Available online at www.sciencedirect.com
SCIENCE DIRECT®

Ecological Modelling 193 (2006) 602–614

ECOLOGICAL MODELLING

www.elsevier.com/locate/ecolmodel

Multivariate analysis and self-organizing mapping applied to analysis of nest-site selection in Black-tailed Gulls

Who-Seung Lee [a], Young-Soo Kwon [a], Jeong-Chil Yoo [a,*], Mi-Young Song [b], Tae-Soo Chon [b]

[a] *The Korea Institute of Ornithology and Department of Biology, KyungHee University, Seoul 130-701, Republic of Korea*
[b] *Division of Biological Sciences, Pusan National University, Pusan 609-735, Republic of Korea*

Received 4 September 2004; received in revised form 3 August 2005; accepted 31 August 2005
Available online 3 November 2005

Abstract

The factors affecting nest-site selection and breeding success of Black-tailed Gulls (*Larus crassirostris*) were studied in Hongdo Island in Korea during the breeding seasons in 2002 and 2003. Two analyzing methods, Principal Component Analysis (PCA) and Self-Organizing Map (SOM) – an unsupervised learning method in artificial neural networks, were applied to multivariable datasets characterizing nest-sites of the gulls. Both methods provided insights on the major trends in nest-site selection by Black-tailed Gulls. PCA showed that the variables regarding the "wall" effect such as rock cover and nest-wall (positively), and the nearest distance between neighbors (negatively) were related to breeding success of Black-tailed Gulls. SOM confirmed ordination of the sample sites by PCA and efficiently classified nest-sites according to environmental condition for breeding. Grouping based on the "wall" effect on PCA was more finely revealed in subdivision on SOM regarding the variables of slope and the nearest distance between neighbors. The use of techniques in ecological informatics such as SOM would be an efficient tool in analyzing data for breeding behavior of birds.

Keywords: Artificial neural network; Black-tailed Gulls; Nest-site characteristics; Nest-site selection; Principal Component Analysis; Self-Organizing Map

1. Introduction

The suitable selection of nest-sites is critical for the success of a species (Buckley and Buckley, 1980) since nest-sites influence the risk of nest-predation in various ways (Ricklefs, 1969; Martin, 1993). Numerous environmental and biological factors are involved in determining nest-site selection (Lack, 1968; Partridge, 1978; Cody, 1985; Burger, 1985; Good, 2002). Nest-predation rates have been shown to vary with nest-site characteristics such as vegetation cover, vegetation height (Burger and Gochfeld, 1981, 1987), and slope

* Corresponding author. Tel.: +82 2 961 0849; fax: +82 2 964 0591.
E-mail addresses: whoseung@khu.ac.kr (W.-S. Lee), jcyoo@khu.ac.kr (J.-C. Yoo).

0304-3800/$ – see front matter © 2005 Elsevier B.V. All rights reserved.
doi:10.1016/j.ecolmodel.2005.08.043

(Burger and Gochfeld, 1988, 1991) for a wide range of bird species (Ricklefs, 1969; Collias and Collias, 1984; Martin, 1993). Pierotti (1982) reported that Herring Gulls (*Larus argentatus*) preferred rocky habitats in Newfoundland Island since rocks provided the shelters against the attack by predators. Additionally parental breeding experience (Lack, 1968; Burger and Gochfeld, 1991) and quality (Nisbet et al., 1998; Good, 2002) have been reported to be essential for selection of suitable nest-sites. Parents with much experience and better quality should select nest-sites with complex characteristics (Lack, 1968) that could provide more reliable protection for higher breeding success such as natural screens around nests and boundaries among territorial pairs, etc. (Cezilly and Quenette, 1988; Good, 2002). Martin and Roper (1988) also suggested that dense vegetation cover was useful for nest protection from predators in a majority of ground-nesting birds.

In this study, we intend to reveal the important factors on breeding success of the Black-tailed Gull (*Larus crassirostris*) residing in Korea. The Black-tailed Gull is one of the most abundant birds in Korea, Japan, East Russia, and China, and breeds colonially on islets, islands, or rocky cliffs (Paek and Yoo, 1996; Kwon, 2004). Hongdo Island, one of the main breeding sites in Korea, accommodates the largest breeding colony of Black-tailed Gulls, numbering about 10,000 pairs, and is designated as a natural monument for conservation in Korea. However nest-site characteristics regarding suitability of breeding success in birds have not been closely examined.

In order to reveal nest-site characteristics for determination of nesting success of Black-tailed Gulls, we used two processes for analysis. Firstly, we checked nest-site characteristics that are more suitable for nest-site selection by comparing with random points. And secondly we further evaluated nest-site characteristics contributing to the birds' breeding success by examining successful and unsuccessful nests.

So as to analyze the datasets for the factors characterizing habitats, we employed a relatively new method, Self-Organizing Map (SOM) in artificial neural networks in behavioral study along with conventional Principal Component Analysis (PCA). Artificial neural networks have been extensively used for patterning and analyzing data in ecological sciences recently (Zurada, 1992; Smith, 1996; Lusk et al., 2003), and have been especially useful for dealing with complex and non-linear data in various ways (Lek and Guégan, 1999): classifying groups (Chon et al., 1996; Levine et al., 1996; Park et al., 2003a,b), revealing complex relationships (Lek et al., 1996; Tuma et al., 1996), predicting population and community development (Tan and Smeins, 1996; Recknagel et al., 1997; Chon et al., 2000), modeling habitat suitability (Paruelo and Tomasel, 1997; Özesmi and Özesmi, 1999), and patterning community parameters (Park et al., 2003a). To the field of breeding biology, however, artificial neural networks have been rarely used except a few cases for prediction of the presence of small-bodied fish (Mastrorillo et al., 1997) and determination of nest-site selection and breeding success for birds (Lusk et al., 2003). We utilized SOM in revealing breeding behaviors of a bird population, Black-tailed Gulls, in association with nest-site characterization in this study. Information on breeding success extracted from behavioral data would be useful for delivering habitat quality of the gulls and would be further available for establishing conservation policies for management of Black-tailed Gulls.

2. Methods

2.1. Study site

Fieldwork was carried out in Hongdo Island during the breeding seasons (April–August) of 2002 and 2003. Hongdo Island (34°31′87″N, 128°43′88″E) is located about 23 km south of Geojedo, Gyeongsangnam-Do, Republic of Korea (Fig. 1). The area of the island is 98,380 m^2 while the highest point of the island is about 115 m above sea level. Cliffs surround the coastline of Hongdo Island and are inclined >45°. The maximum ambient temperature and the amount of rainfall from April to July are approximately 30.5 °C and 340 mm, respectively (Korea Meteorological Administration, 2004). The flora on Hongdo Island is dominated by the sedge (*Carex boottiana*) which has been reported to be used as nest material for the birds (Fig. 1) (Kwon, 2004; Lee, 2004). The island is additionally vegetated with *Camellia japonica*, *Opuntia ficus-indica*, *Aster spathulifolius*, and *Taraxacum mongolicum* (Cultural Properties Administration, 2003).

We selected 130 (and 128) nest-sites and 60 (and 60) random points in 2002 (and 2003). Nest-sites and

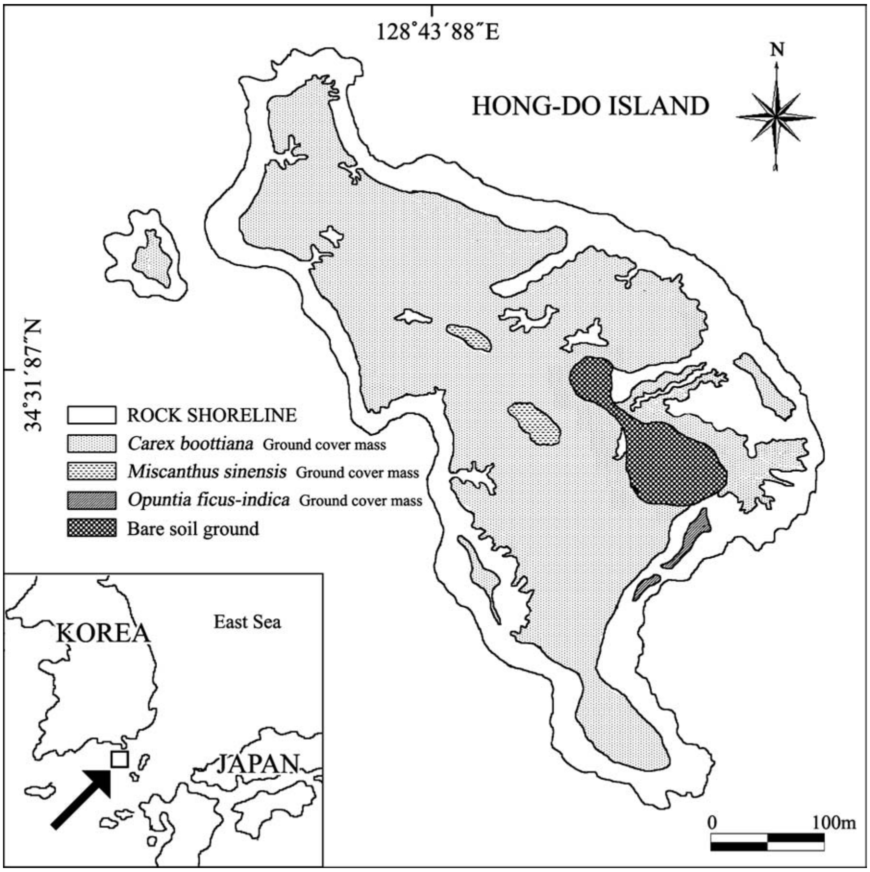

Fig. 1. A map of Hongdo Island.

random points were chosen by selecting x and y coordinates from a table of random numbers (Burger and Gochfeld, 1986; Bosch and Sol, 1998).

2.2. Biological data

During the breeding season (April–August) of 2002 and 2003, we visited the nests either every day to check hatching until 15 days after hatching. Chicks' mortality rapidly decreased from 15 days on after hatching (Kwon, 2004). We defined "Surviving" chicks if they survived until 15 days after hatching and "Non-surviving" if they are dead within this period (Hunt and Hunt, 1976; Kwon, 2004). We further regarded the nests as "Successful" if one or more young chicks fledged, while "Unsuccessful" if breeding failure occurred (Clark and Shutler, 1999). Subsequently, breeding success was separately quantified by the following formula:

$$\text{breeding success (\%)} = \frac{\text{no. of fledged chicks}}{\text{total number of eggs}} \times 100 \tag{1}$$

We evaluated suitability of the nest-sites according to the rate of breeding success. If breeding success was obtained above mean (i.e., 42.88% and 61.72% in 2002 and 2003, respectively), the sites were considered as "Good", while the nests were "Poor" if breeding success was below mean.

2.3. Environment data

To reveal association of nest-site characteristics with breeding success we measured six variables (percent of vegetation cover, height of the nearest vegetation, percent of rock cover, number of rock or vegetation walls, distance from each nest, and slope of the ground) characterizing physical structure of nest-sites: (1) percent of vegetation cover was obtained by laying down a 0.5 m^2 quadrate frame over each nest and estimating the percentage of vegetation or bare ground present (Burger and Gochfeld, 1988); (2) height of the nearest vegetation was measured from the base of nest to the top of vegetation. If the nearest vegetation was greater than 100 cm away from the nest cup, the height was recorded as not applicable since it was assumed to offer little or no cover to the nest (Burger and Gochfeld, 1988); (3) percent of rock cover was recorded by laying down a 0.5 m^2 quadrate frame (Saliva and Burger, 1989); (4) number of rock or vegetation walls (nest-wall) around the nest were measured including the amount of overhang of the nearest wall. In this study, walls were defined as moderately steep areas of substrate providing some cover to the nest or restricting its visibility from above (Burger and Gochfeld, 1988); (5) distance from each nest to its nearest neighbor was measured as the distance from the center of the two nests (Good, 2002); (6) slope of the ground supporting the nest-site was recorded in degrees using a 'Suunto' clinometer. From an imaginary line running through the nest center, we made slope measurements on the quadrants of the nest, and the averages on the quadrants were given as the slope of the nest-site (Siegel-Causey and Hunt, 1981).

2.4. Principal Component Analysis

Statistical methods have been applied to analysis of association of the variables in multivariate measures in behavioral data. In this study, we carried out PCA in revealing association of the variables for characterizing nest-sites for Black-tailed Gulls' breeding success. We used the correlation matrix consisting of the variables for PCA as stated before: vegetation cover, height of vegetation, rock cover, nest-walls, the nearest distance between neighbors, and slope. Before the analyses, all variables were transformed into normalized variables which have a mean of zero. PCA was conducted through SPSS (SPSS, 2002).

2.5. Self-Organizing Map

SOM has properties of neighborhood preservation and local resolution of the input space proportional to the data distribution (Kohonen, 1982; Park et al., 2003c). SOM consists of two layers: an input layer formed by a set of units (or neurons), and an output layer formed by units arranged in a two-dimensional grid (Fig. 2). In this study, each input unit accounts for values of each nest-site characteristic, leading to an input layer made of six neurons. The output layer was made of 250 (=25 × 10) output units in the hexagonal lattice. The SOM algorithm maps a set of input vectors (i.e., nest-sites) onto a set of vectors of output units according to the characteristics of the input vector components (i.e., nest-site characteristics). It can be interpreted as a non-linear projection of the high dimensional input data onto an output array of nodes (Park et al., 2003c). In this study, each nest-site has been assigned to one output unit as a result of SOM algorithm calculation. Each output unit has a vector of coefficients associated with input data. The coefficient vector is referred to as weight (or connection intensity) vector $\boldsymbol{W}$ between input and output layers. The weights establish a link between the input units and their associated output units. The algorithm can be described as follows: when an input vector $\boldsymbol{X}$ (in this case, the values of six nest-site characteristics) is presented to the network, the neurons in the output layer compete with each other and the winner (whose weight has the minimum distance from the input vector) is chosen. The winner and its neighbors predefined in the algorithm update their weight vectors according to the SOM learning rules as follows:

$$w_{ij}(t+1) = w_{ij}(t) + \alpha(t) \cdot h_{jc}(t)[x_i(t) - w_{ij}(t)] \quad (2)$$

where $w_{ij}(t)$ is a weight between a neuron i in the input layer and a neuron j in the output layer at iteration time t, $\alpha(t)$ is a learning rate factor which is a decreasing function of the iteration time t, and $h_{jc}(t)$ is a neighborhood function (a smoothing kernel defined over the lattice points) that defines the size of neighborhood of the winner neuron (c) to be updated during the learning process.

This learning process is continued until a stopping criterion is met, usually, when weight vectors stabilize or when a number of iterations are completed. This learning process results in training the network to pattern the input vectors and preserves the connection intensities in the weight vectors. For a detailed algorithm for SOM, refer to Kohonen (1982), Lippmann

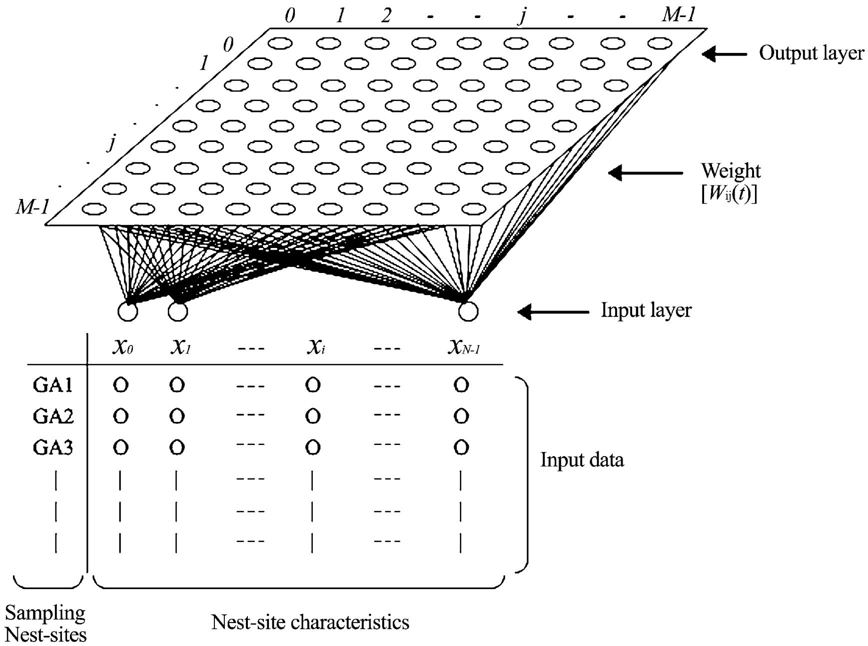

Fig. 2. A two-dimensional Kohonen Self-Organizing Map. The input datasets were shown below the self-organizing mapping. $X_0, \ldots, X_{N-1}$ are the input variables (nest-site characteristics), while GA1, GA2, etc., are the observation data (sampling nest-sites).

(1987), Chon et al. (1996), and Lek and Guégan (1999).

After training with SOM, the cluster boundaries on the trained maps were further determined by the unified-matrix (U-matrix) algorithm (Ultsch and Siemon, 1990; Ultsch, 1993). The U-matrix calculates distance of the weights in the network between neighbor map nodes, and displays the degree of association in the cluster structure of the map, i.e., how the patterned nodes were close among different groupings. High values of the U-matrix indicate the cluster boundaries, while low values reveal clusters themselves. SOM and U-matrix were conducted through the SOM toolbox (Alhoniemi et al., 2000) available in Matlab (The Math Works, 1999).

2.6. Statistical methods

Regarding the data are complex and generally do not follow the patterns of normality and homoscedasticity, we used a non-parametric statistical method, Kruskal–Wallis χ^2-test, to determine differences in characteristics in different sites: between nest-sites and random points, and between successful and unsuccessful nests (Burger and Gochfeld, 1987; Saliva and Burger, 1989; Ramos, 1998). All the means were presented with the range from minimum to maximum. Linear regression analysis was carried out to examine the relationships between principal components and breeding success (Rising and Somers, 1989; Freeman and Jackson, 1990; Yang et al., 2003). Statistical analyses were conducted through SPSS (SPSS, 2002).

3. Results and discussion

3.1. Environmental conditions for nests

All nest-site characteristics were significantly different between nest-sites and random points (Table 1). The higher values of the variables (e.g. larger vegetation cover, higher vegetation height, etc.) appeared to improve environmental conditions for nest-site selection. The contribution of these variables to habitat suitability has been confirmed by other research (Cody, 1985; Burger and Gochfeld, 1986, 1987, 1991). The nearest distance was significantly shorter in the nest-sites than in the random sites, indicating that spacing would contribute to characterization of the nest-sites.

Table 1
Comparison of variables characterizing nests and random points on Hongdo Island

Nest-site characteristics	Nest ($n = 258$)	Random point ($n = 120$)	χ^2
Vegetation cover (%)	52.5 (0–100)	40.0 (0–100)	13.45***
Height of vegetation (cm)	57.3 (0–78.5)	46.5 (0–79.5)	12.36***
Rock cover (%)	40.0 (5.0–100)	25.0 (0–100)	29.30***
Nest-wall (%)	50.0 (0–100)	25.0 (0–75.0)	81.66***
Nearest distance between neighbor (cm)	87.7 (42.0–186.0)	91.0 (57.0–190.0)	3.93*
Slope at nest (°)	4.0 (−14.0 to 15.0)	−2.0 (−9.5 to 7.5)	121.56***

Given are median, range and Kruskal–Wallis χ^2-values.
* $P < 0.05$.
*** $P < 0.001$.

We further checked other variables relating to spatial dimension such as the sea-level height of the nest, *x* and *y* coordinates, distance from sea, etc. in preliminary survey. Contribution of these variables to nest-site selection, however, was not clearly revealed, indicating that these factors relating to spatial dimension were not as critical as the factors relating to nest structure.

Subsequently, significant differences in the nest-site characteristics were also observed between successful and unsuccessful nests (Table 2). All the variables, except for slope, were different between successful and unsuccessful nests. The level of significance showed the highest levels in vegetation cover and nest-wall (Table 2). The nearest distance was also significantly different between the successful and unsuccessful nests.

However, slope at nests was not significantly different between the successful and unsuccessful nests. This indicated that slope would not contribute to breeding success considerably, although the slope appeared to be important with the initial determination of the nest-sites (Table 1). Slope is in general known to be related to prevention of eggs from being rolled off from cliffs or into the sea, thus it would be an important factor in selecting nest-site from breeding habitats (Table 1). However, once the nest-site was chosen, the rolling effect on the eggs would not be decisive in characterizing the successful nests, since all the nest-sites used for determining the successful or unsuccessful nests could be already chosen to have the higher levels of slope according to Table 1. In this regard, topographic conditions of Hongdo Island may be more closely investigated in the future.

The other variables characterizing structure of nests were still effective in differentiating the successful nests from the unsuccessful nests (Table 2). These findings were matched to previous results. Nest-site characteristics directly influenced survival of offspring and incubating adults (Burger and Gochfeld, 1991). Variations in environmental factors at the selected nests were related to the presence of predators as well as weather constraints (Cody, 1985; Bukacinska and Bukacinski, 1993). For instance, Mew Gull (*Larus canus*) nest-sites differed from random points with respect to exogenous factors (e.g. rock cover, slope, and vegetation cover) in addition to endogenous factors (e.g. hatching success

Table 2
Comparison of variables characterizing successful and unsuccessful nests on Hongdo Island

Nest-site characteristics	Successful nest ($n = 174$)	Unsuccessful nest ($n = 84$)	χ^2
Vegetation cover (%)	65.0 (0–100.0)	40.0 (0–100)	18.36***
Height of vegetation (cm)	51.0 (0–78.5)	48.0 (0–71.0)	5.40*
Rock cover (%)	45.0 (5–100.0)	35.0 (5–100.0)	5.38*
Nest-wall (%)	75.0 (0–100.0)	25.0 (0–75.0)	83.96***
Nearest distance between neighbor (cm)	84.0 (42.0–137.0)	93.0 (53.2–186.0)	6.91**
Slope at nest (°)	4.0 (−14.0 to 14.0)	4.5 (−12.0 to 15.0)	2.77

Given are median, range and Kruskal–Wallis χ^2-values.
* $P < 0.05$.
** $P < 0.01$.
*** $P < 0.001$.

and fledgling, etc.) (Burger and Gochfeld, 1987). This study confirmed importance of environmental variables for characterizing morphometrics of nest-site such as vegetation cover, nest-wall, rock cover, etc. in determining the successful nest-sites.

Although the variables were evaluated in determining success of the nest-sites (Tables 1 and 2), each variable was separately checked in this case. Association among the variables was not collectively revealed. In other words, the habitat suitability would be more efficiently addressed if association of all the variables is analyzed in one dataset. Multivariate analyses were further carried out to characterize the variables for the nest-sites.

3.2. Ordination of nest-sites with PCA

PCA revealed association among the variables representing nest-site characteristics. We used the three main components with the percent of the total variance 27.23%, 22.25% and 16.40% for the Components I, II and III respectively (Fig. 3). The first principal component appeared to be mainly related to the factors covering the nests. Vegetation cover and height of vegetation were located at the positive side, while rock cover was placed at the negative side on the Component I. Rock cover also appeared to be important on the Component II, being located at the positive side along with nest-wall. The nearest distance between neighbors was placed at the negative side on the Component II. For the Component III, the slope was solely appeared at the positive side.

Overall, vegetation cover and height of vegetation showed higher levels of association in PCA. These results are understandable, regarding that abundant covering would be advantageous for providing protection for the gulls as screens against avian predators (i.e., Peregrine Falcons) and as shades from the sun (Cody, 1985). The second principal component additionally represented the variables related to the "wall" effect. The rates of rock cover and nest-wall were positively related to the "wall" effect, while the nearest distance between the neighbors was negatively associated with the "wall" effect. Although rock cover played a negative role for covering the gulls in Component I, rock cover was presented as a positive factor regarding the "wall" effect on Component II. The rock cover would provide blocking for eggs and chicks against the

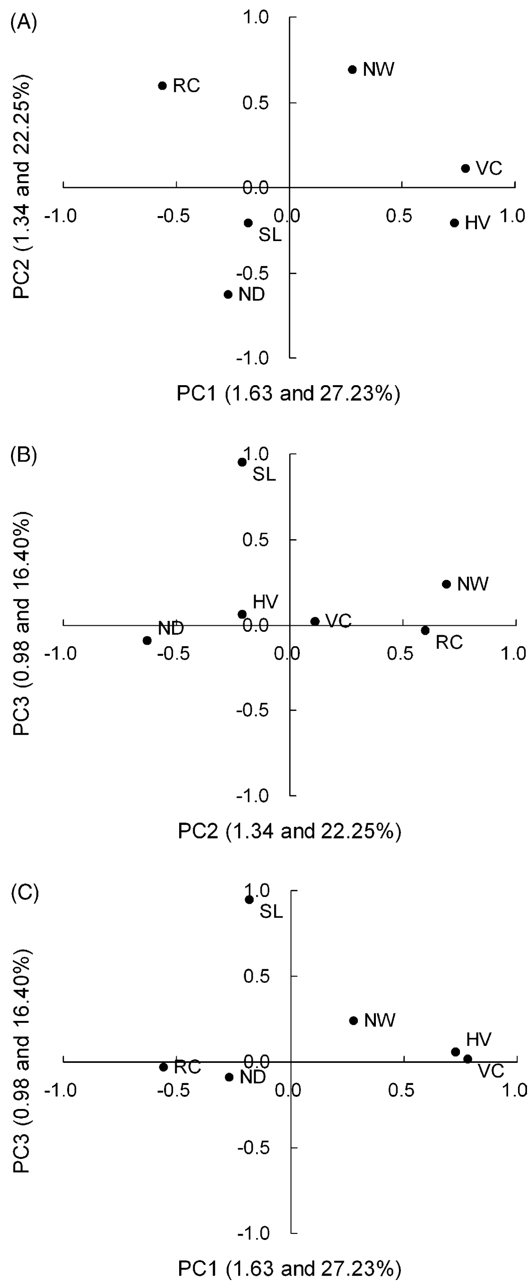

Fig. 3. Principal components analysis of nest-site characteristics (VC = vegetation cover, HV = height of vegetation, RC = rock cover, NW = nest wall, ND = the nearest distance between neighbors, SL = slope) measured in 258 nest-sites on Hongdo Island: (A) PC1 and PC2; (B) PC2 and PC3; (C) PC3 and PC1. The parenthesis in each axis was the eigenvalue and variation percentage in PCA.

attack by neighbor adults in this regard. The third principal component was associated with slope. The higher values were advantageous for the gulls choosing nest locations with higher slope as stated above (Table 1).

The variables for breeding success were significantly and positively associated with Component II (linear regression, $R^2 = 27.89\%$, $t_{254} = 9.98$, $P < 0.001$), while the variables were not significantly correlated with Components I and III (Component I: $R^2 = 0.93\%$, $t_{254} = 1.83$, $P = 0.07$; Component III: $R^2 = 0.02\%$, $t_{254} = -0.29$, $P = 0.78$, Table 3). This indicated that the variables related with nest-wall (i.e., the "wall" effect), rock cover and the nearest distance between neighbors, were more selectively advantageous for breeding success.

Importance of the "wall" effect for breeding success was previously reported. Lee (2004) reported that mortality of eggs and chicks was mostly caused by

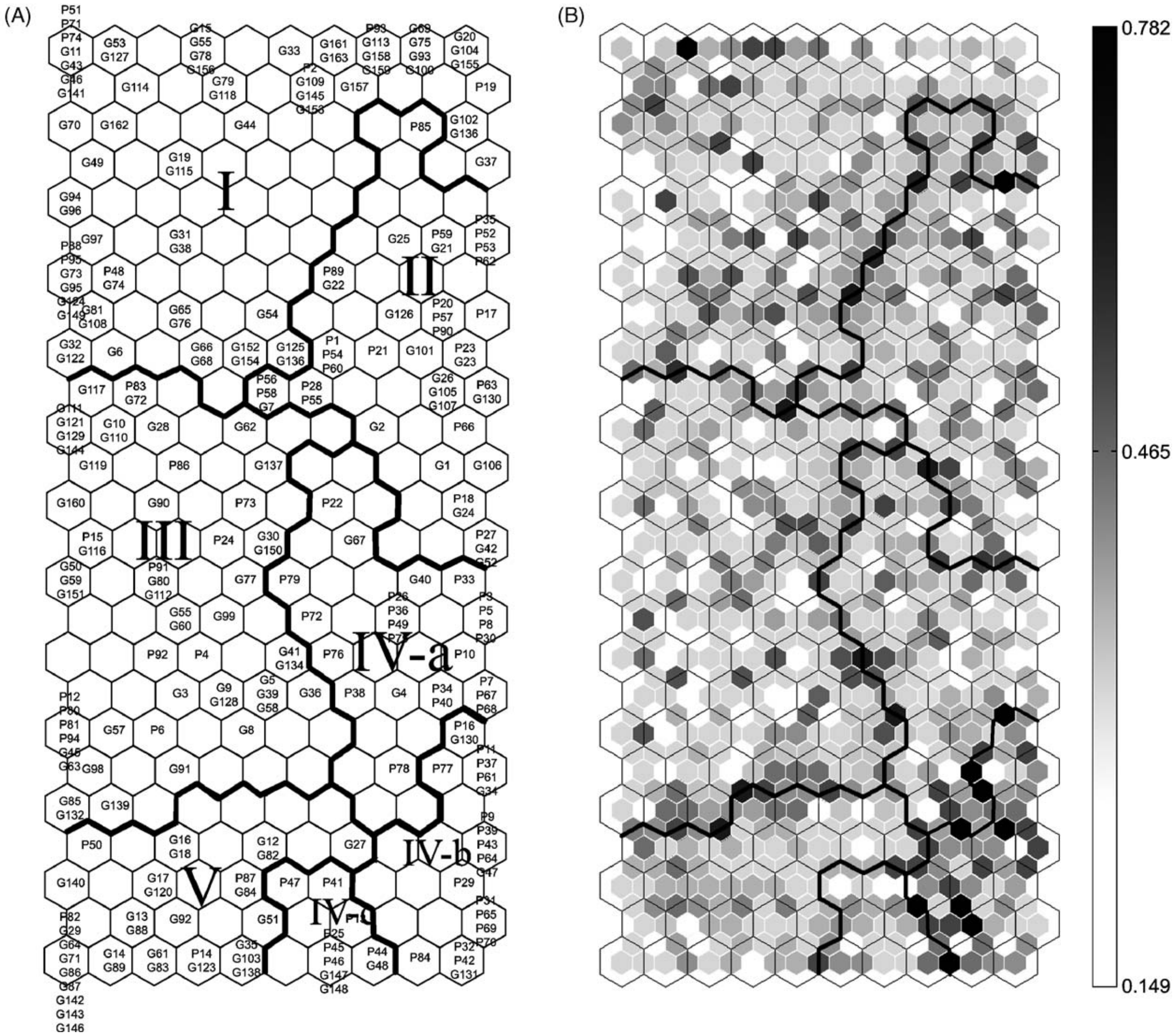

Fig. 4. Classification of the nest sites on the SOM based on nest-site characteristics (A), and the U-matrix clustering SOM units (B). The Latin numbers (I–V) on SOM (A) display clusters defined based on the U-matrix. The codes in each unit of the map represent the sampling nest-sites (G = good success, P = poor success; see Section 2.2).

Table 3
Linear regression analysis of breeding success in Black-tailed Gulls on nest-site characteristics on Hongdo Island

Model	Coefficient[a]	R^2 (%)	t	P
PC1	4.074	0.93	1.826	0.069
PC2	22.262	27.89	9.977	<0.001
PC3	−0.636	0.02	−0.285	0.776

[a] Coefficient is the slope of regression.

"disappearance" and "killing" from the attack by neighboring adults. The likelihood of nest failure increases as distance between neighboring nests decreases because conspecific adults destroy eggs and kill chicks. Our results on the "wall" effect regarding the nearest distance between neighbors confirmed this observation.

It is notable that the variables strongly related to Component I, vegetation cover and height of vegetation, did not show significant relationships with breeding success, while the variables associated with Component II, rock cover and the nearest distance between neighbors, were statistically associated with breeding success according to Table 1. Although rock cover and the nearest distance between neighbors were associated with the nest-site selection, the variables' relationships to breeding success per se were not clearly revealed. Regarding that 'breeding success' was defined as the percents of the number of fledged birds to the total number of eggs (formula (1)), the survival of young chicks was more closely related to the "wall" effect rather than to overall morphological structure of the nest-sites such as vegetation cover and height of vegetation. This suggested that, in the phase of rearing chicks in the latter part of nest establishment, the factors for protecting young birds against other neighbors may be critical for breeding success. The factors related to overall morphology of the nest-sites such as vegetation cover and height of vegetation may be related to early phase in nest establishment for providing protection for adults from other predators for instance. At this point the causality mechanism between the variables and nest-site selection by the birds in different phases of nest establishment could not be discussed in detail, and will be dealt with elsewhere. In this study, however, the results indicated that different phases may exist in association of variables with nest-site selection in the procedure of the nest-site establishment.

Although a general trend on association of characterizing variables with breeding success was disclosed,

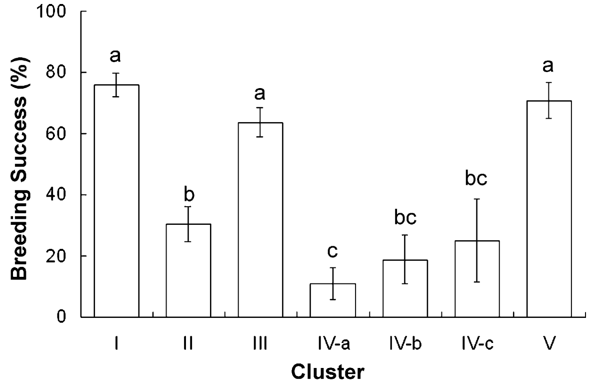

Fig. 5. Difference of breeding success among clusters (I–V). Error bars are standard errors. The same characters on the bars indicate no significant difference at the 5% level of confidence by using the LSD multiple comparison test.

PCA is in general limited in its application to non-linear data for dimension compression (Melssen et al., 1993). The data collected in field, however, are mostly complex and non-linear (James and McCulloch, 1990). We utilized SOM to extract information from the data for the nest-site characterization used in this study.

3.3. Classification of nest-sites with SOM

Fig. 4 showed grouping of nest-sites after training with SOM. Groups were produced according to "Good" or "Poor" sites for breeding as defined in Section 2.2. "Good" sites were mostly located on the left area on the map, while "Poor" sites were placed on the right area of the map (Fig. 4). Using the U-matrix, nest-sites were further classified into seven groups in two levels (I–III, IV-a, IV-b, IV-c and V) (Fig. 4A and B). In the U-matrix, the distances were rescaled from 0.149 to 0.782. Bright areas with low values in the U-matrix depict shorter distances, while dark areas with high values represent longer distances to the surrounding neighbors (Fig. 4B). We chose "0.500" as the criteria for establishing boundaries between clusters on SOM for clustering. Clusters I, III and V accommodated mainly "Good" sites, while Clusters II, IV-a, IV-b and IV-c covered "Poor" sites (Figs. 4A and 5).

To understand the effects of environmental variables on breeding success, the values of nest-site characteristics were visualized from weight vectors of trained SOM (Fig. 4A). Variables were characteristically pre-

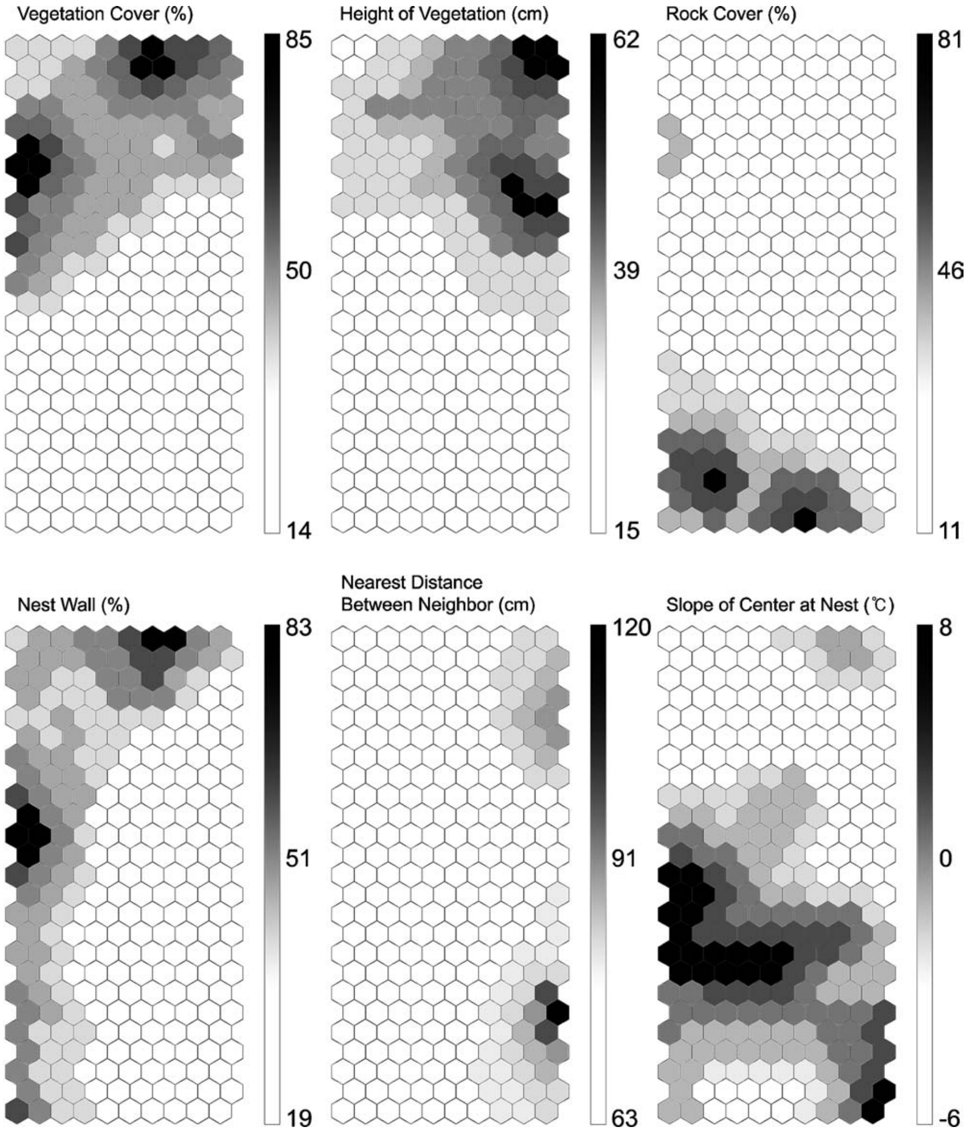

Fig. 6. Variables characterizing nest-sites matching to the nodes of sample sites grouped on Fig. 4A. The values were visualized from weight vectors of trained SOM. Dark represents a high value for each variable, whereas light is for low values.

sented in various ways in different clusters (Fig. 6). In Cluster I, vegetation cover and nest-wall showed higher values. In Cluster III, higher values were observed for nest-wall and rock cover, while rock cover separately showed higher values in Cluster V. Lower values of the nearest distance between neighbors were widely presented in the Clusters covering I, III and V.

In Clusters II and IV, the variables associated with negative effects on breeding success were presented (Fig. 4). Cluster II was associated with lower values of vegetation cover and nest-wall, while Cluster IV was related to lower values of rock cover and nest-wall. Cluster II was additionally incorporated with higher values of the nearest distance between neighbors and lower values of slope. Although height of vegetation was higher in Cluster II, vegetation cover was comparatively lower in the same compared with other clusters. Consequently high values of vegetation height could

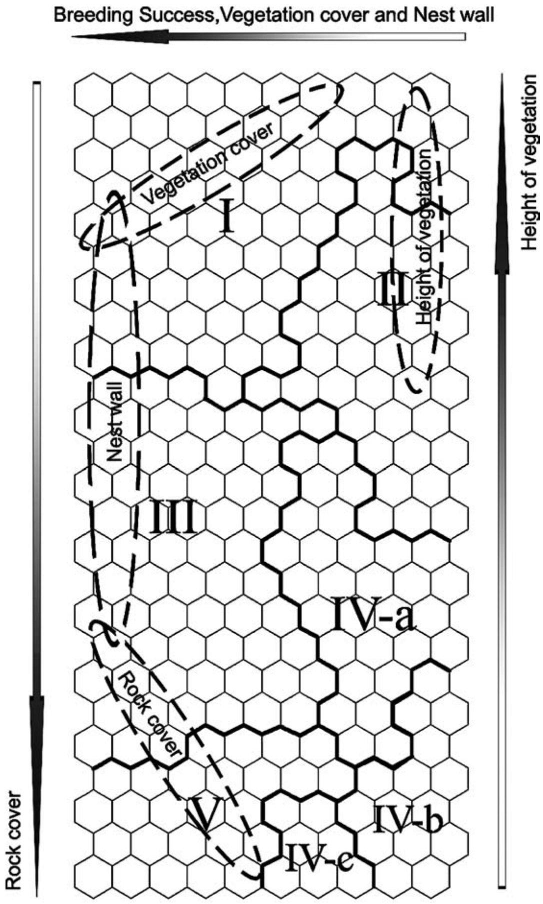

Fig. 7. Comparison of relationships between nest-sites and environmental factors on SOM. Each parameter from Figs. 4 and 6 is overlaid on the trained SOM.

not fully contribute to protection of nests due to low values of vegetation cover (Figs. 4 and 6).

Sub-division was additionally observed in Cluster IV (IV-a, IV-b and IV-c). (1) Sub-cluster IV-b was associated with high values of slope and the nearest distance between neighbors, (2) Sub-cluster IV-c was related with intermediately high values of slope and rock cover, and (3) Sub-cluster IV-a was incorporated with an intermediately high values of slope and with low values of the nearest distance between neighbors (Figs. 4 and 6). Cluster IV on SOM was in analogy with Component II in PCA: the grouped samples in Cluster IV were collectively associated with the "wall" effect, rock cover and nest-wall (Figs. 3 and 7). In comparison with PCA, however, SOM revealed detailed grouping within Cluster IV based on the variables such as the slope and the nearest distance between neighbors (Fig. 7). Cluster IV-b was located in between Cluster IV-a and IV-c. Cluster IV-b was separated from Cluster IV-a regarding the nearest neighbor distance (longer in Cluster IV-b), while slope in Cluster IV-b was different from Cluster IV-c (higher in Cluster IV-b). As stated before, slope appeared to be important on Component III in PCA (Fig. 3), although it did not show statistical contribution in determining breeding success (Table 3). The lower level of data grouping was more precisely visualized with sub-divisions in Cluster IV on SOM. Clustering in other groups was also efficiently visualized as stated previously (Fig. 7). This demonstrated that SOM could efficiently reveal finer level of clustering and would be useful for characterizing the variables in a hierarchical manner.

Regarding that conventional statistical methods on multi-variables including PCA are in general limited to linear data (James and McCulloch, 1990; Melssen et al., 1993), methods in artificial neural networks are more feasible in patterning the complex ecological and behavioral data, and this advantage in information extraction was confirmed in this study. Through classification in different levels according to U-matrix, the differences in association among the variables were efficiently visualized for the success of nest-site selection for birds.

4. Conclusions

Association among environmental factors for characterizing nest-sites for breeding success was analyzed by using PCA and SOM. PCA showed that the variables regarding the "wall" effect such as rock cover and nest-wall were associated in determining breeding success of Black-tailed Gulls. SOM was efficient in extracting information from complex habitat data. Grouping by SOM was in general in accordance with PCA, while the patterned nodes on SOM were further classified according to the gradient of "Good" to "Poor" sites. SOM was additionally useful for finely classifying the groups. Grouping based on the "wall" effect on PCA was further subdivided by SOM depending upon the variables, slope and the nearest distance between neighbors. In addition to the conventional multivariate statistical methods, the use of techniques in ecological informatics such as SOM would be an efficient tool in analyzing data for breeding behavior of birds and would be further available for establishing conservation policies for management of bird populations.

Acknowledgments

We thank all colleagues of the Laboratory of Animal Ecology at KyungHee University, Seoul, Korea, and the Laboratory of Ecology and Behavior System at Pusan National University, Pusan, Korea. We also would like to express our thanks to Dr. Wayne E. Thogmartin and Ms. Karen Hyun for translation of the manuscript. Special thanks also given to Professor Robert E. Ricklefs and Dr. Thierry Boulinier for providing helpful comments on earlier drafts of this manuscript.

References

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

Bosch, M., Sol, D., 1998. Habitat selection and breeding success in yellow legged gulls (*Larus cachinnans*). Ibis 140, 415–421.

Buckley, F.G., Buckley, P.A., 1980. Habitat selection and marine birds. In: Burger, J., Olla, B.L., Winn, H.E. (Eds.), Behaviour of Marine Animals. Plenum Press, New York, pp. 69–112.

Bukacinska, M., Bukacinski, D., 1993. The effect of habitat structure and density of nests on territory size and territorial behaviour in the Black-headed Gull (*Larus ridibundus*). Ethology 94, 306–316.

Burger, J., 1985. Habitat selection in marsh-nesting birds. In: Cody, M. (Ed.), Habitat Selection in Birds. Academic Press, New York, pp. 253–281.

Burger, J., Gochfeld, M., 1981. Colony and habitat selection of six Kelp Gull (*Larus dominicanus*) colonies in South Africa. Ibis 123, 298–310.

Burger, J., Gochfeld, M., 1986. Nest-site selection in Sooty Terns (*Sterna fuscata*) in Puerto Rico and Hawaii. Colonial Waterbirds 9, 31–45.

Burger, J., Gochfeld, M., 1987. Nest-site selection by Mew Gulls (*Larus canu*): a comparison of marsh and dry-land colonies. Wilson Bull. 99, 673–687.

Burger, J., Gochfeld, M., 1988. Nest-site selection by Roseate Terns in two tropical colonies on Culebra. Puerto Rico. Condor 90, 843–851.

Burger, J., Gochfeld, M., 1991. The Common Tern – Its Breeding Biology and Social Behavior. toExcel, New York, pp. 72–73.

Cezilly, F., Quenette, P.Y., 1988. Role des ecrans naturels attenant au nid chez le Goeland Leucophee (*Larus cachinaans michahellis*). Alauda 56, 41–50.

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.

Chon, T.-S., Park, Y.S., Park, J.H., 2000. Determining temporal pattern of community dynamics by using unsupervised learning algorithms. Ecol. Model. 132, 151–166.

Clark, R.G., Shutler, D., 1999. Avian habitat selection: pattern from process in nest-site use by ducks? Ecology 80, 272–287.

Cody, M.L., 1985. An introduction to habitat selection in birds. In: Cody, M. (Ed.), Habitat Selection in Birds. Academic Press, New York, pp. 3–56.

Collias, N.E., Collias, E.C., 1984. Nest Building and Bird Behavior. Princeton University Press, Princeton, 336 pp.

Cultural Properties Administration, 2003. The research on the actual condition and management plan of avian breeding habitat, a natural monument. Seoul.

Freeman, S., Jackson, W.M., 1990. Univariate metrics are not adequate to measure avian body size. Auk 107, 69–74.

Good, T.P., 2002. Breeding success in the western Gull × Glaucous winged Gull complex: the influence of habitat and nest-site characteristics. Condor 104, 353–365.

Hunt Jr., G.L., Hunt, M.W., 1976. Gull chick survival: the significance of growth rates, timing of breeding and territory size. Ecology 57, 62–75.

James, F.C., McCulloch, C.E., 1990. Multivariate analysis in ecology and systematics: panacea or Pandora's box? Annu. Rev. Ecol. Syst. 21, 129–166.

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

Korea Meteorological Administration, 2004. Annual climatological report. Online: http://www.kma.go.kr/weather/climate/summ/sum_year_frame.html.

Kwon, Y.-S., 2004. Some aspects of the breeding biology of the Black-tailed Gull (*Larus crassirostris*). Ph.D. Thesis. Kyung Hee University, Seoul.

Lack, D., 1968. Ecological Adaptations for Breeding in Birds. Methuen, London, pp. 124–133.

Lee, W.-S., 2004. Habitat selection in Black-tailed Gulls *Laurs crassirostris* at Hond-Do Island, Korea. M.Sc.Thesis. Kyung Hee University, Seoul.

Lek, S., Guégan, J.F., 1999. Artificial neural networks as a tool in Ecological Modelling. An introduction. Ecol. Model. 120, 65–73.

Lek, S., Delacoste, M., Baran, P., Dimopoulos, I., Lauga, J., Aulagnier, S., 1996. Application of neural networks to modelling nonlinear 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., 1987. An introduction to computing with neural nets. IEEE ASSP Mag., 4–22.

Lusk, J.J., Wells, K.S., Guthery, F.S., Fuhlendorf, S.D., 2003. Lark sparrow (*Chondestes grammacus*) nest-site selection and success in a mixed-grass prairie. Auk 120, 120–129.

Martin, T.E., Roper, J.J., 1988. Nest predation and nest site selection in a western population of the Hermit Thrush. Condor 90, 51–57.

Martin, T.E., 1993. Nest predation and nest sites: new perspectives on old patterns. BioScience 43, 523–532.

Mastrorillo, S., Lek, S., Dauba, F., Belaud, A., 1997. The use of artificial neural networks to predict the presence of small-bodied fish in a river. Freshwater Biol. 38, 237–246.

Melssen, W.J., Smits, J.R.M., Rolf, G.H., Kateman, G., 1993. Two-dimensional mapping of IR spectra using a parallel implemented self-organizing feature map. Chemom. Intell. Lab. Syst. 18, 195–204.

Nisbet, I.C.T., Spendelow, J.A., Hatfield, J.S., Zingo, J.M., Gough, G.M., 1998. Variations in growth of Roseate Tern chicks. II. Early growth as an index of parental quality. Condor 100, 305–315.

Özesmi, S.L., Özesmi, U., 1999. An artificial neural network approach to spatial habitat modeling with interspecific interaction. Ecol. Model. 96, 11–28.

Paek, W.K., Yoo, J.-C., 1996. Time budgets of the Black-tailed Gull (*Larus crassirostris*) in the daytime of the breeding season. Kor. J. Orni. 3, 1–9.

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

Park, Y.S., Verdonschot, P.F.M., Chon, T.-S., Lek, S., 2003b. Patterning and predicting aquatic macroinvertebrate diversities using artificial neural network. Water Res. 37, 1749–1758.

Park, Y.S., Chang, J., Lek, S., Cao, W., Brosse, S., 2003c. Conservation strategies for endemic fish species threatened by the Three Gorges Dam. Conserv. Biol. 17, 1748–1758.

Partridge, L., 1978. Habitat selection. In: Krebs, J.R., Davies, N.B. (Eds.), Behavioural Ecology: An Evolutionary Approach. Blackwell Scientific Publications, Oxford, pp. 351–376.

Paruelo, J.M., Tomasel, F., 1997. Prediction of functional characteristics of ecosystems: a comparison of artificial neural networks and regression models. Ecol. Model. 98, 173–186.

Pierotti, R., 1982. Habitat selection and its effect on reproductive output in the Herring Gull in Newfoundland. Ecology 63, 854–868.

Ramos, J.A., 1998. Nest-site selection by Roseate Terns breeding on Aride Island, Seychelles. Colonial Waterbirds 21, 438–443.

Recknagel, F., French, M., Harkonenen, P., Yabunaka, K.L., 1997. Artificial neural network approach for modeling and prediction of algal blooms. Ecol. Model. 96, 11–28.

Ricklefs, R.E., 1969. An analysis of nesting mortality in birds. Smithson. Contrib. Zool. 9, 1–48.

Rising, J.D., Somers, K.M., 1989. The measurement of overall body size in birds. Auk 106, 666–674.

Saliva, J.E., Burger, J., 1989. Effect of experimental manipulation of vegetation density on nest-site selection in Sooty Terns. Condor 91, 689–698.

Siegel-Causey, D., Hunt Jr., G.L., 1981. Colonial defense behavior in Double-Crested and Pelagic Cormorants. Auk 98, 522–531.

Smith, M., 1996. Neural Networks for Statistical Modeling. International Thomson Computer Press, London.

SPSS Inc., 2002. SPSS for Windows, Rel. 11.5.0. Chicago.

Tan, S.S., Smeins, F.E., 1996. Predicting grassland community changes with an artificial neural network model. Ecol. Model. 84, 91–97.

The Math Works, 1999. Matlab 5.3. Massachusetts.

Tuma, A., Haasis, H.D., Rentz, O., 1996. A comparison of fuzzy expert systems, neural networks and neuro-fuzzy approaches controlling energy and material flows. Ecol. Model. 85, 93–98.

Ultsch, A., Siemon, H.P., 1990. Kohonen's self-organizing feature maps for exploratory data analysis. In: Proceedings of the INNC'90, International Neural Network Conference, Dordrecht, Netherlands. Kluwer, pp. 305–308.

Ultsch, A., 1993. Self-organizing neural networks for visualization and classification. In: Opitz, O., Lausen, B., Klar, R. (Eds.), Information and Classification. Springer-Verlag, Berlin, pp. 307–313.

Yang, W.K., Qiao, J.F., Combreau, O., Gao, X.Y., Zhong, W.Q., 2003. Breeding habitat selection by the Houbara Bustard (*Chlamydotis* [*undulate*] *macqueenii*) in Mori, Xinjiang, China. Zool. Stud. 42, 470–475.

Zurada, J.M., 1992. Introduction to Artificial Neural Systems. West Publishing Company, New York.