Retrieving forest stand parameters from SAR backscatter data using a neural network trained by a canopy backscatter model

Y. Wang & D. Dong

whatsoever or howsoever caused arising directly or indirectly in connection with or arising out of the use of this material.
Retrieving forest stand parameters from SAR backscatter data using a neural network trained by a canopy backscatter model

Y. WANG
Department of Geography, East Carolina University, Greenville, NC 287588, U.S.A.

and D. DONG
Department of Mathematics, Henan University, Kaifeng, Henan 475001, China

(Received 24 January 1995; in final form 24 July 1995)

Abstract. It was possible to retrieve the stand mean dbh (tree trunk diameter at breast height) and stand density from the Jet Propulsion Laboratory (JPL) Airborne Synthetic Aperture Radar (AIRSAR) backscatter data by using three-layered perceptron neural networks (NNs). Two sets of NNs were trained by the Santa Barbara microwave canopy backscatter model. One set of the trained NNs was used to retrieve the stand mean dbh, and the other to retrieve the stand density. Each set of the NNs consisted of seven individual NNs for all possible combinations of one, two, and three radar wavelengths. Ground and multiple-wavelength AIRSAR backscatter data from two ponderosa pine forest stands near Mt. Shasta, California (U.S.A.) were used to evaluate the accuracy of the retrievals. The r.m.s. and relative errors of the retrieval for stand mean dbh were 6.1 cm and 15.6 per cent for one stand (St2), and 3.1 cm and 6.7 per cent for the other stand (St11). The r.m.s. and relative errors of the retrieval for stand density were 7.12 trees ha⁻¹ and 23.0 per cent for St2, and 49.7 trees ha⁻¹ and 21.3 per cent for St11.

1. Introduction

Distributions and rates of change of forests around the world are important in understanding phenomena such as the global carbon cycle, hydrologic cycle, and energy balance (Woodwell 1984). The demand on biotic resources by an increasing population and current deforestation around the world requires accurate and timely information about the distribution and the rate of change of forests. A method of gathering this information is by remote sensing, as has been demonstrated using Advanced Very High Resolution Radiometer (AVHRR) data (Justice et al. 1985, Tucker et al. 1986). However, visible and infrared instruments depend on solar irradiance and visibility conditions. Pervasive cloud cover in the tropics severely limits the utility of optical sensors. Microwave sensors penetrate clouds and can provide all-weather capability. Existing research has demonstrated that microwave remote sensing of vegetation is a powerful tool for acquiring biophysical data. Recent successful launches of ERS-1 (by ESA of Europe) and JERS-1 (by Japan) satellites make Synthetic Aperture Radar (SAR) data covering the globe available. Future launches of ERS-2 (by ESA) and RADARSAT (by Canada) satellites will provide further SAR data.
To understand forest scattering mechanisms and interpret SAR backscatter measurements from forests, and to retrieve forest biophysical parameters from SAR backscatter data, analytical canopy backscatter models (e.g., Ulaby et al. 1990, Sun et al. 1993a) have been developed. These models contribute to the understanding of radar backscatter over forested regions to the extent that they capture the basic interactions between microwave radiation and tree canopies, understories, and ground layers as functions of radar incidence angle, wavelength, and polarization.

The Santa Barbara microwave canopy backscatter model, developed for forest stands with discontinuous tree canopies, consist of a single-tree scattering model, and a gap probability model. The modelling approach is to treat individual tree crowns as scatterers and attenuators, using the probabilities of scattering and attenuation to compute total backscatter. The model input includes stand mean tree trunk diameter at breast height (dbh) and stand density, and other parameters such as branch dimension and angular orientation, leaf dimension and angular orientation. Main model outputs are $HH$, $HV$, and $VV$ backscatter (Sun et al. 1991, Wang et al. 1993a). Comparison of the model predictions with Jet Propulsion Laboratory (JPL) Airborne Synthetic Aperture Radar (AIRSAR) data for boreal forests (Want et al. 1993b) and for ponderosa pine forest (Sun et al. 1991, Wang et al. 1993c) are promising. The successes in collecting the SAR data and ground data and in predicting backscattering from forests by this canopy backscatter model have encouraged us to pursue inversion: retrieving forest biophysical parameters from the SAR backscatter data using the canopy backscatter model.

Because the scattering and interaction mechanisms from forests are complex and because our canopy backscatter model is also complex and requires multiple inputs (Sun et al. 1991, Want et al. 1993a), the extraction of forest parameters from SAR backscatter data through direct inversion of the models is impossible. Alternative approaches are being sought. A neural network (NN) trained with a canopy backscatter model is one such alternative. Recent work (Tsang et al. 1992) supported the rationale of using a multiple-layered perceptron (MLP) NN in the inversion. With the MLP NN trained by a scattering model, Tsang et al. successfully inverted the mean-grain size of ice particles in snow, snow density, snow temperature, and snow depth from passive microwave data. Their success encourages us to investigate the feasibility of retrieving the stand mean dbh and stand density of ponderosa pine forests from SAR backscatter data by using an MLP NN trained by our canopy backscatter model. The dbh and density are important parameters to describe a forest stand. From the dbh, one can derive tree height, tree crown depth and width, and tree biomass, based on allometric equations. Combining this with stand density makes it possible to estimate stand biomass. If the retrieval of the dbh and density is possible, one could use SAR data to estimate forest biomass.

2. A three-layered perceptron neural network

Neural networks have been studied in the fields of speech and image recognition in the hope of accomplishing human-like performance (Rosenblatt 1959, Rumelhart et al. 1986, Lippmann 1987, McClelland and Rumelhart 1988). An NN consists of many nonlinear (or linear in some simple cases) computational elements operating in parallel and arranged in patterns representing biological neural cells. Computational elements or nodes are connected via weights that are typically
adapted during use to improve performance. There are typically six types of NNs (Lippmann 1987). For this study, a three-layered perceptron (TLP) was chosen. The TLP NN was trained with the back-propagation training algorithm (Rumelhart et al. 1986, Lippmann 1987, McClelland and Rumelhart 1988). This algorithm is an interactive gradient descent algorithm designed to minimize the sum of squared error between the actual outputs of a TLP feed-forward perceptron and the desired outputs. Because the error is considered as a smooth error function in the weight space, the iterative gradient descent algorithm can theoretically find the global minima of the error function, if properly executed.

2.1. Two sets of three-layered perceptron neural networks

Using the back-propagation program (McClelland and Rumelhart 1988) as a basis, we have designed two sets of TLP NNs, one set of NNs to retrieve the stand mean dbh and the other the stand density. Both sets of NNs have an input layer, a hidden layer, and an output layer (figure 1). In the output layer there is one node that outputs the dbh or density. The selection of one hidden layer simplifies the design of the NNs, and greatly reduces the time needed to train the NNs. Also, any NN with multiple hidden layers can be replaced by an NN with a single hidden layer (Hornik 1993). The number of nodes in the hidden layer is 30, empirically determined on the basis of the computer time needed to train an NN. The more nodes in the hidden layer, the longer the computer time required to finish the training. The number of input nodes varies because multiple-polarization and multiple-wavelength AIRSAR data are used as inputs to the NNs. The combinations of one wavelength (C-band, 5.6 cm wavelength, L-band, 23.5 cm wavelength; or P-band, 68 cm wavelength), two wavelengths (C- and L-band, C- and P-band, or L- and P-band), and three wavelengths (C-, L-, and P-band) provide a set of seven individual NNs. Because backscatter of three SAR polarizations (HH, HV, and VV) are available for each wavelength, the numbers of input nodes are three for the NNs using a single-wavelength multipolarization data, six for the NNs using two-wavelength (in parallel) multipolarization data, and nine for the NNs using three-wavelength (in parallel) multipolarization data (table 1).

![Figure 1. A three-layer neural network.](image-url)
Table 1. Number of nodes of seven neural networks.

<table>
<thead>
<tr>
<th>Combinations of wavelengths</th>
<th>C</th>
<th>L</th>
<th>P</th>
<th>C &amp; L</th>
<th>C &amp; P</th>
<th>L &amp; P</th>
<th>C &amp; L &amp; P</th>
</tr>
</thead>
<tbody>
<tr>
<td>No. of input nodes</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>6</td>
<td>6</td>
<td>6</td>
<td>9</td>
</tr>
<tr>
<td>No. of hidden nodes</td>
<td>30</td>
<td>30</td>
<td>30</td>
<td>60</td>
<td>60</td>
<td>60</td>
<td>90</td>
</tr>
<tr>
<td>No. of output nodes</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
</tbody>
</table>

3. Method

Although retrieval of the dbh or density is handled independently (i.e. two separate sets of TLP NNs are used), the approach is similar for both. Therefore, we describe the approach only for the dbh retrieval.

3.1. Inversion of the stand mean dbh

(i) We run the canopy backscatter model to simulate an array of outputs of $HH$, $HV$, and $VV$ backscatter for a range of stand mean dbh. In general, we express the output array as $[x_i] = [\sigma_{HH}^0, \sigma_{HV}^0, \sigma_{VV}^0] (i = 1, 2, \ldots, Q)$ and the dbh array as $[y_i] = [dbh_i] (i = 1, 2, \ldots, Q)$. $x_i$ is the output of simulated pixel $i$, and $y_i$ the mean dbh in the $i$-th pixel. $Q$ is the number of pixels simulated.

(ii) We pair (or geo-reference) the output array $[x_i]$ and dbh array $[y_i]$ to produce a data set $[x_i, y_i]$ for training. The $[x_i]$ are inputs to the NN, and each $x_i$ is called as an input pattern to the NN. The $[y_i]$ is a target array used to estimate the error of the NN.

(iii) We train a TLP NN with the $[x_i, y_i]$ data set by using the back-propagation algorithm (Rumelhart et al. 1986, Lippmann 1987, McClelland and Rumelhart 1988). Five steps are used in the training process (Lippmann 1987):

(a) Initializing weights and error tolerance;
(b) Presenting input and desired outputs;
(c) Calculating outputs;
(d) Adapting weight;
(e) Repeating by going to (b) as needed, or halting the training when the error tolerance is reached.

For an input pattern $x_i$ and a TLP NN (of one output node), the squared error of the NN is

$$e_i = (y_i - o_i)^2$$

where $o_i$ is the output of the NN given $x_i$ as an input. For a total of $Q$ patterns, the r.m.s. error of the NN is

$$E_{dbh} = \sqrt{\frac{1}{Q} \sum_{i=1}^{Q} e_i}$$

(iv) We extract the SAR $HH$, $HV$, and $VV$ backscatter data from the stands whose stand mean dbhs are within the range of the dbhs used in the training of the NN, and input the SAR data to the trained TLP NN to invert the dbh.
Table 2. Training r.m.s. errors of NNs for retrieving the stand mean dbh (cm).

<table>
<thead>
<tr>
<th>Combinations of wavelengths</th>
<th>C</th>
<th>L</th>
<th>P</th>
<th>C &amp; L</th>
<th>C &amp; P</th>
<th>L &amp; P</th>
<th>C &amp; L &amp; P</th>
</tr>
</thead>
<tbody>
<tr>
<td>r.m.s. errors</td>
<td>6.7</td>
<td>6.5</td>
<td>6.8</td>
<td>6.5</td>
<td>6.6</td>
<td>6.7</td>
<td>6.5</td>
</tr>
</tbody>
</table>

Table 3. Training r.m.s. errors of NNs for retrieving the stand density (trees ha⁻¹).

<table>
<thead>
<tr>
<th>Combinations of wavelengths</th>
<th>C</th>
<th>L</th>
<th>P</th>
<th>C &amp; L</th>
<th>C &amp; P</th>
<th>L &amp; P</th>
<th>C &amp; L &amp; P</th>
</tr>
</thead>
<tbody>
<tr>
<td>r.m.s. errors</td>
<td>5.1</td>
<td>5.0</td>
<td>4.8</td>
<td>5.0</td>
<td>4.9</td>
<td>4.8</td>
<td>4.8</td>
</tr>
</tbody>
</table>

(v) The inverted and measured dbhs are compared to estimate the error. The relative error is

\[ \delta_{dbh} = \frac{E_{dbh}}{dbh} \]  

where \( dbh \) is the actual measured dbh.

4. Results

4.1. Training a three-layered perceptron (TLP) neural network

Using ground data of ponderosa pine forest stands near Mt. Shasta, California (Wang et al. 1993c), we modelled \( HH, HV, \) and \( VV \) backscatter for a range of stand dbhs, and paired the backscatter and dbhs as \( [\sigma_{HH}^0, \sigma_{HV}^0, \sigma_{VV}^0, \sigma_{dbh}^0] \) \((i = 1, 2, \ldots, 14000)\) at C-, L-, and P-band, respectively. Detailed descriptions of the ground conditions and frosts can be found in Sun et al. (1991) and Wang et al. (1993c). The range of the dbhs is 34–57 cm, resulting from 14000 random samples from the dbh distribution collected in the field. We then designed three TLP NNs, one for C-band data, one for L-band data, and one for P-band data. Each TLP NN has three input nodes, 30 hidden nodes, and one output node (see also table 1). We next trained the TLP NNs with the 14000 patterns. The r.m.s. errors of trainings (equation (2) in §3.1) are 6.8 cm (table 2).

By combining in parallel the simulated \( HH, HV, \) and \( VV \) backscatter at two and three wavelengths, we have four more training data sets: \( [\sigma_{HH}^0, \sigma_{HV}^0, \sigma_{HH}^0, \sigma_{HV}^0, \sigma_{VV}^0, \sigma_{dbh}^0, \sigma_{HH}^0, \sigma_{HV}^0, \sigma_{VV}^0, \sigma_{dbh}^0, \sigma_{HH}^0, \sigma_{HV}^0, \sigma_{HH}^0, \sigma_{HV}^0, \sigma_{VV}^0, \sigma_{dbh}^0] \) for C and L bands, \( [\sigma_{HH}^0, \sigma_{HV}^0, \sigma_{HH}^0, \sigma_{HV}^0, \sigma_{VV}^0, \sigma_{dbh}^0] \) for C and P bands, \( [\sigma_{HH}^0, \sigma_{HV}^0, \sigma_{HH}^0, \sigma_{HV}^0, \sigma_{HH}^0, \sigma_{HV}^0, \sigma_{VV}^0, \sigma_{dbh}^0] \) for L and P, and \( [\sigma_{HH}^0, \sigma_{HV}^0, \sigma_{VV}^0, \sigma_{dbh}^0] \) for C and L and P bands. The superscript c, l, or p stands for the C, L, or P-band. We then designed and trained another four NNs (see also table 1 for the structures of the NNs). The r.m.s. errors of training are 6.7 cm (table 2).

A similar procedure is followed for the retrieval of the stand density. In the training data sets, the range of the stand density is between 192 and 364 trees ha⁻¹. The training r.m.s. errors are between 4.8 trees ha⁻¹ and 5.1 trees ha⁻¹ (table 3).
4.2. Retrieving the stand mean dbh and stand density

After the NNs are trained, they are ready to be evaluated and to invert the dbh from the SAR backscatter data. The evaluation of the inversion is done by extracting the SAR $HH$, $HV$, and $VV$ backscatter data of the ponderosa pine forest stands from SAR imagery, filtering the data to reduce the speckle by a median filter ($3 \times 3$), and inputting the data to the trained NNs. Eight JPL AIRSAR multi-frequency data takes, acquired on 6 September 1989 after a dry summer, were used. The SAR data were processed and calibrated by JPL. The estimated calibration uncertainty of backscatter is $\pm 1\text{ dB}$ at C-, L-, and P-band. The data are standard 4-look, with a compressed data format (10-byte per image pixel). The pixel resolution is 12.1 m in azimuth, and 6.7 m in slant range (Wang et al. 1993c). Two ponderosa pine stands (St2 and St11) are used in the retrieval. For trees with dbh $\leq 7$ cm, the stand mean dbhs and standard deviations of the dbhs are 39 cm and 23 cm for stand St2, and 46 cm and 29 cm for stand St11. The stand densities (for trees with dbh $\geq 7$ cm) are 309 trees ha$^{-1}$ for stand St2, and 233 trees ha$^{-1}$ for stand St11. Other stand parameters can be found in Wang et al. (1993c).

The r.m.s. errors of the inverted dbhs are 6.1 cm for stand St2, and 3.1 cm for stand St11 (table 4). (A possible explanation why the r.m.s. errors for the inverted dbhs are smaller than those for the trained dbhs is discussed in § 5). The relative errors are between 12.3 per cent and 15.6 per cent for St2, and between 4.3 per cent and 6.7 per cent for St11.

Retrieving the stand densities of stands St2 and St11 was also carried. The r.m.s. errors of the inverted stand densities are from 39.8 trees ha$^{-1}$ to 71.2 trees ha$^{-1}$ for St2 (with a stand density of 309 trees ha$^{-1}$), and from 16.6 trees ha$^{-1}$ to 49.7 trees ha$^{-1}$ for St11 (with a stand density of 223 trees ha$^{-1}$) (table 5). The relative errors are 12.9–23.0 per cent for St2, and 7.1–21.3 per cent for St11 (table 5). Because of large r.m.s. and relative errors, these NNs may not invert the stand density accurately.

Table 4. Errors (r.m.s.) of retrieved stand mean dbh (cm) from SAR data by the trained NNs.

<table>
<thead>
<tr>
<th>Combinations of wavelengths</th>
<th>C</th>
<th>L</th>
<th>P</th>
<th>C &amp; L</th>
<th>C &amp; P</th>
<th>L &amp; P</th>
<th>C &amp; L &amp; P</th>
</tr>
</thead>
<tbody>
<tr>
<td>For St2</td>
<td>5.1</td>
<td>5.7</td>
<td>4.8</td>
<td>6.1</td>
<td>5.5</td>
<td>5.9</td>
<td>6.0</td>
</tr>
<tr>
<td>For St11</td>
<td>3.1</td>
<td>2.0</td>
<td>3.1</td>
<td>2.5</td>
<td>2.6</td>
<td>2.5</td>
<td>2.5</td>
</tr>
</tbody>
</table>

The relative errors are between 12.3 per cent and 15.6 per cent for stand St2, and between 4.3 per cent and 6.7 per cent for stand St11.

Table 5. Errors (r.m.s.) of retrieved stand density (trees ha$^{-1}$) from SAR data by the trained NNs.

<table>
<thead>
<tr>
<th>Combinations of wavelengths</th>
<th>C</th>
<th>L</th>
<th>P</th>
<th>C &amp; L</th>
<th>C &amp; P</th>
<th>L &amp; P</th>
<th>C &amp; L &amp; P</th>
</tr>
</thead>
<tbody>
<tr>
<td>For St2</td>
<td>39.8</td>
<td>61.5</td>
<td>71.2</td>
<td>50.0</td>
<td>44.0</td>
<td>67.2</td>
<td>67.6</td>
</tr>
<tr>
<td>For St11</td>
<td>42.0</td>
<td>48.0</td>
<td>16.6</td>
<td>46.0</td>
<td>49.7</td>
<td>26.9</td>
<td>29.4</td>
</tr>
</tbody>
</table>

The relative errors range from 12.9 per cent to 23.0 per cent for stand St2, and 7.1 per cent to 21.3 per cent for stand St11.
5. Concluding remarks

The direct method to invert forest biophysical parameters such as the stand mean dbh from the SAR backscatter data using a neural network (NN) is to train the NN with the SAR and ground data. The SAR data from the interested area are input to the NN, and the ground data of the area are used to estimate the training error. To train the NN successfully, a large amount of ground and SAR data are required, and the ground and SAR data must go geo-referenced. For instance, in our case, 14000 training patterns were used; this would have required 14000 pairs of dbh and SAR measurements. Because limited resources can be spent in the field to collect the ground data, map the data, and geo-reference the ground and SAR data, this method becomes infeasible. Alternative approaches are being sought. One alternative is to train the NN by a validated canopy backscatter model such as the one in this study. The model has been slightly modified by simulating the ‘geo-referencing’ the HH, HV, and VV backscatter data and the ground data. This study shows that the alternative training approach is feasible, and that it is possible to retrieve the stand dbh and density of ponderosa pine forests from SAR backscatter data by using the trained NN. For two ponderosa pine forest stands tested, the r.m.s. and relative errors of the retrieval for the stand mean dbh are 6.1 cm and 15.6 per cent. The r.m.s. and relative errors of the retrieval for the stand density are 71.2 trees ha$^{-1}$ and 23.0 per cent. The r.m.s. and relative errors of the retrieved forest stand parameters (stand mean dbh and density) may be large. To fully understand the causes of the errors further analysis of a larger data set is needed and funding for this analysis will be pursued in the near future. However, this study moves one step forward and shows the feasibility to extract the stand parameters from SAR backscatter data. Because the retrieval approach is based on the neural networks trained by a radar canopy backscatter model, and is verified by SAR backscatter data from forest stands this approach could be potentially applicable to other types of forests in other locations.

Because the simulated backscatter data and ground data are used to train an NN, the trained NN can approximate well the relationship between the simulated backscatter and ground data if the training r.m.s. error is tolerable. When the trained NN is used to retrieve the ground data from the SAR data, the inverted ground data may differ from the ground measured data if the model does not predict the backscatter well. Therefore, to ensure a successful inversion by this approach it is crucial that the canopy backscatter model must be valid. It also should be noted that, for a trained NN, the approximated relationship (by the NN) between the simulated backscatter and ground data may be off from the relationship between the simulated backscatter and ground data. It is possible that the approximated relationship may be closer to the relationship between the observed SAR backscatter and ground data than the relationship between the simulated backscatter and ground data. That is why, in some cases (e.g., the retrieval of the stand mean dbh in §4.2), the errors for the inverted values are smaller than those for the trained values.

SAR HH, HV, and VV backscatter data at one wavelength (C, L, or P), two wavelengths (C and L, C and P, or L and P), and three wavelengths (C and L and P) are used. The training r.m.s. errors of the NNs may be smaller at multiple wavelengths than at a single wavelength. However, the errors of the inverted parameters from the three-wavelength SAR data may not be less than those from two- or one-wavelength SAR data; further studies will be needed to understand why.

In the training, we noticed that the error function of an NN in the weight space
is complex, and the error function typically has multiple minima. Many of the minima are local ones. The purpose of the training is to find the global minima. Unfortunately, the back-propagation algorithm or the iterative gradient descent algorithm has a great potential to be trapped inside local minima. A learning rate coefficient to adjust the step of each move is introduced to help the algorithm out of the local minima (McClelland and Rumelhart 1988). The coefficient depends on the size of the training patterns (Eaton and Olivier 1992). Our experience on training the NN suggests that the coefficient is roughly inversely proportional to the number of training patterns, and increasing the coefficient can help the algorithm out of the local minimum traps. However, we generally do not know whether the algorithm is trapped within local minima or has found the global minima. We judge by trial and error. In the training, we also notice that when there are local maxima around the global minima of the error function, the NN becomes unstable as the training approaches the global minima. Here, reducing the learning rate coefficient can generally stabilize the training to reach the global minima.

Acknowledgments
This research was funded by California Space Institute (contract # CS-6-93) and NASA SIR-C/X-SAR project through the Jet Propulsion Laboratory (contract # 958468). Both fundings were managed by the Center for Remote Sensing and Environmental Optics, University of California at Santa Barbara. The ground data and AIRSAR data were obtained from the NASA SIR-C/X-SAR project.

References
