Training and Generalization of Experimental Values Of Ice Scour Event By A Neural-Network

S.Kioka and A.Kubouchi
Civil Engineering Research Institute of Hokkaido, Sapporo, Japan
H.Saeki
Hokkaido University, Sapporo, Japan

ABSTRACT

This paper mainly reports on the training/generalization of experimental data on a scour-curve, which is a two dimensional locus remained when ice keel ploughs the seabed, by a Neural-Network (NN). As a result of training, the multiple correlation between estimated value by NN after training and experimental values was more than 0.99. We proposed that this method could replace a nonlinear multiple-regression analysis, which is very difficult to be applied when unknown parameters are independent. The NN model driven by an extensive data set will be useful tool for developing a practical method for estimation of scour depth by combining it with a mechanical ice scour model that we had already developed.

KEY WORDS: Ice Scour; Scour depth; Neural-Network;

INTRODUCTION

Ice-Scour Event is a phenomenon that occurs when ice comes into contact with seabed. Ice-Scour has been reported to cause damage to communication cables and water intake pipelines (Duval, 1975; Grass, 1984). While we have studied the scour-processes by many experiments on ice-scour event including the medium-scale model test, we have also developed the mechanical model [Kioka et al., 2000], which consists of the equation of motion concerning an ice and the simple model of interaction between the seabed and the ice keel. Also, we indirectly validated the mechanical model using the experimental results obtained in small- and medium-scale model tests [Kioka et al., 2001a; 2001b; 2002]. Our final goal is to develop the method for estimation of scour depth, or optimal depth of a buried structure such as a pipeline in order not to contact an ice keel. We believe that we can establish the method by combining the mechanical model that we had already developed and the scour curve. However, we need to establish the scour curve experimentally, which is difficult to be estimated by a theoretical approach. The method is based on both theoretical and experimental approach, that is, we take advantage of information of experimental results with respect to the lack of a theoretical model. In a practical use, while we have to establish the characteristics of experimental results concerning the scour curve under various conditions including its scale-effect, we also have to generalize the experimental data. In this study, we attempt to generalize the experimental values by “Layered Neural-Network”, and validate its availability. Finally, we outline the approach to estimate the scour depth by an ice keel.

REPRESENTATIVE PARAMETERS OF SCOUR CURVE

Fig.1(a)(b) shows the scouring process (scour curve, $\zeta(X)$) in case of a small-scale model test as one example. Although we have examined the characteristics of scour curve (digging line, that is, locus of motion of a keel), we have found that scour curve has the linearity without exception. This means that the keel does not plough the seabed horizontally, but moves upward as scouring. A medium-scale test (Kioka et al., 2001a) also showed the same tendency, and Shearer (1986) proposed the similar scouring processes on a full-scale event. The curve are thought to form by the following mechanisms; (1) ice/keel moves temporarily along a shear failure surface of seabed, (2) slip between the surface of the keel and sand, (3) effects of dilatancy of soil, and (4) settling of keel due to the decrease of its buoyancy. It will also depend on the sea bottom gradient, the characteristics/conditions of seabed, the velocity of ice drift, the attack angle of the keel and so on. Thus, scour curve forms by various complicated factors, and it is very complex.
difficult to define the scour curve theoretically. Here, we attempt to estimate the curve from experimental results. Although we must consider the similarity or scale-effect between small- and full-scale events, we showed that a scour curve followed the geometric similarity law from the results of the medium scale model test (field test) (Kioka et al., 2001a). As a preparation for generalization of the experimental results of the curve, we simplify the curve and propose representative parameters to describe the scour curve. As shown in Fig.1 (b), let $L_m$ be the length of the horizontal segment of a line $(C1)$, and let $\phi_m$ be an angle between the adjacent straight line $(C2)$ and the horizontal axis. So, $L_m$, $\phi_m$, and variations around $C2$-line, $C(0)$, can be proposed as the parameters.

TRAINING AND ESTIMATION OF PARAMETERS OF SCOUR CURVE USING NEURAL NETWORK

Neural network is a new and advanced artificial intelligence technology that mimics the brain’s training and decision making process. A neural network consists of a number of nodes with “neuron” connections between the nodes. When a training process is being conducted, the neural network learns from the given input/output data such as experimental data, what we call, “teacher data” or “training data” and gradually adjusts its neurons to reflect the desired outputs. In other word, “training” is equivalent to pattern recognition of complicated information. Fig.2 shows a schematic diagram of a neural network.

![Schematic diagram of neural network](image)

In this study, training and estimation of the above-mentioned representative parameters of scour curves were discussed using a three-layered neural network as an analysis method equivalent to nonlinear multi-regression analysis, and validity of this method was verified. In this case, however, four explanatory variable (Input layer parameters) - attack angle ($\theta$), ice drift velocity ($V_0$), sand property (mean grain size as a representative property, $d$) and sea-bottom gradient ($tani$) - were utilized, and the scour parameters was used as objective variable (Output layer parameters) as shown in Table 1. In this case, the average value under the same experimental conditions was used to remove noises of the teacher data (experimental data).

<table>
<thead>
<tr>
<th>Input layer parameters (Experimental conditions)</th>
<th>Output layer parameters (parameters of scour curve)</th>
</tr>
</thead>
<tbody>
<tr>
<td>attack angle ($\theta$)</td>
<td>$L_m$</td>
</tr>
<tr>
<td>ice drift velocity ($V_0$)</td>
<td>$C(0)$</td>
</tr>
<tr>
<td>mean grain size of sand ($d$)</td>
<td>$\tan\phi_m$</td>
</tr>
<tr>
<td>sea-bottom gradient ($\tan\theta$)</td>
<td>$C(0)$; variation around $C2$-line in fig.1(b)</td>
</tr>
</tbody>
</table>

Notes: $L_m$: length of the horizontal segment of a line ($C1$) in fig.1(b) $\phi_m$: angle between $C2$-line and the horizontal axis in fig.1(b) $C(0)$; variation around $C2$-line in fig.1(b)

Design of the network and training method

Data set

As mentioned before, the input layer had four neurons (explanatory variable) and the output layer had one neuron (objective variable), which was the average value of cases under the same experimental conditions (sample number: 5 – 10) as the teacher data. The scour parameters as output value was set as the above $\tan\phi_m$, $L_m$ and $C(0)$. In this case, there were 66 groups of teacher data (experimental data) corresponding to the number of experimental conditions.

Preprocessing of data

Adjustment of the data in advance was necessary because output from each neuron was limited to 0 – 1. Each data was preprocessed as follows:

1) Ice drift velocity: Because there was a significant deviation/bias, the minus signs of the logarithms were first taken and then normalized to be 0.1 to 0.9 taking the respective maximum and minimum values within the range of this experiment into account.

2) Mean grain size of sand: Normalized to be 0.1 to 0.9 taking the respective maximum and minimum values within the range of this experiment into account, without logarithmic transform.

3) Sea-bottom gradient: The sea-bottom gradient was expressed as $\tan\theta$; its upper limit was 1/25 in this experiment and its lower limit was set for values up to 1/1,000, although it was outside the scope of the experiment. It was also eventually normalized to be 0.1 to 1 by taking the minus signs of the logarithms of $\tan\theta$.

4) Attack angle: Normalized to be 0 – 1 so as to correspond to values up to 20 deg. by expressing the angle as $\frac{\tan\theta}{\tan\theta}$.

5) $\tan\phi_m$: Normalized to be 0 – 1 to make the minimum value 0 and the maximum value the sea-bottom gradient.

6) $L_m$: Normalized to be 0 – 0.9 to make the minimum value 0 and the maximum value this experiment.

7) $C(0)$: Eventually normalized to be 0 – 1 by giving a single-digit margin to the maximum and minimum values, respectively, and taking the minus signs of the logarithms of $C(0)$.

Training algorithm

The back propagation method was used for training, and the inertia method, which is one of the steepest descent methods, was used for renewal of the connection weights equivalent to synapses and threshold in all neurons. The upper limit of the training number of repetitions was set at 100,000. Since the most standard method was used for this algorithm, explanation of the details is omitted.
**Determination of number of neurons in intermediate layer**

It is usually possible to construct a NN model faithful to the teacher data according to the number of neurons in the intermediate layer. It is, however, necessary to set a statistically appropriate number when there is a possibility that the teacher data contain errors (noises) as in the case of this experiment, or when reliability of the model may be lowered by an increase in unknown parameters. In this case, the maximum log-likelihood was calculated taking measurement noises (assumed to be errors of average values caused by the sample size in this case) of the teacher data into account, and the number of intermediate neurons was determined in accordance with the Akaike’s information criteria (AIC). Although it is technically necessary to take into account conditional variances and systematic errors that vary by experimental conditions related to the measurement noises, average values of the noise/errors were used for this study by assuming that they would not vary by experimental conditions. In this case, the calculation method shown below was used. First, the relationship between teacher data and true teacher data was considered as follows:

**System equation:**

\[ x = f(z_a; q) + e_z \]  

**Measurement equation:**

\[ z = x + e_0 \]

Where, \( x \): true teacher data (true measured value vector), \( z \): used teacher data including noise/error (measured value vector), \( f(\cdot) \): constructed NN model, \( z_a \): input vector, \( q \): estimated values of unknown parameters (connection weight matrix, threshold), \( e_z \): system noise, which was assumed to follow the normal distribution of mean, \( \theta \), and variance-covariance matrix, \( \sigma^2 I \), as shown eq. (2).

\[ e_z \sim N[0, \sigma^2 I] \]

(2)

\( e_0 \): measurement noise, which is assumed to be statistically independent, and to follow the normal distribution of mean, \( \theta \), and variance, \( V_0 \). Here, we assume that there are not systematic errors due to the experimental conditions. We also assume that \( e_0, e_z \) and the error between \( z \) and output value from NN model, \( e_s (\equiv z - f(z_a; \theta)) \), are non-correlated. The likelihood of true teacher data;

\[ \lambda(x) = \frac{1}{(2\pi)^{n/2}} \exp \left\{ -\frac{1}{2\sigma^2} \left[ z - e_0 - f(z_a; q) \right] \right\} \]

(3)

Here, we take logarithms of both members of eq.(3), and \( \sigma \) to maximize the log-likelihood:

\[ \hat{\sigma}^2 = \frac{1}{n}[e_z - e_0][e_z - e_0]^T \]

(4)

Therefore, we have the maximum log-likelihood including measurement noise as follows;

\[ L_{\text{max}} = \ln \lambda(x; \hat{\sigma}^2) = -\frac{n}{2} - \frac{n}{2} \ln [2\pi V_0^2 + V_z^2] \]

(5)

Where, \( V_0 / V_z \): variance of \( e_0 / e_z \), (variance-covariance matrix, \( V / J \)

**Identification of NN model**

Even if a NN model has been constructed through above procedures, it is still necessary to confirm that the model is essentially recognized. The amount of information, which is a factor of row j, column k in Fisher’s information matrix, can be expressed by the equation below.

\[ J_{jk} = \sum_{n=1}^{N} \frac{\partial^2 \log \lambda(x_n; q)}{\partial q_j \partial q_k} \]

(7)

If this information matrix has a sufficient number of stimuli (\( n \)) and is full rank, an inverse matrix can be calculated and used as an asymptotic covariance matrix of parameters. On the other hand, because standard errors of parameters cannot be evaluated when the model is not recognized, an information matrix that is not full rank can mean that the model is not recognized.

**Training results**

As an example, Fig.3 (a) shows changes in AIC by the number of intermediate neurons (studied for 1 – 11) for the case in which the average value of \( \tan \phi m \), which is the most important output quantity, was used as the teacher data in output layer. In this case, training parameters were fixed at a certain values that were confirmed to be appropriate through advance trial-and-error efforts. The upper limit of the training number of repetitions was set at 100,000 times. In general, when the number of intermediate neurons (\( N \)) becomes larger, multiple correlation coefficient also becomes larger and the error between the estimated value (output from NN system) and measured data tends to become smaller. However, because the decline in reliability (penalty)
of the model due to an increase in unknown parameters is taken into account in accordance with AIC, the minimum value of \( N \) exists and a certain number of neurons become the optimum. In this case, \( N = 5 \).

Optimum values also existed for other output quantities (\( C(0), L_m \)). Fig. 3 (b) shows the relationship between the estimated from NN model after training and measured values by also using the case of \( 	an \phi_m \).

Since the multiple correlation coefficient is 0.995 in this case, the degree of recognition is obviously high. Table 2 shows a summary of training results for the three output quantities.

![Fig.3 (b) Relationship between the estimated values from NN model after training and measured values (\( \tan \phi_m \))](image)

Table 2. Summary of training results (with respect to average values of \( \tan \phi_m, L_m \) and \( C(0) \))

<table>
<thead>
<tr>
<th></th>
<th>( \tan \phi_m )</th>
<th>( L_m )</th>
<th>( C(0) )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( N_e )</td>
<td>66</td>
<td>66</td>
<td>66</td>
</tr>
<tr>
<td>( N_n )</td>
<td>5</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>( N_p )</td>
<td>31</td>
<td>25</td>
<td>25</td>
</tr>
<tr>
<td>S.D.</td>
<td>0.00106</td>
<td>0.0991</td>
<td>5.87E-05</td>
</tr>
<tr>
<td>Multiple correlation</td>
<td>0.995</td>
<td>0.976</td>
<td>0.819</td>
</tr>
<tr>
<td>( n, \alpha, T )</td>
<td>0.14,0.06,0.6</td>
<td>0.25,0.07,0.9</td>
<td>0.16,0.09,0.6</td>
</tr>
<tr>
<td>Recognition of NN system</td>
<td>Ok</td>
<td>ok</td>
<td>ok</td>
</tr>
</tbody>
</table>

Notes:
- \( N_e \): Sample number (number of the experimental conditions)
- \( N_n \): Number of neurons in the intermediate layer
- \( N_p \): Number of unknown population parameters of NN-model
- S.D.: Standard deviation of differences between the measured and estimated value.
- \( n, \alpha, T \): Training parameters

Table 2 shows that all the output values were recognized and were well learned in general. Looking at the multiple correlation, \( \tan \phi_m > L_m > C(0) \), it can be seen that, especially for \( C(0) \), the degree of measurement errors that were not actually significant was increasing. Therefore, due to the reliability of the teacher data, the number of intermediate neurons remained at 4.

Next, Figs. 4 (a) – (c) show comparisons between measured and estimated values from the obtained network model, using \( \tan \phi_m \) as an example. These figures reveal that the measured and estimated values corresponded well with each other and that good training was achieved at a reasonable level. What is important to note here is that this estimation method seems to bear comparison with the least square method, rather than just having a faithful interpolation function of measured values. As previously mentioned, this was due to the assumption that the teacher data, which were measured values, contained errors, and because the model was constructed with a certain degree of robustness. This procedure is important in cases where measured values contain errors, and the method of controlling the number of intermediate neurons in accordance with the above-mentioned AIC was thought to be appropriate. Next, Fig. 4(c) indicates the sea-bottom gradient as the horizontal axis. While the range of it in this experiment was 0.04 to 0.01, the NN model was structured to be applicable up to 0.001 as mentioned earlier. However, since the estimated value at 0.001 is of course outside of the scope of the experiment, its validity cannot be studied. However, the result seems to be appropriate, because the sea-bottom gradient gradually approach 0 when gradient came closer to 0. This means that, if normalization using a maximum or minimum value for which expansion is desirable is successful, it is possible to apply the method to conditions outside of the scope of the experiment to some extent.
The above results revealed validity of the application of a neural network to training/generalization of experimental data. Although only the average values were discussed above and it is possible to construct a NN model for the standard deviation or coefficient of variation using a similar method, it is thought that subjective values must be given for the coefficient of variation in practical applications. In addition, only the mean grain size was used as a representative property of sand in this study. If other properties (e.g., void ratio, changes in the internal friction angle of sand) are known, they should also be used as input values.

**OUTLINE OF APPROACH TO ESTIMATE SCOUR DEPTH BY ICE KEEL.**

As described at the beginning, our final goal is to develop the method for estimation of the scour depth, or the optimal depth of a buried structure such as a pipeline in order not be in contact with an ice keel. Under our experimental conditions, we give the model keel forced horizontal displacement with constant velocity. In a real event, driving forces including wind, current and interactions of sea ice (level ice) surrounding an ice ridge, are finite. So, after the ice ridge makes contact with the seabed and ploughs to some extent, it will reduce its speed to stop finally. In our previous study, we indirectly validated the mechanical model by mean of the bulldozing force that could be measured easily, and we also had the basis that it could be applied to full-scale events (Kioka, 2001a; 2002). Therefore, if the scour curve is given, we can estimate the stopping distance of the keel by replacing the bulldozing force with the driving force on the motion equation. Finally, we can estimate the scour depth in a certain water depth by the geometric relationship. Furthermore, as described earlier, we can simplify the scour curve by means of the assumption of its linearity and extraction of its representative parameters. In order to study and generalize the previously obtained experimental values under the various conditions, we proposed that Neural-Network could be used as one method. A system can thereby be built to produce values of representative parameters of scour curve from input data of certain environmental conditions (input-output operation system). Here, the procedure for the estimation of scour depth is shown as follows (see Fig.5).

1) Determination of a project such as installation of a pipeline
2) Perform experiments concerning ice scour event under similar conditions to the field as objects.
3) Determine population parameters and p.d.f of environment parameters as a random variable
4) Generation of random number with respect to each environmental parameter as a random variable
5) Connect to Neural-Network system after training (Output of parameters of the scour curve from Neural-Network)
6) Calculation of scour depths using the mechanical model of scour event (solve the motion equation of an ice/keel)
7) Repeat 1)-6) up to a certain number and obtain distributions of scour depths

Figure 6 shows an example of the simulation of scour depths, which was actually conducted using the above procedure, although it was still at an experimental stage. This case involves the frequency distribution of scour depths, which generated 100,000 ridge keels when the water depth was 15 – 20 m and the sea-bottom gradient was 1/500 (almost the same results were obtained when the gradient was 1/100). The calculation was performed using the keel draft, ridge shape, scour curve parameters, initial velocity of ridge drift, strength parameters of the seabed, thickness of pack/level ice surrounding the ridge and other factors as random variables. Figure 7 shows the frequency distribution of scour depths observed at water depth of 15 – 20 m in the Alaskan Beaufort Sea (Weeks et al., 1983) as a reference.
Although it is difficult to make a simple comparison of these figures, the distribution forms and ranges of scour depths seemed to roughly correspond to each other. The frequency of scour depths of 0 – 30 cm was 0 in the observed value, which was probably due to the observation accuracy, the disappearance or deterioration of scour depths or other reasons. As a result of calculations made under various conditions, orders similar to those in the field observation data – the scour length of several tens of meters to several kilometers and the scour depth of several tens of centimeters to several meters – were found. Furthermore, due to the similarity to (reproducibility of) the comprehensive tendency/property of actual scour events, results that compared favorably with field observation results were obtained. Such results included 1) the distribution form of scour depths, 2) the fact that the scour depth increased with an increase in water depth but remained constant after a certain water depth (approx. 20 m in this calculation) (Lewis, 1978; Blasco, 1998), 3) the fact that the scour depth did not exceed 1 m when the water depth was 6 m or less (Machemehl, 1989) and 4) the distribution of the frequency of events (impacts) in relation to the water depth (Blasco, 1998).

CONCLUSION

This paper discussed the training/generalization of experimental data, which are important subsystems for the development of practical methods for estimation of scour depths. We confirmed that it would be possible to use a training algorithm by a neural network effectively. Also, we suggested a method for designing a neural network on the assumption that teacher data contained errors, and its validity was indicated. It should be noted that this method is an analysis method equivalent to nonlinear multi-regression analysis, which is considered to be extremely difficult when the unknown parameters are in a nonlinear relationship. A simulation method of scour depths, which combined this method with mechanical model of ice scour events that we had already developed, was also presented briefly. Although these study results are still at an experimental stage, we leave detailed discussion for another opportunity including further improvement.

REFERENCES