OPTICAL SENSING OF PARTICLE SIZE DISTRIBUTION BY NEURAL NETWORK TECHNIQUE

Akira Ishimaru, Robert J. Marks, II, Leung Tsang, Chi M. Lam, and Dong C. Park
Department of Electrical Engineering
University of Washington
Seattle, Washington 98195

Shinzo Kitamura
Department of Instrumentation Engineering
Kobe University
Kobe 657, Japan

ABSTRACT

We present an inverse technique to determine particle size distributions by training a layered perceptron neural network with optical backscattering measurements at three wavelengths. An advantage of this approach is that, once the neural network is trained, the inverse problem of obtaining size distributions can be solved speedily and efficiently.

INTRODUCTION

We consider the inverse problem of finding the particle size distribution from the measurements of backscattered light on an optically thin medium containing particles. The first-order scattering approximation is used. The measured quantity is the backscattered intensity $\beta(\lambda_i)$ at three different wavelengths λ_i(i=1,2,3), and it is related to the size distribution function $n(r)$ by a Fredholm integral equation of the first kind as

$$\beta(\lambda_i) = \int_{r_1}^{r_2} K(\lambda_i, m, r) n(r) dr, \tag{1}$$

where m is the particle refractive index, r is the radius of particle and $K(\lambda_i, m, r)$ is the backscattering cross section [1]. We assume that the particles are spherical so that the backscattering cross section can be computed by the Mie solution. The inversion problem is to find the distribution, $n(r)$, from $\beta(\lambda_i)$ measurements.

Previously, methods such as smoothing, statistical and Backus-Gilbert inversion techniques have been used in finding profiles of particle distributions [2–5]. The smoothing technique requires a judicious choice of two parameters which control the smoothness of the solution. Statistical inversion technique requires the knowledge of statistical properties of the unknown function and the measurement errors. The Backus-Gilbert technique, however, requires a good compromise between the spread and the variance.

In this paper, we utilize a layered perceptron neural network to determine particle size distributions [6]. A multilayer perceptron neural network trained with the backpropagation algorithm is used. A major advantage of this approach is that, once the neural network is trained, the inverse problem of estimating size distributions can be performed speedily and efficiently.

MULTI-LAYER PERCEPTRON TYPE ARTIFICIAL NEURAL NETWORK

An artificial neural network can be defined as a highly connected array of elementary processors called neurons. In this paper, we consider the multi-layer perceptron (MLP) type artificial neural network [7–9].

As shown in Fig. 1, the MLP type neural network consists of one input layer, one or more hidden layers and one output layer. Each layer employs several neurons and each neuron in the same layer is connected to the neurons in the adjacent layer with different weights. A schematic diagram of this model is depicted in Fig. 1. We use 3 inputs $(\beta(\lambda_1), \beta(\lambda_2), \beta(\lambda_3))$ and 2 outputs (r_m, σ) neurons. Signals pass from the input layer, through the hidden layers, to the output layer. Except for the input layer, each neuron receives a signal which is a linearly weighted sum of all the outputs from the neurons of the former layer. The neuron then produces its output signal by passing the summed signal through the sigmoid function $1/(1+e^{-x})$.

The backpropagation learning algorithm is employed for training the neural network. Basically this algorithm uses the gradient descent algorithm to get the best estimates of the interconnected weights, and to make the output of the network as close to the desired value as possible for the given input. More detailed descriptions on the backpropagation algorithm can be found in [10] and [11].

LOG-NORMAL SIZE DISTRIBUTION

We consider the backscattering of light from a volume distribution of spherical particles with 31 radii ranging from 0.01 to 40μm. We assume that the size distribution function $n(r)$ is governed by the log-normal function so that it is characterized by two quantities: the mean radius r_m and the standard deviation σ. Therefore, it is given by

$$\mathrm{n(r)r} = \frac{\mathrm{dN(r)}}{\mathrm{d}\log \mathrm{r}} = \frac{\mathrm{N}}{\sqrt{2\pi}\log(\sigma)} \exp\left\{-\frac{[\log(\mathrm{r}) - \log(\mathrm{r_m})]^2}{2\,[\log(\sigma)]^2}\right\}. \tag{2}$$

CH2825-8/90/0000-02129/$01.00

Thus the relation between $\beta(\lambda_i)$ and the parameters r_m and σ is nonlinear. The inverse problem becomes that of finding the output r_m and σ for given input $\beta(\lambda_i)$.

We first conduct a study of forward problem of finding $\beta(\lambda_i)$ for various r_m and σ. Since the radius of particles varies from 0.01 to 40μm, i.e., $-2 \leq \log(r) \leq 1.66$, the ranges for r_m and σ are chosen such that $-1 \leq \log(r_m) \leq 0.44$, and $0.03 \leq \log(\sigma) \leq 1$. Thus the actual size of particles ranging from (r_m/σ) to $r_m\sigma$ will be within the range for r. Both $\log(r_m)$ and $\log(\sigma)$ are divided into 10 intervals for generating the training and testing data. We chose the refractive index of the particle to be $m = 1.53 - j0.008$ and, the wavelengths to be $\lambda_1 = 0.53\mu$m, $\lambda_2 = 1.06\mu$m and $\lambda_3 = 2.12\mu$m. The study of $\beta(\lambda_i)$ reveals that for some values of $\log(r_m)$ and $\log(\sigma)$, they are close to each other. This may create a problem of getting nonunique solution for r_m and σ with such $\beta(\lambda_i)$. In order to obtain unique solution of r_m and σ for given $\beta(\lambda_i)$, the change in $\beta(\lambda_i)$ for given change of r_m and σ must be sufficiently large. Therefore we define the distance D, a measure of separation of $\beta(\lambda_i)$, as

$$D = \sqrt{\sum_{i=1}^{3} [\beta(\lambda_i, \sigma_j, r_{ml}) - \beta(\lambda_i, \sigma_k, r_{mn})]^2}. \qquad (3)$$

Here we have divided $\log(r_m)$ and $\log(\sigma)$ into a number of intervals such that

$$0.03 = \log(\sigma_1) < \log(\sigma_2) < \ldots < \log(\sigma_M) = 1,$$

and

$$-1 = \log(r_{m1}) < \log(r_{m2}) < \ldots < \log(r_{mN}) = 0.44.$$

In order to ensure the $\beta(\lambda_i)$'s are sufficiently separated, we require that D exceeds a minimum distance D_m. To find D_m, we first notice that there is a large difference in magnitude between $\beta(\lambda_i, \sigma_j, r_{ml})$ and $\beta(\lambda_i, \sigma_k, r_{ml})$ for $k > j$. For instance we have $\beta(\lambda_i, \sigma_1, r_{m1}) \sim 10^{-15}$ and $\beta(\lambda_i, \sigma_M, r_{m1}) \sim 10^{-6}$. Thus D_m cannot be fixed for all σ_j but should vary according to σ_j. In addition, for the same σ_j, the value for $\beta(\lambda_i, \sigma_j, r_{ml})$ increases from $l = 1$ to $l = N$. The lowest value occurs when $l = 1$. Hence, the minimum distance D_m is chosen proportional to $\beta(\lambda_i)$ obtained from the first mean radius r_{m1}. Specifically

$$D_m = D_1\sqrt{\beta^2(\lambda_1, \sigma_j, r_{m1}) + \beta^2(\lambda_2, \sigma_j, r_{m1}) + \beta^2(\lambda_3, \sigma_j, r_{m1})} \quad (4)$$

where D_1 is a constant. Thus D_m is a fixed quantity when D_1 and σ_j are fixed. Therefore, we can determine the allowable range of $\log(r_m)$ for that particular $\log(\sigma_j)$, the lower and upper bounds of $\log(r_m)$, by enforcing the requirement of $D \geq D_m$. Similary for each $\log(\sigma_j), j = 1, 2, \ldots$, we compute the corresponding allowable ranges of $\log(r_m)$. From the diagram of all the allowable ranges for $\log(r_m)$, we can estimate the desired region for $\log(r_m)$ and $\log(\sigma)$.

RESULTS AND DISCUSSIONS

The constant D_1 in (4) controls the size of the allowable region for $\log(r_m)$ and $\log(\sigma)$. A large value of D_1 will generally create a small allowable region, but the values of $\beta(\lambda_i)$ are reasonably separated and therefore unique sets of $\beta(\lambda_i)$ can be obtained. On the other hand, a small value of D_1 will create a large allowable region, but the sets of $\beta(\lambda_i)$ are close to each other. Unique sets of $\beta(\lambda_i)$ are thus difficult to obtain resulting in a large percentage of error in obtaining the unknown size distribution. A value of D_1 ranging from 0.1 to 50 has been tested in finding the suitable D_1. It is found that a value of 10 for D_1 is a good compromise between the percentage error and the size of the allowable region for $\log(r_m)$ and $\log(\sigma)$. With such a value, the allowable region is found to be $-0.328 \leq \log(r_m) \leq 0.44$ and $0.03 \leq \log(\sigma) \leq 0.5$ as shown in Fig. 2.

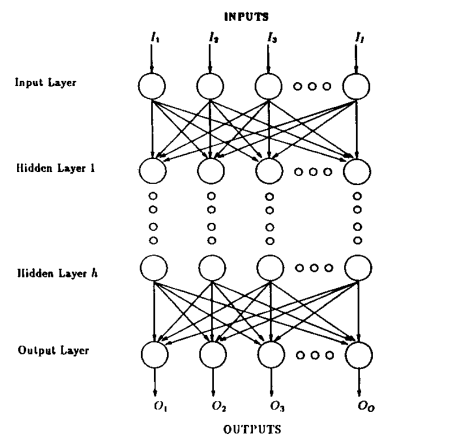

Figure 1 Structure of a multi-layered perceptron type artificial neural network.

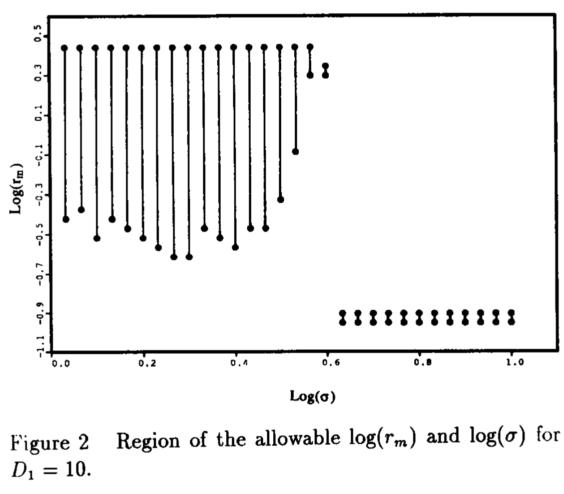

Figure 2 Region of the allowable $\log(r_m)$ and $\log(\sigma)$ for $D_1 = 10$.

Based on the allowable region discussed above, a group of 480 sets of data was generated from (1). In order to maximize the computing accuracy of the neural network, all the data are first normalized from zero to unity. We use 462 data sets to trained the neural network. The remaining 18 sets are used to test the system. Finally, as shown in Figs. 3 to 6, the results are converted back to the original values.

Figures 3 and 5 show the performance of the neural network in obtaining the size distribution function $n(r)$ by computing the desired output $\log(r_m)$ and $\log(\sigma)$ for each testing data set. The solid line is the line of the true value that the computed results should be as close to it as possible. In the process of training the neural network, different number of iterations are used. It is shown in Figs. 3 and 5 that increasing number of iterations tends to converge to the true values. Figures 4 and 6 show the performance of neural network in terms of absolute percentage error for $\log(r_m)$ and $\log(\sigma)$, respectively. Again, it is clear from Figs. 4 and 6 that increasing number of iterations tends to converge to the real value and hence lowers the absolute percentage errors. Except that the desired output $\log(r_m)$ and $\log(\sigma)$ are small, the neural network yields good results for most of the testing data with the absolute percentage of error less than 10%.

CONCLUSION

In this paper, we present an inverse technique of finding particle size distribution by using neural network. Size distribution function is assumed to be a log-normal function so that it is characterized by the mean radius r_m and standard deviation σ. We first train the neural network by inputting the backscattered intensities $\beta(\lambda_i), i = 1,2,3$, at three wavelengths with known particle refractive index, r_m and σ. A group of 462 data sets is used to train the neural network. Another group of 18 data sets is used to test the neural network in obtaining the desired output r_m and σ. It is shown that the neural network yields good results for the testing data with absolute percentage of errors less than 10% for most of the testing input $\beta(\lambda_i)$. A major advantage of this technique is that, once the neural network is trained, the inverse problem of obtaining the size distributions can be solved speedily and efficiently.

ACKNOWLEDGEMENT

This work is partially supported by the Grant-in-Aid for Scientific Research by the Ministry of Education, Science, and Culture of Japan (Monbusho), NSF and NASA.

REFERENCE

[1] A. Ishimaru, Wave Propagation and Scattering in Random Media, vol. 2, Academic Press, 1978.

[2] D.L. Phillips, "A Technique for the Numerical Solution of Certain Integral Equations of the First Kind," JACM 9, pp.84-97, 1962.

[3] P. Edenhofer, J.N. Franklin and C.H. Papas, "A New Inversion Method in Electromagnetic Wave Propagation," IEEE Trans. Antennas Propagat. AP-21, pp.260-263, 1973.

[4] G. Backus and F. Gilbert, "Uniqueness in the Inversion of Inaccurate Gross Earth Data," Phil. Trans. Roy. Soc. London Ser. A266, pp.123- 192, 1970.

[5] E.R. Westwater and A. Cohen, "Application of Backus-Gilbert Inversion Technique in Determination of Aerosol Size Distributions from Optical Scattering Measurements," Appl. Opt. 12, pp.1340-1348, 1973.

[6] S. Kitamura and P. Qing, "Neural Network Application to Solve Fredholm Integral Equations of the First Kind," International Joint Conference on Neural Network, Washington D.C., June 1989.

[7] M. El-Sharkawi, R. Marks II, M.E. Aggoune, D.C. Park, M.J. Damborg and L. Atlas, "Dynamic Security Assessment of Power Systems Using Back Error Propagation Artifical Neural Network," Proceedings of the 2nd Annual Symposium on Expert System Application to Power Systems, Seattle, WA, pp.366-370, July, 1987.

[8] L. Atlas, R. Cole, Y. Muthusamy, A. Lippman, G. Connor, D. Park, M. El-Sharkawi and R. Marks II, "A Performance Comparison of Trained Multi-layer Perceptron and Trained Classification Trees," submitted to Proceedings of IEEE Special Issue on Neural Networks, Aug. 1990.

[9] D. Park, M. El-Sharkawi, R. Marks II, L. Atlas and M. Damborg, "Electrical Load Forecasting Using an Artifical Neural Network," submitted to IEEE Power Engineering Society Winter Meeting, 1990.

[10] D.E. Rumelhart, G.E. Hinton and R.J. Williams, Learning Internal Representation by Error Propagation, in D.E. Rumelhart and J.L. McClelland (Eds.), Parallel Distributed Processing, Cambridge, MA, MIT Press, 1986.

[11] R.P. Lippmann, "Introduction to Computing with Neural Nets," IEEE ASSP Magazine, pp.4-22, April 1987.

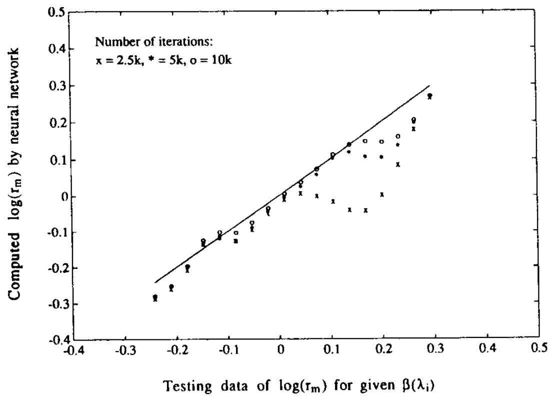

Figure 3 Performance of the neural network in generating the mean particle size, $\log(r_m)$, from the given backscattered intensities, $\beta(\lambda_i)$. The solid line is the line of true value that the computed results should be as close to it as possible. A different number of iterations was used to train the neural network.

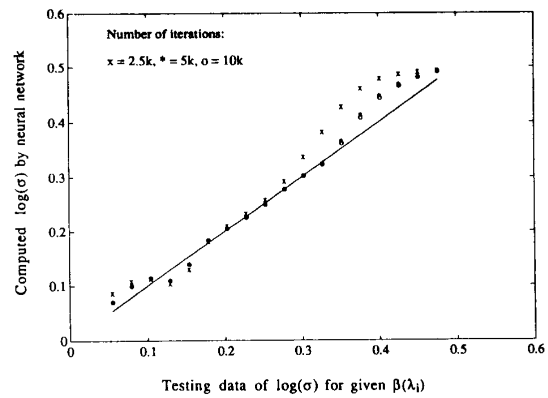

Figure 5 Performance of the neural network in generating the standard deviation of particle size, $\log(\sigma)$, from the given backscattered intensities, $\beta(\lambda_i)$. The solid line is the line of true value that the computed results should be as close to it as possible. A different number of iterations was used to train the neural network.

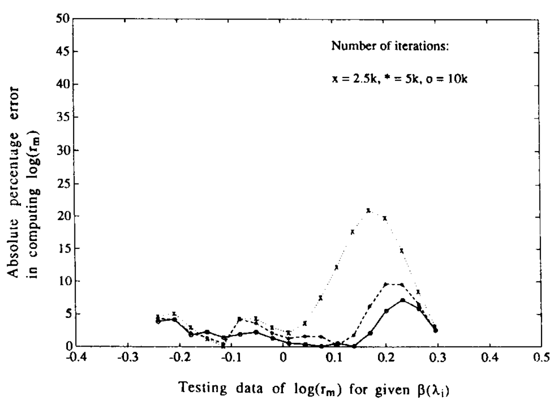

Figure 4 Performance of the neural network in generating the mean particle size, $\log(r_m)$, from the given backscattered intensities, $\beta(\lambda_i)$, in terms of absolute percentage error. Increasing number of iterations tends to converge to the true value and hence lowers the absolute percentage error.

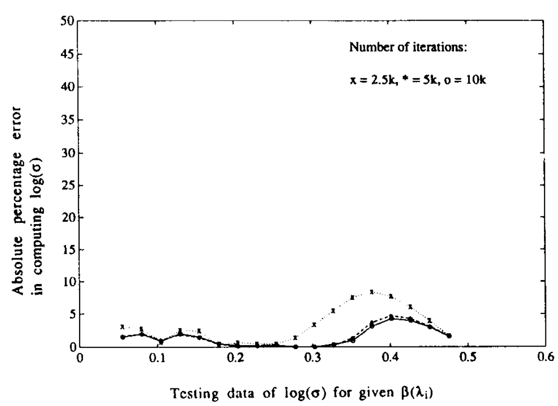

Figure 6 Performance of the neural network in generating the standard deviation of particle size, $\log(\sigma)$, from the given backscattered intensities, $\beta(\lambda_i)$, in terms of absolute percentage error. Increasing number of iterations tends to converge to the true value and hence lowers the absolute percentage error.