, chemical oxygen demand; G, ammonia; H, nitrate; I, nitrate; J, phosphate). The blue, turquoise, yellow and red lines represent SOM groups 1, 2, 3 and 4, respectively. See online publication.



**Fig. 6.** The characteristics of environmental parameters in each group defined on the Self-Organizing Maps SOM map. Error bars indicate the standard error of each variable. Different alphabets (a,b,c,d) on the bars display significant differences ( $P < 0.05$ ) between groups based on Tukey's multiple comparison test (Temp, temperature; Sal, salinity; DO, dissolved oxygen; SS, suspended sediment; COD, chemical oxygen demand); shared alphabets between groups indicate no significant difference.

output from cluster analysis or multidimensional scaling (Richardson *et al.*, 2002). Figures 3 and 5 present the SOM topology maps, which have  $20 \times 10$  grids (neurons) with each neuron representing a cluster of similar input patterns; in fact, results are more robust with large data sets because the SOM can learn from more data. The SOM could be used on data sets with thousands of profiles: these would be more difficult to analyse with PCA, cluster analysis or multidimensional scaling (Richardson *et al.*, 2002). For ordination on a simple

output space, SOM had the advantage over PCA, independent component analysis and multidimensional scaling, of visualizing the distance compression on the projected space (Ultsch & Morchen, 2006). This interesting property was used to define our 4 SOM groups. Another advantage was to cluster and ordinate in a single analysis, which is not possible with classical multivariate analysis like DCA.

A limitation of the SOM technique is that it is not underpinned by a rigorous statistical framework, as is PCA. Thus,

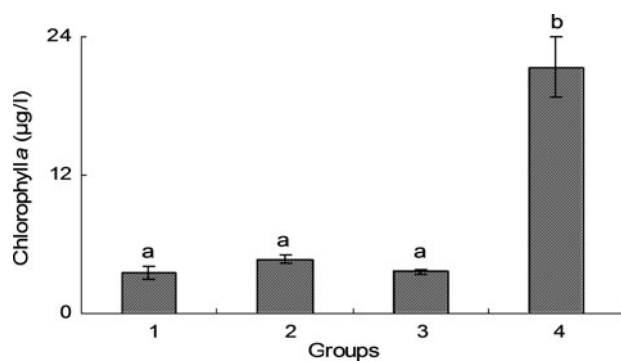


Fig. 7. Chlorophyll-*a* concentration of each group for determination of environmental parameters. Error bars indicate the standard error. Different alphabets (a,b,c,d) on the bars display significant differences ( $P < 0.05$ ) between groups based on Tukey's multiple comparison test; shared alphabets between groups indicate no significant difference.

the SOM provides no significance level for the patterns and does not give the proportion of variance explained by the patterns. Therefore, additional analysis (e.g. *post hoc* permutation and randomization tests) have to be computed as we did in this study. A perceived difficulty in using SOM to identify patterns is that the number of patterns chosen is arbitrary, as the researcher chooses the dimensions for the output maps. Moreover, as Bowden *et al.* (2005 a, b) pointed out, there are several disadvantages with this approach, including the increase of computational complexity and memory requirements, difficulty in learning, increase of the complexity of the model and consequently, a difficulty in understanding the model as well as increasing noise due to inclusion of spurious input variables.

### Relating SOM pattern to seasonal and spatial variability in environmental parameters and phytoplankton biomass

In Figure 4, the four SOM groups matched with the geographical distribution of the sampling sites and described that spatial variation with site-specific physical–chemical oceanographic parameters was the main factor for characterizing estuarine phytoplankton distribution in GB on a large scale. Geographical location was effectively identified with the clusters according to the trained SOM. Noteworthy, phytoplankton variability may result from changes in the physical characteristics of a system (e.g. hydrology, wind-driven resuspension and tidal mixing), biological interactions (e.g. reduced grazing), or an increased organic and inorganic nutrient loading; all these processes vary between ecosystems and change over time within ecosystems (Rabalais *et al.*, 2009). Most of the sites, however, in each SOM group were located

within a distinct geographical area. The SOM 4 group (summer), for example, located at SL and Upper GB, is a more eutrophic state characterized by algal blooms, enhanced nutrients, temperature, lower salinity, suspended solids and dissolved oxygen. In contrast, the other three SOM groups (groups 1–3), located at broad macrotidal of the GB region, are a non-eutrophic state characterized by lower chlorophylls, nutrients, higher salinity and suspended solids. Therefore, these two alternative states demonstrated the SOM mapping feasibility regarding provision of information on geographical distribution and algal blooming pattern at the same time.

Even if most of the SOM group-subgroups could be distinguished from each other in topological maps (Figures 3 & 5), some of the sub-group stations were still overlapped by others (Figure 4). This might be explained by site-specific environmental values and different coastal hydrological processes. The SOM 4 subgroups (4I–III; summer: upper GB and SL) showed, for example, meaningful ecological interpretations by salinity stratification and were markedly differentiated from well-mixed macrotidal SOM 3 (summer: lower GB). Generally, the bulk of sea surface salinity during summer is low by comparison with the freshwater fluxes into the surface layer because of precipitation and rainfall resulting in vertical salinity contrasts ( $< 5$  psu) (NFRDI, 2008). Moreover, stratification at upper GB and SL might also cause rapid settling of suspended particulate by forming flocs at the time of mixing of freshwater with saline seawater during downstream transport (Postma, 1967).

Efficiency of mapping was further demonstrated in the clusters designated within the same summer season. In contrast to SOM 4, the SOM 3 systems generally exhibit lower levels of chlorophyll-*a* and nutrients, and are also characterized by macro-tidal activities through advection and diffusion processes responsible for this different response. Fundamentally, well-mixed circulation in an estuary occurs where the tidal prism is significantly larger than river discharge, and the tidal currents retard any tendency toward stratification of fresh and saltwater, and this increased mixing therefore experienced lower photosynthetic activity and chlorophyll-*a* because of a reduction of the residence time of the algae in the photic zone (Monbet, 1992). To the extent that GB lower coastal ecosystems (SOM 3) also differ from Upper GB and SL (SOM 4) by the presence of salinity gradients, horizontally and vertically, and some other inherent physical (i.e. tide, wind, basin geography and river flow) these attributes operate in concert to set the sensitivity of this ecosystem to nutrient enrichment. Neither the lower GB nor the San Francisco Estuary (Hager & Schemel, 1996; Lucas & Cloern, 2002) and the Delaware Bay (Sharp, 1994) have a major problem with nutrient eutrophication, largely because neither shows summer stratification, which makes them unlike the Chesapeake Bay.

Table 1. Correlation coefficient between the phytoplankton biomass of each group and the environmental parameter: \*,  $P < 0.05$ ; \*\*,  $P < 0.01$ .

	Temp.	Sal.	DO	SS	Ammonia	Nitrite	Nitrate	Phosphate
Group 1	0.08	-0.27**	-0.04	-0.02	0.05	0.13	0.05	0.15*
Group 2	-0.22*	0.01	-0.07	0.06	0.04	0.44**	-0.18*	0.03
Group 3	0.24**	-0.28**	0.41**	0.08	-0.08	-0.10	-0.03	0.00
Group 4	0.26**	-0.40**	0.71**	-0.27**	-0.19*	0.09	-0.24**	-0.28**

Temp., temperature; Sal., salinity; DO, dissolved oxygen; SS, suspended sediment.



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, although statistical test revealed homogenized features in SOM groups 1, 2 and 3 based on nutrient and chlorophyll scenarios. In GB, winter season (SOM 1) has distinct features over spring and autumn (SOM 2) by its adverse hydro-regions. Note that, GB's winter season is characterized by high SS resulting from resuspension of sediments through the tidal mixing (convective mixing being due to the heat exchange and evaporation) and winter mixing due to the strong cold north-west wind (Choi & Shim, 1986b). The effects of suspended solids on phytoplankton are generally not direct; rather the effects are mediated through the agency of light fluctuations. Intense SS concentrations can limit light penetration and suppress cell growth as well. Light limitation, for example, is expected to result in decrease in the half saturation constant for nutrient-limited growth (Flynn, 2003), affecting the kinetics of resource acquisitions, and hence affecting competition between species. Moreover, light as another 'nutrient' has also been subjected to a Tilman'esque resource competition treatment (e.g. Passarge *et al.*, 2006; Caputo *et al.*, 2008) and light-P as a resource pair have been found not to follow standard resource-competition expectations (Passarge *et al.*, 2006), though given the role of P in cellular energetics (Flynn *et al.*, 2010) that is perhaps not unexpected. In GB, the irradiance during winter season ranged from 1.84  $\text{mw cm}^{-2}$  to 4.66  $\text{mw cm}^{-2}$  at the surface and is lower than optimum irradiance (Choi & Shim, 1986b).

During GB's winter season, diatoms are more prominent than dinoflagellates. Turbulence can negatively influence dinoflagellates blooming by three mechanisms: physical damage; physiological impairment (aggregation); phototaxis; and diel vertical migration (Smayda, 1997). In contrast, diatoms are better adapted to intense mixing conditions as they have low respiration to photosynthesis ratios and high growth rates (Cushing, 1989). Given that there are dramatic alternations of diatom species composition, *Thalassiosira nordenskiöldii*, for example, is the only dominant species during winter blooms since 1998, while tychoipelagica diatom (i.e. *Paralia sulcata*) was the most dominant species during 1981–1982 (Choi & Shim, 1986c), suggesting that the factors causing recent change in the phytoplankton community structure in favour of *T. nordenskiöldii* were not exhibited in the past.

## Mechanisms of phytoplankton summer blooms

### GYEONGGI BAY

GB's summer blooms (group 4) support the phase I eutrophication model (Cloern, 2001) emphasizing that changing nutrient input acts as a signal and response to that signal as increased phytoplankton biomass. In outside the SL region, blooms are detected after Shihwa Lake discharge (15 million tons  $\text{day}^{-1}$ ) since 1999 (Park & Park, 2000). Summer dense blooms (chlorophyll  $>7 \text{ mg m}^{-3}$ ) are usually dominant in the nutrient-rich upwelling/eddy north-west Pacific systems caused by the dinoflagellates and diatoms (Shanmugam *et al.*, 2008) and also in Chesapeake Bay (Breitburg, 1990; Harding, 1994), Tolo Harbour, Hong Kong (30  $\mu\text{g l}^{-1}$ ; Xu *et al.*, 2010) and the Mississippi River Plume (Grimes & Finucane, 1991).

GB's summer blooms appeared to be more dependent on physical processes rather than nutrients (Table 1), as indicated by higher correlation between chlorophyll-*a* and environmental parameters (i.e.  $r = 0.26$  SST,  $r = -0.40$  salinity,  $r = 0.71$  DO and  $r = -0.27$  SS). The huge discharge of the Han River not only delivers nutrients to the upper GB but also determines the hydrological properties of the water column, including high temperature, low salinity, vertical thermal stability, low turbidity as well as high light conditions. All of these properties triggered phytoplankton growth most likely by supplying proper temperature, increasing the light intensity and retaining the algal cells in the euphotic zone. Smayda (2008) recently concluded that bloom potential in response to nitrification is mediated by the accompanying irradiance and flushing characteristics. During 1995–2004, GB's chlorophyll profiles experienced a significant increasing trend of about twofold which is consistent with the sharp increase in the global ocean trend (4.13%) during 1998–2003 (Gregg *et al.*, 2005) and with the increase in dissolved inorganic nitrogen trend in GB during the past four decades (1981–2008) (Park & Park, 2000; NFRDI, 2008).

Gyeonggi Bay's summer blooms were mostly dominated by the nano-size ( $<20 \mu\text{m}$ ) phytoplankton. During 2000–2004, cryptomonads ( $<5 \mu\text{m}$ ) was the most dominant phytoplankton species associated with co-dominant diatoms, whereas the diatoms *Skeletonema costatum* and *Chaetoceros* spp. were the only dominant group in the past (1981–1982) (Choi & Shim, 1986c). Progressive changes in these phytoplankton species composition has coincided with the tenfold increased N:P ratio from the Redfield ratio during 1986–2004. A 21-years series of measurements from the western Wadden Sea, for example, provides strong empirical evidence that human-induced changes in nutrient (N:P) ratios can cause changes in phytoplankton species composition (Philippart *et al.*, 2000), and off the coast of Germany, a fourfold increase in the N:Si ratio coincided with decreased abundance of diatoms and an increase in Haptophyceae (*Phaeocystis*) blooms (Radach *et al.*, 1990). A number of variables could contribute to changes in the phytoplankton community over time (Livingston, 2001). These include: (1) the exact timing of nutrient delivery; (2) which nutrient (or nutrients) was (were) being loaded at a given time; (3) interactions among the various nutrients; (4) bay habitat conditions relative to the interannual drought–flood sequence; and (5) the nutrient requirements of the species present at the time of the nutrient loading.

### SHIHWA LAKE

The proposed underlying mechanism for summer blooms (group 4) in SL, is consistent with earlier studies (Choi *et al.*, 1997; Han & Park, 1999; Kim *et al.*, 2004) which concluded that huge fresh water inputs from the neighbouring municipal and industrial complexes through six major streams, are large enough to offset the effects of tidal and wind stirring. As a result the water column remains stratified at a depth range of 6–8 m, for a sufficiently long period that phytoplankton can grow and reach higher levels (167  $\mu\text{g chlorophyll-}a \text{ l}^{-1}$  in 1996) with frequent red-tides indicating a hypertrophic condition. In SL, extreme summer blooms supply organic matter to bottom water and sediment thereby generating oxygen consumption, and accordingly strong stratification limits oxygenation of bottom waters leading to hypoxia (Han & Park, 1999). The hypoxia

(anoxia) in the SL is analogous to that of the Black Sea (Sorokin, 1983; Mee, 1992; Tuncer *et al.*, 1998) and the Changjiang and Mississippi margins (Rabouille *et al.*, 2008). Hypoxia is one of the significant reasons for the unstable benthic ecosystem in SL (Ryn *et al.*, 1997).

Summer harmful algal blooms are frequently caused by dinoflagellates (*Prorocentrum minimum*), cryptomonads and Chrysophyceae, whereas diatoms (*Cyclotella atomas*, *Nitzschia* sp. and *Chaetoceros* sp.) are dominant in autumn and winter (Choi *et al.*, 1997). It is important to note that dinoflagellate blooms (i.e. *Heterocapsa triquetra*) are sometimes also found under ice in SL (HAN, 2011). It is well documented that the physiological flexibility of dinoflagellates in response to changing environmental parameters (e.g. light, temperature and salinity) as well as its ability to utilize both inorganic and organic nitrogen, phosphorus, and carbon nutrient sources suggest that increasing dinoflagellate blooms are a response to increasing eutrophication (Glibert *et al.*, 2005; Heil *et al.*, 2005). Note that dinoflagellate blooms did not develop before dike construction.

## CONCLUSION

In the present study, the Self-Organizing Map model gave satisfactory results for the ordination and clustering of environmental parameters and phytoplankton biomass that revealed four distinct seasonal patterns (SOM 1, winter; SOM 2, spring and autumn; SOM 3, summer; and SOM 4, summer), belonging to different geographical regions of the Gyeonggi Bay and Shihwa Lake. The interpretation of the SOM algorithm enables easy visualization of the patterns in the same form as the large input datasets, something that is not possible with the standard output from cluster analysis or multidimensional scaling. In this study, efficiency of SOM mapping had been demonstrated in the last two clusters designated with the same summer season. The SOM 4 group restricted at Shihwa Lake and Upper Gyeonggi Bay is a more eutrophic state characterized by algal blooms, enhanced nutrients and temperature, and, conversely, the SOM 3 group located at broad lower Gyeonggi Bay regions is a non-eutrophic state and is also characterized by macrotidal activities through advection and diffusion processes responsible for this different response. Therefore, the strengths of our SOM model are the recognition of blooming regions (SOM 4: upper GB and SL) with appropriate ecological explanations (i.e. nutrient, stratification, low salinity and SS) and their linkage to provide a comprehensive view on the eutrophication process in the macrotidal Gyeonggi Bay. So, these results are easy to interpret and useful to environmental decision-makers for sustainable management of estuarine ecosystems. By using other biological and physical oceanographic factors, SOM can offer a better resolution of the complexity of relationships between variables in ecological processes. Finally, once the description of the existing environmental parameters and of their obvious relationship with the environmental pollution is done, the prediction of phytoplankton blooms is demonstrated to be necessary and therefore should be seriously considered.

## ACKNOWLEDGEMENTS

This research was part of the project 'Development of forecasting technology for seawater circulation and ecosystem

change' and it is funded by the Ministry of Land, Transport and Maritime Affairs, Korea.

## REFERENCES

- Alhoniemi E., Himberg J., Parhankangas J. and Vesanto J. (2000) SOM toolbox. <http://www.cis.hut.fi/projects/somtoolbox> (accessed 13 March 2012).
- Anderson T.R. (2005) Plankton functional type modeling: running before we can walk? *Journal of Plankton Research* 27, 1073–1081.
- Ainsworth E.J. (1999) Visualization of ocean color and temperature from multispectral imagery captured by the Japanese ADEOS satellite. *Journal of Visualization* 2, 195–204.
- Ainsworth E.J. and Jones S.E. (1999) Radiance spectra classification from the ocean color and temperature scanner on ADEOS. *IEEE Transactions on Geoscience and Remote Sensing* 37, 1645–1659.
- Bowden G.J., Dandy G.C. and Maier H.R. (2005a) Input determination for neural network models in water resources applications. Part 1. Background and methodology. *Journal of Hydrology* 301, 75–92.
- Bowden G.J., Maier H.R. and Dandy G.C. (2005b) Input determination for natural network models in water resources applications. Part 2. Case study: forecasting salinity in a river. *Journal of Hydrology* 301, 93–107.
- Breitburg D.L. (1990) Nearshore hypoxia in the Chesapeake Bay: patterns and relationships among physical factors. *Estuarine, Coastal and Shelf Science* 30, 593–609.
- Caputo L., Naselli-Flores L., Ordonez J. and Armengol J. (2008) Phytoplankton distribution along trophic gradients within and among reservoirs in Catalonia (Spain). *Freshwater Biology* 53, 2543–2556.
- Choi J.K., Lee E.H., Noh J.H. and Huh S.H. (1997) The study on the phytoplankton bloom and primary productivity in Lake Shihwa and adjacent coastal areas. *The Sea, Journal of the Korean Society of Oceanography* 2, 78–86.
- Choi J.K. and Shim J.H. (1986a) The ecological study of phytoplankton in Kyeonggi Bay, Yellow Sea I. environmental characteristics. *Journal of the Korean Society of Oceanography* 21, 56–71.
- Choi J.K. and Shim J.H. (1986b) The ecological study of phytoplankton in Kyeonggi Bay, Yellow Sea. II. Light intensity, transparency, suspended substances. *Journal of the Korean Society of Oceanography* 21, 101–109.
- Choi J.K. and Shim J.H. (1986c) The ecological study of phytoplankton in Kyeonggi Bay, Yellow Sea. III. Phytoplankton composition, standing crops, tychopeagic plankton. *Journal of the Korean Society of Oceanography* 21, 156–170.
- Chon T.S. (2011) Self-Organizing Maps applied to ecological sciences. *Ecological Informatics* 6, 50–61.
- Cloern J.E. (2001) Our evolving conceptual model of the coastal eutrophication problem. *Marine Ecology Progress Series* 210, 223–253.
- Cloern J.E. and Jassby A.D. (2008) Complex seasonal patterns of primary producers at the land–sea interface. *Ecological Letters* 11, 1294–1303.
- Cushing D.H. (1989) A difference in structure between ecosystems in strongly stratified waters and in those that are only weakly stratified. *Journal of Plankton Research* 11, 1–13.
- Flynn K.J. (2003) Modelling multi-nutrient interactions in phytoplankton; balancing simplicity and realism. *Progress in Oceanography* 56, 249–279.
- Flynn K.J., Raven J.A., Rees T.A.V., Finkel Z., Quigg A. and Beardall J. (2010) Is the growth rate hypothesis applicable to microalgae? *Journal of Phycology* 46, 1–12.

- Franks P.J.S.** (2009) Planktonic ecosystem models: perplexing parameterizations and a failure to fail. *Journal of Plankton Research* 31, 1299–1306.
- Glibert P.M., Anderson D.M., Gentien P., Graneli E. and Sellner K.G.** (2005) The global, complex phenomena of harmful algal blooms. *Oceanography* 18, 136–147.
- Gregg W.W., Casey N.W. and McClain C.R.** (2005) Recent trends in global ocean chlorophyll. *Geophysical Research Letters* 32, L03606. doi: 10.1029/2004GL021808.
- Grimes C.B. and Finucane J.H.** (1991) Spatial distribution and abundances of larval and juvenile fish, chlorophyll and macrozooplankton around the Mississippi River discharge plume, and the role of the plume in fish recruitment. *Marine Ecology Progress Series* 75, 109–119.
- Hager S.W. and Schemel L.E.** (1996) Dissolved inorganic nitrogen, phosphorus, and silicon in South San Francisco Bay. I. Major factors affecting distributions. In Hollibaugh J.T. (eds) *San Francisco Bay: the ecosystem*. San Francisco, CA: Association for the Advancement of Science, pp. 1986–2216.
- HAN (Harmful Algae News)** (2011) *Special issue: the 14th international conference on harmful algae*. Hersonissos, Crete: UNESCO. www.ioc-unesco.org/hab (accessed 13 March 2012).
- Han M.W. and Park Y.C.** (1999) The development of anoxia in the artificial Lake Shihwa, Korea, as a consequence of intertidal reclamation. *Marine Pollution Bulletin* 38, 1194–1199.
- Harding L.W. Jr** (1994) Long-term trends in the distribution of phytoplankton in Chesapeake Bay, 1950–2001: long-term changes in relation to nutrient loading and river flow. *Estuaries* 27, 634–658.
- Hardman-Mountford N.J., Richardson A.J., Boyer D.C., Kreiner A. and Boyer H.J.** (2003) Relating sardine recruitment in the Northern Benguela to satellite-derived sea surface height using a neural network pattern recognition approach. *Progress in Oceanography* 59, 241–255.
- Heil C.A., Glibert P.M. and Fan C.** (2005) *Prorocentrum minimum* (Pavillard) Schiller: a review of a harmful algal bloom species of growing worldwide importance. *Harmful Algae* 4, 449–470.
- Kaski S., Kangas J. and Kohonen T.** (1998) Bibliography of self-organizing map (SOM) papers: 1981–1997. *Neural Computer Survey* 1, 102–350.
- Kim T., Park Y., Lee H. and Kim D.** (2004) The environmental impacts of seasonal variation on characteristics of geochemical parameters in Lake Shihwa, Korea. *Journal of Environmental Science* 13, 1089–1102.
- Kiviluoto K.** (1996) Topology preservation in self-organizing maps. In *Proceedings of the International Conference on Neural Networks (ICNN'96), volume 1, Piscataway, New Jersey, USA, June 1996*. IEEE Neural Networks Council, pp. 294–299.
- KMA (Korea Meteorological Administration)** (2010) Available at <http://web.kma.go.kr> (accessed 13 March 2012).
- Kohonen T.** (1982) Self-organizing information of topologically correct features maps. *Biological Cybernetics* 43, 59–69.
- Kohonen T.** (2001) *Self-Organizing Map, Springer Series in Information Sciences*. 3rd edition. Berlin: Springer-Verlag.
- Latif B.A. and Mercier G.** (2010) Self-Organizing Maps for processing of data with missing values and outliers: application to remote sensing images. In Matsopoulos G.K. (ed.) *Self-Organizing Maps*. InTech, pp. 189–210.
- Livingston R.J.** (2001) *Eutrophication processes in coastal systems*. Boca Raton, FL: CRC Press.
- Longhurst A.** (1995) Seasonal cycles of pelagic production and consumption. *Progress in Oceanography* 36, 77–167.
- Lucas L.V. and Cloern J.E.** (2002) Effects of tidal swallowing and deepening on phytoplankton production dynamics: a modeling study. *Estuaries* 25, 497–507.
- The Mathworks** (2001) *The Mathworks Inc., 2001*. MATLAB Version 6.1, Massachusetts.
- Mee L.** (1992) The Black Sea in crisis: a need for concerted international action. *Ambio* 21, 278–286.
- MOMAF (Ministry of Maritime Affairs and Fisheries)** (2006) *PEMSEA parallel site for integrated coastal management in Republic of Korea: Shihwa Lake*. Seoul, Korea: MOMAF.
- Monbet Y.** (1992) Control of phytoplankton biomass in estuaries: a comparative analysis of microtidal and macrotidal estuaries. *Estuaries* 15, 563–571.
- NFRDI (National Fisheries Research and Development Institute)** (2008) Available at <http://portal.nfrdi.re.kr/> (accessed 13 March 2012).
- Oh S., Kim M.K., Yi S.M. and Zoh K.D.** (2010) Distributions of total mercury and methylmercury in surface sediments and fishes in Lake Shihwa, Korea. *Science of the Total Environment* 408, 1059–1068.
- Oja M., Kaski S. and Kohonen T.** (2002) Bibliography of self-organizing map (SOM) papers: 1998–2001 addendum. *Neural Computer Survey* 3, 1–156.
- Park G.S. and Park S.Y.** (2000) Long-term trends and temporal heterogeneity of water quality in tidally mixed estuarine waters. *Marine Pollution Bulletin* 40, 1201–1209.
- Park Y.C., Kim S.J. and Han M.W.** (2000) Nutrients loading and its impact on the coastal environment: Kyeonggi Bay, Korea. *The Yellow Sea* 6, 73–76.
- Park J.K., Kim E.S., Cho S.P., Kim K.T. and Park Y.C.** (2003a) Annual variation of water quality in the Shihwa Lake. *Ocean and Polar Research* 25, 459–468.
- Park Y.S., Cereghino R., Compin A. and Lek S.** (2003b) Applications of artificial neural networks for patterning and predicting aquatic insect species richness waters. *Ecological Modelling* 160, 265–280.
- Park Y.S., Chon T.S., Kwak I.S. and Lek S.** (2004) Hierarchical community classification and assessment of aquatic ecosystem using artificial neural networks. *Science of the Total Environment* 327, 105–122.
- Park Y.S., Song M.Y., Park Y.C., Oh K.H., Cho E. and Chon T.S.** (2007) Community patterns of benthic macroinvertebrates collected on the national scale in Korea. *Ecological Modelling* 203, 26–33.
- Parsons T.R., Maita Y. and Lalli C.M.** (1984) *A manual of chemical and biological methods for seawater analysis*. Oxford: Pergamon Press.
- Passarge J., Hol S., Escher M. and Huisman J.** (2006) Competition for nutrients and light: stable coexistence, alternative stable states, or competitive exclusion? *Ecological Monographs* 76, 57–72.
- Philippart C.J.M., Cadee G.C., van Raaphorst W. and Riegman R.** (2000) Long-term phytoplankton–nutrient interactions in a shallow coastal sea: algal community structure, nutrient budgets, and denitrification potential. *Limnology and Oceanography* 45, 131–144.
- Postma H.** (1967) Sediment transport and sedimentation in the estuarine environment. In Lauff G.H. (ed.) *Estuaries*. Washington, DC: American Association for the Advancement of Science Publications, pp. 158–179.
- Rabalais N.N., Turner R.E., Diaz R.J. and Justic D.** (2009) Global change and eutrophication of coastal waters. *ICES Journal of Marine Science* 66, 1528–1537.
- Rabouille C., Conley D.J., Dai M.H., Cai W.-J., Chen C.T.A., Lansard B., Green R., Yin, K., Harrison P.J., Dagg M. and Mckee B.** (2008) Comparison of hypoxia among four river dominated ocean

- margins: the Changjiang (Yangtze), Mississippi, Pearl and Rhone Rivers. *Continental Shelf Research* 28, 1527–1537.
- Radach G., Berg J. and Hagmeier E.** (1990) Long-term changes of the annual cycles of meteorological, hydrographic, nutrient and phytoplankton time series at Helgoland and at V ELBE 1 in the German Bight. *Continental Shelf Research* 10, 305–328.
- Richardson A.J., Pfaff M.C., Field J.G., Silulwane N.F. and Shillington F.A.** (2002) Identifying characteristic chlorophyll-a profiles in the coastal domain using an artificial neural network. *Journal of Plankton Research* 24, 1289–1303.
- Richardson A.J., Risien C. and Shillington F.A.** (2003) Using self-organizing maps to identify patterns in satellite imagery. *Progress in Oceanography* 59, 223–239.
- Risien C.M., Reason C.J.C., Shillington F.A. and Chelton D.B.** (2004) Variability in satellite winds over the Benguela upwelling system during 1999–2000. *Journal of Geophysical Research* 109, C03010. doi:10.1029/2003JC001880.
- Ryn J.S., Choi J.W., Kang S.G., Koh C.H. and Huh S.H.** (1997) Temporal and spatial changes in the species composition and abundances of benthic polychaetes after the construction of Shihwa Dyke. *The Sea, Journal of the Korean Society of Oceanography* 2, 101–109.
- Shanmugam P., Ahn Y.H. and Ram P.S.** (2008) SeaWiFS sensing of hazardous algal blooms and their underlying mechanisms in shelf-slope waters of the Northwest Pacific during summer. *Remote Sensing of Environment* 112, 3248–3270.
- Sharp J.H.** (1994) What not to do about nutrients in the Delaware Estuary. In Dyer K.R. and Orth R.J. (eds) *Changes in fluxes in estuaries: implications from science to management*. Fredensborg, Denmark: Olsen and Olsen, pp. 423–428.
- Silulwane N.F., Richardson A.J., Shillington F.A. and Mitchell-Innes B.A.** (2001) Identification and classification of vertical chlorophyll patterns in the Benguela upwelling system and Angola–Benguela Front using an artificial neural network. *South African Journal of Marine Science* 23, 37–51.
- Smayda T.J.** (1997) Harmful phytoplankton blooms: their ecophysiology and general relevance. *Limnology and Oceanography* 42, 1137–1153.
- Smayda T.J.** (2008) Complexity in the eutrophication–harmful algal bloom relationship, with comment on the importance of grazing. *Harmful Algae* 8, 140–151.
- Song M.Y., Hwang H.J., Kwak I.S., Ji C.W., Oh Y.N., Youn B.J. and Chon T.S.** (2007) Self-organizing mapping of benthic macroinvertebrate communities implemented to community assessment and water quality evaluation. In *Ecological Modelling: Special issue on ecological informatics: biologically-inspired machine learning. 4th conference of the International Society for Ecological Informatics, Volume 203*, pp. 18–25.
- Sorokin Y.I.** (1983) The Black Sea. In Ketchum B.H. (ed.) *Ecosystems of the world: estuaries and enclosed seas*. New York: Elsevier, pp. 253–292.
- Su S., Zhi J., Lou L., Huang F., Chen X. and Wu J.** (2011) Spatio-temporal patterns and source apportionment of pollution in Qiantang River (China) using neural-based modeling and multivariate statistical techniques. *Physics and Chemistry of the Earth, Parts A/B/C* 39, 379–386.
- Totterdell I.J., Armstrong R.A., Drange H., Parslow J.S., Powell T.M. and Taylor A.H.** (1993) Trophic resolution. In Evans G.T. and Fasham M.J.R. (eds) *Towards a model of ocean biogeochemical processes*. Berlin: Springer-Verlag, pp. 71–92.
- Tuncer G., Karakas T., Balkas T.I., Gokcay C.F., Aygnn S., Yurteri C. and Tuncel G.** (1998) Land-based sources of pollution along the Black Sea coast of Turkey: concentrations and annual loads to the Black Sea. *Marine Pollution Bulletin* 36, 409–423.
- Ultsch A. and Morchen F.** (2006) U-Maps: topographic visualization techniques for projections of high dimensional data. In *Proceedings of the 30th annual conference of the German Classification Society*, pp. 1–7.
- Ultsch A. and Roske F.** (2002) Self-organizing feature maps predicting sea levels. *Information Sciences* 144, 91–125.
- Vesanto J., Himberg J., Alhoniemi E. and 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.
- Xu J., Yin K., Liu H., Lee J.H.W., Anderson D.M., Ho A.Y.T. and Harrison P.J.** (2010) A comparison of eutrophication impacts in two harbours in Hong Kong with different hydrodynamics. *Journal of Marine Systems* 83, 276–286.
- Yang E.J., Choi J.K. and Hyun J.H.** (2008) Seasonal variation in the community and size structure of nano and microzooplankton in Gyeonggi Bay, Yellow Sea. *Estuarine, Coastal and Shelf Science* 77, 320–330.
- and
- Yoo H., Yamashita N., Taniyasu S., Lee K.T., Jones P.D., Newsted J.L., Khim J.S. and Giesy J.P.** (2009) Perfluoroalkyl acids in marine organisms from Lake Shihwa, Korea. *Archives of Environmental Contamination and Toxicology* 57, 552–560.

#### Correspondence should be addressed to:

Joong Ki Choi  
5N 230, Plankton Laboratory  
Department of Oceanography, INHA University  
253, Yonghyun-Dong, Nam-gu, Incheon 402-751, Korea  
email: jkchoi@inha.ac.kr