Temperature Profiling with Neural Network Inversion of Microwave Radiometer Data

JAMES H. CHURNSIDE, THOMAS A. STERMITZ, AND JUDITH A. SCHROEDER

NOAA Wave Propagation Laboratory, Boulder, Colorado

(Manuscript received 8 December 1992, in final form 24 May 1993)

ABSTRACT

A neural network is used to obtain vertical profiles of temperature from microwave radiometer data. The overall rms error in the retrieved profiles of a test dataset was only about 8% worse than the overall error using an optimized statistical retrieval. In certain cases, such as one with a large temperature inversion, the neural network produced better reproductions of the profiles than did the statistical inversion.

1. Introduction

The radiance measured by a vertically pointing microwave radiometer is an integral function of height above the radiometer, the ground-level pressure, and the vertical profiles of temperature and water vapor. If clouds are present, the radiance also depends on liquid water content of the clouds. The details of the integral function depend on the wavelength at which the radiometer is operating. Therefore, a radiometer operating at several wavelengths can provide information about the vertical profile of the temperature structure in the atmosphere. Obtaining an estimate of that profile requires inversion of a set of integral equations.

One technique for retrieving temperature profiles from microwave radiances is a linear statistical inversion (Strand and Westwater 1968; Hogg et al. 1983). A six-channel radiometer has been operated in Denver, Colorado, since about 1981. It uses a linear statistical inversion to obtain continuous soundings of temperature in real time (Hogg et al. 1983). The rms difference between these temperature profiles and the U.S. National Weather Service operational rawinsonde profiles at the same site is generally 1°–3°C throughout the troposphere (Westwater et al. 1984).

Attempts to improve on the accuracy of the microwave-derived temperature profiles have concentrated on including other sources of information. One approach has been to combine ground-based microwave data with microwave and infrared radiometer data collected from satellites (Westwater and Grody 1980; Westwater et al. 1984; Westwater et al. 1985; Schroeder 1990). Another source of data that has been considered is a radio acoustic sounding system, which infers air temperature from a Doppler radar measurement of the speed of a transmitted sound wave (Schroeder 1990).

There are several similarities between a linear statistical inversion and a neural-network retrieval. In both cases, some input vector is operated on to produce an output vector. For our purposes, the input vector is a set of microwave radiances plus the ground-level pressure, temperature, and humidity. The output vector is the vertical profile of temperature. In both cases, the operator is generated using an existing set of paired input and output vectors. In the linear statistical inversion, this set is used to obtain the pertinent statistics (i.e., the covariance matrices) that are used to derive the inversion coefficients. In the neural-network retrieval, this set is used to train the network.

Despite these similarities, neural networks have not been used extensively for data retrieval. Typical applications include classification of image features (Barnard and Casasent 1989; Lee et al. 1990; Kamata et al. 1991) and image feature detection (Fitch et al. 1991; Ryan et al. 1991). Measure et al. (1991) considered radiometric inversions using a two-layer neural network, and this paper is an extension of their work.

2. Neural network

A schematic diagram of the type of neural network that was used for this work is presented in Fig. 1. The net has an input layer of L nodes to which an input vector $\mathbf{X}$ of length L is applied. Each input node is connected to all M nodes in a hidden layer. Each node in the hidden layer performs a weighted sum over all the input values to produce an output vector $\mathbf{Y}$. Each node in the hidden layer is connected to each node in an output layer, which performs a weighted sum over all of the results of the hidden-layer calculations. The N values from the output-layer nodes create the output vector $\mathbf{Z}$.

Corresponding author address: Dr. James H. Churnside, NOAA/ERL, Mail Code R/W/WP1, 325 Broadway, Boulder, CO 80303.

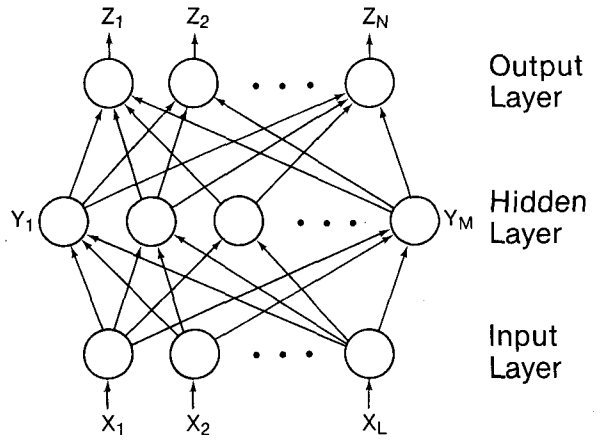

FIG. 1. Schematic of neural network showing input vector **X** and output vector **Z**.

The neural-network calculations in this paper were made with a network simulator called NETS (Baffes et al. 1991) running on a computer workstation. Generally, in this simulator all input and output values are scaled so that they lie between 0 and 1. For atmospheric inversions, the inputs and outputs are unbounded. No matter what the range of values in the training set, they will eventually be exceeded by the atmosphere. For this reason, we scaled all values so that the training set values were between 0.15 and 0.9. We then tested the performance of the network using data that occasionally fell outside of the range of the training set.

In NETS, a neuron performs a weighted sum of its inputs and applies a sigmoidal function to the result to produce an output. For the ith node in the hidden layer, this can be expressed as

$$y_i = S\left(\sum_{j=0}^{M} w_{ij}x_j\right), \tag{1}$$

where S is the sigmoidal function

$$S(a) = \frac{1}{1 + \exp(-a)}, \tag{2}$$

and w_{ij} is the weighting of the connection between the jth input neuron and the ith hidden neuron. Operation of a neuron in the output layer is similar.

The weights w_{ij} are determined during the training process. In NETS, they are obtained using a backpropagation algorithm that is described in detail by Rumelhart et al. (1986). This algorithm adjusts the weights iteratively to reduce the difference between the actual training set output vectors and the estimated output vectors calculated by the network using the input vectors of the training set.

3. Data

The training dataset was based on 10 years (1970–79) of routine National Weather Service rawinsonde data taken at Stapleton International Airport in Denver. Soundings are made twice daily (1100 and 2300 UTC). To allow comparison, we used the same January, February, and March data as the winter dataset of Schroeder (1990). Also, the data were interpolated to obtain values at the same levels as that paper. This was done so that rms errors could be calculated in the same way, and the two techniques could be compared directly.

The input vector comprised nine elements. Three of these were the surface-level temperature, pressure, and relative humidity. The other six were the brightness temperatures at the six frequencies of the radiometer (20.60, 31.65, 52.85, 53.85, 55.45, and 58.80 GHz).

For the training dataset, brightness temperatures were calculated theoretically from the rawinsonde data using a form of the radiative transfer equation in Askne and Westwater (1986). Since rawinsondes do not measure cloud water content, cloudy conditions were modeled as in Decker et al. (1978). When the relative humidity measured by the rawinsonde was above 95%, a cloud was assumed, and three input vectors were calculated with different assumptions about the total water content. The amounts of water assumed were based on cloud thickness, and a fraction of the total water was assumed to be ice, with the fraction depending on temperature. The attenuation due to these clouds was calculated, and these values were used in the radiative transfer calculation. To simulate the effects of instrument noise, a zero-mean Gaussian random error with 0.5-K standard deviation was added to each calculated brightness temperature.

The output vector was a 45-element vertical profile of temperature. The first 17 elements were evenly spaced in height above the ground from 50 to 2150 m. The other 28 were evenly spaced in pressure from 625 to 10 mb.

A test dataset was used that comprised 15 rawinsondes in February 1989, and actual radiometric data taken at the same time and the same location. This case was chosen because it contained an extremely cold surface layer with a strong inversion (Neiman et al. 1989) and is a severe test of any inversion technique.

4. Results

We first considered the case of a neural network with no hidden layer. Figure 2 is a typical temperature retrieval from the test data. From this figure, we see that the agreement between the radiosonde data and the retrieved profiles are similar for the statistical inversion and the neural-network inversion. The major difference is a fairly large error in the neural-network retrieval at high altitudes.

Figure 3 is a similar plot for a case with a large temperature inversion. In this case, the neural network reproduces the magnitude of the temperature inversion and approximates the profile much better than the sta-

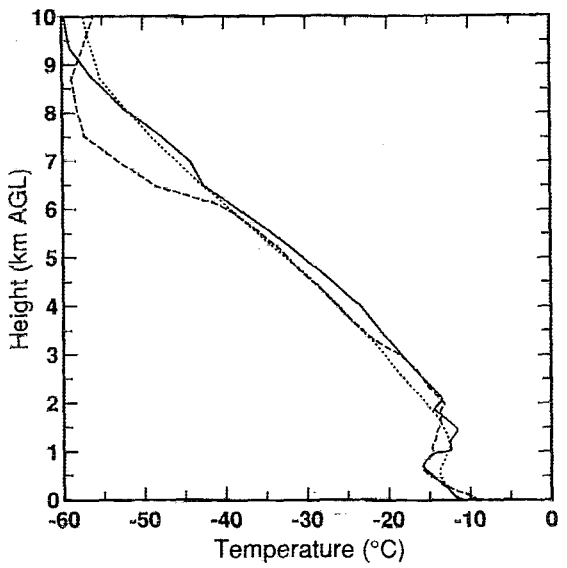

FIG. 2. Temperature as a function of height for a typical case from the radiosonde (solid line), the statistical retrieval (dotted line), and the neural network with no hidden layers (dashed line).

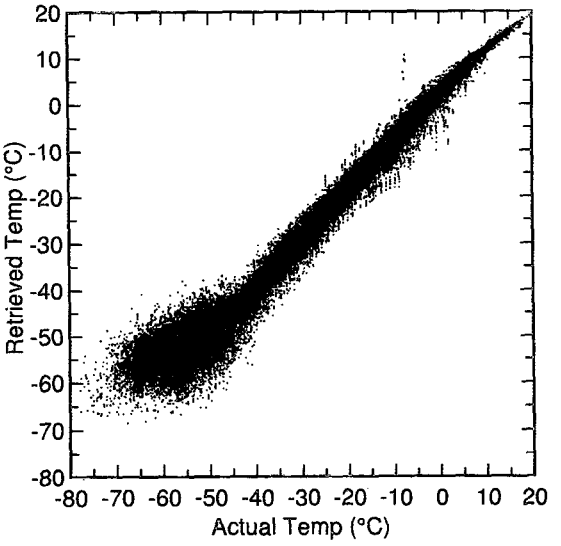

FIG. 4. Scatterplot of retrieved temperatures versus actual temperatures for the training dataset.

tistical retrieval in the lower portion of the atmosphere. As before, a substantial error appears at higher altitudes.

Overall performance of the two-layer neural network is summarized in Fig. 4, which plots the estimated temperature as a function of the actual temperature for all points in the training dataset, and Fig. 5, which is the same plot for the test dataset. All points from each profile were plotted; in Fig. 5, one can clearly see the traces of individual profiles because adjacent points are not statistically independent. As one might expect, agreement is better for the training set than for the test set. The total rms difference between the radiosonde data and the neural-network retrievals was 2.53°C for the training set. The equivalent value for the test set was 4.93°C. This is 24% worse than the value of 3.98°C obtained by the statistical inversion (Schroeder 1990).

The values presented above were for a training session of 10 000 iterations. The rms error on the training dataset is plotted as a function of iterations in Fig. 6.

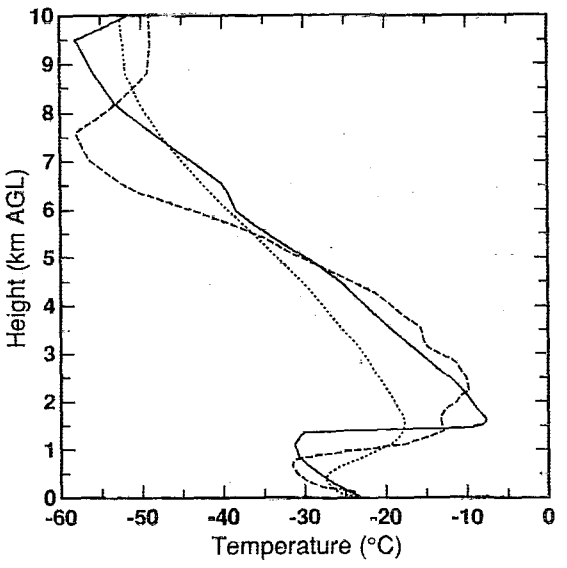

FIG. 3. Temperature as a function of height for an extreme case from the radiosonde (solid line), the statistical retrieval (dotted line), and the neural network with no hidden layers (dashed line).

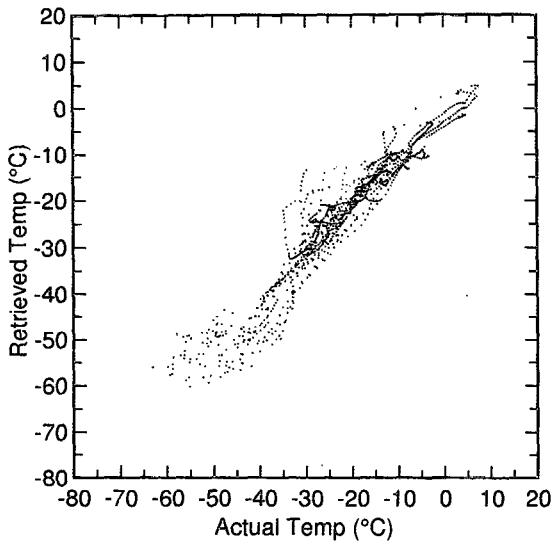

FIG. 5. Scatterplot of retrieved temperatures versus actual temperatures for the test dataset.

The error drops fairly rapidly for the first few hundred iterations, and then the performance improves more slowly.

We next considered neural networks with a hidden layer. Twenty neurons in the hidden layer were used. Increasing the number up to 100 did not seem to increase the performance significantly. Figure 7 is a plot of the hidden-layer retrieval of the strong-temperature-inversion case of Fig. 3. We see that the retrieval is a little better with the hidden layer, especially at higher altitudes. The total rms error on the test set using this network was 4.28°C, which is only about 8% worse than the optimized statistical inversion.

Finally, we considered the addition of more input neurons that represented nonlinear combinations of the data, such as squares and cross products. None of the combinations we tried increased the performance of the network significantly.

5. Conclusions

A neural-network retrieval with 20 neurons in the hidden layer, trained for 10 000 iterations, can provide temperature retrievals from microwave radiometer data that are nearly as good as an optimized statistical retrieval in terms of overall rms error. We would not expect a neural network to do better than the statistical inversion, because the latter is designed to minimize rms error based on knowledge of the statistics. If the neural network is given the same information, as was done here, it could do no better.

Although the neural network did slightly worse than the statistical retrieval overall, it seemed to do better in certain cases. In difficult cases, such as the strong temperature inversion of Fig. 7, the neural network reproduced the essential features of the profiles better than the statistical retrieval. It also seemed to generally do better near the ground and worse at higher altitudes. Thus, a neural network inversion might be preferred despite a slightly worse rms error because it seems to more faithfully recover the more meteorologically interesting extreme cases.

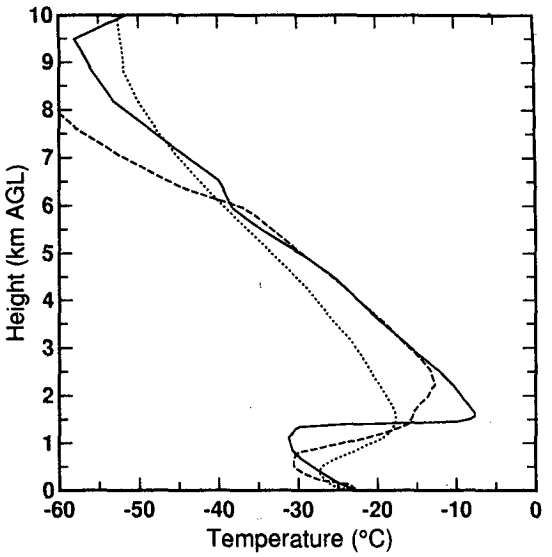

FIG. 7. Temperature as a function of height for an extreme case from the radiosonde (solid line), the statistical retrieval (dotted line), and the neural network with 20 hidden-layer neurons (dashed line).

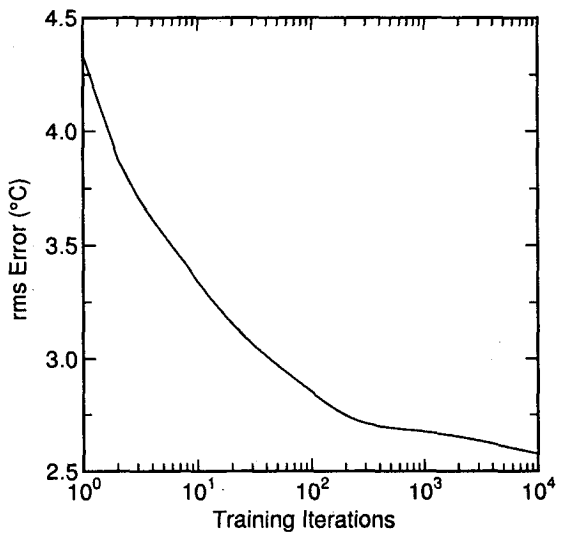

FIG. 6. Root-mean-square error in training set as a function of number of training iterations.

Finally, we note that the performance of the neural network seemed very insensitive to details of the implementation. As long as we used a hidden layer of 20 neurons or more and trained the network for more than a few hundred iterations, the results were very similar. This, plus the closeness of the results to the statistical retrieval results, suggests that we are getting nearly all of the useful information out of the data.

The neural-network technique lends itself to retrievals based on multiple sensors. It would be easy to add data from a radio acoustic sounding system, data from satelliteborne sounders, and radar wind profiles. One only needs a suitable training set; the network should adjust the weighting of the various data according to their predictive capability. Radiosonde data from other locations or previous times could also be used as long as the separations or time delays were either constant or provided as input variables. Various combinations of these types of data are currently being investigated.

REFERENCES

Askne, J. I. H., and E. R. Westwater, 1986: A review of ground-based remote sensing of temperature and moisture by passive microwave radiometers. *IEEE Trans. Geosci. Remote Sens.*, **24**, 340–352.

Baffes, P. T., R. O. Shelton, and T. A. Phillips, 1991: *NETS User's Guide.* NASA Johnson Space Flight Center. [Available from

COSMIC, University of Georgia, 382 East Broad St., Athens, GA 30602.] 103 pp.

Barnard, E., and D. P. Casasent, 1989: Optical neural net for classifying imaging spectrometer data. *Appl. Opt.*, **28**, 3129–3133.

Decker, M. T., E. R. Westwater, and F. O. Guiraud, 1978: Experimental evaluation of ground-based microwave radiometric sensing of atmospheric temperature and water vapor profiles. *J. Appl. Meteor.*, **17**, 1788–1795.

Fitch, J. P., S. K. Lehman, F. U. Dowla, S. Y. Lu, E. M. Johansson, and D. M. Goodman, 1991: Ship wake-detection procedure using conjugate gradient trained artificial neural networks. *IEEE Trans. Geosci. Remote Sens.*, **29**, 718–726.

Hogg, D. C., M. T. Decker, F. O. Guiraud, K. B. Earnshaw, D. A. Merritt, K. P. Moran, W. B. Sweezy, R. G. Strauch, E. R. Westwater, and C. G. Little, 1983: An automatic profiler of the temperature, wind, and humidity in the troposphere. *J. Appl. Meteor.*, **22**, 807–831.

Kamata, S., R. O. Eason, and E. Kawaguchi, 1991: Classification of Landsat image data and its evaluation using a neural network approach. *Trans. Soc. Instrum. Control Eng.*, **27**, 1302–1306.

Lee, J., R. C. Weger, S. K. Sengupta, and R. M. Welch, 1990: A neural network approach to cloud classification. *IEEE Trans. Geosci. Remote Sens.*, **28**, 846–855.

Measure, E. M., Y. P. Yee, J. M. Balding, and W. R. Watkins, 1991: Inverting radiometric measurements with a neural network. *AGARD Conf. Proc. 502, Remote Sensing of the Propagation Environment*, Cesme, Turkey, NATO, 30-1–30-6.

Neiman, P. J., P. T. May, B. B. Stankov, and M. A. Shapiro, 1989: Radio acoustic sounding system observations of an arctic front. *J. Appl. Meteor.*, **30**, 881–892.

Rumelhart, D. E., G. E. Hinton, and R. J. Williams, 1986: Learning internal representations by error propagation. *Parallel Distributed Processing*, D. E. Rumelhart and J. L. McClelland, Eds., The MIT Press, 45–76.

Ryan, T. W., P. J. Sementilli, P. Yuan, and B. R. Hunt, 1991: Extraction of shoreline features by neural nets and image processing. *Photo. Eng. Remote Sens.*, **57**, 947–955.

Schroeder, J. A., 1990: A comparison of temperature soundings obtained from simultaneous radiometric, radio-acoustic, and rawinsonde measurements. *J. Atmos. Oceanic Technol.*, **7**, 495–503.

Strand, O. N., and E. R. Westwater, 1968: Minimum rms estimation of the numerical solution of a Fredholm integral equation of the first kind. *J. Numer. Anal.*, **5**, 287–295.

Westwater, E. R., and N. C. Grody, 1980: Combined surface- and satellite-based microwave temperature profile retrieval. *J. Appl. Meteor.*, **19**, 1438–1444.

——, W. Sweezy, L. McMillin, and C. Dean, 1984: Determination of atmospheric temperature profiles from a statistical combination of ground-based profiler and operational *NOAA 6/7* satellite retrievals. *J. Climate Appl. Meteor.*, **23**, 689–703.

——, W. Zhenhui, N. C. Grody, and L. McMillin, 1985: Remote sensing of temperature profiles from a combination of observations from the satellite-based microwave sounding unit and the ground-based profiler. *J. Atmos. Oceanic. Technol.*, **2**, 97–109.