Artificial Neural Networks to Extract Optical Properties of Marine Microorganisms From Their Mueller Scattering Matrix

Patricia G. Hull
Tennessee State University
Department of Physics and Mathematics
3500 John Merritt Blvd.
Nashville, TN 37209-1561
Phone/FAX: 615-963-5846 e-mail address: phull@harpo.tnstate.edu
Award No: N00014-96-1-0307

LONG-TERM GOAL

The long term goals of this project are to understand and quantify light scattering from ensembles of both spherical and non-spherical objects in ocean water, to characterize the effect of ensembles of micro-organisms and inorganic particulates on the propagation of polarized light through sea water, and to assess the feasibility of computer simulated artificial neural network to extract optical properties of marine particulates from polarized light scattering measurements.

SCIENTIFIC OBJECTIVES

The scientific objectives are to develop a numerical or analytical model that predicts angle-dependent scattering of polarized light from ensembles of non-spherical marine organisms, detritus, and inorganic particulates, and to verify and examine the validity and range of applications of the model by comparison with exact calculations and/or experimental results as appropriate. Specific tasks toward these objectives are:

(1) to develop an artificial neural network to recognize features in the Mueller matrix elements associated with the optical properties, size distribution, and irregular shape of ocean scatterers,

(2) to make experimental measurements in the laboratory of light scattering from samples of micro-organisms and inorganic particles in ocean water, and

(3) to continue to refine and enhance analytical models such as the coupled-dipole method for predicting polarized light scattering from non-spherical particles.

APPROACH

Experimental measurements and mathematical modeling continue to be important, however, much of the current work is concentrated on an inverse problem. That is, given the values of the Mueller matrix elements as a function of scattering angle, what are the optical properties and size distribution of the particles that scatter the light? The complexity and nature of this task suggests the use of artificial neural networks, computer systems comprised of a number of simple, interconnected processing elements, called neurons or nodes, operating in parallel.

Analytical calculations of the scattering matrix elements rather than experimental data were used in order to provide well-characterized training data for the neural networks. Mie calculations were made for a log-normal distribution of sphere sizes, a physically realistic distribution for marine microorganisms. The equation below shows the equation of the log-normal distribution used in the

Report Documentation Page

Form Approved
OMB No. 0704-0188

1. REPORT DATE	2. REPORT TYPE	3. DATES COVERED
1998		**00-00-1998 to 00-00-1998**

4. TITLE AND SUBTITLE	5a. CONTRACT NUMBER
Artificial Neural Networks to Extract Optical Properties of Marine Microorganisms From Their Mueller Scattering Matrix	5b. GRANT NUMBER
	5c. PROGRAM ELEMENT NUMBER
6. AUTHOR(S)	5d. PROJECT NUMBER
	5e. TASK NUMBER
	5f. WORK UNIT NUMBER
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) **Tennessee State University,Department of Physics and Mathematics,3500 John Merritt Blvd,Nashville,TN,37209-1561**	8. PERFORMING ORGANIZATION REPORT NUMBER
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)	10. SPONSOR/MONITOR'S ACRONYM(S)
	11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for public release; distribution unlimited

13. SUPPLEMENTARY NOTES
See also ADM002252.

14. ABSTRACT

15. SUBJECT TERMS

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT	b. ABSTRACT	c. THIS PAGE			
unclassified	**unclassified**	**unclassified**	**Same as Report (SAR)**	**7**	

calculations. In this equation, x is the particle size parameter and $N(x)$ is the fraction of particles having a size parameter, x. The quantity x_m is the average value of the particle size parameter, x_p is the value of the size parameter that has maximum probability of occurrence, and σ_g is the standard deviation of log$x$.

$$N(x) = \frac{1}{x\sigma_g^2\sqrt{2\pi}} \exp\left[\frac{-(\log x - \log x_m)^2}{2\sigma_g^2}\right]$$

where,

$$\log x_p = \log x_m - \frac{\sigma_g^2}{2.302585}$$

Four values were needed to make a light scattering calculation, 1) the size parameter at which the log-normal distribution has its maximum value (SP), 2) the standard deviation from the mean for the log of the size parameter, sigma (SG), 3) the real part of the complex index of refraction (NX), and 4) the imaginary part of the complex index of refraction (the absorption coefficient), (AB). It is these four parameters, SP, SG, NX, and AB, that we want the neural network to retrieve from the light scattering calculations. The size parameter is defined here as $2\pi r/\lambda$ where r is the radius of the spherical particle and λ is the wavelength of the incident light in the medium. For example, a size parameter of 10 corresponds to a particle diameter of slightly less than two microns for incident green light. The complex index of refraction is written as $m = n + i\kappa$, where n is the relative index of refraction and κ is the absorption coefficient.

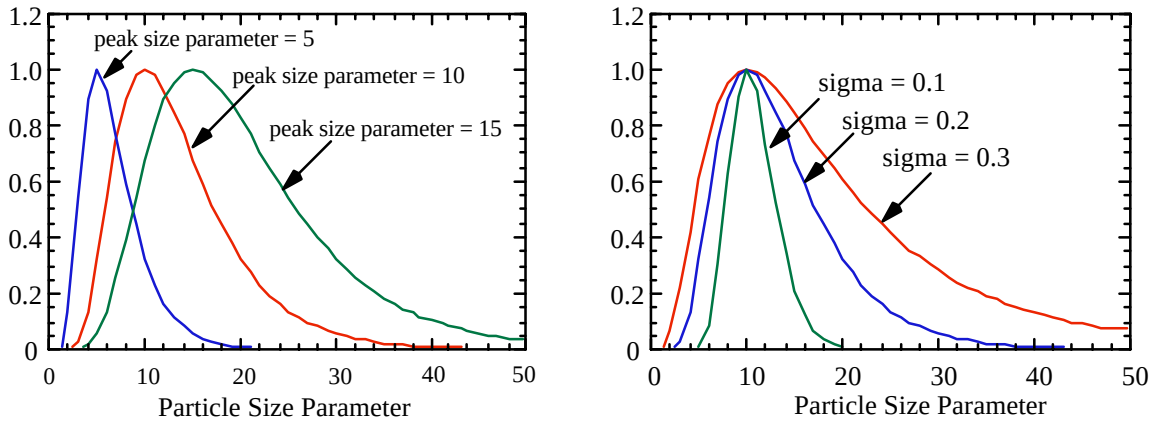

Figure 1. Log normal distribution functions used in Mie calculations. Graph on the left shows peak values in size parameter of 5, 10 and 15 each with the same deviation, sigma of 0.20. The graph on the right shows three different values of sigma, 0.10, 0.20 and 0.30, each with a peak size parameter of 10.

In order to construct networks that could be evaluated for different learning strategies and error determination methods, it was important to keep the number of input data points as small as possible so that the network training could be carried out on a desktop computer. Experience has shown that at least 3 or 4 processing elements for each input data point is required for a network to have sufficient

power to solve a problem of the type of interest here. The four independent Mueller matrix elements calculated from Mie theory, S_{11}, S_{12}, S_{33}, and S_{34} each have a period of 2π and are even functions of the scattering angle. The number of input data points can be reduced if the functional form of each of these elements is described by a simple Fourier series of cosine terms such as those as shown below for S_{34}.

$$S_{34}(\theta) = c_o + \sum_{k=1}^{\infty} c_k \cos(k\theta)$$

where,

$$c_o = \frac{1}{\pi}\int_o^{\pi} S_{34}(\theta)\,d\theta \qquad \text{and} \qquad c_k = \frac{2}{\pi}\int_o^{\pi} S_{34}(\theta)\cos(k\theta)\,d\theta$$

As much useful information about the functional form of each matrix element can be supplied to the network using 12 to 16 Fourier coefficients as 90 or more data points from the original graph. Training and testing sets for the networks were constructed from Mie calculations by varying the peak size parameter from 1.0 to 20 in steps of 1.0, the deviation from 0.10 to 0.30 in steps of 0.05, the relative index of refraction from 1.02 to 1.12 in steps of 0.02, and the absorption coefficient from 0.10 to 0.20 in steps of 0.02. Calculations for all 3600 possible combinations of the four parameters were not necessary. Instead, about six hundred Mie calculations were made using randomly selected combinations of the four parameters, SP, SG, NX and AB. A set of 16 Fourier coefficients for each of the scattering matrix elements was generated at the end of each Mie calculation although only those for S_{12} and S_{34} were used in the current work. Three vectors (column matrices) were constructed for each calculation, 1) a vector of the 16 coefficients of S_{34}, 2) a vector of the first 8 coefficients of S_{12} then the first eight coefficients of S_{34} for a total of 16, and 3) a vector of the first 8 elements of S_{34}. A training or testing set consisted of a matrix made up of many of these vectors. The resulting large matrices were used either to train the networks to the desired value of the sum-squared error or to test a network with vectors not used in the training. Back propagation with momentum, a slow method requiring a long calculation time but very little computer memory was used for training larger networks. Levenburg-Marquardt optimization, a gradient descent method, was used to train the smaller networks. This method is extremely fast but it requires a lot of memory (RAM). It could only be used for small networks on the desktop computer.

Computations such as selection of initial weight matrices, summing weighted inputs (matrix inner product), calculations of the transfer functions, applying learning rules, and assessing the network's learning rate and performance were carried out using algorithms in the MATLAB library and its associated Neural Network Toolbox, products of the Math Works, inc.

WORK COMPLETED

We completed the design and initial training of a set interconnecting artificial neural networks that predict the peak size parameter in the range between 1.0 and 20, the deviation in size distribution for values between 0.10 and 0.30 (see Figure 1), relative index of refraction for values between 1.02 and 1.12, and absorption coefficient for values between 0.10 and 0.20, of a light scattering medium given its S_{12} and S_{34} matrix elements. The range for values for each of the four parameters was limited to order to train the networks with a desktop computer.

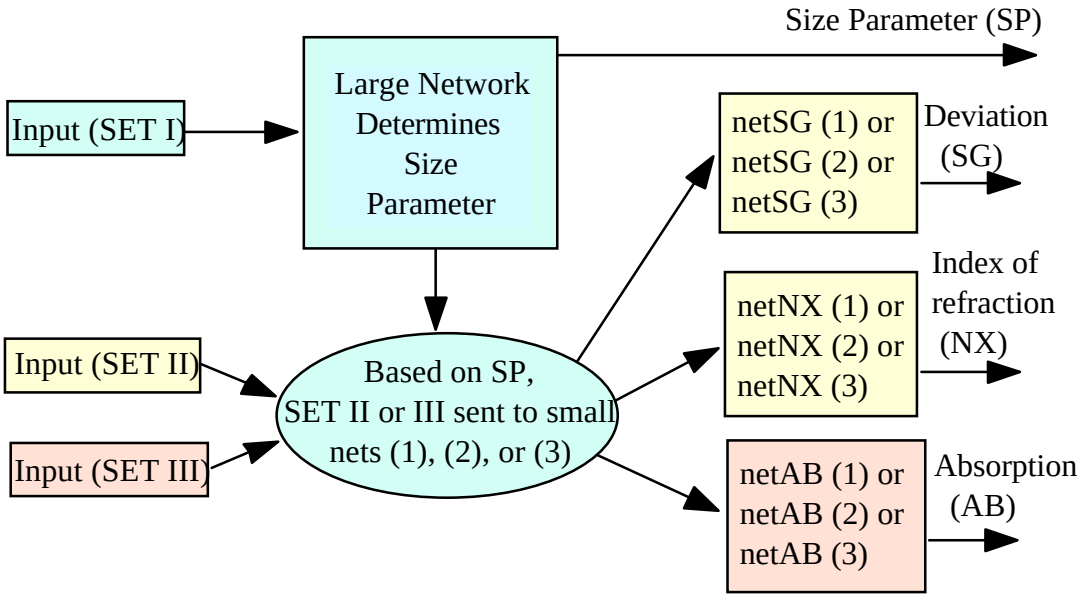

Figure 2. Diagram of network design. Input set I is sent to the large network to determine peak size parameter. Then, based on range of the output size parameter, set II is sent to one of the small networks in each of the two blocks 'deviation' and 'index of refraction' and set III is sent to one of the small networks in the block 'absorption' networks.

RESULTS

It was not possible using only a desktop computer to train a single network that could recognize all four of the desired parameters in one input data set. The large number of neurons necessary to provide a network with sufficient power for this task required made the training beyond the scope of our computer. Instead of one large network, we used several smaller ones, each trained to recognize only one of the four parameters. For example, netSP, a network of 80 neurons in the first hidden layer and 20 neurons in the second hidden layer was successful in predicting size parameters between 1.0 to 20.0. This network was trained to recognize SP using an input matrix of 240 of the 16-element vectors of Fourier coefficients of S_{34}. The strong effects of variations in size parameter tended to mask the more subtle changes in the Fourier coefficients of the matrix elements due to index of refraction, sigma, and absorption. Generally, networks for predicting these parameters were successful only for small ranges of size parameter. The networks for predicting SG, NX, and AB were subdivided into three separate networks, each trained for small ranges of size parameter. For example the networks called, netSG(1), netNX(1), and netAB(1) were trained using only data sets with size parameters between 1 and 5. The networks labeled netSG(2), netNX(2), and netAB(2) were trained for size parameters from 6 to 12 and those labeled netSG(3), netNX(3), and netAB(3) were trained for size parameters from 13 to 20. These nine networks each had 48 neurons in the first hidden layer and 12 neurons in the seconds hidden layer were small enough to train using Levenburg-Marquardt optimization. The networks netSG and netNX were trained to recognize SG and NX using an input matrix of 120 of the 16-element vectors of Fourier coefficients of both S_{34} and S_{12}. The networks netAB(1), (2) and (3) were trained using an input matrix 60 of the 8-element vectors of coefficients of S_{34}. Some of these networks reached the error goal in fewer than 30 iterations using this training method.

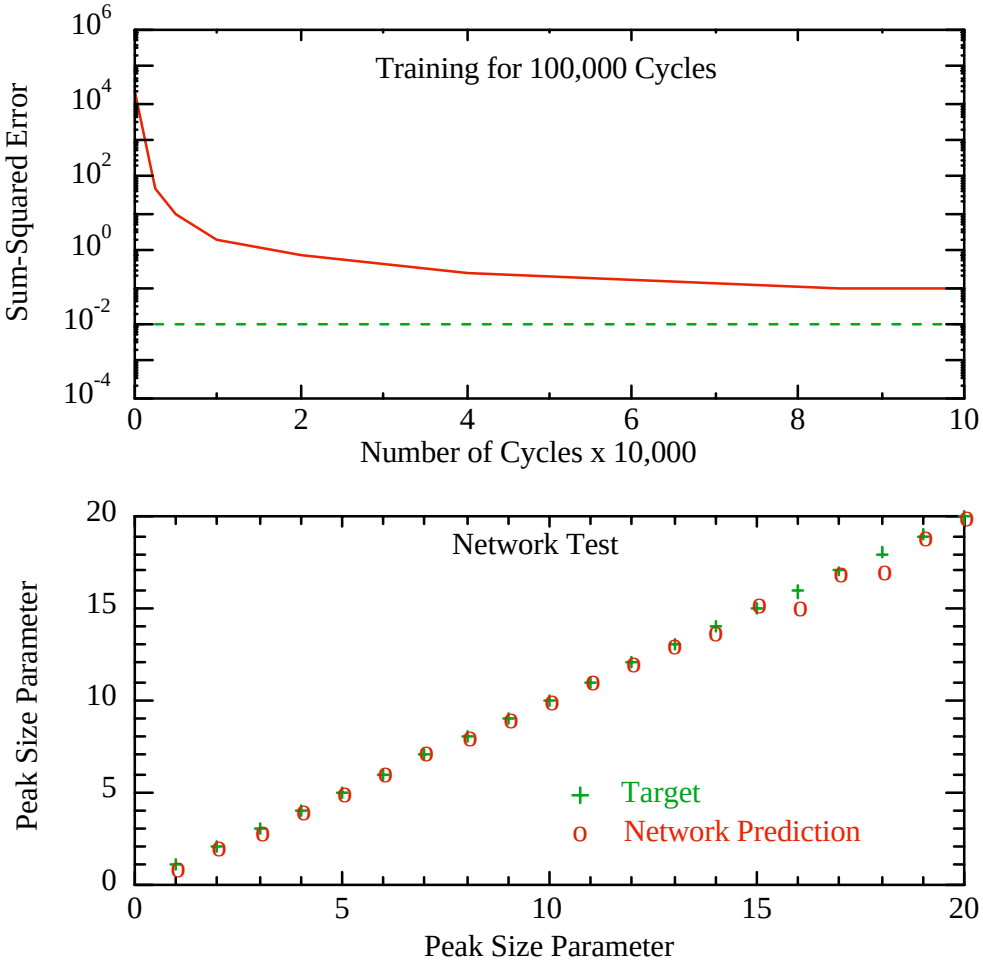

Figure 3. Training and testing of large network for peak size parameter (SP). The top graph shows the training for 100,000 cycles of back propagation with momentum for 240 training sets. The network failed to meet the error goal of 0.01. The graph below shows this network's prediction of the peak size parameter from data it had not seen in the training.

Figure 3 shows the results of a typical training session for the network, netSP. In this example, the network failed to met its error goal, arbitrarily set at 0.01, in the first 100,000 cycles. Although the network reached a sum-squared error of only 0.08 in this session, it nevertheless produced very good agreement in predicting the peak size parameter for a data set of twenty input vectors that it had not seen in the training. With additional training, the network reached its error goal, and difference between the target and network prediction was not detectable on a graph with the scale of Figure 3.

Figure 2 illustrates the procedure for obtaining the values of SP, NX, SG and AB from light scattering data using the group of networks netSP, netSG, netNX and netAB. An input vector of 16 Fourier coefficients of S_{34} (set I) is presented to the network, netSP which returns a value for the peak size parameter, SP, between 1 and 20. The value of the size parameter determines which small network

receives the second input set (set II), a vector of the Fourier coefficients of S_{12} and S_{34} and set III a vector of the first 8 coefficients of S_{34}. Set II is sent to both netNX and netSG, but the two groups of networks have been trained to recognize different features in the set. Set III is sent to netAB. For example, if netSP returns a value of 7, set II is sent to both netSG(2) and netNX(2), and set III is sent to netAB(2). The networks netSG(2), netNX(2), and netAB(2) return the values of standard deviation, index of refraction, and absorption, respectively.

IMPACT/APPLICATION

Currently, artificial neural networks are being applied to large classes of problems. Conventional computer techniques require detailed, explicit specification of the rules governing each application, must be virtually perfect in order to work, and require programming changes for every new or changed situation. This has, in the past, made computers of limited help in many challenging problems. Artificial neural networks do not require detailed, explicit specification of the rules governing each application, but rather they learn by example. They do not require perfection in order to give a good solution, and require no programming changes for every new or changed situation. One of the more successful training methods, Levenberg-Marquart optimization, resulted in very short training times, but the memory requirements for this method were often excessive for a desktop computer. This approach is promising, however, for training networks having a small number of neurons, or for training with more powerful mainframe computers or parallel processing machines. It is important to note that a mainframe computer or parallel processing machines may be necessary for training a network, using the trained network, even a very large one, is well within the power of a typical desktop computer.

TRANSITIONS

The following projects make use of computer software and/or experimental methods developed in this research for measuring and modeling the light scattering by irregularly shaped particles:

1 - The coupled-dipole approximation in which the anisotropy of hemoglobin is described by ellipsoidal polarizability tensors at each dipole site is being used in a study of sickle cell hemoglobin at Wake Forest University. NIH funds the study.

2 - A version of the coupled-dipole approximation is being used in a project funded by DOE at Berkeley Lab for modeling soot particles in diesel exhaust. The goal of that project is to develop an instrument, possibly using a neural network, for measuring sizes of the soot particles.

3 - In another DOE funded project at Berkeley Lab, experimental methods developed by this project are being used to help design an instrument for on-line measurement of the alignment of the fibers in paper production.

RELATED PROJECTS

This work was a component of past ONR-sponsored projects and is related to a current project by Hunt and Quinby-Hunt at Berkeley Lab for the measurement of light scattering in both seawater and the sea-air boundary layer, and calculations of light scattering from non-spherical aerosol particles using parallel processing.

REFERENCES

D.Kim-Shapiro, and P.Hull. 1997. Polarized light scattering from sickle hemoglobin polymers, J. Chem. Phys. August 1997.

B. Bronk, W.Van De Merwe, and M.Stanley. 1992: In vivo Measurement of Average Bacterial Cell Size from a Polarized Light Scattering Function, Cytometry, Vol.13, 155.

P. Smith, D.Green, and R.Naimimohasses. 1994: Some Investigations on neural processing have scattered light in water quality assessment. IEE Proc.-Vis. Image Signal Process., Vol. 141, No.4, 261.

P. Hull and M. Quinby-Hunt. 1996. "A neural network to extract size parameter from light scattering data." *Ocean Optics III*, Halifax Novia Scotia.

Z. Ulanowski, Z. Wang,, P. Kaye, and I. Ludlow. 1998. "Application of neural networks to the inverse light scattering problem for spheres." Applied Optics, Vol. 37, No. 18, 4027-4033.

M. McCord and W. Illiangworth. 1991. *A Practical Guide to Neural Nets*, Addison-Wesley Publishing Co., Reading, Mass.

R. Lippman. 1987. "An Introduction to Computing with Neural Nets", IEEE ASSP Magazine Vol.4 (2) 4.

R. Beale and T. Jackson. 1990. *Neural Computing: an Introduction.* Institute of Physics Publishing, Philadelphia, PA.

M. Hagan, H. Demuth, and M. Beale. 1996. *Neural Network Design.* PWS Publishing Co., Boston, MA.

H. Demuth and M. Beal. 1993. *Neural Network Toolbox User's Guide*, The Math Works, Inc., Natwick, Mass. MATLAB is a registered trademark of The Math Works, Inc.