Prediction Method of Underwater Acoustic Transmission Loss Based on Deep Belief Net Neural Network

Yihao Zhao 1, Maofa Wang 1,*, Huanhuan Xue 1, Youping Gong 2 and Baochun Qiu 2

Abstract: The prediction of underwater acoustic transmission loss in the sea plays a key role in generating situational awareness in complex naval battles and assisting underwater operations. However, the traditional classical underwater acoustic transmission loss models do not consider the regional hydrological elements, and the performance of underwater acoustic transmission loss prediction under complex environmental conditions in a wide range of sea areas is limited. In order to solve this problem, we propose a deep learning-based underwater acoustic transmission loss prediction method. First, we studied the application domains of typical underwater acoustic transmission loss models (ray model, normal model, fast field program model, parabolic equation model), analyzed the constraint rules of its characteristic parameters, and constructed a dataset according to the rules. Then, according to the characteristics of the dataset, we built a DBN (deep belief net) neural network model and used DBN to train and learn the dataset. Through the DBN method, the adaptation and calculation of the underwater acoustic transmission loss model under different regional hydrological elements were carried out in a simulation environment. Finally, the new method was verified with the measured transmission loss data of acoustic sea trials in a certain sea area. The results show that the RMSE error between the underwater acoustic transmission loss calculated by the new method and the measured data was less than 6.5 dB, the accuracy was higher than that of the traditional method, and the prediction speed was faster, the result was more accurate, and had a wide range of adaptability in complex seas.

Keywords: deep learning; DBN; underwater acoustic transmission loss; hydroacoustic survey

1. Introduction

The ocean contains abundant natural resources, and all countries are intensifying their development of the ocean and enhancing the maintenance of maritime rights and interests in those territorial waters. Future naval warfare will have a high degree of informatization. Obtaining information superiority on the battlefield is the basis for winning modern wars and defending maritime rights and interests. Since the Germans published their first paper on underwater acoustics in 1919 [1], people have gradually established a theoretical system of underwater acoustic modeling and constructed a classic underwater acoustic transmission loss model, mainly including the ray model (RM), normal model (NM), fast field program (FFP) model, parabolic equation (PE) model, and some hybrid algorithms [1–3]. Each hydroacoustic model algorithm has limitations in the applicable situations. For example, the NM and FFP models are suitable for distance-independent sound field calculations; the RM and PE model are suitable for distance-dependent sound field calculation. In addition, the RM is suitable for high-frequency calculations, and it is more difficult to handle low-frequency and caustic calculations; the PE model is suitable for low-frequency and narrow-angle calculations, and the calculation efficiency is not very high for high-frequency and deep-sea environments.
regional physical parameters and signal frequency ranges may change. It is necessary to select an adapted hydroacoustic model to calculate the most accurate hydroacoustic propagation loss prediction value. This is an urgent problem in the research of marine underwater acoustics.

Jensen (1982, 1984) proposed an underwater acoustic transmission loss model matching scheme that classified five acoustic models according to the depth (deep and shallow), frequency (high and low), and the relationship between ocean environment and distance (related and unrelated) adaptation [3]. This method has no restrictions on frequency and water depth, has few judged features, and relies on a large amount of prior experience and expert knowledge for feature extraction. Based on the ray model and the parabolic equation model, Chen Zhiguo and others [4] studied the energy propagation mode of ocean fronts and vortices and made more accurate predictions for four different ocean sound fields. Li Jiaxun and others [5] designed a system with the Kraken normal model as the core including the Bellhop ray model, the Scooter finite element model, and the Sparc time-domain stepping model to calculate several typical sound field transmission losses. The above forecasting method still relies on expert experience to realize it. However, the large dynamic changes of the marine environment make the hydrological conditions have complex time–space–frequency characteristics, and the underwater acoustic field also presents complicated time–space random fluctuations, the environment is uncertain, and the parameters are uncertain. In a dynamically changing environment, it is difficult for traditional methods that rely on expert experience to accurately extract features so that the acoustic model can effectively and adaptively match the changing marine environment in real time. Therefore, when predicting the underwater acoustic transmission loss, the adaptive matching of the underwater acoustic transmission loss model needs to find a new method. In recent years, in the context of big data and deep learning, deep learning has been successfully applied to multiple research fields such as speech recognition, image processing, wireless communication, medical treatment, and social decision-making [6–8], bringing new ideas in solving traditional underwater acoustic transmission loss prediction problems. The deep belief net (DBN) proposed by Hinton is a neural network containing multiple hidden layers [9]. Compared with shallow models, it can effectively avoid problems such as overfitting and insufficient classification capabilities [10–12]. DBN is composed of several layers of restricted Boltzmann machine (RBM) and a layer of back propagation (BP) neural network [13–15]. It not only has a good classification effect but also has an efficient feature extraction ability. We propose a prediction method of underwater acoustic transmission loss based on deep learning. Based on the structure of the DBN network, according to the marine environment and signal characteristics, we designed a deep belief neural network model (DBN) and generated a training dataset to train the DBN neural network model and evaluate the accuracy of the model. The new method could accurately adapt the underwater acoustic transmission loss model and accurately predict the underwater acoustic transmission loss in the absence of prior knowledge, multiple frequency bands, and complex marine environmental conditions. In order to verify the practicability of the underwater acoustic transmission loss prediction method proposed in this paper, we used sea test data in a certain sea area in May 2020 to calculate the measured underwater acoustic transmission loss. We compared and analyzed the measured underwater acoustic transmission loss with the underwater acoustic transmission loss predicted by the new method. The results show that the RMSE error of the predicted transmission loss and the sea trial transmission loss was stable below 6.5 dB, which was lower than the transmission loss prediction error predicted by the classical method, which verifies the reliability and accuracy of the new method.

2. Deep Belief Net (DBN)

2.1. Principle of DBN Algorithm

DBN is a generative model of deep learning, which can be used for unsupervised learning and as an autoencoder, and can also be used for supervised learning and used as a classifier.
From the perspective of unsupervised learning, its purpose is to retain the characteristics of the original features as much as possible while reducing the dimension of the features; from the perspective of supervised learning, its purpose is to make the classification error rate as small as possible. As the basic building block of DBN, RBM is a probabilistic modeling method based on energy function. It has strong unsupervised learning ability and is a key model in the process of constructing and training DBN. The model performs continuous state transitions on the input features of the sample, so that the model can finally fit the distribution characteristics of the sample set, can learn arbitrarily complex rules in the data, and is suitable for the training and learning of large-scale dynamic data features in complex sea areas. The core of the algorithm is to optimize the connection weights between layers with a layer-by-layer greedy learning algorithm [16–18]. As shown in Figure 1, each RBM contains a pair of visible layers \((v_1, v_2, \ldots, v_m)\) and hidden layer \((h_1, h_2, \ldots, h_n)\), where each layer is connected by a fully connected layer [19] and \(w\) is the weight value of the connection between the two layers. It is a random neural network model without feedback.

Visible Layer

![Schematic diagram of restricted Boltzmann machine.](image)

Figure 1. Schematic diagram of restricted Boltzmann machine.

Suppose that in a restricted Boltzmann machine network, the number of visible units is \(m\), the state of visible units set in a certain layer is represented by a vector \(v\), and the state of hidden units in a certain layer is represented by a vector \(h\). \(V_i\) and \(h_j\) represent the states of the \(i\) visible unit and the \(j\) hidden unit, respectively. When the state of the nodes in the visible layer and the hidden layer is determined, the energy function of the RBM [20] can be expressed as:

\[
E(v, h|\theta) = -\sum_{i=1}^{m} a_i v_i - \sum_{j=1}^{n} b_j h_j - \sum_{i=1}^{m} \sum_{j=1}^{n} v_i w_{ij} h_j
\]  

(1)

in which \(\theta = (w, a, b)\) is the parameter of RBM: \(W_{ij}\) is the connection weight between the visible unit \(i\) and the hidden unit \(j\); \(a_i\) is the bias of the visible unit \(i\); and \(b_j\) is the bias of the hidden unit \(j\). When \(\theta\) is known, since the state of RBM obeys the regular distribution, the joint probability distribution of any set of visible and hidden unit node states \((v, h)\) is:

\[
P(v, h|\theta) = \frac{e^{-E(v, h|\theta)}}{Z(\theta)}
\]  

(2)

in which, \(Z(\theta) = \sum_{v,h} e^{-E(v,h|\theta)}\) is the normalization factor.
For the distribution $P(v|\theta)$ of the input data $v$, it can be determined by the likelihood function of the joint probability distribution $P(v, h|\theta)$:

$$P(v, h|\theta) = \frac{\sum_h e^{-E(v, h|\theta)}}{Z(\theta)}$$  \hspace{1cm} (3)

Because the connection between RBM layers is restricted, regardless of whether it is a visible layer or a hidden layer, there is no connection between nodes in the layer, and its activation state is also independent. Therefore, when the visible unit is given, the activation probability of the $j$ hidden unit is:

$$P(h_j = 1|v, \theta) = \delta(h_j + \sum_i h_i W_{ij})$$  \hspace{1cm} (4)

In which, $\delta(x) = \frac{1}{1+e^{-x}}$ is the sigmoid activation function. Similarly, when the hidden unit is given, the activation probability of the $i$ visible element is:

$$P(v_i = 1|h, \theta) = \delta(a_i + \sum_j v_j W_{ij})$$  \hspace{1cm} (5)

When the state $v$ of the visible layer node is known, the above formula can be used to calculate $P(h|v)$ to determine the state of the hidden layer $h$; then the state of the visible layer $V_1$ can be obtained by calculating $P(v|h)$.

Take $S = \{v_1, v_2, \ldots, v_s\}$ (capacity $s$) as the training sample, and obtain the model parameter $\theta$ by maximizing the log likelihood function $L(\theta)$ of the RBM on the sample and fit the training sample. The hidden layer can be regarded as a feature of the input data of the visual layer.

$$\theta^* = \arg \max_{\theta} L(\theta) = \arg \max_{\theta} \sum_{k=1}^{s} \ln P((v_k|h, \theta))$$  \hspace{1cm} (6)

The log-likelihood function is differentiated by the contrast divergence (CD) [14] algorithm. For example, the current weight is $w_{ij}$:

$$\frac{\partial \ln P(v|\theta)}{\partial w_{ij}} = E_{v|h, data} - E_{v|h, model}$$  \hspace{1cm} (7)

Gibbs sampling is used to determine the expectation $E_{v|h, data}$ of the training samples and the expectation $E_{v|h, model}$ defined by the RBM model. Up to now, $w_i$ can be updated as follows:

$$w_{ij} = \theta w_{ij} + \eta \left( E_{v|h, data} - E_{v|h, model} \right)$$  \hspace{1cm} (8)

In the formula, $\theta$ is the momentum and $\eta$ is the learning rate. Similarly, $a_j$ and $b_j$ can also be updated in this way.

### 2.2. DBN Model Structure

The structure of the DBN model constructed in this paper is shown in Figure 2. The first visible layer and the second hidden layer constitute the first RBM, and the second hidden layer and the third hidden layer form the second RBM, by analogy, a total of five RBM are formed. Input the data of the fifth RBM into a supervised classifier, and use BP to fine-tune the relevant parameters of the DBN to form a complete DBN to classify the input data. $X = \{x_1, x_2, \ldots, x_r\}$ is the input of the first visual layer, representing the current marine environment and signal characteristic parameters (as shown in Table 1). $h_{x,y}$ are the states of $y$ unit in the $x$ hidden layer; $\{x_1, x_2, x_3, x_4, x_5\}$ are the connection weight vectors between adjacent layers. The top-level classifier network structure uses the BP backpropagation neural network and $w_{hi}$ is the weight vector of the BP neural network. If the output result cannot get the expected tag data, the BP backpropagation adjustment is used to continuously approach the expected output of the representative label.
As shown in Figure 3, the training DBN was divided into three stages. The first stage is the model initialization stage. Regarding any two adjacent layers in the DBN as an RBM, input the output vector of the previous RBM to the next RBM, and initialize the parameters in the DBN layer by layer. The second stage is the unsupervised forward learning stage. The preprocessed marine environment features and signal features are input to the first visual layer. The greedy unsupervised learning method is used to train layer by layer, and the activation function of each unit is sigmoid. In the second stage, the shallow original features are expressed at a high level. Since the unsupervised algorithm cannot optimize the entire network parameters, it is necessary to enter the third stage, the reverse supervised learning stage, which uses the BP supervised algorithm to reversely fine-tune the DBN to optimize the underwater acoustic transmission loss prediction under the large-scale marine environment and the changeable signal frequency band. This training mode directly maps data from input to output through unsupervised training, shortening the time of supervised training, and is also the key to a powerful feature rule extraction capability.

### Table 1. Characteristic parameters.

| Characteristic Parameters | Symbol | Range            | Definition                                                  |
|---------------------------|--------|------------------|-------------------------------------------------------------|
| Frequency /(Hz)           | Fs     | 10–30,000        | The frequency of the signal emitted by the sound source.   |
| Sound source depth/(m)    | SD     | 100–4000         | Depth of emitted sound source.                             |
| Sound source opening angle/(°) | SA  | 90–90            | The opening angle of the emitted sound source.              |
| Horizontal distance/(km)  | R      | 0.5–500          | Distance between sound source and receiving array.         |
| Sea-depth/(m)             | H      | 10–4000          | The depth of the sea.                                      |
| Distance-independent/re-  | RI/RD  | 0/1              | Is the ocean characteristic only related to depth, or is it |
| lated                      |        |                  | related to receiver azimuth and distance.                  |
| Deposition sound velocity/(m·s⁻¹) | Cs | 1520–1836        | Sound propagation velocity of seabed sediments             |
| Deposit density/(g·cm⁻³)  | p      | 1.421–2.034      | Density of seabed sediments                                |
| Deposition layer thickness/(m) | Mt | 5–50             | The thickness of the deposited layer                        |
3. Dataset Construction and Model Training

3.1. Feature Parameter Rules and Sample Set Construction

This article mainly studied the prediction method of underwater acoustic transmission loss, so it focused on the influence of different underwater acoustic characteristic parameters on the prediction results of the acoustic propagation model. The problem of sound transmission under water is mainly studied by two methods: one method is the wave theory, which studies the changes in the amplitude and phase of the acoustic signal in the sound field; and the other method is the ray theory, which can be used in a high-frequency range, sound wave is regarded as a beam of rays, and the change of the sound intensity in the sound field with the beam of rays is usually studied. From these two theories, several classic underwater acoustic transmission models have been developed to characterize the underwater acoustic propagation process including NM, RM, PE model, etc. [3].

In underwater acoustics, the standard measure of sound field signal strength versus distance is the transmission loss (TL), which is the ratio of the sound intensity $I(r, z)$ at a certain point in the sound field to the sound intensity $I_0$ at a distance of 1 m from the sound source. It can be expressed as Equation (9):

$$TL = -10 \log \frac{I(r, z)}{I_0} = -20 \log \frac{|p(r, z)|}{P_0}$$  \hspace{1cm} (9)

where $p(r, z)$ is the sound pressure at a certain point in the sound field; $r$ is the distance from the sound source position to a certain point in the sound field; $z$ is the receiving depth of a certain point in the sound field; $P_0$ is the sound pressure value at a distance of 1 m from the sound source; and TL is the value of the transmission loss of the sound field in decibels (dB).

This paper selected the Bellhop model in ray mode for research. This model was written by Porter and Bueker [21] in 1987 using the Gaussian beam tracking method to calculate the sound field in a horizontally non-uniform environment (conditions related to distance). The experimental data of this model in the frequency range of 600 Hz–30 KHz were consistent with the theoretical model [22]. It is suitable for near-field conditions. The important characteristic length of the water depth or sound velocity profile should be greater than 20 wavelengths [23].

For the normal model, this paper selected the Kraken model for research. This model was jointly developed by the U.S. Naval Ocean System Center (NOSC) and the U.S. Naval Research Laboratory (NRL) using the finite difference method to solve the normal wave equation. It is suitable for low-frequency (lower than 500 Hz) shallow sea and far-field calculation under uniform horizontal (distance-independent) marine environmental conditions, and the seabed structure must be given in the calculation, that is, the parameters of $C_s$, $K_d$, $p$, and $M_t$ in Table 1.

For the fast field program model (FFP model), we choose the Scooter model for research. This model applies fast Fourier transform (FFT) technology to solve the wave equation. It is limited to the condition of a horizontally layered homogeneous medium.
that is not related to distance, near-field environment, and is suitable for calculations in the low-frequency range (below 500 Hz).

For the parabolic equation (PE) model, the RAMGeo model was selected. This model is an approximate solution of the wave equation with a narrow angle (<±20°) considering the coupling effect of acoustic diffraction and normal waves of various numbers. It is suitable for calculating narrow angles. The calculation speed for low frequency (lower than 500 Hz) and shallow sea is very fast. When the frequency increases and the depth increases, the calculation time increases rapidly, therefore, this model is not applicable.

The above far-field and near-field rules are constrained by the characteristic parameter horizontal distance (R). Suppose that the total spacing of the uniform linear array element D = m × d; m is the number of array element spacing; the highest frequency wavelength of the sound source (the minimum wavelength of the sound source) is \( \lambda_{\text{min}} \), if the sound source is at the level of the array Distance \( R > \frac{2D}{\lambda_{\text{min}}} \), it is the far-field model, otherwise it is the near-field model. Under the near-field model, sound waves are regarded as spherical waves, and under the far-field model, sound waves are regarded as plane waves. The calculation of underwater acoustic transmission loss mainly includes characteristic parameters such as signal frequency, sea depth, and seabed environment. The specific parameters and parameter ranges are shown in Table 1.

The optimal characteristic parameter ranges of the above four types of underwater acoustic transmission loss models are shown in Table 2.

**Table 2.** Underwater acoustic transmission loss mode characteristic parameter range.

| Mode        | Fs (Hz) | SD (m)   | SA (°)  | R (Km)  | H (m)  | RI/RD (0/1) | Cs     | p       | Mt     |
|-------------|---------|----------|---------|---------|--------|-------------|--------|---------|--------|
| Ray Mode    | 600~30,000 | 210~4000 | −90~−90 | 0.5~500 | 200~4000 | 1           | 1520~1836 | 1.421~2.034 | 5~50   |
| Normal Mode | 0~500   | 10~800   | −90~−90 | 100~500 | 10~800  | 0           | 1520~1836 | 1.421~2.034 | 5~50   |
| FFP Model   | 0~500   | 210~4000 | −90~−90 | 0.5~100 | 200~4000 | 0           | 1520~1836 | 1.421~2.034 | 5~50   |
| PE Model    | 0~500   | 10~500   | −20~−20 | 0.5~500 | 10~500  | 0/1         | 1520~1836 | 1.421~2.034 | 5~50   |

Sample data generation is very important for neural network training, and the distribution and statistical characteristics of the sample will also directly affect the training efficiency and prediction effect of the neural network. Based on the above constraints, we built a dataset with a normal distribution to train the neural network model. The training set generation method is as follows:

1. Select an underwater acoustic propagation model, and generate the data on the principle of a normal distribution within the constraint range according to the above-mentioned constraint conditions of the parameters of the selected model.
2. Set labels for the samples in the established training set. The labels of each sample are different. Usually, a combination of 0–K (K is the number of models) is used to set the corresponding label. The labels are formed into a new set corresponding to the input data.
3. Combine different marine environment characteristics and signal characteristics according to the above steps to form a large training set. Then, after making a one-to-one correspondence between the input data and the label, randomly shuffle the order of the input data, and recombine the training set. The purpose of shuffling the data is to prevent too much noise data of the same type from clustering together, causing over-fitting.

Different feature parameters have different weights in the adaptation model. In the DBN neural network, the weight of each feature parameter is automatically optimized according to the training data, and finally, the optimal weight is achieved through iterative training. If we want the neural network to learn these rules in-depth, we will build more data for the network’s learning of the feature parameters with larger weight, and the rest of the feature parameters can be iterated by the neural network to obtain the optimal weight.

As shown in Table 2, we extracted a total of nine feature parameters. Compared with other
models, Fs is a significant feature parameter in the ray mode and has a larger weight, so we increased the number of Fs parameters in the ray mode training. For the ray model, we increased the number of samples for the feature parameter Fs with a larger weight, and constructed 50 different values within the constraint range. The remaining eight different characteristic parameters were used as a group, and a total of 20 groups of different values were generated within the constraint range. Through the combination of the characteristic parameter Fs and other characteristic parameters, the total number of data samples of the ray model was 1000. For normal mode, it can be seen that H and R had a larger weight. We first generated 50 data with different values for the characteristic parameter H, and the remaining eight characteristic parameters generated 20 sets of sample data within the constraint range, and combined them to generate 1000 sets of data. Then, we generated 50 different values for the characteristic parameter R, and the remaining eight characteristic parameters generated 20 sets of sample data within the constraint range, and combined them to generate 1000 sets of data. A total of 2000 sets of sample data were generated for training the normal wave model data. For the fast field model, like the normal model, the H parameter and the R parameter had greater weight, but the constraint range was different, so 2000 sets of sample data were also generated. For the parabolic model (PE), it can be seen that the characteristic parameter SA had a greater weight, the same as the ray model, and 1000 sets of data were generated. Finally, 6000 sets of sample data were generated for DBN training.

3.2. Model Initialization and Training

First, we initialized the model to determine the momentum factor, learning rate, maximum number of iterations, and the number of hidden layer nodes related to the RBM training process. In addition, because DBN does not clearly define the most suitable layer number, the determination of the layer number needs to determine the depth of the DBN model through experimental evaluation results. Then, we set the relevant parameters of RBM, set the initial value of learning rate to 0.1 according to experience, and set the initial momentum to 0.5, where the number of input nodes is the number of characteristic parameters in Table 1, and the number of output nodes is four nodes equal to the number of acoustic models. The number of pre-training iterations (epochs) was selected as 200, and the number of global reverse fine-tuning was 100. In order to prevent over-fitting, we randomly selected 3/4 of each type of data as training samples, and the rest as test samples. The number of the random sample training batch size was 20. Table 3 shows the performance of the DBN neural network at different depths. It can be seen that as the depth of the DBN model increases, the time cost also doubles. When the depth is greater than 5, the classification accuracy does not increase significantly. The training time has been increasing by a large margin, so for comprehensive training time and classification accuracy, five layers is a more reasonable depth choice.

| DBN Layers | Accuracy (%) | Training Time (s) |
|------------|--------------|-------------------|
| 3          | 75.135       | 210.752           |
| 4          | 93.456       | 420.456           |
| 5          | 95.876       | 590.456           |
| 6          | 94.123       | 760.356           |
| 7          | 95.756       | 900.456           |

The training of DBN mainly consists of two parts. The first part is unsupervised bottom-up layer-by-layer pre-training, which transfers the feature vector from the input to the output. The second part is the supervised top-down reverse fine-tuning. The output value is compared with the given data label to get the error, and the error is back propagated from the output end to the input end to achieve the purpose of adjusting the DBN parameters. The overall training process is shown in Figure 4:
The specific training steps are as follows:

Step 1: Preprocess the dataset constructed in Section 3.1 and input data into the model in batches for training.

Step 2: Set the bias of the visible layer and the hidden layer and the initial value of the weight between each neuron $\theta = \{b, c, w\}$. $b$ represents the bias of the input data layer or visual layer, $c$ represents the bias of each hidden layer or data output layer, and $w$ represents the connection weight of each layer. According to the research of Hinton et al., the connection weight $w$ of each layer is initialized to obey the normal distribution $N(0,1)$, and the initial value of the bias $b$ of the visible layer and the bias $c$ of the hidden layer is set to 0.

Step 3: Next, we perform iterative training on DBN. In the DBN model, each layer of RBM learns the characteristics of the input data forward and unsupervised, and uses the greedy unsupervised learning algorithm to update the bias value of each unit of each layer and the weight between the hidden layer.

Step 4: After completing the network pre-training, we add the Softmax output layer, and use fine-tuning to transform the entire network from a generative model to a discriminative model. This network loss function can be defined as:

$$L(\lambda) = -\frac{1}{N} \sum_{n=1}^{N} 1\{\text{label}(n) = 1\} \log \left( \exp \left( W_0^{(n)} x + b^{(n)} \right) \right) / \sum_{n=1}^{N} \exp(W_0 x + b) \quad (10)$$

In the formula, $1\{\text{label}(n) = 1\}$ means that the condition in brackets is set to 1 when the condition is satisfied, and 0 is not satisfied. $W_0^{(n)}$ and $b^{(n)}$ are the connection weight vector of the nth node in the output layer and its offset; $x$ is the response of the upper node. The fine-tuning process generally uses the gradient descent method, the goal is to minimize the loss function and use the error back propagation (BP) to fine-tune the network. The fine-tune process is to compare the correct classification labels corresponding to different feature parameters (Table 1) in turn. If the expected output value is not obtained, the difference between the actual output and the expected output (error) is calculated, and the error is successively modified by modifying the weight of each neuron propagate to the input layer.

Step 5: When iterations end, the training of the entire model is completed. After the training of the DBN model is completed, all training samples can be accurately classified under the correct label assigned, that is, the error between the predicted label and the real label converges to 0. Use loss $L(\lambda)$ to evaluate the degree of convergence of the model: Loss tending to 0 means that the model...
converges well. As shown in the loss curve of Figure 5, as the number of iterations increases, the loss tends to 0.

![Figure 5. DBN training accuracy and loss curve.](image)

This training uses 75% of the data to train the model, and its loss function and accuracy are shown in Figure 5. It can be seen from the figure that as the number of training increases, the model prediction accuracy gets higher and higher, and the error loss curve also declines, and finally remains at a stable value. At this time, an increase in the number of training may lead to overfitting. Finally, the 25% data in the sample set is used for testing, and the accuracy of model adaptation is obtained, which is 94.86% in combination. It shows that the recognition accuracy is already very high, close to the distribution of the sample data, and then the underwater acoustic transmission loss is calculated through the model matching and adaptive underwater acoustic propagation model.

4. Experimental Results and Discussion

4.1. Sources of Data

To verify the accuracy of the new forecasting method, we conducted a sea trial experiment in May 2020 to analyze and process the obtained sea trial data to obtain the actual underwater acoustic transmission loss data in this sea area. Then, we compared and analyzed the measured transmission loss data with the data predicted by this method. The sea trial site is shown in Figure 6.

The experimental area is shown in Figure 7. The experimental ship traveled from Q1 (Figure 7) to Q2 (Figure 7) for the experiments. The ship carried the towed array sonar and continuously emitted a set of signals, and the 20-element vertical receiving array at the ZS (Figure 7) station received the signals emitted by the towed array sonar. We placed a CTD at the receiving array for sound velocity profile measurement. The measured sound velocity profile is shown in Figure 7. The planned total voyage of the experimental ship was 160 km, the starting Q1 was 274 km away from the receiving array, and the ending Q2 point was 420 km away from the receiving array. Among them, we selected one of the sea trial data with a voyage to establish transmission loss data for verification.
The signals continuously transmitted by the experimental ship during the experiment are shown in Table 4, where Continuous Wave (CW) represents a single-frequency signal, and Linear Frequency Modulated (LFM) represents a frequency-modulation signal. In each group of signals, single-frequency or FM signals are transmitted according to the law shown in Table 4, and a group of signals is transmitted with a period of 160 S. As shown in Figure 8, it is a time-frequency domain diagram of the transmitted signal received at a certain moment for the 311 m depth array element in the receiving array. As shown in Figure 8a, although it is impossible to analyze useful information on the time domain diagram, it can be seen in the frequency domain diagram that there are significantly higher peaks at the frequency of 360 Hz than the other frequency points. This phenomenon indicates that the 360 Hz signal transmitted by the experimental ship is received on the receiving array element, and the energy intensity of the received signal can be obtained in the frequency domain. Similarly, as shown in Figure 8b, the 320–400 Hz LFM signal can be found in the frequency domain, and its center frequency can be selected to obtain the energy intensity of the received signal. The remaining CW and LFM signals use the same method to obtain the received signal energy intensity. Next, according to the known transmitted signal energy intensity and received signal energy intensity, we can calculate the actual underwater acoustic transmission loss for this distance and generate the underwater acoustic transmission loss dataset.
Table 4. Periodic signal.

| Signal Type | Signal Frequency | Duration (S) | Interval (S) |
|-------------|------------------|--------------|--------------|
| CW          | 245              | 5            | 15           |
| CW          | 360              | 5            | 15           |
| LFM         | 220–270          | 10           | 30           |
| LFM         | 320–400          | 10           | 30           |
| LFM         | 600–800          | 10           | 30           |

(Continuous Wave (CW) Linear Frequency Modulated (LFM)).

Figure 8. Time–frequency domain diagram.
Based on the transmitted signals in Table 4, this paper constructed a sea trial transmission loss dataset. The sound source level (SL) of the CW signal was 192 dB, the SL of the LFM signal was 174 dB, and the emission depth was 140 m. As shown in Figure 9, it can be clearly seen in the LOFAR spectrum that the high-brightness area represents the transmitted signal received at the current moment. The LOFAR spectrum of the 360 Hz CW signal is shown in Figure 9a. It can be seen that the duration was 10 S, which was the same as the description in Table 4, indicating that the transmitted CW signal was received at this time. The LOFAR spectrum of the 320–400 Hz LFM signal is shown in Figure 9b. It can be seen that its duration was also 10 S, which was the same as the description in Table 4, indicating that the transmitted LFM signal was received at this time. The rest of the signal can be found in the same way, and the received energy can be analyzed. For this experiment, the calculation of the center frequency of the CW signal and the LFM signal was used. According to the definition of the sound source level and signal level, the transmission loss calculated by Equation (9) can be simplified as:

$$TL_i = SL - S_i$$ (11)

where $TL_i$ represents the transmission loss of the $i$ transmitted pulse, and represents the signal level of the transmitted pulse at the receiving element end. At each signal frequency, each array element collects 30 data points, so each signal frequency forms a sea trial transmission loss data set of 250 data points. The five transmitted signals form a sea trial transmission loss dataset of 1250 data points in total.

(a) 360 Hz LOFAR spectrum  (b) 320–400 Hz LOFAR spectrum

**Figure 9.** Low-Frequency acquisition and ranging (LOFAR) spectrum.

4.2. Sea Area Characteristic Parameters

The currently known environmental parameters are as follows: the depth of the sound source was 141 m, the distance between the sound source and the receiving array was 274–340 km, and 200 sampling points were taken at equal intervals in this range. The signal frequency was 360 Hz, the sea depth was 2000 m, the sedimentary layer thickness was 50 m, the sea bottom was sand-mud-clay, that is, the sedimentary sound velocity was 1579 m/s, the sedimentary density was 1.596 g/cm³, and the sedimentary attenuation was 0.113 dB/m-kHz. We placed a CTD at the receiving array to measure the sound velocity profile. The measured sound velocity profile is shown in Figure 10.
Figure 10. Sound speed profile.

Figure 10 shows that it is a typical deep-sea sound speed profile curve. It can be seen that below 1000 m, the temperature distribution was mainly affected by sunlight, which presents a relatively stable negative gradient. Above 1000 m, the seawater temperature is constant, which is called the deep-sea isothermal layer. The sound velocity distribution of this layer mainly depends on the increase in static pressure, presenting an almost linear positive gradient (0.16 m/s/m) distribution.

4.3. Forecast Result Analysis

To make a more scientific assessment of the overall performance of the underwater acoustic transmission loss prediction by the new method, we applied the root mean square error (RMSE) to evaluate the transmission loss and the transmission loss of the underwater acoustic model mismatch prediction by this method. The actual transmission loss of sea trials was compared with errors to evaluate the forecast accuracy of this method. The specific definition is as Equation (12):

$$RMSE = \sqrt{\frac{\sum_{i=1}^{N} (\hat{y}_i - y_i)^2}{N}}$$  (12)

where $\hat{y}_i$ represents the predicted transmission loss; $y_i$ represents the actual transmission loss; and $N$ represents the predicted transmission loss point. The smaller the value of RMSE, the smaller the difference between the predicted transmission loss data and the sea trial transmission loss data, that is, the more accurate the forecasting ability.

According to the above-mentioned marine environmental characteristic parameters, we selected the characteristic parameters mentioned in Table 1, and input them into the DBN neural network constructed in this article for underwater acoustic model adaptation and underwater acoustic transmission loss prediction. As shown in Figure 10, Comparing the prediction results of the model selected by the new method with the prediction results of other underwater acoustic models, it can be seen that the prediction results of the model selected by the new method were the most consistent. In the figure, the $x$-axis is the transmission distance in m, and the $y$-axis is the energy intensity in dB.

Figure 11a shows that the ocean transmission loss predicted by the new method using this system is basically in line with the actual transmission loss. Although there are error fluctuations between the predicted curve and the actual curve, it can be seen in Figure 11 that the system has the smallest prediction error. The calculated RMSE error is shown in
Table 5. The average RMSE error of the transmission loss data under each array element was 6.31 dB, which was the smallest error compared to other models.

Comparing Figure 11c,d, it is obvious that these two models are not suitable for the current marine environmental conditions. If this model is selected incorrectly for calculation, it will cause the calculation result to deviate too much from the actual result. The results verified by actual data showed that the new method had higher matching for the model selection of transmission loss prediction, and the prediction of transmission loss had higher accuracy.

Figure 11b shows that the transmission loss curve predicted by Bellhop had a certain error deviation from the actual transmission loss curve. The calculated RMSE error is shown in Table 5. The average RMSE error of the transmission loss data under each array element was 11.56 dB.

Comparing Figure 11a with Figure 11b, it can be seen that the predicted transmission loss curve was actually very close, and in Figure 12, it can be seen that the error curve was also very close. In this case, in actual application, if there is no rich expert experience, it is difficult to choose the best matching underwater acoustic propagation model. The RMSE of the transmission loss result predicted by this method was compared with the
transmission loss RMSE predicted by Bellhop, and the error was reduced by 5.25 dB. The results show that this method can easily match a suitable underwater acoustic propagation model and calculate accurate underwater acoustic transmission loss prediction in a complex ocean environment.

Comparing Figure 11c,d, it is obvious that these two models are not suitable for the current marine environmental conditions. If this model is selected incorrectly for calculation, it will cause the calculation result to deviate too much from the actual result. The results verified by actual data showed that the new method had higher matching for the model selection of transmission loss prediction, and the prediction of transmission loss had higher accuracy.

As shown in Figure 11c, it can be seen that there were obvious errors between the actual sea test data and the data predicted by the RAMGeo model, and there was a model mismatch. Figure 12 shows that the error curve was obviously too large. The calculated RMSE error is shown in Table 5, and the average RMSE error of the transmission loss data under each array element was 17.594 dB.

As shown in Figure 11d, it can be seen that there was a large error between the actual sea test data and the data predicted by the Scooter model. It is obvious that this model does not fit under the current marine environmental conditions, and the calculated average RMSE error was 55.334 dB.

Comparing Figure 11c,d, it is obvious that these two models are not suitable for the current marine environmental conditions. If this model is selected incorrectly for calculation, it will cause the calculation result to deviate too much from the actual result.

The results verified by actual data showed that the new method had higher matching for the model selection of transmission loss prediction, and the prediction of transmission loss had higher accuracy.

4.4. Discussion

Aiming at the prediction of underwater acoustic transmission loss, we proposed a new method for the prediction of underwater acoustic transmission loss. Under ideal marine environment and constraint conditions, we constructed sample data to train and test the accuracy of the new method and finally verified by sea trial transmission loss data. Under the ideal marine environment and constraint conditions, the parameter changes are controllable, the model applicable parameter boundaries are clear, and simulation testing is easy. In the simulation test, the adaptation accuracy of the underwater acoustic transmission loss model in the new method reached 94.86%, indicating that the recognition rate is already very high. Under constrained conditions, the sample dataset generated based on the normal distribution conforms to the actual marine environment parameter distribution, which is conducive to training stable and reliable results.

Compared with the ideal environment condition where the characteristic parameter boundary during simulation training is clear, the characteristic parameter of the sea test environment has a certain degree of complexity, and the boundary overlaps each other, which increases the difficulty in adapting the acoustic model. Without in-depth research on
marine acoustics, it is difficult to quickly adapt a suitable underwater acoustic transmission loss model. Using the new method proposed in this paper, only the currently measurable characteristic parameters of the marine environment (shown in Table 1) are needed, then a suitable underwater acoustic transmission loss model can be quickly matched to calculate the transmission loss prediction. We used the measured data of a certain sea area in 2020 to verify, calculate the measured transmission loss under different depth array elements, and compare it with the transmission loss predicted by the new method. Finally, all underwater acoustic transmission loss models were used for calculation, and RMSE was used to evaluate the accuracy. The RMSE results are shown in Tables 5–9. It can be seen that the transmission loss error of the acoustic model selected by the new method under different signals and elements was the smallest. In practical applications, there are still many complicated marine environments, and the new method needs to be tested for reliability in more areas.

Table 6. RMSE error comparison (360 Hz).

| Array Element Depth (m) | Systematic Prediction Error (dB) | Bellhop Prediction Error (dB) | RAMGeo Prediction Error (dB) | Scooter Prediction Error (dB) |
|-------------------------|---------------------------------|------------------------------|------------------------------|------------------------------|
| 181                     | 6.56                            | 11.86                        | 17.75                        | 55.32                        |
| 261                     | 6.15                            | 12.75                        | 16.45                        | 54.75                        |
| 311                     | 6.45                            | 12.34                        | 18.45                        | 55.69                        |
| 411                     | 6.35                            | 11.65                        | 19.56                        | 56.23                        |
| 511                     | 6.04                            | 11.56                        | 15.76                        | 54.68                        |
| Average RMSE            | 6.31                            | 12.032                       | 17.594                       | 55.334                       |

Table 7. RMSE error comparison (220–270 Hz).

| Array Element Depth (m) | Systematic Prediction Error (dB) | Bellhop Prediction Error (dB) | RAMGeo Prediction Error (dB) | Scooter Prediction Error (dB) |
|-------------------------|---------------------------------|------------------------------|------------------------------|------------------------------|
| 181                     | 5.58                            | 12.56                        | 15.73                        | 54.58                        |
| 261                     | 5.75                            | 12.73                        | 16.55                        | 55.59                        |
| 311                     | 6.65                            | 11.45                        | 16.44                        | 57.78                        |
| 411                     | 5.74                            | 12.65                        | 14.53                        | 56.65                        |
| 511                     | 6.34                            | 11.46                        | 17.79                        | 57.79                        |
| Average RMSE            | 6.012                           | 12.41                        | 16.208                       | 56.478                       |

Table 8. RMSE error comparison (320–400 Hz).

| Array Element Depth (m) | Systematic Prediction Error (dB) | Bellhop Prediction Error (dB) | RAMGeo Prediction Error (dB) | Scooter Prediction Error (dB) |
|-------------------------|---------------------------------|------------------------------|------------------------------|------------------------------|
| 181                     | 5.68                            | 10.86                        | 18.65                        | 54.22                        |
| 261                     | 5.35                            | 11.35                        | 18.37                        | 53.37                        |
| 311                     | 5.85                            | 10.45                        | 17.03                        | 54.45                        |
| 411                     | 6.05                            | 12.78                        | 18.45                        | 54.23                        |
| 511                     | 5.04                            | 10.46                        | 17.33                        | 55.67                        |
| Average RMSE            | 5.594                           | 11.18                        | 17.966                       | 54.388                       |

Table 9. RMSE error comparison (600–800 Hz).

| Array Element Depth (m) | Systematic Prediction Error (dB) | KRAKEN Prediction Error (dB) | RAMGeo Prediction Error (dB) | Scooter Prediction Error (dB) |
|-------------------------|---------------------------------|------------------------------|------------------------------|------------------------------|
| 181                     | 5.56                            | 15.74                        | 20.15                        | 58.45                        |
| 261                     | 5.93                            | 14.78                        | 19.45                        | 57.77                        |
| 311                     | 6.35                            | 16.35                        | 19.78                        | 58.61                        |
| 411                     | 6.48                            | 15.45                        | 19.56                        | 57.28                        |
| 511                     | 5.78                            | 16.96                        | 18.96                        | 59.67                        |
| Average RMSE            | 6.06                            | 15.942                       | 19.58                        | 58.356                       |
5. Conclusions

In this paper, we constructed a new prediction underwater acoustic transmission loss method based on the DBN neural network. The new method solves the problem that the traditional underwater acoustic transmission loss prediction cannot accurately and adaptively select the underwater acoustic transmission loss model in the complex and changeable marine environment.

First of all, based on the characteristics of the DBN network, we designed a DBN model according to the marine environment and signal characteristics. We studied the applicable boundary conditions of various underwater acoustic transmission loss models (ray model, normal model, fast field program model, parabolic equation model). Based on these conditions, we constructed rules to generate datasets for training and testing DBN neural networks. Then, we evaluated the accuracy of the DBN neural network model and obtained the accuracy of the model adaptation, which was 94.86%. It shows that the recognition rate was already very high. Finally, the new method was verified by a certain sea trial data in 2020. The results showed that the new method achieved accurate prediction of the underwater acoustic propagation model in a complex ocean environment, and the average RMSE error of the prediction results under different array elements was less than 6.5 dB. The new method reduces the difficulty in matching underwater acoustic transmission loss models in complex marine environments and improves the speed and accuracy of underwater acoustic transmission loss calculations. The new method can accurately calculate the transmission loss of multi-band signals under different marine environmental conditions and has the generalization ability for the calculation of transmission loss in a wide range of sea areas.

Author Contributions: Conceptualization, Y.Z. and M.W.; Data curation, H.X. and B.Q.; Investigation, Y.Z. and B.Q.; Methodology, Y.Z. and Y.G.; Project administration, Y.Z. and H.X.; Resources, M.W., Y.G., and H.X.; Software: Y.Z., H.X., and B.Q.; Supervision, Y.Z. and H.X.; Validation, Y.G., M.W., and H.X.; Writing—original draft, Y.Z. and B.Q.; Writing—review & editing, Y.G. and M.W. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data that support the findings of this study are available from the corresponding author upon reasonable request.

Conflicts of Interest: The authors declare no conflict of interest.

References
1. Wang, D.Z.; Shang, E.C. Underwater Acoustic, 2nd ed.; Science Press: Beijing, China, 2013; pp. 59–149.
2. Jensen, F.B.; Kuperman, W.A.; Porter, M.B. Computational Ocean Acoustics, 1st ed.; Computers in Physics: New York, NY, USA, 1995; pp. 233–322.
3. Etter, P.C. Underwater Acoustic Modeling and Simulation, 3rd ed.; Publishing House of Electronics Industry: Beijing, China, 2005; pp. 112–155.
4. Chen, Z.G.; Zhang, X.; Zhang, X.H. Ocean acoustic field prediction and tactical application. Ship Sci. Technol. 2005, 27, 65–69.
5. Li, J.X.; Zhang, R.; Wang, X.L. Kraken marine acoustic model and its numerical experiment for acoustic propagation and decay. Adv. Mar. Sci. 2009, 27, 51–58.
6. Hinton, G.L.; Deng, D.; Yu, D. Deep Neural Networks for Acoustic Modeling in Speech Recognition: The Shared Views of Four Research Groups. IEEE Signal Process. Mag. 2012, 29, 82–97. [CrossRef]
7. Collobert, R.; Weston, J. A unified architecture for natural language processing: Deep neural networks with multitask learning. Int. Conf. Mach. Learn. 2008, 61, 160–167.
8. Oquab, M.; Bottou, L.; Laptev, I.; Sivic, J. Learning and transferring mid-level image representations using convolutional neural networks. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Columbus, OH, USA, 1 March–14 April 2014; pp. 1717-1724.
9. Hinton, G.E.; Salakhutdinov, R.R. Reducing the dimensionality of data with neural networks. Science 2006, 313, 504. [CrossRef] [PubMed]
10. Balouji, E.; Gu, I.Y.H.; Bollen, M.H.J.; Bagheri, A.; Nazari, M. A LSTM-based deep learning method with application to voltage dip classification. In Proceedings of the 2018 International Conference on Harmonics and Quality of Power (ICHQP), Ljubljana, Slovenia, 13–16 May 2018; pp. 1–5.
11. Hollesen, P.; Connors, W.A.; Trappenberg, T. Comparison of Learned versus Engineered Features for Classification of Mine Like Objects from Raw Sonar Images; Springer: Berlin/Heidelberg, Germany, 2011; pp. 174–185.
12. Althobiani, F.; Ball, A. An approach to fault diagnosis of reciprocating compressor valves using Teager–Kaiser energy operator and deep belief networks. Expert Syst. Appl. 2014, 41, 4113–4122.
13. Shao, H.; Jiang, H.; Zhang, X.; Niu, M. Rolling bearing fault diagnosis using an optimization deep belief network. Meas. Sci. Technol. 2015, 26, 115002. [CrossRef]
14. Li, C.; Sanchez, R.-V.; Zurita, G.; Cerrada, M.; Cabrera, D.; Vásquez, R.E. Multimodal deep support vector classification with homologous features and its application to gearbox fault diagnosis. Neurocomputing 2015, 168, 119–127. [CrossRef]
15. Hinton, G. A practical guide to training restricted boltzmann machines. Momentum 2010, 9, 926–947.
16. Bengio, Y. Learning Deep Architectures for AI. Found. Trends Mach. Learn. 2009, 2, 1–127. [CrossRef]
17. Shui, H.H.; Xu, Y.L. Target recognition method based on deep belief network. J. Comput. Appl. 2014, 34, 3314–3317. Available online: http://www.joca.cn/EN/abstract/abstract17651.shtml (accessed on 24 May 2021).
18. Chen, E.; Yang, X.; Zha, H.; Zhang, R.; Zhang, W. Learning object classes from image thumbnails through deep neural networks. In Proceedings of the 2008 IEEE International Conference on Acoustics, Speech and Signal Processing, Caesar Palace, Las Vegas, NV, USA, 31 March–4 April 2008; pp. 829–832.
19. Salakhutdinov, R.; Hinton, G. Using Deep Belief Nets to Learn Covariance Kernels for Gaussian Processes. In Proceedings of the 2007 International Conference on Neural Information Processing, Vancouver, BC, Canada, 3 December 2007; pp. 1249–1256.
20. Jolliffe, I.T. Principal Component Analysis. J. Mark. Res. 2002, 87, 513.
21. Porter, M.B. The Bellhop Manual and User’s Guide: Preliminary Draft; Heat, Light, and Sound Research: La Jolla, CA, USA, 2011.
22. Gul, S.; Zaidi, S.S.H.; Khan, R.; Wala, A.B. Underwater acoustic channel modeling using BELLHOP ray tracing method. In Proceeding of the International Bhurban Conference on Applied Sciences & Technology, Islamabad, Pakistan, 10–14 January 2017; pp. 665–670.
23. Xiong, C.L.; Wang, X.; Xia, Q.F. Bellhop Model’s Application on Efficiency Analysis of Towed Line Array Sonar. Int. Bhurban Conf. Appl. Sci. Technol. 2017, 3, 135–139.