FIFTH INTERNATIONAL CONGRESS ON SOUND AND VIBRATION
DECEMBER 15-18, 1997
ADELAIDE, SOUTH AUSTRALIA

AN EVALUATION METHOD OF THE COMPLICATED ACOUSTIC SYSTEM BASED ON THE NEURAL NETWORKS REFLECTING THE STATISTICAL STRUCTURES

Noboru Nakasako[1)], Yasuo Mitani[2)] & Mitsuo Ohta[3)]
1) Faculty of Biology-Oriented Sc. & Tech., Kinki Univ., Wakayama, 649-64 Japan.
2) Faculty of Eng., Fukuyama Univ., Fukuyama, 729-02 Japan.
3) Faculty of Eng., Kinki Univ., Higashi-Hiroshima, 739-21 Japan.

Abstract An actual sound environment often shows complicated fluctuation patterns apart from a usual Gaussian type. Various evaluation procedures for the sound environment were methodologically proposed, owing to the variety of phenomena and the complexity of human response to them. It is very difficult to evaluate precisely the proper characteristics of complicated systems only from a physical viewpoint based on the structural mechanism.

In such situation, a regression analysis function is usually employed between the input and the output fluctuations especially under the assumption of Gaussian property for the input-output fluctuations and/or the usual least squares error criterion. Furthermore, the extended regression analysis method which we previously reported was necessarily reduced in a complicated mathematical form, though it could utilize the lower and higher order correlations on the basis of Bayes' theorem.

This paper describes a regression analysis matched to the prediction of the output response probability for complicated sound environmental systems by introducing a hierarchical neural network with statistical structure. Then, based on the estimated result, the output probability can be easily predicted for the same system with arbitrary input signal. Finally, the effectiveness of the present method is experimentally confirmed by applying it to the actually observed data.

1. INTRODUCTION

An actual sound environment surrounding us often shows a complicated fluctuation pattern apart from a usual Gaussian type. Furthermore, various evaluation procedures for the sound environment were methodologically proposed from many different viewpoints, owing to the variety of phenomena and the complexity of human response to them. In such a situation, it is very difficult to evaluate precisely the proper characteristics of complicated sound environmental systems only from a bottom-up way viewpoint using physical (or structural) mechanism. Even when confining the problem to the sound insulation system, it is impossible to evaluate precisely the acoustic characteristics of complicated objective systems (e.g., non-parallel double-wall type sound insulation systems, sound-bridge type insulation systems, other actual acoustic systems, etc.), by using the traditional structural methods like the wave equation[1], the S.E.A. method[2] and the sound ray method[3]. On the other hand, a noise evaluation index L_x which is a $(100 - x)$ percentile of sound level probability distribution is widely used in evaluating an environmental noise. This means that it is necessary to identify the objective system or to predict the output response fluctuation in a form of the whole probability distribution instead of only its averaged form.

A kind of regression analysis method is usually employed between the input and the output fluctuations especially from a problem-oriented viewpoint of functional approach. The standard regression analysis artificially treats the fluctuations around the regression curve as some meaningless information[4, 5] under the assumption of Gaussian property of the input and output fluctuations and/or the usual least squares error criterion. The usual system identification requires the pre-experiment with the ideal input signal to predict the output response with a different type (actual) input signal, for instance, through some deterministic transformation based on the regression curve. However, the actual sound environmental system cannot be exactly modeled only by using such a simple regression analysis method owing to various environmental factors. Moreover, it should be noticed that there exist not only linear correlation but also higher order correlations between input and output acoustic signals. These lower and higher order correlations play an essential role in finding the system characteristics precisely.

In the previous research, the regression analysis method was extended on the basis of Bayes' theorem[6], using the above lower and higher order correlation informations. This extended regression analysis[7] results in a complicated mathematical form, though it is strictly derived from a theoretical viewpoint. The problem is how to realize the higher order statistics and how to utilize them in the probabilistic prediction.

In this paper, a regression analysis especially matched to the prediction of the output response probability for complicated actual sound environmental systems is proposed from two different viewpoints, by introducing a hierarchical neural network with statistical structure and the practical probability distribution. More concretely, in the first method, since the regression analysis is based on the conditional probability distribution, the conditional statistics on its distribution are constructed by use of multilayer neural networks. In the second method, so-called Parzen-window is introduced as the joint probability density function of input-output fluctuations because of its simple form and the conditional probability density function with respect to the output response is estimated on the basis of the Bayes' theorem. Then, based on the estimated result of sound environmental systems, the output probability can be easily predicted for the same system with arbitrary input signal. Finally, the effectiveness of the present method is experimentally confirmed by applying it to the prediction problem of output response in the actual sound environments.

2. INPUT-OUTPUT RELATIONSHIP IN A COMPLICATED SOUND ENVIRONMENT

Let us consider an actual case when a response variable y depends on the independent variable x fluctuating in arbitrary type probability distribution forms (See Fig.1). In the

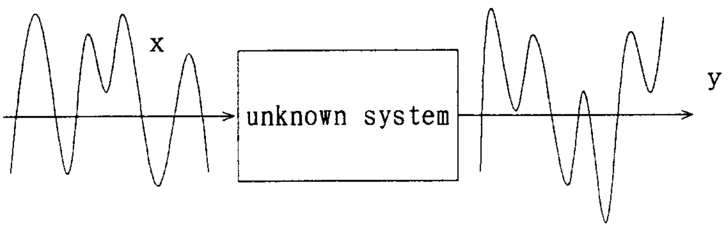


Fig.1 A complicated sound environmental system with input x.

usual regression analysis, only the conditionally expected value of y can be utilized under the definition of regression curve as follows:

$$\hat{y}(x) = \int y p(y|x) dy. \tag{1}$$

Here, the problem is how to utilize the linear and non-linear correlations between x and y, because the conditional probability density function includes all of correlation informations.

Until now, though theoretically strict methods were proposed on the basis of Bayes' theorem to extract the lower and higher order correlations between x and y, and utilize them in the prediction, most of their methods were given in a complicated mathematical form. So, it is necessary to find a new method to predict the output probability distribution from a practical viewpoint.

3. ESTIMATION OF INPUT AND OUTPUT RELATIONSHIP

As mentioned in the previous section, the conditional probability distribution $p(y|x)$ plays an essential role in predicting the response fluctuation not only in an averaged form but also in a whole probability distribution form. As is well-known, there are some ways to treat concretely the conditional probability distribution $p(y|x)$.

3.1 Neural Network to Express Conditional Statistics (Method 1)

First, let us introduce the statistical Hermite expression[8] as the conditional probability density function $p(y|x)$ which is generally applicable to arbitrarily random phenomena, as follows:

$$\begin{aligned} p(y|x) &= \frac{1}{\sqrt{2\pi}\sigma_y(x)} \exp\left\{-\frac{(y-\mu_y(x))^2}{2\sigma_y(x)^2}\right\} \sum_{n=0}^{\infty} a_n(x) \frac{1}{\sqrt{n!}} H_n\left(\frac{y-\mu_y(x)}{\sigma_y(x)}\right), \qquad (2) \\ \mu_y(x) &= \langle y|x\rangle,\ \sigma_y^2(x) = \langle (y-\mu_y(x))^2|x\rangle, \\ a_0(x) &= 1,\ a_1(x) = a_2(x) = 0,\ a_n(x) = \frac{1}{\sqrt{n!}} \left\langle H_n\left(\frac{y-\mu_y(x)}{\sigma_y(x)}\right) \middle| x \right\rangle (n \geq 3), \end{aligned}$$

where $H_n(\cdot)$ denotes the n-th order Hermite polynomial. In Eq.(2), it should be noticed that each distribution parameter is generally a function of x.

On the other hand, recently, multilayer neural networks are very often used in many kinds of engineering fields, especially for practical use. Though many kinds of learning methods for the networks are also proposed, most of them are fundamentally based on the least squares error criteria (e.g., well-known back-propagation learning). This means that the estimate from learning can be resultantly given in a conditional statistics. So, the multilayer neural network (perceptron) can be adopted to express each distribution parameter $\mu_y(x)$,$\sigma_y^2(x)$ and $a_n(x)$ in Eq.(2), as shown Fig.2.

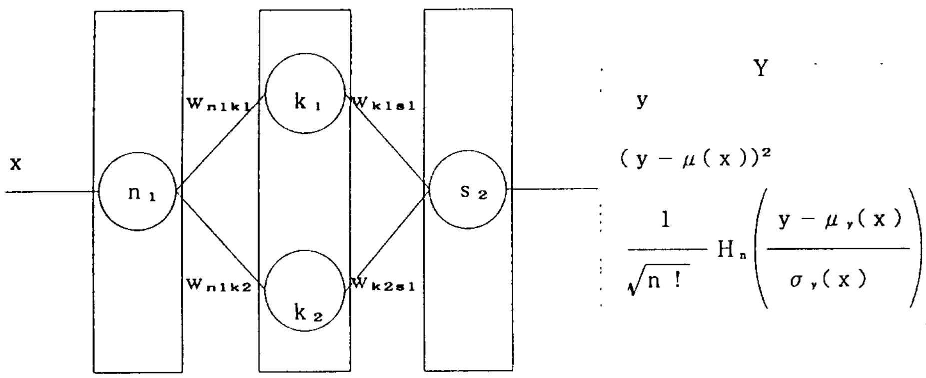


Fig.2 Multilayer neural network to express the conditional statistics.

A multilayer perceptron is a feed forward neural network that has one or more layers of hidden neurons between the input and output layers. Several results[9, 10] have shown that two-layer (one-hidden layer) perceptron with sigmoidal nodes can in principle represent any continuous function. The training of a multilayer perceptron very often uses a back-propagation learning for practical use. Learnings of $\mu_y(x)$, $\sigma_y^2(x)$ and $a_n(x)$ require training sets (x_k, y_k), $(x_k, (y_k - \mu_y(x_k))^2)$ and $(x_k, \frac{1}{\sqrt{n!}}H_n(\frac{y_k-\mu_y(x_k)}{\sigma_y(x_k)}))$, respectively. Concludingly, the generalized capability makes it possible to learn the complicated nonlinear relationship between x and functions of y.

3.2 Parzen Window (Method 2)

In order to grasp a complicated non-linear relationship between the input x and output y, a neural network of layered structure can be introduced and the back propagation algorithm is used for training the networks. It is pointed out that it takes a large number of iterations to converge to the desired solution . This section introduces the practical probability density function of input-output fluctuations for predicting the output response probability distribution $p(y)$ by use of the input data x.

First, the joint probability density function $p(x, y)$ must be estimated from samples of simultaneous observation set of x and y. Here, the estimation for $p(x, y)$ is based on sample value x_i and y_i of the random variables x and y first by introducing so-called Parzen window[11] especially for practical use as follows:

$$p(x,y) = \frac{1}{2\pi\sigma^2}\frac{1}{n}\sum_{i=1}^{n}\exp\left\{-\frac{(x-x_i)^2}{2\sigma^2}\right\}\exp\left\{-\frac{(y-y_i)^2}{2\sigma^2}\right\}, \tag{3}$$

where n and σ denote the number of samples and a constant, respectively.

Equation (3) yields the desired conditional probability distribution as:

$$p(y|x) = \frac{p(x,y)}{\int_{-\infty}^{\infty} p(x,y)dy} = \frac{1}{\sqrt{2\pi}\sigma} \frac{\sum_{i=1}^{n} \exp\left\{-\frac{(x-x_i)^2}{2\sigma^2}\right\} \exp\left\{-\frac{(y-y_i)^2}{2\sigma^2}\right\}}{\sum_{i=1}^{n} \exp\left\{-\frac{(x-x_i)^2}{2\sigma^2}\right\}}. \tag{4}$$

It should be noticed that only n sample values x_i and y_i lead to the complicated regression relationship between x and y . This conditional probability density function can be applied to the actual situation with the non-Gaussian and non-linear properties.

4. PREDICTION METHOD FOR OUTPUT RESPONSE PROBABILITY DISTRIBUTION

By use of the leaning result of perceptrons for each conditional statistics in Eq.(2) or the estimated result of conditional probability density function in Eq.(4), the objective output response probability $p_y(y)$ and its cumulative distribution function can be predicted as:

$$p_y(y) = < p(y|x) >_{x'}, \tag{5}$$

$$Q_y(y) = \int_{-\infty}^{y} p_y(\xi)d\xi = < p(y|x) >_{x'} . \tag{6}$$

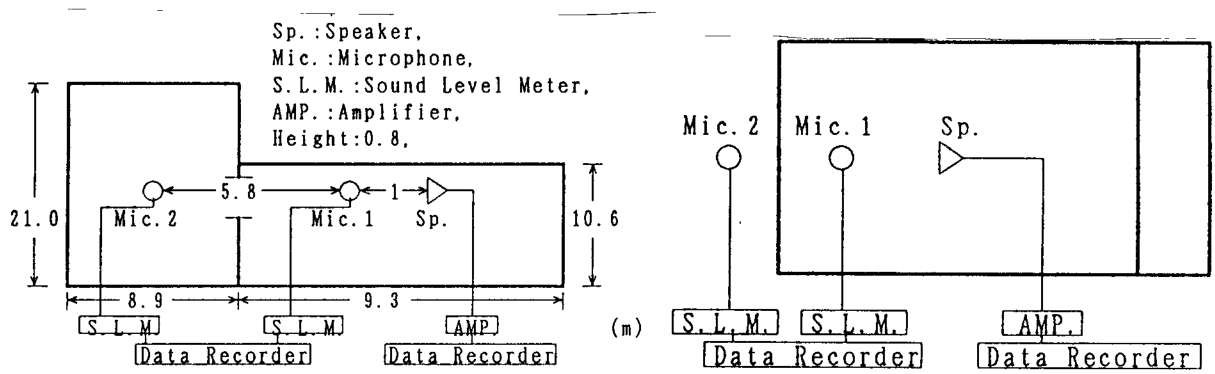
Concludingly, the output probability distribution for the sound environmental systems with the other arbitrary input signal can be predicted. So, we can construct the cumulative probability distribution directly connected with an evaluation index Lx.

5. APPLICATION TO ACTUAL SOUND ENVIRONMENTS

The proposed method has been experimentally confirmed by applying it to the complicated indoor and indoor-outdoor systems. Here, our method has been employed to estimate the system characteristics (i.e., input-output relationship) by using the input and output data. Then, the output probability distribution for the input in a different time interval from that for estimation process based on the estimated relationship has been predicted.

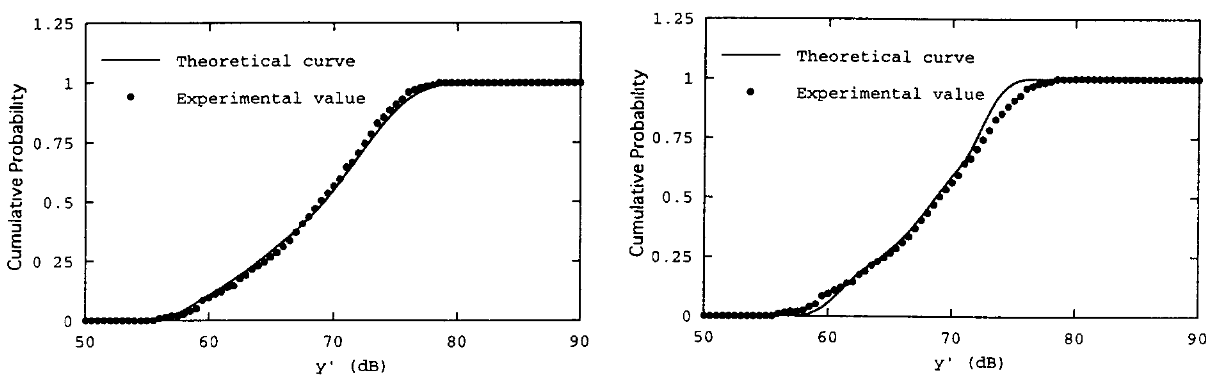
First, we have taken two laboratory rooms shown in Fig.3(a) as the actual case of the indoor system. The actual traffic noise radiated from the loud speaker has been used as an input sound source. The observed level data have been measured through the microphones 1 and 2 as an input $x(t)$ and an output $y(t)$, respectively. 3,000 of input-output data have been synchronously sampled with the sampling interval of 1 sec. For Method 1, the conditional statistics have been learned by using the first 2,000 of input and output data. For Method 2, the number of samples has been set as 25 of input-output data. Then, the latter 1,000 of input-output data have been divided into 5 parts and the output probability distribution has been predicted.

Figure 4 shows a comparison of the experimentally sampled point and the theoretically predicted curve for the output probability distribution by using the estimated input-output relationship and the latter 200 of input data. From this result, it is obvious that the theoretical curve agrees well with the experimental values.



(a) (b)

Fig.3 A block diagram of the experimental arrangement in (a) two laboratory rooms and (b) indoor-outdoor system.

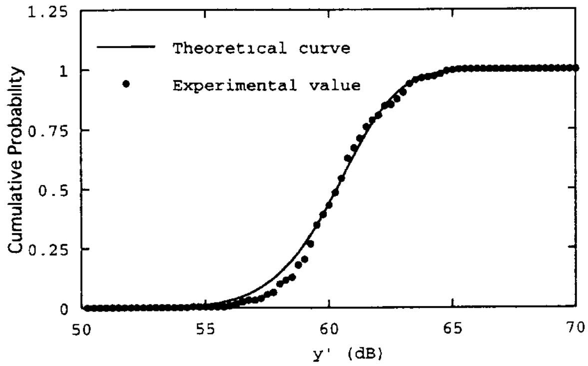


(a) (b)

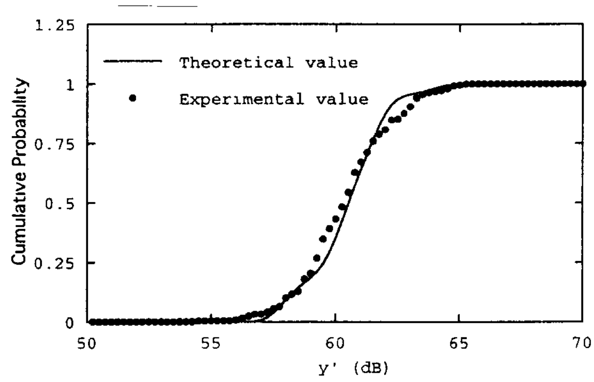
Fig.4 A comparison between cumulative probability distributions of experimentally sampled points and theoretically predicted curve by use of the proposed evaluation methods ((a) Method 1 and (b) Method 2) for the indoor sound system with the road traffic noise input.

Next, we have taken the actual building shown in Fig.3(b) as the indoor-outdoor system. The rock music radiated from the loud speaker has been used as an input sound source. The observed level data have been measured through the microphones 1 and 2 as an input $x(t)$ and an output $y(t)$, respectively. 500 of input-output data have been synchronously sampled with the sampling interval of 1 sec. In this case, the system has been identified by using the first 250 for Method 1. The number of samples has been set as 25 of input-output data for Method 2. Then, the output probability distribution has been predicted by using the latter 250 of input data.

Figure 5 shows a comparison of the experimentally sampled point and the theoretically predicted curve for the output probability distribution by using the estimated input-output relationship and the latter 250 of input data. The theoretical curve catches successfully the whole shape of the experimentally obtained cumulative distribution function.



(a)



(b)

Fig.5 A comparison between cumulative probability distributions of experimentally sampled points and theoretically predicted curve by use of the proposed evaluation methods ((a) Method 1 and (b) Method 2) for the indoor-outdoor sound system with the rock music input.

6. CONCLUDING REMARKS

In this paper, a regression analysis especially matched to the prediction of the output response probability of complicated actual sound environmental systems has been proposed by introducing a hierarchical neural network. More concretely, in the first method, since the regression analysis is based on the conditional probability distribution, the conditional statistics on its distribution have been constructed by use of multilayer neural networks. In the second method, so-called Parzen-window has been introduced as the joint probability density function of input-output fluctuations because of its simple form and the conditional probability density function with respect to the output response has been estimated on the basis of the Bayes' theorem. Then, based on the estimated result of sound environmental systems, the output probability has been easily predicted for the same system with arbitrary input signal. Finally, the effectiveness of the present method has been experimentally confirmed by applying it to the prediction problem of output response in the actual sound environments.

Needless to say, since this research is at an earlier stage of study for the regression analysis and probabilistic response prediction by introducing the multilayer networks or the practical probability distribution, there remain many future problems to be solved such as;
1) to apply the proposed method to various kinds of data in many other actual fields,
2) to investigate theoretically the proposed method in more detail and find more simplified practical method through the approximation of this method, especially in comparison with the traditional regression analysis method,
3) to find appropriate ways to determine the optimal order of series expansion in Method 1 and to determine the optimal parameters n and σ in Method 2,
4) to reflect the actual structure of the sound environment (e.g., memory effect of the system, etc.) into the theory,
and so on.

ACKNOWLEDGMENTS

We would like to express our cordial thanks to Mr. T. Matsumoto and Mr. T. Sera for their helpful assistance.

References

[1] A. London: " Transmission of Reverberant Sound through Double Walls", J. Acoust. Soc. Am., **22**, pp. 270-279 (1950).

[2] M. J. Crocker and A. J. Price: " Sound Transmission Using Statistical Energy Analysis", J. Sound Vib. **9**, 469-486 (1969).

[3] Kuttruff: "Room Acoustics (2nd Ed.)", Applied Science Publisher, London (1970).

[4] N. R. Draper and H. Smith: " Applied Regression Analysis", John Wiley & Sons, New York (1981).

[5] F. Acton: " Analysis of Straight-Line Data", John Wiley & Sons, New York (1959) .

[6] H. Cramer: "Mathematical Methods of Statistics", Princeton Univ. Press., Princeton (1974).

[7] M. Ohta, B. Chang and Y. Xiao: " Regression Analysis with Less Information Loss under the Existence Background Noise", IEICE Trans. on Fundamentals, **E72**, pp. 531-538 (1989).

[8] M. Ohta and T. Koizumi: " General Statistical Treatment of a Non-linear Rectifying Device to a Stationary Random Input", IEEE Trans. Inf. Theory, **IT-14**, pp. 595-598 (1968).

[9] R. Hecht-Nielsen: "Theory of Back-propagation Neural Networks", Proc. Int'l Joint Conf. on Neural Networks, pp. 593-608 (1989).

[10] K. Homik: "Multilayer Feedforward Networks Are Universal Approximators", Neural Networks, **2**, pp. 359-366 (1989).

[11] E. Specht: "On Estimation of a Probability Density Function and Mode", Ann. Inst. Statist., **33**, pp. 1065-1076 (1962).