Retrieval of Atmospheric Temperature Profiles from AMSU-A Measurement Using a Neural Network Approach

LEI SHI

SeaSpace Corporation, Poway, California

(Manuscript received 4 February 2000, in final form 15 May 2000)

ABSTRACT

Backpropagation neural networks are applied to retrieve atmospheric temperature profiles and tropopause variables from the *NOAA-15* Advanced Microwave Sounding Unit-A (AMSU-A) measurement based on two different data sources. The first case uses direct acquisition of 15-channel AMSU-A data over the eastern United States and western Atlantic Ocean for the months of July 1998 and January 1999, and the second case uses recorded global AMSU-A data for several days of January 2000. The corresponding global analysis data from the National Centers for Environmental Prediction are employed to build the neural network training sets. The retrievals yield excellent results in the atmospheric temperature profiles from the surface to the 10-hPa pressure level. For the more generalized global data retrieval case, the root-mean-square (rms) deviation of temperature retrieval is 3.2°C at the surface, only 1.0° to 1.2°C in the midtroposphere, less than 1.5°C around the tropopause, and between 1.0° and 1.5°C in the stratosphere. Simultaneous retrieval of tropopause temperature, height, and pressure yields the rms deviations of 1.9°C, 0.58 km, and 18.1 hPa, respectively, for these variables. Within the scope of regional data, the trained neural network results in smaller values of temperature profile rms deviations than those of the global-data case. When compared to a linear regression approach, the neural network retrieval yields significantly better results for all the atmospheric levels. The neural network with parameters obtained from the network training optimizations can be easily applied to AMSU-A retrieval operationally.

1. Introduction

The Advanced Microwave Sounding Unit-A (AMSU-A) is a 15-channel microwave radiometer on board the new generation of National Oceanic and Atmospheric Administration (NOAA) polar-orbiting satellites. The first of this new series of satellites, *NOAA-15,* was launched on 13 May 1998. The *NOAA-15* is in a circular sun-synchronous near-polar orbit at an altitude of 833 km. The orbit period is 101.35 min, and the equator crossing time is at 0730 local solar time (LST) for ascending node, and at 1930 LST for descending node. The AMSU-A is composed of two separate units: AMSU-A1 and AMSU-A2. The AMSU-A1 consists of 12 channels in the 50–60-GHz oxygen absorption band and one channel at 89-GHz frequency. The AMSU-A2 contains two lower-frequency channels, with channel 1 at 23.8 GHz and channel 2 at 31.4 GHz (Kramer 1996). Taking the channel specification values from Goodrum et al. (1999), a list of the AMSU-A 15-channel frequencies is provided in Table 1. A description of atmospheric microwave characteristics in the AMSU-A frequency range can also be found in Goodrum et al. (1999). The AMSU-A is a cross-track, line-scanned instrument. The sensor scans in a stepped-scan fashion and covers 30 discrete scene resolution cells. The scan pattern and geometric resolution correspond to a 50-km-diameter cell at nadir and a 2343-km swath. The sensor is designed to provide measurement for deriving atmospheric temperature profiles from the surface up to 45 km. The retrieval of temperature profiles from AMSU-A using a neural network technique is investigated in this study.

Neural networks have been used in an increasing number of meteorological applications in recent years. Stogryn et al. (1994) presented their work of using fully connected, feed-forward neural networks to retrieve ocean surface wind speed based on the Special Sensor Microwave Imager (SSM/I) on board the Defense Meteorological Satellite Program (DMSP) satellites. Tsintikidis et al. (1997) examined a neural network approach to estimate rainfall from SSM/I data. They showed that the neural network was able to represent more accurately the underlying relationship between the SSM/I brightness temperature and rain rate than a regression model. Yang et al. (1997) used a neural network to estimate soil temperature. Hsieh and Tang (1998) reviewed the obstacles to adapting the neural network technique to meteorological and oceanographic prediction and data

Corresponding author address: Dr. Lei Shi, SeaSpace Corporation, 12120 Kear Place, Poway, CA 92064.
E-mail: lshi@seaspace.com

TABLE 1. Channel frequencies of AMSU-A.

Channel number	Channel frequency (MHz)	Bandwidth (MHz)
1	23 800	251.02
2	31 400	161.20
3	50 300	161.14
4	52 800	380.52
5	53 596 ± 155	168.20
6	54 400	380.54
7	54 940	380.56
8	55 500	310.34
9	57 290.344	310.42
10	57 290.344 ± 217	76.58
11	57 290.344 ± 322.2 ± 48	35.11
12	57 290.344 ± 322.2 ± 22	15.29
13	57 290.344 ± 322.2 ± 10	7.93
14	57 290.344 ± 322.2 ± 4.5	2.94
15	89 000	1998.98

analysis, and discussed methods for overcoming the obstacles. More recently, Jones et al. (1999) examined the neural network approach to derive ocean surface specific humidity and air temperature from remotely sensed precipitable water and analysis of sea surface temperature. They showed good agreement of the neural network method with surface marine observations. Ward and Redfern (1999) used a neural network model to predict the bulk–skin temperature difference at the sea surface. Neural network approaches have also been successfully used in retrieving ocean winds from satellite scatterometers (Chen et al. 1999; Cornford et al. 1999). These studies have demonstrated the great potential of neural networks in a large variety of meteorological applications.

Despite recent developments in neural network applications, and neural network studies on microwave temperature forward modeling (e.g., Rieu et al. 1996), there are only a few studies on applying neural networks in retrieving atmospheric temperature profiles. Churnside et al. (1994) studied the retrieval of temperature profiles from ground-based vertically pointing microwave radiometer measurements with a neural network. They noticed that the neural network produced better retrieval of temperature profiles than a statistical method in certain cases, such as one with a large temperature inversion, though their neural network model did not improve the overall retrieval result. In a study of neural network retrieval of atmospheric temperature profiles from a DMSP Special Sensor Microwave Temperature Sounder (SSM/T1) measurement, Butler et al. (1996) showed that backpropagation neural networks yield excellent results in retrieving temperature profiles from 1000 to 10 hPa pressure levels together with tropopause temperature and pressure estimates. The root-mean-square (rms) deviation tested with matched SSM/T1 measurements and conventional soundings were substantially less than 2 K from 500 to 30 hPa. They also examined the retrieval accuracy under possible failure modes of the SSM/T1 instrument.

Most of the satellite microwave sensor temperature profile retrievals have been based on statistical methods. The Air Force Global Weather Center operationally uses a statistically based algorithm for the SSM/T1 retrieval. This method is referred to as the D-matrix method. Descriptions of D-matrix approaches can be found in Rigone and Stogryn (1977), Grody et al. (1985), and Reale (1993). Other methods include a unified retrieval scheme that allows for the incorporation of all the DMSP meteorological sensor measurements in the retrieval procedure (Isaacs 1989).

Several AMSU-A simulation studies were conducted before the launch of *NOAA-15*. Rosenkranz et al. (1994) examined the temperature retrieval accuracy in the middle and lower atmosphere. The error performance of cross-track and conical scan modes was compared. Rosenkranz (1994) developed a rapid transmittance algorithm for AMSU-A sounding frequencies. Subsequently, Rosenkranz et al. (1997) discussed the impact of satellite microwave sounder incidence angle and scan geometry on the accuracy of temperature profile retrievals. A detailed AMSU-A antenna pattern correction scheme was later presented by Mo (1999).

In this paper, the retrieval of atmospheric temperature profiles from AMSU-A measurements using a neural network technique is discussed. The conversion of AMSU-A channel measurements to temperature profiles contains nonlinear processes. Neural networks generally have advantages over traditional statistical methods in that they can build nonlinear models based on the data used to train them, and thus potentially provide more accurate results over a wide range of data. In this study, two cases are examined based on two different data sources. The first case uses direct acquisition of regional AMSU-A data over the eastern United States and western Atlantic Ocean, and the second case uses recorded global AMSU-A data.

2. Regional temperature profile retrieval

a. Regional dataset

To include a wide range of temperature conditions, AMSU-A measurements from July 1998 and January 1999 have been selected. These two months represent summer and winter temperature conditions, respectively, for the Northern Hemisphere. The *NOAA-15* passes were taken at Louisiana State University. The AMSU-A data were extracted from the High Resolution Picture Transmission telemetry and processed into brightness temperatures.

The derivation and quality control of AMSU-A data are performed following the procedure of Goodrum et al. (1999). For each radiometric count measured from the earth target C_s, the scene radiance R_s can be expressed as

$$R_s = a_0 + a_1 C_s + a_2 C_s^2. \quad (1)$$

The coefficients a_i (where $i = 0$, 1, and 2) are functions of internal blackbody temperature, effective cold space temperature, the averaged blackbody and space counts, and an instrument parameter. The coefficients and parameters needed to compute the dependent variables of a_i are available from the National Environmental Satellite, Data, and Information Service. The brightness temperature is calculated based on the inverse of the Planck function for R_s.

The area covered by the AMSU-A data ranges from 17° to 50°N latitude, and 51° to 100°W longitude. All 15 channels of the AMSU-A data are used in the retrieval of temperature profile. These channels represent the atmospheric temperature conditions from the surface (channels 1, 2, and 15) to the lower and upper atmosphere (channels 3–14).

The analysis data of July 1998 and January 1999 from the National Centers for Environmental Prediction (NCEP) global model GDAS at 0000 and 1200 UTC are used in constructing the training set. The NCEP data consist of the temperature values at the surface, temperature profile values, and tropopause values. The temperature profile values are provided at the pressure levels of 1000, 975, 950, 925, 900, 850, 800, 750, 700, 650, 600, 550, 500, 450, 400, 350, 300, 250, 200, 150, 100, 70, 50, 30, 20, and 10 hPa. The tropopause values include the tropopause temperature, height, and pressure. The training set is constructed based on the subset of AMSU-A data that are within 1 h of 0000 and 1200 UTC and the subset of NCEP analysis data that are within the corresponding AMSU-A coverage. Collocation of AMSU-A and NCEP data is performed by matching each NCEP grid dataset with the closest AMSU-A pixel in the corresponding pass. The collocation generates a total of 2535 matched patterns.

b. Neural network

In general, a neural network is a computer model composed of individual processing elements called neurons. The neurons are connected by links in terms of weights. A neural network may consist of multiple layers of neurons interconnected with other neurons in the different layers. These layers are referred to as input layer, hidden layer, or output layer. The inputs and the interconnection weights are processed by a weighted summation function to produce a sum that is passed to a transfer function. The output of the transfer function is the output of the neurons. A neural network is trained with input and output pattern examples. It then constructs a nonlinear numerical model of a physical process in terms of network parameters.

For the AMSU-A temperature profile retrieval, the use of backpropagation neural networks is examined. As discussed in Butler et al. (1996), it is desirable to choose a network that is capable of modeling nonlinear data from examples and is able to generalize and interpolate, so that each profile retrieved by the network during actual operation is a unique, new profile and not simply the learned training example most closely approximating the required profile. These criteria are well met by fully connected, feed-forward, backpropagation networks. A commercial package NeuroShell 2 (Ward Systems Group, Inc. 1996) is employed in the network training. General descriptions of backpropagation networks can be found in a number of publications (e.g., Dayhoff 1990; Hertz et al. 1991; Cichocki and Unbehauen 1993; Rojas 1996). A more focused presentation of the backpropagation theory, architectures, and applications is given by Chauvin and Rumelhart (1995). In a backpropagation network, each layer is generally fully connected to the layers below and above. When the network is given an input, the updating of activation values propagates forward from the input layer of neurons, through each internal layer, to the output layer of neurons. Each neuron in the output layer produces an output, which is compared to the target output defined in the training set. An error value is calculated for each neuron in the output layer. The network corrects its internal parameters (weights) to lessen the errors. The correction mechanism starts with the output neurons and propagates backward through each internal layer to the input layer. The iteration continues, and as a result, future responses of the network are more likely to be correct.

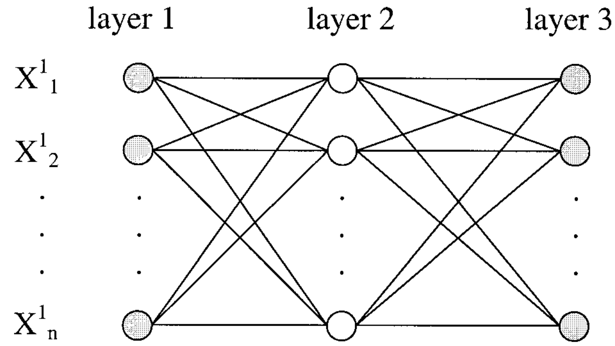

FIG. 1. Schematic diagram of a three-layer backpropagation neural network.

The structure of a three-layer backpropagation neural network is illustrated in Fig. 1. In the figure, the layers 1, 2, and 3 represent the input layer, the hidden layer, and the output layer, respectively. The neurons of the input layer are represented by vector $\mathbf{X}^1$ ($\mathbf{X}^1_1$, $\mathbf{X}^1_2$, $\mathbf{X}^1_3$, . . . , $\mathbf{X}^1_n$, where n is the number of the input neurons). In the AMSU-A retrieval training set, $n = 17$. The 17 elements of vector $\mathbf{X}^1$ correspond to the brightness temperatures (in kelvins) of the 15 AMSU-A channels (T1, T2, T3, . . . , T15), secant of satellite viewing angle, and position of the pixel in a scan line. The secant of satellite viewing angle and position of the pixel in a scan line are introduced to the input neurons to account for the sidelobe effect and any possible scan-dependent bias. Layer 2 and layer 3 contain a different number of neurons. The number of neurons in layer 2 is determined during network architecture design and adjusted to achieve best network performance. The number of neurons in layer 3 is the number of the output variables in the retrieval.

Among the total of 2535 collocated patterns, 20% (507) are randomly extracted to construct a testing set, and another 20% are randomly extracted and set aside for later statistical studies. As a result, there are 1521 patterns remaining in the training set. The testing set is used in the training calibration to prevent overtraining the network so that it will generalize well on new data.

Different transfer functions are examined in constructing the backpropagation network architectures. It is found that using a hyperbolic tangent (tanh) transfer function to propagate to the hidden layer and a logistic transfer function to propagate to the output layer in a three-layer backpropagation architecture gives the optimum network performance for the type of data studied. The definitions of the hyperbolic tangent and logistic transfer functions are

$$f(x) = \tanh(x) \quad \text{and} \tag{2}$$

$$f(x) = \frac{1}{[1 + \exp(-x)]}, \tag{3}$$

respectively. Four-layer backpropagation architectures are also examined. When using combinations of Gaussian, tanh, and logistic transfer functions in the four-layer backpropagation architectures, no significant difference is found in the performance of the networks compared with the above three-layer backpropagation network. To simplify the retrieval data processing, the three-layer backpropagation network is chosen for AMSU-A retrieval.

To effectively process the data in the neural network, the data variables are scaled linearly into the range [−1, 1]. The input layer of the network contains 17 processing elements (neurons) as described previously. For the hidden layer, generally speaking, as more neurons are added to a backpropagation neural network, more degrees of freedom are obtained, and the network is able to store more complex patterns. However, adding more neurons to the network may produce tighter data fits to the training set and cause poor generalization for new cases. For the AMSU-A retrieval, experiments are done for a large range of hidden neuron numbers. Good network performances are found when the numbers of hidden neurons are within the range of 45 to 80 neurons. The best performance is obtained by setting 70 neurons in the hidden layer. The output layer of the network contains 30 variables. They are defined as the temperature values at the 26 pressure levels of NCEP analysis, the surface temperature, and the tropopause temperature, height, and pressure.

A backpropagation network is trained by ‘‘supervised learning.’’ The network is presented with a series of pattern pairs, each consisting of an input pattern and an output pattern. The target output pattern is the desired response to the input pattern and is used to determine the error values in the network when the weights are adjusted. Both the training set and testing set of data are used in the training procedure. The neural network package processes the training automatically based on specified training criteria. Each pattern, which contains both input elements and output elements, in the training set is presented to the network in random order until the predetermined convergence criteria are met. At this time the network presents the input elements in the testing set and retrieves the output elements. Then the retrieved output elements are compared with the output elements in the testing set, and the averaged rms deviation of all the output elements are computed. The network parameters are saved if the averaged rms deviation is less than that computed previously. This process is repeated until no improvement is found for a specified number of test trials.

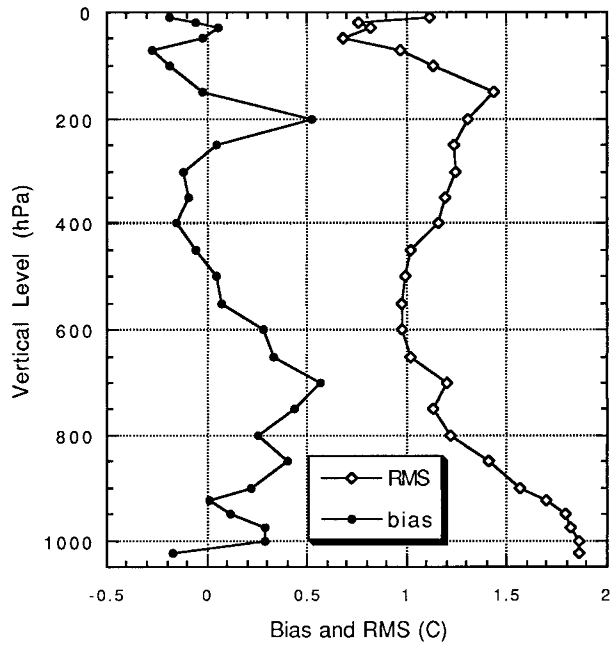

FIG. 2. The bias (dots) and rms deviation (diamonds) of temperature retrieval based on a regional dataset at levels from the surface to 10 hPa.

c. Regional temperature retrieval result

Before doing the network training, 20% of collocated patterns have been randomly extracted and set aside. This set of data has not been ‘‘seen’’ by the network during the training and testing. It is considered as a validation dataset to examine the performance of the network. Using the saved network parameters, the retrievals corresponding to each input pattern in the validation set are calculated. The bias and rms deviation are computed using the output patterns in the validation dataset as truth data. Figure 2 illustrates the resulting bias and rms deviation for the temperature profiles. The values for the surface temperatures are plotted below the 1000-hPa level, regardless of its surface pressure. In the figure, the dots represent the bias values and the diamonds represent the rms values.

Figure 2 shows that the biases of the retrievals are all within a small value of 0.6°C compared with the temperature profiles in the validation set. For the rms deviation plot, the largest values (1.9°C) are found at the surface and at the lowest level. For the lowest levels there are large horizontal gradients of temperature, and thus a broad temperature range is sampled. The complex surface radiative properties represent the most diverse temperature features and are major factors leading to the largest values of the rms deviations. The rms deviation decreases steadily toward the middle troposphere. The values decrease to 1°C throughout the 650–450-hPa levels. Contributed by the relatively more frequent temperature fluctuation near the tropopause, the rms deviations increase slightly up to 1.4°C around these levels. The rms values remain small in the stratosphere (between 0.7° and 1.1°C). At the tropopause level, the retrieval bias of tropopause temperature, height, and pressure are 0.47°C, −0.096 km, and 2.26 hPa, respectively, and the retrieval rms deviations of these variables are 3.35°C, 0.859 km, and 21.70 hPa, respectively.

During the early development of the microwave sounder retrieval work, Rigone and Stogryn (1977) used a statistical method referred to as the D-matrix algorithm to retrieve temperature profiles from the DMSP Special Sensor Microwave Temperature Sounder (SSM/T) measurement. They showed the rms deviation profiles for six geographically distinct atmospheres, which include spring profiles for tropic, midlatitude, and Arctic regions of the Northern Hemisphere, and fall profiles for tropic, midlatitude, and Antarctic regions of the Southern Hemisphere. Among these six types of atmospheric profiles, the Northern Hemisphere midlatitude profile is geographically closest to the samples of profiles used in the present study. The Northern Hemisphere midlatitude rms profiles presented by Rigone and Stogryn (1977) also showed a pattern of largest rms deviation at the 1000-hPa level, smaller deviation in the middle atmosphere, increasing values around the tropopause, and smaller values in the stratosphere. However, their rms values were much larger, with 7.2°C at 1000 hPa, around 2°C in the middle atmosphere, close to 3°C at the tropopause, and about 1.5°C in the stratosphere.

Recently, Butler et al. (1996) studied the temperature retrieval from DMSP SSM/T1 using neural networks. They compared the rms deviations of a neural network method with those of the standard D-matrix method. They found that the neural network retrievals are more accurate at all atmospheric levels. Their network rms profiles show the values of 4°C at the surface, 1.5°C in the middle atmosphere, approximately 2°C at the tropopause, and between 1° and 2°C in the stratosphere. These values represent a substantial improvement from the D-Matrix method.

Compared to the DMSP microwave sounder retrieval studies of Rigone and Stogryn (1977) and Butler et al. (1996), the *NOAA-15* AMSU-A retrieval of this study using the neural network significantly improves the retrieval accuracy at all levels. The better retrieval accuracy is attributed to the improvement of the microwave sensor and to the use of the nonlinear, multilayer neural network. Compared with the 7-channel SSM/T1 instruments, the 15-channel AMSU-A is able to provide much better vertical resolution. Furthermore, the reduced footprint size of 50 km of the AMSU-A produces much finer horizontal resolution than the 175-km footprint of the SSM/T1. Use of the neural network connects the complex, nonlinear relationship between channel measurements and atmospheric temperature profiles. Proper design and training of the neural network are key factors in translating the complex relationship into a mathematical function.

3. Global temperature profile retrieval

a. Global data

The neural network retrieval described in the previous section is based on regional data covering the eastern United States and the western Atlantic Ocean. To extend the neural network application to other regions, especially the polar regions, a series of global AMSU-A data are collected. This time series, taken during 9–14 January 2000, contains 43 randomly collected full orbits.

The AMSU-A data are processed into brightness temperatures and collocated with the corresponding NCEP global model Global Data Assimilation System analyses at 0000, 0600, 1200, and 1800 UTC. The neural network training dataset is constructed based on the subset of AMSU-A data that are within 0.5 h of the four NCEP analysis times and the subset of NCEP analysis data that are within the corresponding AMSU-A coverage. A total of 29 685 matched patterns are generated from the global dataset.

In building the retrieval database, the NCEP analysis is selected instead of radiosonde measurements for several reasons. First, while the high-vertical resolution radiosonde measurement well represents the atmosphere observed, the 15-channel AMSU-A measurement approximates a much smoother feature of the atmosphere. This smoother feature is better represented by the 26-level NCEP analysis. In the NCEP analysis structure, the grid values are adjusted to quality-controlled radiosondes, therefore the radiosondes have their presence in NCEP analysis, in a smoother fashion. Second, the majority of regular radiosonde measurement is taken over land surfaces. Very few are available over the oceans. Such a database would limit the retrieval scheme to land temperature profile retrieval only. Furthermore, over land, the unevenly distributed radiosonde network may lead to bias of the retrieval scheme toward the environment where denser radiosonde sites are located. And third, the location of the radiosonde measurement in the upper atmosphere usually drifts away from the balloon launch site. The AMSU-A measures the temperatures in an atmospheric column. This vertical column is better simulated by the NCEP analysis.

Similar to the regional data processing, the collocated data are divided into three datasets. Among the 29 685 collocated patterns, 20% (5937) are randomly extracted to construct a testing set for use during the network training phase. Another 20% are randomly extracted to construct a final validation dataset. The remaining 17 811 patterns builds the neural network training set.

b. Neural network for the global dataset

Different backpropagation architectures are examined. Unlike for the regional data case, in which using three-layer or four-layer networks were found to yield no significant difference in the performance of the network, for the global dataset, using a four-layer backpropagation network exhibits remarkable performance improvement from a three-layer network. The global dataset consists of data collected over a full range of geographical locations, from the polar regions to the Tropics, from mountain ridges to the oceans, and from deserts to water surfaces. An increase of a hidden layer in the neural network facilitates reserving the different characteristics of data sources. Further investigation of five-layer backpropagation networks reveals that they perform notably better than the four-layer networks. Therefore, five-layer backpropagation networks, with one input layer, three hidden layers, and one output layer, are chosen for the global data retrieval.

In constructing the five-layer backpropagation network architectures, it is found that using a hyperbolic tangent transfer function to propagate to each of the three hidden layers and a logistic transfer function to propagate to the output layers gives the optimum network performance. The data variables are scaled linearly into the range $[-1, 1]$ for effectiveness in data processing. The numbers of processing elements in the input and output layers are the same as those in the neural network of regional retrieval. For the hidden layers, setting the number of each layer's neurons in the range of 30 to 70 results in good network performance. The best performance is obtained by setting 50, 52, and 50 neurons, respectively, for the first, second, and third hidden layer.

In this neural network structure, the output layer contains 30 variables, including 26 temperature values corresponding to the NCEP analysis pressure levels, one surface temperature value, and three tropopause variables. Experiments are carried out to structure the output layer of the neural network to contain 26 temperature profile values and one surface temperature value only. The resulting neural network performance of temperature profiles is slightly better than the network that includes the tropopause variables in the output layer. However, when a network is structured with only three tropopause variables in the output layer, the network performance of tropopause variables is not as good as the network that contains both temperature profiles and tropopause variables in the output layer. This indicates that including temperature profile values in the neural network output layer plays a role in better defining and restricting the tropopause variable value ranges.

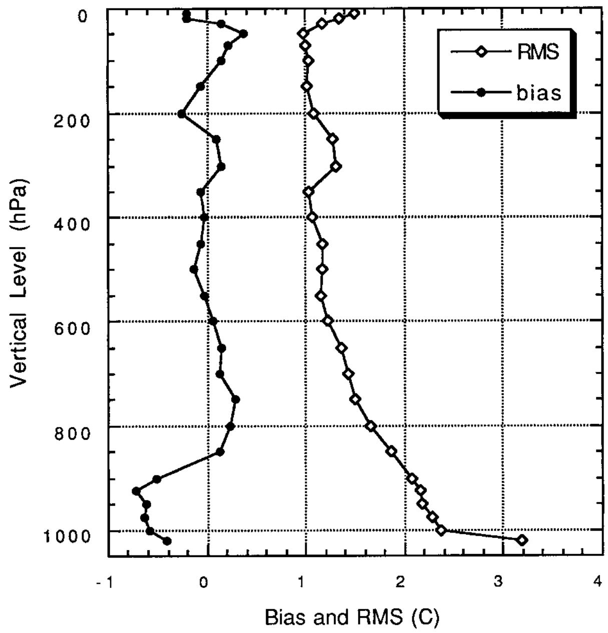

FIG. 3. The bias (dots) and rms deviation (diamonds) of temperature retrieval based on a global dataset at levels from the surface to 10 hPa.

c. Neural network retrieval result

Figure 3 illustrates the bias and rms deviation of the global-based temperature profile retrieval. The statistics are obtained by applying the trained neural network to the 20% of the collocated dataset that is not used during training. The collocated NCEP temperature profiles in this validation dataset are considered as truth data in calculating the bias and rms deviation. In the figure, the values for the surface temperature are plotted below the 1000-hPa level. The dots represent the bias values and the diamonds represent the rms values.

The figure shows that the bias values of global-based retrieval are within ± 0.4°C for temperature profiles at all the levels at and above 850 hPa. At surface and near-surface levels, there is a negative bias trend with values in the range of -0.4° to -0.7°C. The surface has the largest rms value of 3.2°C. The rms value reduces to 2.4°C at the 1000-hPa level and steadily decreases to between 1.0° and 1.2°C in the middle of the troposphere. At the tropopause the rms value increases slightly to

1.3°C, then decreases to 1.0°C in the lower stratosphere. At the top levels, the rms values range from 1.0° to 1.5°C. For the tropopause variables, the rms deviation values for the tropopause temperature, height, and pressure are 1.9°C, 0.58 km, and 18.1 hPa, respectively. Compared with the regional retrieval results as discussed previously, the global temperature profile rms deviations at most of the levels are slightly greater than the regional rms deviations. The increase of deviations is mainly due to the much more different atmospheric conditions that are included in the database. The diversity of atmospheric conditions in the training database is an important factor in determining the generalized applicability of the trained neural network. Therefore, despite the slight increase of the rms deviations, the neural network trained by the global dataset is considered a better retrieval network for operational applications than the network trained by the regional dataset.

d. Comparison with linear regression approach

The linear regression approach is a traditional method in retrieving geophysical variables from satellite measurements. To examine the difference of a neural network approach from a linear regression approach, the temperature profile retrieval using a linear regression method is carried out.

The same global data training set is used to derive the linear regression coefficients. In a linear regression relationship,

$$y_i = C_{i,0} + \sum_{j=1}^{n} C_{i,j} x_j, \qquad (4)$$

in which y_i represents one of the 30 retrieval variables, $C_{i,j}$ are linear regression coefficients, $C_{i,0}$ is the constant coefficient term corresponding to y_i, x_j is the input variable, and n (=17) is the number of input variables. The input variable contains the 15 AMSU-A channel brightness temperatures, secant of satellite viewing angle, and position of the pixel in a scan line. The 30 retrieval variables are the temperature values at the 26 pressure levels; the surface temperature; and the tropopause temperature, height, and pressure. For each retrieval variable, the linear regression coefficients are determined by minimizing the sum of the square deviations of the predicted values of y_i from the NCEP analysis values, a procedure referred to as a least squares fit.

After obtaining the coefficients for each retrieval variable, Eq. (4) is applied to the same validation dataset that was used to validate the global neural network retrieval. The linear regression rms deviations are calculated, and the results are plotted as triangles in Fig. 4. For comparison, the rms deviations of the global-based neural network approach (diamonds) are also plotted in Fig. 4. For the linear regression approach, the rms deviation of surface temperature is 4.7°C. The rms deviations decrease to around 2°C in the middle and upper troposphere. At the tropopause levels, the values increase to 2.7°C. The values decrease to around 1.5°C in the stratosphere. Compared with the linear regression approach, the rms deviation of the neural network is 1.5°C smaller at the surface and 2.2°C smaller at the 1000-hPa level. From 800 through 300 hPa, the rms deviation differences of the two approaches are in the range of 0.7° to 0.9°C. The neural network performs significantly better around the tropopause levels, at which the differences are approximately 1.5°C. In the stratosphere, the linear regression rms deviations are larger than those of the neural network approach at all the levels, but the differences become small at the highest levels. For the tropopause variable retrieval, the linear regression rms deviations of tropopause temperature, height, and pressure are 3.5°C, 0.92 km, and 28.0 hPa, respectively. These rms deviation values are significantly larger than those of the neural network retrieval.

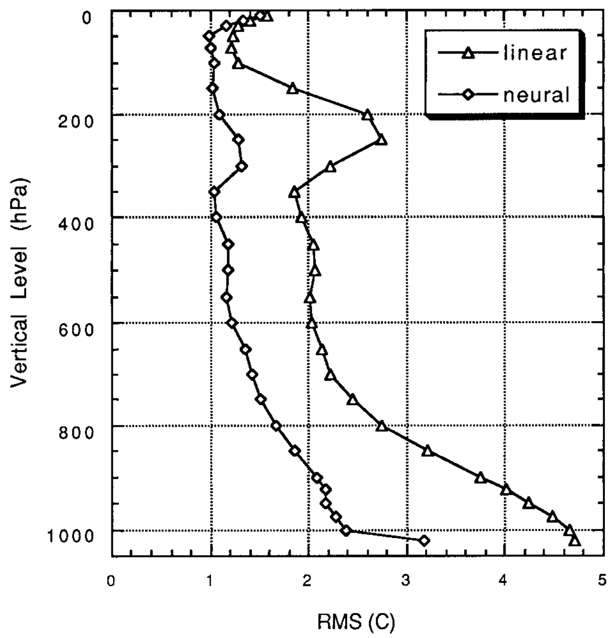

FIG. 4. Comparison of rms deviations between the neural network approach and the linear regression approach.

4. Conclusions

Backpropagation neural networks are found capable of connecting the nonlinear relation between AMSU-A channel measurements and atmospheric temperatures. The retrieval result shows significantly better accuracy than the operational DMSP SSM/T1 retrieval (Rigone and Stogryn 1977; Butler et al. 1996) due to the use of neural network method and improvement in the satellite microwave sounder. The rms deviations of the temperature profiles achieved by using backpropagation neural networks are less than 3.2°C at the surface, in the range of 1.0° to 1.2°C in the middle troposphere, less than 1.5°C at the tropopause, and in the range of 1.0° to 1.5°C throughout the stratosphere. The neural network is also capable of simultaneously retrieving tropopause temperature, height, and pressure. The optimized sets of backpropagation network parameters are saved during network training, and they can be easily applied to AMSU-A retrieval operationally. Using the neural networks examined, the retrieval of temperature profiles and tropopause variables are based on 15-channel AMSU-A measurement, the satellite viewing angle, and the relative position of the pixel in a scan line only, without requiring additional datasets.

The regional retrieval is based on the datasets from July 1998 and January 1999 covering the eastern United States and western Atlantic Ocean, and the global retrieval uses the global coverage data taken from six days of January 2000. Though good retrieval performances are obtained from the neural networks, the data collected in the training sets have not covered all the atmospheric conditions. For example, seasonal variation of the atmosphere has not been sufficiently represented in the training sets. To generalize the retrieval for all seasons, additional data for temperature conditions not represented by these collected data are needed to augment the training database. Based on the performance of the cases studied, neural networks show great promise for use in operational retrieving of temperature profiles from the AMSU-A measurement.

Acknowledgments. The *NOAA-15* data were kindly provided by Prof. Oscar Huh and his coworkers at the Earth Scan Laboratory, Coastal Studies Institute, Louisiana State University. I wish to thank Peter Nowak for assistance in data processing, Robert Bernstein for encouragement, and the journal reviewers for critical reviews and helpful suggestions.

REFERENCES

Butler, C. T., R. V. Z. Meredith, and A. P. Stogryn, 1996: Retrieving atmospheric temperature parameters from DMSP SSM/T-1 data with a neural network. *J. Geophys. Res.,* **101,** 7075–7083.

Chauvin, Y., and D. E. Rumelhart, 1995: *Backpropagation: Theory, Architectures, and Applications.* Lawrence Erlbaum Associates, 561 pp.

Chen, K. S., Y. C. Tzeng, and P. C. Chen, 1999: Retrieval of ocean winds from satellite scatterometer by a neural network. *IEEE Trans. Geosci. Remote Sens.,* **37,** 247–256.

Churnside, J. H., T. A. Stermitz, and J. A. Schroeder, 1994: Temperature profiling with neural network inversion of microwave radiometer data. *J. Atmos. Oceanic Technol.,* **11,** 105–109.

Cichocki, A., and R. Unbehauen, 1993: *Neural Networks for Optimization and Signal Processing.* John Wiley and Sons, 526 pp.

Cornford, D., I. T. Nabney, and C. M. Bishop, 1999: Neural network–based wind vector retrieval from satellite scatterometer data. *Neural Comput. Appl.,* **8,** 206–217.

Dayhoff, J. E., 1990: *Neural Network Architectures—An Introduction.* Van Nostrand Reinhold, 259 pp.

Goodrum, G., K. B. Kidwell, and W. Winston, 1999: NOAA KLM user's guide. National Oceanic and Atmospheric Administration, 870 pp.

Grody, N. C., D. G. Gray, C. S. Novak, J. S. Prasad, M. Piepgrass, and C. A. Dean, 1985: Temperature soundings from the DMSP microwave sounder. *RSRM'87 Advances in Remote Sensing Retrieval Methods,* A. Deepak, H. E. Fleming, and M. T. Chahine, Eds., A. Deepak Publishing, 249–265.

Hertz, J., A. Krogh, and R. G. Palmer, 1991: *Introduction to the Theory of Neural Computation.* Addison-Wesley Publishing Company, 327 pp.

Hsieh, W. W., and B. Tang, 1998: Applying neural network models to prediction and data analysis in meteorology and oceanography. *Bull. Amer. Meteor. Soc.,* **79,** 1855–1870.

Isaacs, R. G., 1989: A unified retrieval methodology for the DMSP meteorological sensors. *RSRM'87 Advances in Remote Sensing Retrieval Methods,* A. Deepak, H. E. Fleming, and M. T. Chahine, Eds., A. Deepak Publishing, 203–214.

Jones, C., P. Peterson, and C. Gautier, 1999: A new method for deriving ocean surface specific humidity and air temperature: An artificial neural network approach. *J. Appl. Meteor.,* **38,** 1229–1245.

Kramer, H. J., 1996: *Observation of the Earth and Its Environment: Survey of Missions and Sensors.* Springer, 466 pp.

Mo, T., 1999: AMSU-A antenna pattern corrections. *IEEE Trans. Geosci. Remote Sens.,* **37,** 103–112.

Reale, A. L., 1993: Sea surface temperature estimation from the DMSP operational linescan system using a SSM/I-derived water vapor correction. *Geophys. Res. Lett.,* **20,** 583–586.

Rieu, H., J. Escobar, N. A. Scott, and A. Chedin, 1996: SSM/T forward modeling using neural networks. *J. Quant. Spectrosc. Radiat. Transfer,* **56,** 821–833.

Rigone, J. L., and A. P. Stogryn, 1977: Data processing for the DMSP microwave radiometer system. *Proceedings of the Eleventh International Symposium on Remote Sensing of the Environment,* University of Michigan Press, 1599–1608.

Rojas, R., 1996: *Neural Networks—A Systematic Introduction.* Springer-Verlag, 502 pp.

Rosenkranz, P. W., 1994: A rapid transmittance algorithm for microwave sounding frequencies. Preprints, *Int. Geoscience and Remote Sensing Symp.,* Pasadena, CA, IEEE Geoscience and Remote Sensing Society, 1–10.

——, K. D. Hutchison, and K. R. Hardy, 1994: Simulations of temperature retrieval accuracy in the middle and lower atmosphere using oxygen microwave lines. Preprints, *Seventh Conf. on Satellite Meteorology and Oceanography,* Monterey, CA, Amer. Meteor. Soc., J11–J14.

——, ——, ——, and M. S. Davis, 1997: An assessment of the impact of satellite microwave sounder incidence angle and scan geometry on the accuracy of atmospheric temperature profile retrievals. *J. Atmos. Oceanic Technol.,* **14,** 488–494.

Stogryn, A. P., C. T. Butler, and T. J. Bartolac, 1994: Ocean surface wind retrievals from special sensor microwave imager data with neural networks. *J. Geophys. Res.,* **99,** 981–984.

Tsintikidis, D., J. L. Haferman, E. N. Anagnostou, W. F. Krajewski, and T. F. Smith, 1997: A neural network approach to estimating rainfall from spaceborne microwave data. *IEEE Trans. Geosci. Remote Sens.,* **35,** 1079–1093.

Ward, B., and S. Redfern, 1999: A neural network model for predicting the bulk–skin temperature difference at the sea surface. *Int. J. Remote Sens.,* **20,** 3533–3548.

Ward Systems Group, Inc., 1996: *NeuroShell 2.* Ward Systems Group, Inc., 298 pp.

Yang, C. C., S. O. Prasher, and G. R. Mehuys, 1997: An artificial neural network to estimate soil temperature. *Can. J. Soil Sci.,* **77,** 421–429.