Analysis and prediction of the fluctuation of sardine abundance using a neural network

Ichiro AOKI and Teruhisa KOMATSU

Ocean Research Institute, University of Tokyo, Minamidai, Nakano, Tokyo 164, Japan.

Received in revised form 31/07/96, accepted 01/08/96.

ABSTRACT

This paper examines the use of a neural network to analyse and predict the winter catch, in the Joban-Boso Seas off the Pacific coast of central Japan, of young Japanese sardine (Sardinops melanostictus) representing the index of recruits in the sardine stock. The supervised learning paradigm, a three-layer network and a back-propagation algorithm were employed in constructing the neural net. A number of biological, hydrographical and climatic factors constituted an input vector, the output being the catch of young sardine. The association of sardine abundance with environmental factors was quantified in the form of the trained neural network, which indicated important associations with the Southern Oscillation Index, with patterns of the Kuroshio and the Oyashio currents, and with SST and zooplankton densities in the northeastern sea area of Japan. Changes in the sardine abundance during the past two decades could be predicted from environmental conditions. Neural networks can serve as a useful tool for the analysis and prediction of such marine ecosystem dynamics including fish population dynamics which are typically non linear multivariate problems.

RÉSUMÉ

Utilisation d’un réseau neuronal pour analyser et prévoir les variations d’abondance de la sardine.

Cet article examine l’utilisation d’un réseau neuronal pour analyser et prévoir la prise hivernale de jeunes sardines (Sardinops melanostictus) qui est l’indice de recrutement du stock de sardines dans la mer de Joban-Boso, au milieu de la côte Pacifique du Japon. La méthode utilise un réseau à trois couches et un algorithme de rétro-propagation. Un ensemble de facteurs biologiques, hydrologiques et climatiques constitue le vecteur d’entrée, tandis que la prise de jeunes sardines est le signal de sortie. L’association de l’abondance des sardines et des facteurs environnementaux est quantifiée dans le réseau neuronal qui montre l’importance de l’Oscillation Australe, des courants Kuroshio et Oyashio, de la température superficielle de l’eau et des densités du zooplancton dans le nord-est de la mer du Japon. Les variations de l’abondance des sardines pendant les deux dernières décennies sont prédictibles à partir des conditions de l’environnement. Le réseau neuronal peut être un outil performant pour résoudre les problèmes non-linéaires multivariables posés par l’analyse et la prévision de la dynamique des écosystèmes marins, en particulier la dynamique des populations de poissons.

INTRODUCTION

The Japanese sardine population *Sardinops melanostictus* has undergone long-term and substantial changes in abundance. The population was large in the 1930s, then fell to low levels during three decades before beginning to increase in the early 1970s. Annual catches in Japan reached the highest level of four million tons from 1984 to 1989 (Fig. 1). Subsequently, the sardine stock entered a period of decline, and landings were about 1,700 thousand tons in 1993.

![Figure 1](image)

**Figure 1**

*Total annual catch of Japanese sardine in Japan and catch of young sardine in the Joban-Boso Seas off the Pacific coast of central Japan from December to February of the following year.*

It has been indicated by several authors (Kawasaki, 1991; Lluch-Belda et al., 1989; Crawford et al., 1991) that the three Pacific sardine populations (*Sardinops* spp.) around Japan and off the western coasts of North and South America are fluctuating in phase on long-term scales; and that the European pilchard (*Sardina pilchardus*) was also abundant in the 1970s and 1980s. Moreover, there are similarities between these fluctuations in abundance and some large-scale environmental changes involving, for example, solar radiation, air temperature, and sea-surface temperature. Mechanisms relating the abundance of the sardine population to environmental variables may exist in the processes of spawning and survival during the first year of life.

Young Japanese sardine are caught by purse seine fisheries in winter in the Joban-Boso Seas off the Pacific coast of central Japan (Fig. 1). The volume of the catch reflects trends in recruits of the year or year-class strength, because it consists mainly of fish hatching the previous winter and spring, and because the sardine fishery is prosecuted without special regulation.

In this paper, using a neural network, we show that the catch of young sardine can be predicted by climatic, hydrographic, and biological factors on regional scales. The neural network has the advantages of being able to handle nonlinear problems and to process correlated variables.

DATA AND ANALYTICAL METHOD

Recruitment of the Japanese sardine

The location of the fishing grounds of the wintering young sardine and the oceanic environments studied here are given in Figure 2.

![Figure 2](image)

**Figure 2**

*Study area and schematic oceanic environments: 1. First Intrusion of the Oyashio; 2. Second Intrusion of the Oyashio; 3. Kuroshio Extension. A solid fine rectangle indicates the area for which sea-surface temperature was used in this study (TB, Table 2). A broken fine rectangle indicates the area for which zooplankton densities were obtained from Odate (1994): this area was divided into the Oyashio, transition and Kuroshio regions by Odate (1994).*

Off the Pacific coast of Japan, adult sardine spawn in winter and spring in the sea areas close to the Kuroshio current. Larvae and juveniles are transported by the current and dispersed widely in northeastern Japanese waters in summer. This sea area is complicated under the influences of the Kuroshio and Oyashio current dynamics. In autumn, young-of-the-year sardine migrate southward; in winter, they aggregate off the coast of central Japan. The wintering young fish become the main target of purse seiners for the first time.

Therefore, we considered the Kuroshio and Oyashio currents, together with sea-surface temperatures and zooplankton densities in the northeastern sea area of Japan, as factors affecting young fish catch or recruitment (Table 1). In addition, some climatic indices were taken into account.

Data sources

Data for the period 1972 to 1992 were used in this analysis, with the exception of zooplankton data, available up to 1990.

For young sardine catch data, we used total landings during the period December-February from Onahama, Fukushima Prefecture to Choshi, Chiba Prefecture, compiled by Ibaragi Prefectural Fisheries Experimental Station.
Table 1

List of input variables used in the neural network. See text for full explanation.

<table>
<thead>
<tr>
<th>Variables (Abbreviation)</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hydrographic</td>
<td>Southern limit of the Oyashio (OY)</td>
</tr>
<tr>
<td>Path type of Kuroshio (KR)</td>
<td>KR - A, B, C, D and N</td>
</tr>
<tr>
<td>Northern limit of the Kuroshio Extension (KN)</td>
<td></td>
</tr>
<tr>
<td>Sea surface temperature in the north eastern sea area of Japan (TB)</td>
<td></td>
</tr>
<tr>
<td>Sea surface temperature in the fishing ground (TF)</td>
<td></td>
</tr>
<tr>
<td>Biological</td>
<td>Zooplankton in the Oyashio region in the north eastern sea area of Japan (PO)</td>
</tr>
<tr>
<td></td>
<td>Zooplankton in the transition region in the north-eastern sea area of Japan (PT)</td>
</tr>
<tr>
<td></td>
<td>Zooplankton in the Kuroshio region in the north-eastern sea area of Japan (PK)</td>
</tr>
<tr>
<td>Climatic</td>
<td>Southern Oscillation Index (SO)</td>
</tr>
<tr>
<td></td>
<td>Far East Zonal Index (ZI)</td>
</tr>
<tr>
<td></td>
<td>East Sea Index (ES)</td>
</tr>
</tbody>
</table>

All of the following variables with the exception of zooplankton were averaged over the twelve months of the year to represent average overall conditions during the first year of life, because this study aimed to examine the general effects of forcing variables.

The hydrographic variables used in the analysis were as follows:

1. Southern limit of the Oyashio (OY); the southernmost latitude of the First Intrusion of the Oyashio was determined from monthly 100 m depth temperature charts by the Japan Meteorological Agency (JMA), according to the indicative temperature of the Oyashio Front defined by Kawai (1972).
2. Path types of the Kuroshio (KR); these were classified into A, B, C, D and N types from the semi-monthly Bulletin of Ocean Conditions issued by the Department of Maritime Safety Agency of Japan, according to Yoshida (1961), Nitani (1969) and Kawai (1972). Frequencies of each path type were counted over the twelve months of the year.
3. Northern limit of the Kuroshio Extension (KN); the northernmost latitude of the Kuroshio Extension axis was determined from the same chart as in determination of OY according to the indicative temperature of the Kuroshio Extension axis defined by Murakami (1993).
4. Sea-surface temperature in the northeastern sea area of Japan (TB); we used ten-day mean sea-surface temperature anomaly data by JMA for the area indicated in Figure 2.
5. Sea-surface temperature in the fishing grounds (TF); we used monthly hydrographic data on sea-surface temperature obtained at a representative station of the Kashima-nada Sea by the Ibaragi Prefectural Fisheries Experimental Station.

After Odate (1994), zooplankton abundance was represented by mean zooplankton densities (wet weight/m²) in June for the Oyashio region (PO), the transition region (PT) and the Kuroshio region (PK) in the northeastern sea area of Japan indicated in Figure 2. Zooplankton samples were collected most extensively in June every year to cover this area. June is the high abundance season for planktonic organisms and the dispersed period for sardine juveniles in the area.

Climatic variables were represented by the following three atmospheric circulation indices, postulated to affect oceanic circulation patterns and sea temperature:

1. Southern Oscillation Index (SO); monthly data compiled by JMA were used. This indexes ENSO, the El Niño-southern oscillation, events. The value of the Southern Oscillation Index becomes negative when the El Niño occurs.
2. Far East Zonal Index (ZI); difference in monthly mean 500 hPa height anomalies between 40° N and 60° N in the range of 90° E to 170° E, compiled by JMA, which indicates the amplitude of the meander of the westerlies.
3. East Sea Index (ES); monthly mean anomaly of 500 hPa height at 40° N in the range of 140° E to 170° E. This represents the strength of the high pressure east of Japan.

**Neural network structure**

Feed-forward neural networks are one of a class of flexible non linear regression methods (Ripley, 1994). Unlike commonly used conventional regression methods, neural networks do not require a particular functional relationship and distributional assumptions for the data. Hence, the neural network can be easily applied by non-statisticians and novice users (Cherkassky et al., 1994).

We used a commercial neural network simulator, RHINE (CRC Inc.) for a personal computer. A feed-forward layered neural network was employed, having three layers: an input layer, a hidden layer, and an output layer (Fig. 3). Since it has been shown that neural networks with at least one hidden layer can approximate arbitrary conditions function (Funahashi, 1989), we chose one hidden layer for simplicity.

![Figure 3](image)

*Principle of a feed-forward multi-layered neural network.*

The units are connected to the next layer units with adjustable weights. When one unit receives input signals from others, these are weighted by different values and
summed. Then, the unit outputs the signal according to a non-linear sigmoid transfer function. A back-propagation algorithm was used to adjust the weight values so as to minimize the difference between the output signal from the output unit and the teaching (correct) signal. The neural network becomes precise after a number of learning cycles on a set of learning data. This process is called supervised learning or training. Thus the feed-forward multi-layer network is characterized by non-linear units, at least one hidden layer and a back-propagation algorithm efficient for determining network parameters.

The number of units in the input layer was 15, representing the climatic, hydrographic and biological variables shown in Table 1. One unit was assigned to the output layer to represent the catch of young sardine. Five units were used in the hidden layer. Fifteen variables which appeared to be related to sardine catch were defined as input variables, since we were not sure which of them would be important. Input values were transformed to compound variables in hidden layer units, and a selection of input variables made in the form of the interconnecting weight in the network. This process can be viewed as a joint use of non-linear principal component analysis and multiple regression analysis (Hirafuji et al., 1988). The number of the hidden units was chosen empirically as one-third of the number of input units.

The neural network learns a mapping function from a training set of data composed of pairs of input and output vectors for each year. In neurocomputing, input values for each variable were linearly normalized with max=1 and min=0, and initial values for the weights of the connection were set at random in the range of ±0.3. The number of learning cycles was 5,000. The error decreased as the number of learning cycles increased, and showed scarcely any changes after 5,000 cycles.

RESULTS

Some examples of time series data on input variables over the two decades are given, together with the catch of young sardine, in Figure 4. As described above, the catch for each year refers to the catch from December of the year to February of the next year and represents an index of recruits or year-class strength of the year. Values of SO, TB, and OY appear to be low in the 1980s when recruits were high. The occurrence rate of the Kuroshio path type B tends to be high in the 1980s. The neural network can generate the quantitative associations between environmental variables and sardine recruits.

In order to test the performance of the network, it was necessary to reserve a part of the data set aside from the training data. Data for testing predictions cannot be used for the training procedure. Since there were only 19 years with complete data from 1972 to 1990, data (periods) for training and testing were exchanged so that the testing predictions became more reliable.

Figure 5 shows the results of the training and testing predictions. In case 1, we used the data from 1972 to 1986 for the training patterns and the data from 1987 to 1990 (case 1-a) or to 1992 (case 1-b) for the test of the

![Figure 5](image-url)
trained network. In this study, zooplankton data were not available after 1990. Therefore, all 15 variables were used in case 1-a, and the training and testing were performed excluding zooplankton variables, and testing was done up to 1992 in case 1-b. In the two experiments, the neural network could be trained well, and the testing predictions reproduced the decline trend in the late 1980s. In case 1-b, predicted values were somewhat inferior to those in case 1-a, but they represented the temporary increase in 1992.

Conversely, in case 2, data from 1976 to 1990 were assigned to the training and data from 1972 to 1975 were used for the test. The testing period coincides with the early increasing phase of sardine stock. Also, in this case, the testing predictions were good (Fig. 6).

Analysis of the weights of connections between units in the trained neural network can lead to the determination of which input variables impact on the sardine recruitment. As the training progresses, the absolute value of the weight increases between related units, so that the connection is strengthened. The absolute values of the interconnecting weights between hidden units and the output unit were large for three hidden units in cases 1 and 2, and for four in case 3 (Fig. 8). The weights between each of the three or four hidden units and 15 input units are shown in Figure 9. The input units having greater absolute weight values affect the output unit greatly. The plus and minus signs indicate positive and inverse relations, respectively. As the signs of the weight between hidden units 2 and 5 and the output unit were minus (Fig. 8), the signs for hidden units 2 and 5 were reversed for unity in Figure 9. The magnitudes of the weights of the input units were consistent among the three cases.

The Kuroshio path types B and D, and zooplankton densities in the transition region (PT) and the Kuroshio region (PK) had positive greater weight values. The Southern Oscillation Index (SO), the southern limit latitude of the Oyashio (OY), and SSTs in the northeastern sea area (TB) and in the fishing grounds (TF) had greater negative weight values. These results suggest that good environmental conditions for sardine recruitment are created at or during the following times: when the ENSO event occurs frequently; when the Kuroshio meanders in a small scale; when the Oyashio stretches southward; and when SST is lower and zooplankton densities are high off the northeastern coast of Japan. The Far East Zonal index (Z1) exhibited a negative greater weight value in case 2. The greater negative value of Z1 indicates the large amplitude of the meandering of the westerlies.

**DISCUSSION**

This study indicated that the change in the Japanese sardine population showing growth, prosperity and decline during the 1970s and 1980s can be predicted from the climatic, hydrographic and biological environments. The associations of sardine abundance with environmental factors were quantified in the form of the trained neural network.
network. The magnitudes of the weights indicated that the important associations of the recruits are with the Southern Oscillation Index, ocean currents, SST and zooplankton.

The inverse relation of sardine abundance with the value of the Southern Oscillation Index is consistent with the hypothesis (Bakun, 1994) that El Niño events which occurred frequently during the mid-1970s to mid-1980s were responsible for increases in the abundance of several fish species including sardines during the same period. It has been suggested that the Pacific/North American (PNA) pattern, which has a connection with the ENSO phenomenon, leads to the "spin-up" of the subarctic gyre and the southward intrusion of the Oyashio current through the enhancement of westerlies in the North Pacific during the winter season (Hanawa, 1991; Nitta and Yamada, 1989). The trend in the southward intrusion of the Oyashio in the 1980s (Fig. 4) is probably connected with lower SST in the northeastern sea area and the fishing grounds concerned in this study. It has been pointed out (Kodama, 1992; Sasaki and Tsuchiya, 1990) that there was an association of catches of 1-year sardine with the southern intrusion of the Oyashio. In addition, sardine catches in the Joban sea area were inversely correlated with SST in that area (Yokota, 1990). The weight values of OY, TB and TF in the neural networks (Fig. 9) are consistent with these observations.

The Kuroshio path types B and D, the weight values of which were large in the networks, are characterized by a small meander and relatively short duration (Nitani, 1969; Kobayashi et al., 1984). Kondo (1980, 1988) proposed that a shift from one pattern to another of the Kuroshio current created a dominant year-class. Nakata et al. (1994) found that primary productivity in the coastal waters was higher when the Kuroshio path fluctuated. The variability of the path types B and D is likely to result in good conditions for sardine survival.

The weight values of zooplankton densities were much greater in the Kuroshio and the transition regions than that in the Oyashio region. Sardine hatch out in the seas off the southern coast of Japan and grow as they drift or move to the Kuroshio, the Kuroshio Extension, and the transition area, successively. Some juvenile and young fish may reach as far north as the Oyashio waters. Therefore, our results agree with the expectation that key localities for the survival of sardine to the young fish stage in the northeastern sea area may be the Kuroshio and the transition regions rather than the Oyashio region. For our data set on zooplankton, mean zooplankton densities in the Kuroshio and the transition regions were one-quarter and one-half of that in the Oyashio region, respectively. The ratios are similar to those for the period May-July during the past three decades in the northeastern sea area of Japan (Odate, 1994). Zooplankton densities in the Kuroshio and the transition regions varied by a factor of 2-4 over a few decades. Although those zooplankton were less abundant as a whole, the variation seems to influence sardine recruits. The abundance and distribution of zooplankton were probably associated with the Kuroshio and the Oyashio currents. The mechanisms which relate the dynamics of the Kuroshio-Oyashio system, zooplankton abundance and sardine recruits to each other are not yet clear. Further biological oceanographic research must be done to address these mechanisms.

Several authors propose that the decline in sardine populations is attributable to density-dependent rather than environmental factors, while their increase can be attributed to environmental conditions (Kawasaki, 1993; Hiramoto, 1990; Ito, 1983). Kondo (1988) points out the consistent significance of ocean conditions in changes of sardine abundance. Our study does not entirely support the first point of view. It is well known that sardine change their distribution range and spawning area in accordance with changes in abundance. The changes may have an effect on reducing intraspecific competition. For example, the shift of spawning grounds of the Japanese sardine inside the Kuroshio current and to its upstream regions may lead to an intensification of the geographical dispersal of eggs and larvae (Aoki and Murayama, 1993). Such adaptive plasticities in the lifestyles of the sardine could ensure that full advantage is taken of favourable environmental conditions to maintain an extremely large population size. Nevertheless, if favourable environmental conditions for sardine change, the modification will be less effective, and abundance will decline as a consequence.
The object of neural network learning is to generalize from a training set to the systems as a whole. There is a danger that too many input variables lead to overfitting the model and lowering the ability of the generalization. This problem was explored using fewer variables. We selected eight and further four input variables, the magnitudes of whose weights were larger in the trained networks as shown in Figure 9. In the case of four input variables, the number of the hidden units was three. Mean absolute errors of the test predictions are given in Table 2. By using eight determinative variables with greater weights, the predictions were not improved by parsimony. The smallest set of four variables resulted in less successful performance.

Thus, the back-propagation neural network is capable of easy application and good prediction even when a number of factors are assumed to be associated with a problem. However, neural networks lack the probability basis that statistical methods have for measuring confidence intervals and judging the best model.

There have been a few applications of the neural network to oceanography and fisheries research: identification of plankton using image data (Nakano et al., 1991); identification of fish school species in acoustic data (Haralabous and Georgarakos, 1993); and prediction of ocean and fishery conditions (Komatsu et al., 1994a, b). Marine ecosystems are organized by a number of elements and generally involve nonlinear phenomena. Moreover, most of the important elements are closely related. Thus, the artificial neural network can be a useful tool for the analysis and prediction of marine ecosystem as well as fish population dynamics.

REFERENCES


