A Neural-Network Technique for the Retrieval of Atmospheric Temperature and Moisture Profiles From High Spectral Resolution Sounding Data

William J. Blackwell, *Member, IEEE*

***Abstract*—A novel statistical method for the retrieval of atmospheric temperature and moisture profiles has been developed and evaluated with simulated clear-air and observed partially cloudy sounding data from the Atmospheric InfraRed Sounder (AIRS) and the Advanced Microwave Sounding Unit (AMSU). The algorithm is implemented in two stages. First, a projected principal components (PPC) transform is used to reduce the dimensionality of and optimally extract geophysical profile information from the cloud-cleared infrared radiance data. Second, a multilayer feedforward neural network (NN) is used to estimate the desired geophysical parameters from the PPCs. For the first time, NN temperature and moisture retrievals are presented using actual microwave and hyperspectral infrared observations of cloudy atmospheres, over both ocean and land (with variable terrain elevation), and at all sensor scan angles. The performance of the NN retrieval method (henceforth referred to as the PPC/NN method) was evaluated using global Earth Observing System Aqua orbits colocated with European Center for Medium-range Weather Forecasting fields for seven days throughout 2002 and 2003. Over 350 000 partially cloudy footprints were used in the study, and retrieval performance was compared with the AIRS Science Team Level-2 retrieval algorithm (version 3). Performance compares favorably with that obtained with simulated clear-air observations from the NOAA88b radiosonde set of approximately 7500 profiles. The PPC/NN method requires significantly less computation than traditional variational retrieval methods, while achieving comparable performance.**

***Index Terms*—Advanced Microwave Sounding Unit (AMSU), Advanced Technology Microwave Sounder (ATMS), Atmospheric InfraRed Sounder (AIRS), canonical correlations, Crosstrack Infrared Sounder (CrIS), humidity, hyperspectral, infrared, Infrared Atmospheric Sounding Interferometer (IASI), inversion, microwave, moisture, neural networks (NNs), principal components, retrieval, sounding, temperature.**

I. Introduction

MODERN atmospheric sounders measure radiance with unprecedented resolution (spatial, spectral, and temporal) and accuracy. For example, the Atmospheric Infrared Sounder (AIRS), operational since late 2002, provides a spatial resolution of $\sim$15 km, a spectral resolution of $\nu/\Delta\nu \approx 1200$ (with 2378 channels from 650–2675 cm^{-1}), and a radiometric accuracy on the order of ± 0.2 K. Approximately 90% of the Earth's atmosphere is measured (in the horizontal dimension) every 12 h or so. This wealth of data presents two major challenges from the point of view of retrieval algorithm development. The first concerns the robustness of the retrieval operator and involves maximal use of the geophysical content of the radiance data with minimal interference from instrument and atmospheric noise. The second concerns computational efficiency, where it is desirable to implement a robust algorithm within a given computational budget.

In this paper, a nonlinear retrieval algorithm is presented that offers the numerical stability and efficiency of statistical methods without sacrificing the accuracy of physical, model-based methods. The algorithm is implemented in two stages: a linear transform is first used to reduce the dimensionality of and optimally extract geophysical information from the radiance data, and a multilayer feedforward neural network (NN) is subsequently used to estimate the desired geophysical profiles. This approach allows relatively small NNs to be used and consequently yields better retrieval accuracies, reduced sensitivity to input noise, improved resistance to instabilities (overfitting, for example), and reduced computational burden.

This paper is organized into seven sections. First, the application of principal components transforms to hyperspectral sounding data is briefly reviewed, and a new approach is presented where sensor radiances are projected into a subspace that reduces spectral redundancy and maximizes the resulting correlation to a given parameter. This method is very similar to the concept of canonical correlations introduced by Hotelling over 70 years ago [1], but its application in the hyperspectral sounding context is new. Second, the use of multilayer feedforward NNs for geophysical parameter retrieval from hyperspectral measurements (first proposed in 1993 [2]) is reviewed, and an overview of the network parameters used in this work is given. The combination of the projected principal components (PPC) radiance compression operator derived in Section II with the NN described in Section III is discussed in Section IV. Performance analyses comparing the PPC/NN algorithm to traditional retrieval methods are presented in Sections V and VI, including both simulated clear-air and observed partially cloudy data from the AIRS and Advanced Microwave Sounding Unit (AMSU) sensors. While only AIRS/AMSU observations are considered in this paper, the PPC/NN algorithm could easily be applied to other microwave and hyperspectral infrared sensor suites, such as the Crosstrack Infrared Sounder (CrIS) and the Advanced Technology Microwave Sounder (ATMS)

Manuscript received November 1, 2004; revised June 26, 2005. This work was supported by the National Oceanic and Atmospheric Administration under Air Force Contract F19628-00-C-0002. Opinions, interpretations, conclusions, and recommendations are those of the author and are not necessarily endorsed by the U.S. Government.

The author is with the Lincoln Laboratory, Massachusetts Institute of Technology, Lexington, MA 02420 USA (e-mail: wjb@ll.mit.edu).

Digital Object Identifier 10.1109/TGRS.2005.855071

scheduled for launch on the NPP/NPOESS satellite series, the Infrared Atmospheric Sounding Interferometer (IASI) and AMSU scheduled for launch on the EUMETSAT Polar System (EPS), and the Hyperspectral Environmental Suite (HES) scheduled for launch into geostationary orbit beginning with the GOES-R satellite. Section VI also includes retrieval sensitivity analyses; the effects of sensor scan angle, orbit type (ascending or descending), cloud fraction, and training set comprehensiveness on performance are explored. Finally, an overview of future work is given, including a discussion on how the PPC/NN algorithm could be used in operational systems.

II. Dimensionality Reduction With Principal Components Transforms

Principal components (PC) transforms can be used to represent radiance measurements in a statistically compact form, enabling subsequent retrieval operators to be substantially more efficient and robust (see [3], for example). Furthermore, measurement noise can be dramatically reduced through the use of PC filtering [4], [5], and it has also been shown [6] that PC transforms can be used to represent variability in high spectral resolution radiances resulting from clouds. In the following sections, several variants of the PC transform are briefly discussed, with emphasis focused on the ability of each to extract geophysical information from the noisy radiance data.

A. PC Transform

The PC transform is a linear, orthonormal operator[1] $\mathbf{Q}_r^T$ which projects a noisy m-dimensional radiance vector $\widetilde{R} = R + \Psi$ into an r-dimensional ($r \leq m$) subspace. The additive noise vector Ψ is assumed to be uncorrelated with the radiance vector R and is characterized by the noise covariance matrix $\mathbf{C}_{\Psi\Psi}$. The "principal components" of $\widetilde{R}$, i.e., $\widetilde{P} = \mathbf{Q}_r^T\widetilde{R}$, have two desirable properties: 1) the components are statistically uncorrelated and 2) the reduced-rank reconstruction error

$$c_1(\cdot) = E\left[(\widehat{\widetilde{R}}_r - \widetilde{R})^T(\widehat{\widetilde{R}}_r - \widetilde{R})\right] \tag{1}$$

where $\widehat{\widetilde{R}}_r \triangleq \mathbf{G}_r\widetilde{R}$ for some linear operator $\mathbf{G}_r$ with rank r is minimized when $\mathbf{G}_r = \mathbf{Q}_r\mathbf{Q}_r^T$. The rows of $\mathbf{Q}_r^T$ contain the r most significant (ordered by descending eigenvalue) eigenvectors of the noisy data covariance matrix $\mathbf{C}_{\widetilde{R}\widetilde{R}} = \mathbf{C}_{RR} + \mathbf{C}_{\Psi\Psi}$.

B. NAPC Transform

Cost criteria other than (1) are often more suitable for typical hyperspectral compression applications. For example, it might be desirable to reconstruct the noise-free radiances and filter the noise. The cost equation thus becomes

$$c_2(\cdot) = E\left[(\widehat{R}_r - R)^T(\widehat{R}_r - R)\right] \tag{2}$$

where $\widehat{R}_r \triangleq \mathbf{H}_r\widetilde{R}$ for some linear operator $\mathbf{H}_r$ with rank r. The noise-adjusted principal components (NAPC) transform [7], where $\mathbf{H}_r = \mathbf{C}_{\Psi\Psi}^{1/2}\mathbf{W}_r\mathbf{W}_r^T\mathbf{C}_{\Psi\Psi}^{-1/2}$ and $\mathbf{W}_r^T$ contains the r most significant eigenvectors of the whitened noisy covariance matrix $\mathbf{C}_{\widetilde{W}\widetilde{W}} = \mathbf{C}_{\Psi\Psi}^{-1/2}(\mathbf{C}_{RR} + \mathbf{C}_{\Psi\Psi})\mathbf{C}_{\Psi\Psi}^{-1/2}$, maximizes the signal-to-noise ratio of each component and is superior to the PC transform for most noise-filtering applications where the noise statistics are known *a priori*.

C. Projected PC Transform

It is often unnecessary to require that the principal components be uncorrelated, and linear operators can be derived that offer improved performance over PC transforms for minimizing cost functions such as (2). It can be shown [8] that the optimal linear operator with rank r that minimizes (2) is

$$\mathbf{L}_r = \mathbf{E}_r\mathbf{E}_r^T\mathbf{C}_{RR}(\mathbf{C}_{RR} + \mathbf{C}_{\Psi\Psi})^{-1} \tag{3}$$

where $\mathbf{E}_r = [E_1|E_2|\cdots|E_r]$ are the r most significant eigenvectors of $\mathbf{C}_{RR}(\mathbf{C}_{RR} + \mathbf{C}_{\Psi\Psi})^{-1}\mathbf{C}_{RR}$. Examination of (3) reveals that the Wiener-filtered radiances are projected onto the r-dimensional subspace spanned by $\mathbf{E}_r$. It is this projection that motivates the name "projected principal components." An orthonormal basis for this r-dimensional subspace of the original m-dimensional radiance vector space $\mathcal{R}$ is given by the r most significant right eigenvectors $\mathbf{V}_r$ of the reduced-rank linear regression matrix $\mathbf{L}_r$ given in (3). We then define the projected principal components of $\widetilde{R}$ as

$$\widetilde{P} = \mathbf{V}_r^T\widetilde{R}. \tag{4}$$

Note that the elements of $\widetilde{P}$ are correlated, as $\mathbf{V}_r^T(\mathbf{C}_{RR} + \mathbf{C}_{\Psi\Psi})\mathbf{V}_r$ is not a diagonal matrix.

Another useful application of the PPC transform is the compression of spectral radiance information that is correlated with a geophysical parameter, such as the temperature profile. The r-rank linear operator that captures the most radiance information which is correlated to the temperature profile is similar to (3) and is given below

$$\mathbf{L}_r = \mathbf{E}_r\mathbf{E}_r^T\mathbf{C}_{TR}(\mathbf{C}_{RR} + \mathbf{C}_{\Psi\Psi})^{-1} \tag{5}$$

where $\mathbf{E}_r = [E_1|E_2|\cdots|E_r]$ are the r most significant eigenvectors of $\mathbf{C}_{TR}(\mathbf{C}_{RR}+\mathbf{C}_{\Psi\Psi})^{-1}\mathbf{C}_{RT}$, and $\mathbf{C}_{TR}$ is the cross-covariance of the temperature profile and the spectral radiance.

D. Evaluation of Compression Performance Using Two Different Metrics

The compression performance of each of the PC transforms discussed previously was evaluated using two performance metrics. First, we seek the transform that yields the best (in the sum-squared sense) reconstruction of the noise-free radiance spectrum given a noisy spectrum. Thus, we seek the optimal reduced-rank linear filter. The second performance metric is quite different and is based on the temperature retrieval performance in the following way. A radiance spectrum is first compressed using each of the PC transforms for a given number of coefficients. The resulting coefficients are then used in a linear regression to estimate the temperature profile.

The results that follow were obtained using simulated, clear-air radiance intensity spectra from an AIRS-like sounder. Approximately 7500 1750-channel radiance vectors were generated with spectral coverage from approximately 4–15 μm

[1]The following mathematical notation is used in this paper: $(\cdot)^T$ denotes the transpose, $\widetilde{(\cdot)}$ denotes a noisy random vector, and $\widehat{(\cdot)}$ denotes an estimate of a random vector. Matrices are indicated by bold upper case, vectors by upper case, and scalars by lower case.

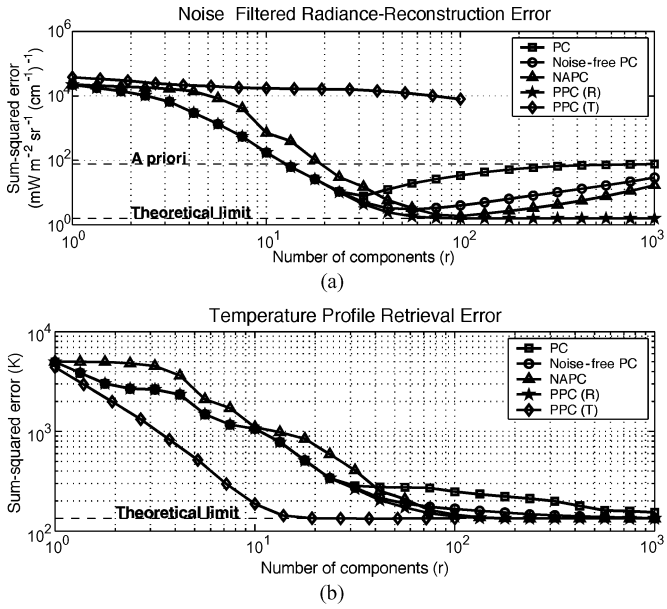

Fig. 1. Performance comparisons of the PC (where the components are derived from both noisy and noise-free radiances), NAPC, and PPC transforms for a hypothetical 1750-channel infrared (4–15 μm) sounder. Two projected principal components transforms were considered: PPC (R) and PPC (T), which are, respectively: 1) maximum representation of noise-free radiance energy and 2) maximum representation of temperature profile energy. The first plot shows the sum-squared error of the reduced-rank reconstruction of the noise-free spectral radiances. The second plot shows the temperature profile retrieval error (trace of the error covariance matrix) obtained using linear regression with r components.

using the NOAA88b radiosonde set (see Section V). The simulated intensities were expressed in spectral radiance units (milliwatts per square meter per steradian per inverse centimeter).

1) PC Filtering: Fig. 1(a) shows the sum-squared radiance distortion (1) as a function of the number of components used in the various PC decomposition techniques. The *a priori* level indicates the sum-squared error due to sensor noise. Results from two variants of the PC transform are plotted, where the first variant (the "PC" curve) uses eigenvectors of $\mathbf{C}_{\widetilde{R}\widetilde{R}}$ as the transform basis vectors, and the second variant (the "noise-free PC" curve) uses eigenvectors of $\mathbf{C}_{RR}$ as the transform basis vectors. It is shown in Fig. 1(a) that the PPC reconstruction of noise-free radiances (PPC[R]) yields lower distortion than both the PC and NAPC transforms for any number of components (r). It is noteworthy that the "PC" and "noise-free PC" curves never reach the theoretically optimal level, defined by the full-rank Wiener filter. Furthermore, the PPC distortion curves decrease monotonically with coefficient number, while all the PC distortion curves exhibit a local minimum, after which the distortion increases with coefficient number as noisy, high-order terms are included. The noise in the high-order PPC terms is effectively zeroed out, because it is uncorrelated with the spectral radiances.

2) PC Regression: The PC coefficients derived in the previous example are now used in a linear regression to estimate the temperature profile. Shown in Fig. 1(b) is the temperature profile error (integrated over all altitude levels) as a function of the number of coefficients used in the linear regression, for each of the PC transforms. To reach the theoretically optimal value achieved by linear regression with all channels requires approximately 20 PPC coefficients, 200 NAPC coefficients, and 1000 PC coefficients. Thus, the PPC transform results in a factor of ten improvement over the NAPC transform when compressing temperature-correlated radiances (20 versus 200 coefficients required), and approximately a factor of 100 improvement over the original spectral radiance vector (20 versus 1750). Note that the first guess in the AIRS Science Team Level-2 retrieval uses a linear regression derived from approximately 60 of the most significant NAPC coefficients of the 2378-channel AIRS spectrum (in units of brightness temperature) [3]. Results for the moisture profile are similar, although more coefficients (typically 35 versus 20 for temperature) are needed because of the higher degree of nonlinearity in the underlying physical relationship between atmospheric moisture and the observed spectral radiance. This substantial compression enables the use of relatively small (and thus very stable and fast) neural-network estimators to retrieve the desired geophysical parameters.

It is interesting to consider the two variants of the PPC transform shown in Fig. 1, namely PPC(R), where the basis for the noise-free radiance subspace is desired, and PPC(T), where the basis for only the temperature profile information is desired. As shown in Fig. 1(a), the PPC(T) transform poorly represents the noise-free radiance space, because there is substantial information that is uncorrelated with temperature (and thus ignored by the PPC(T) transform) but correlated with the noise-free radiance. Conversely, the PPC(R) transform offers a significantly less compact representation of temperature profile information [see Fig. 1(b)], because the transform is representing information that is not correlated with temperature and thus superfluous when retrieving the temperature profile.

III. Neural-Network Inversion of Temperature and Moisture Profiles

A neural network is an interconnection of simple computational elements, or nodes, with activation functions that are usually nonlinear, monotonically increasing, and differentiable. NNs are able to deduce input-output relationships directly from the training ensemble without requiring underlying assumptions about the distribution of the data. Furthermore, a neural network with only a single hidden layer of a sufficient number of nodes with nonlinear activation functions is capable of approximating any real-valued continuous scalar function to a given precision over a finite domain [9].

Consider a multilayer feedforward NN consisting of an input layer, an arbitrary number of hidden layers (usually one or two), and an output layer. The hidden layers typically contain sigmoidal activation functions of the form $z_j = \tanh(a_j)$, where $a_j = \sum_{i=1}^{d} w_{ji} x_i + b_j$. The output layer is typically linear. The weights (w_{ji}) and biases (b_j) for the jth neuron are chosen to minimize a cost function over a set of P training patterns. A common choice for the cost function is the sum-squared error, defined as

$$E(\mathbf{w}) = \frac{1}{2} \sum_{p} \sum_{k} \left(t_k^{(p)} - y_k^{(p)} \right)^2 \tag{6}$$

where $y_k^{(p)}$ and $t_k^{(p)}$ denote the network outputs and target responses, respectively, of each output node k given a pattern p,

and $\mathbf{w}$ is a vector containing all the weights and biases of the network. The "training" process involves iteratively finding the weights and biases that minimize the cost function through some numerical optimization procedure. Second-order methods are commonly used, where the local approximation of the cost function by a quadratic form is given by

$$E(\mathbf{w}+\mathbf{dw}) \approx E(\mathbf{w}) + \nabla E(\mathbf{w})^T \mathbf{dw} + \frac{1}{2}\mathbf{dw}^T \nabla^2 E(\mathbf{w})\mathbf{dw} \tag{7}$$

where $\nabla E(\mathbf{w})$ and $\nabla^2 E(\mathbf{w})$ are the gradient vector and the Hessian matrix of the cost function, respectively. Setting the derivative of (7) to zero and solving for the weight update vector $\mathbf{dw}$ yields

$$\mathbf{dw} = -\left[\nabla^2 E(\mathbf{w})\right]^{-1} \nabla E(\mathbf{w}). \tag{8}$$

Direct application of (8) is difficult in practice, because computation of the Hessian matrix (and its inverse) is nontrivial and usually needs to be repeated at each iteration. For sum-squared error cost functions, it can be shown that

$$\nabla E(\mathbf{w}) = \mathbf{J}^T \mathbf{e} \tag{9}$$

$$\nabla^2 E(\mathbf{w}) = \mathbf{J}^T \mathbf{J} + \mathbf{S} \tag{10}$$

where $\mathbf{J}$ is the Jacobian matrix that contains first derivatives of the network errors with respect to the weights and biases, $\mathbf{e}$ is a vector of network errors, and $\mathbf{S} = \sum_{p=1}^{P} \mathbf{e}_p \nabla^2 \mathbf{e}_p$. The Jacobian matrix can be computed using a standard backpropagation technique [10] that is significantly more computationally efficient than direct calculation of the Hessian matrix [11]. However, an inversion of a square matrix with dimensions equal to the total number of weights and biases in the network is required. For the Gauss–Newton method, it is assumed that $\mathbf{S}$ is zero (a reasonable assumption only near the solution), and the update (8) becomes

$$\mathbf{dw} = -[\mathbf{J}^T\mathbf{J}]^{-1}\mathbf{J}\mathbf{e}. \tag{11}$$

The Levenberg–Marquardt modification [12] to the Gauss–Newton method is

$$\mathbf{dw} = -[\mathbf{J}^T\mathbf{J} + \mu\mathbf{I}]^{-1}\mathbf{J}\mathbf{e}. \tag{12}$$

As μ varies between zero and ∞, $\mathbf{dw}$ varies continuously between the Gauss–Newton step and steepest descent. The Levenberg–Marquardt method is thus an example of a model trust region approach in which the model (in this case the linearized approximation of the error function) is trusted only within some region around the current search point [13]. The size of this region is governed by the value μ.

The use of multilayer feedforward NNs, such as the multilayer perceptron (MLP), to retrieve temperature profiles from hyperspectral radiance measurements has been addressed by several investigators (see [14] and [15], for example). Neural-network retrieval of moisture profiles from hyperspectral data is relatively new [16], but follows the same methodology used to retrieve temperature.

IV. PPC/NN Algorithm

A first attempt to combine the properties of both neural-network estimators and principal component transforms for the inversion of microwave radiometric data to retrieve atmospheric temperature and moisture profiles is reported in [17], and a more recent study with hyperspectral data is presented in [16]. A conceptually similar approach is taken in this work by combining the PPC compression technique described in Section II with the NN estimator discussed in Section III. As shown in Section II, PPC compression offers substantial performance advantages over traditional PCA and is the cornerstone of the present work.

A. Network Topology

All multilayer perceptrons used in the PPC/NN algorithm are composed of one or two hidden layers of nonlinear (hyperbolic tangent) nodes and an output layer of linear nodes. For the temperature retrieval, 25 PPC coefficients are input to six NNs, each with a single hidden layer of 15 nodes. Separate NNs are used for different vertical regions of the atmosphere; a total of six networks are used to estimate the temperature profile at 65 points from the surface to 50 mbar. For the water vapor retrieval, 35 PPC coefficients are input to nine NNs, each with a single hidden layer of 25 nodes. The water vapor profile (mass mixing ratio) is estimated at 58 points from the surface to 75 mbar. These network parameters were determined largely through empirical analyses. Work is underway to dynamically optimize these parameters as the neural network is trained. Separate training and testing datasets are used and are discussed in more detail in Sections V and VI.

B. Network Training

The weights and biases were initialized using the Nguyen–Widrow method [18]. The neural network was trained using the Levenberg–Marquardt backpropagation algorithm discussed in Section III. For each epoch, the μ parameter was initialized to 0.001. If a successful step was taken (i.e., $E(\mathbf{w}+\mathbf{dw}) < E(\mathbf{w})$), then μ was decreased by a factor of ten. If the current step was unsuccessful, the value of μ was increased by a factor of ten until a successful step could be found (or until μ reached 10^{10}). The network training was stopped when the error on a separate dataset did not decrease for ten consecutive epochs. The sensor noise was changed on each training epoch to desensitize the network to radiance measurement errors.

V. Retrieval Performance Comparisons With Simulated Clear-Air AIRS Radiances

In this section, the temperature and moisture profile retrieval performance is compared for three methods using simulated clear-air AIRS radiances. The three retrieval methods are: 1) linear regression; 2) the PPC/NN algorithm; and 3) an iterated minimum-variance (IMV) technique.

A. Simulation of AIRS Radiances

The Atmospheric Infrared Sounder (AIRS) is a high-resolution grating spectrometer with 2378 channels from 650–2675 cm^{-1} [19]. Nominal $\mathrm{NE}\Delta\mathrm{T}$ values used in this study were obtained from the AIRS Science Team, and are shown in Table I. These values were converted to units of

TABLE I
ASSUMED NEΔT NOISE CHARACTERISTICS OF AIRS AT 250 K

ν (cm^{-1})	NEΔT (K)	ν (cm^{-1})	NEΔT (K)
648.93	0.370	980.39	0.100
684.93	0.320	2325.58	0.100
740.19	0.300	2439.02	0.140
740.74	0.180	2564.10	0.200
847.46	0.140	2702.70	0.250
900.90	0.110		

TABLE II
MEANS AND STANDARD DEVIATIONS OF THE SURFACE EMISSIVITY HINGE POINTS

ν (cm^{-1})	Mean	Standard deviation
769	0.9501	0.0102
909	0.9336	0.0170
1111	0.9172	0.0238
2105	0.9007	0.0303
2500	0.8844	0.0367
2857	0.8678	0.0432

NEΔN, and NEΔN values for the full spectrum were linearly interpolated. While the NEΔN values of actual AIRS data exhibit significantly more structure than those used here, subsequent PPC/NN simulation studies yield similar retrieval performance curves for the actual AIRS noise values and the relatively simple model used here.

1) Atmospheric Transmittance and Radiative Transfer Models: All transmittance and radiative transfer computations performed in this study (including both the statistical and physical retrieval algorithms) were based on the algorithm of Strow *et al.* [20]. Surface reflection of solar radiation is included in the model.

2) Radiosonde Dataset: The ground truth for this study was the NOAA88b radiosonde dataset, which contains 7547 radiosonde/rocketsonde profiles, globally distributed seasonally and geographically. Atmospheric temperature, moisture, and ozone are given at 100 discrete levels ranging in pressure from 0.0160–1100.0 mbar. It should be noted that water vapor values above the highest altitude radiosonde report, generally between 500 and 300 mbar, are subject to artifacts resulting from the method used to combine low- and high-altitude data. Therefore, water vapor measurements above approximately 300 mbar are of questionable quality. Skin surface temperature is also recorded. Two mutually exclusive sets were randomly selected from these 7547 profiles: a training set of 6547 profiles, and a validation set of 1000 profiles.

3) Surface Model: The surface pressure was fixed at 1013.25 mbar for all profiles (variable surface pressure is explored in a Section VI), and a Lambertian surface was assumed. Frequency dependence of the surface emissivity was modeled as a piecewise-linear function with six random (Gaussian) hinge points. The means and standard deviations of the hinge points are shown in Table II. The correlation matrix of the hinge points was assumed to be a Toeplitz matrix, with a first column of 1.0000, 0.9193, 0.8348, 0.7501, 0.6697, and 0.5820.

B. Iterated Minimum-Variance Technique for the Retrieval of Atmospheric Profiles

Let X be a vector that represents some atmospheric state variable, such as the temperature profile, and Y be a vector of measurements. The Bayesian retrieval approach seeks to find the most probable value of the atmospheric state X given the measurements Y. In the case of Gaussian error statistics, the most probable solution is that which minimizes the following cost function [21]:

$$\gamma(\cdot) = (X - X_0)^T \mathbf{C}_{XX}^{-1}(X - X_0) + (Y - \mathbf{F}(X))^T \mathbf{C}_{\Psi\Psi}^{-1}(Y - \mathbf{F}(X)) \quad (13)$$

where X_0 is a "first guess" with expected error covariance $\mathbf{C}_{XX}$, $\mathbf{F}(X)$ is the "forward model," and $\mathbf{C}_{\Psi\Psi}$ is the expected covariance of the measurement error.

Rodgers [22] has shown that the forward model can be expanded as a Taylor series about a guessed value X_0, and Newtonian iteration can be used to iteratively converge to a solution. The solution after the nth iteration can be expressed in closed form as

$$X_{n+1} = X_0 + \mathbf{C}_{XX}\mathbf{K}_n^T \left(\mathbf{K}_n \mathbf{C}_{XX} \mathbf{K}_n^T + \mathbf{C}_{\Psi\Psi}\right)^{-1} \times \left[Y - Y_n - \mathbf{K}_n(X_0 - X_n)\right] \quad (14)$$

where $\mathbf{K} = \partial\mathbf{F}/\partial X$ is the Fréchet derivative, which is computed at each iteration.

In this work, the retrievals are carried out in four successive steps after a first guess is computed using linear regression, as suggested by Susskind *et al.* [23]. The retrieval steps are: 1) surface temperature and emissivity (53 channels used); 2) temperature profile (147 channels used); 3) moisture profile (66 channels used); and 4) improved temperature profile (154 channels used). Step 4) is identical to step two, with the exception that seven additional channels in the water vapor band that produce sharp temperature weighting functions are included. The channel selection is identical to that of Susskind *et al.* [23]. Each step is carried to completion before the execution of the next step. The measurement error covariance matrix $\mathbf{C}_{\Psi\Psi}$ includes uncertainties due to surface temperature and emissivity, temperature profile, water vapor profile, ozone profile, and instrument noise. These uncertainties are updated at the completion of each retrieval step.

It should be stated that while the IMV retrieval methodology presented here is similar to that used operationally to process AIRS data [23], there are notable differences, including the regularization method used to stabilize the solution and the use of output kernel functions. Direct comparison of the PPC/NN method with the AIRS Level-2 operational algorithm (version 3) is given in Section VI.

C. Retrieval Performance Comparisons

1) Retrieval Accuracy: Temperature profile retrieval performance results for linear regression [full-rank form of (5)], the PPC/NN method, and the IMV method are shown in Fig. 2; water vapor retrieval results are shown in Fig. 3. Both figures show errors in 1-km layers. The water vapor results are expressed as a percentage, where, for each profile, the mass mixing ratio error in the layer is weighted by the true mass mixing ratio in the layer.

The PPC/NN temperature profile retrieval performance is superior to that of the iterated minimum-variance algorithm over

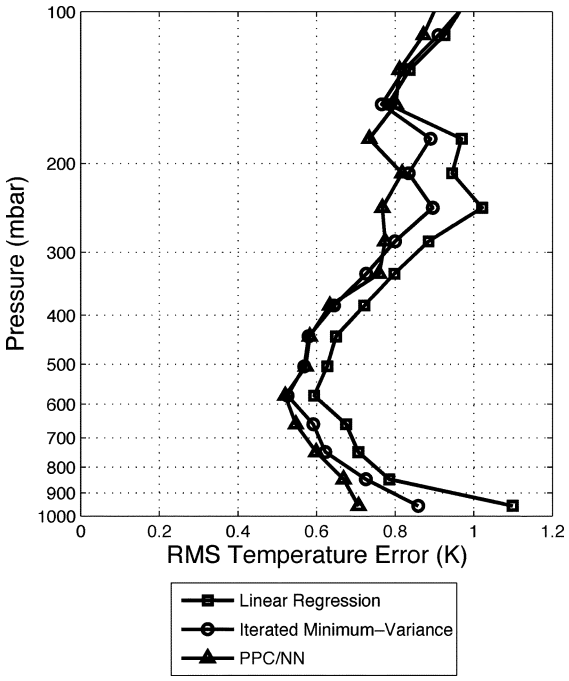

Fig. 2. Temperature retrieval performance of the PPC/NN, IMV, and linear regression methods. RMS errors of 1-km layer means are shown.

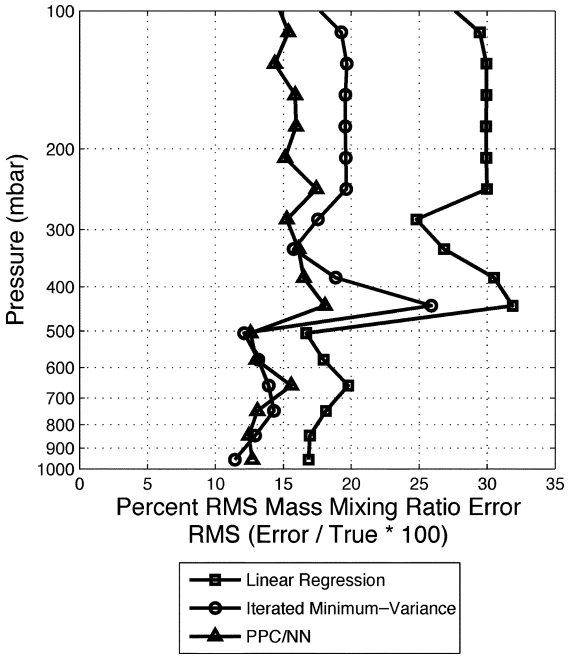

Fig. 3. Water vapor (mass mixing ratio) retrieval performance of the PPC/NN, IMV, and linear regression methods. RMS errors of 1-km layer means are shown. Errors are expressed as a percentage, i.e., $100 \times (q_{\rm true}(P) - q_{\rm est}(P)/q_{\rm true}(P))$.

almost all of the atmosphere and is substantially better near the surface. The PPC/NN water vapor profile retrieval performance is also superior to that of the iterated minimum-variance algorithm over almost all of the atmosphere, and is substantially better in the upper troposphere, although this is probably due to the neural-network fitting to statistical artifacts inherent in the NOAA88b data. The IMV water vapor performance near the surface is better than that of the PPC/NN by approximately 1.5%.

A separate experiment was performed where the PPC/NN estimates were used as the first guess for the IMV algorithm. It was found that the IMV method offered virtually no improvement over the PPC/NN estimates.

2) Computational Efficiency: There are performance benefits of the PPC/NN method other than retrieval accuracy. The high degree of radiance compression afforded by the PPC method allows relatively small NNs (a few thousand free parameters) to be used, thus allowing the use of fast (but memory-intensive) training algorithms, such as Levenberg–Marquardt, on commodity personal computers. The algorithm is extremely fast once trained (retrieval of the temperature and moisture profiles for a single radiance observation takes on the order of a millisecond), and training of the complete algorithm over a global ensemble of 10 000 profiles takes only several hours on a desktop workstation (Pentium IV operating at 3 GHz). The required computation time for most iterated model-based retrieval techniques can exceed that of the PPC/NN approach by two orders of magnitude. Furthermore, changes in instrument behavior (for example, bad or excessively noisy channels and instrument spectral response function shape and/or center changes) can be accommodated by recalculating the PPC coefficients (with relatively light computational burden). It is usually not necessary to retrain the NNs if only a small number (less than about 1%) of the original channels are affected, as their marginal impact on the resulting PPC coefficients is small because of the large degree of spectral redundancy.

D. Discussion

The retrieval accuracy of the PPC/NN method exceeds that of the IMV method over most of the troposphere, and there are several possible reasons for this improved performance. First, the PPC/NN method used all available AIRS channels (2378) whereas the IMV method used only 273. The spectral redundancy afforded by the use of all AIRS channels facilitates the filtering of noise from the radiances via the PPC transform. Second, correlation between temperature and moisture is taken into account during NN training, but not included in the IMV *a priori* statistics because temperature and moisture are retrieved in separate steps. Third, although each step in the IMV retrieval was iterated to convergence, it is possible that the IMV may not have found the overall minimum of the cost function. Finally, minimization of (13) yields the minimum error variance solution (even for non-Gaussian statistics) provided the measurement errors are uncorrelated with the atmospheric state. However, in this work the measurements were expressed as brightness temperatures and not as spectral radiances; therefore, the independent-error assumption may be violated. The neural network is able to fully exploit any non-Gaussian statistical relationships that exist between the radiances and the geophysical parameters.

TABLE III
AIRS SOFTWARE VERSION NUMBERS FOR THE SEVEN DAYS USED IN THE MATCH-UP DATASET

	6 Sep 2002	25 Jan 2003	8 Jun 2003	21 Aug 2003	3 Sep 2003	12 Oct 2003	5 Dec 2003
Cloud-clearing	3.7.0	3.7.0	3.7.0	3.1.9	3.1.9	3.1.9	3.7.0
Level-2	3.0.8	3.0.8	3.0.8	3.0.8	3.0.8	3.0.8	3.0.10

VI. VALIDATION OF THE PPC/NN ALGORITHM WITH AIRS/AMSU OBSERVATIONS OF PARTIALLY CLOUDY SCENES OVER LAND AND OCEAN

In this section, the performance of the PPC/NN algorithm is evaluated using cloud-cleared AIRS data (where the cloud-clearing is performed using both AIRS and AMSU data) and colocated ECMWF forecast fields. The PPC/NN retrieval performance is compared with that obtained using the AIRS Level-2 algorithm. Both ocean and land cases are considered, including elevated surface terrain, and retrievals at all sensor scan angles (out to $\pm 48^\circ$) are derived. Finally, sensitivity analyses of PPC/NN retrieval performance are presented with respect to scan angle, orbit type (ascending or descending), cloud amount, and training set comprehensiveness.

A. Cloud Clearing of AIRS Radiances

The cloud-clearing approach discussed in Susskind *et al.* [23] was applied to the AIRS data by the AIRS Science Team. Version 3 of the algorithm was used in this work (see Table III for a detailed listing of the software version numbers). The algorithm seeks to estimate a clear-column radiance (the radiance that would have been measured if the scene were cloud-free) from a number of adjacent cloud-impacted fields of view.

B. AIRS/AMSU/ECMWF Dataset

The performance of the PPC/NN algorithm was evaluated using 352 903 AIRS/AMSU observations and colocated ECMWF atmospheric fields collected on seven days throughout 2002 and 2003 (see Table III). Various software version changes were made over the course of this work (see Table III for details), but these changes were primarily with regard to quality control and do not significantly affect the results presented here. However, the version 4 release of the AIRS software, which was not available in time to be included in this work, should offer many enhancements over version 3, including improved cloud clearing, retrieval accuracies, quality control, and retrieval yield [24]. Reanalyses of the results presented in this section are therefore planned with the new AIRS software release.

The 352 903 observations were randomly divided into a training set of 302 903 observations (206 061 of which were over ocean), and a separate validation set of 50 000 observations (40 000 of which were over ocean). The *a priori* RMS variation of the temperature and water vapor (mass mixing ratio) profiles contained in the validation set are shown in Fig. 4. The observations in the validation set were matched with AIRS Level-2 retrievals obtained from the Earth Observing System (EOS) Data Gateway (EDG). As advised in the AIRS Version 3.0 L2 Data Release Documentation, only retrievals that met certain quality standards (specifically, $\mathrm{RetQAFlag} = 0$ for ocean and $\mathrm{RetQAFlag} = 256$ for land) were included in the analyses.

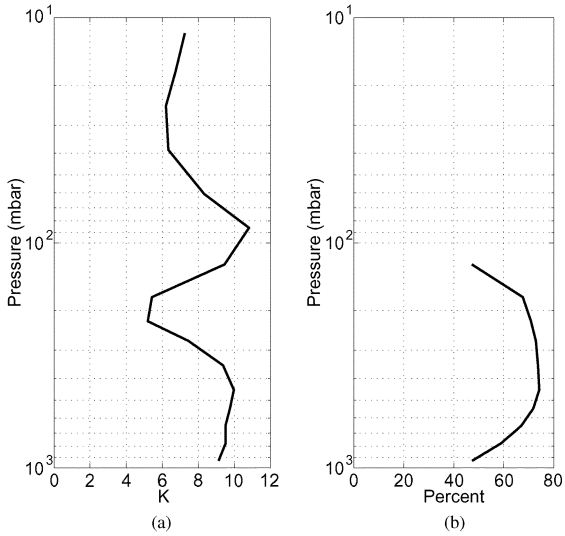

Fig. 4. Temperature and water vapor profile statistics for the validation dataset used in the analysis. See the text for details on how the statistics are computed at each layer. (a) *A priori* temperature error. (b) *A priori* MMR error.

There were 17 856 AIRS Level-2 retrievals (all within $\pm 40^\circ$ latitude) that met this criterion. Reanalysis with AIRS Level-2 version 4 software is planned, as the version 4 products have been validated over both ocean and land at near-polar latitudes.

In order to facilitate comparison with results published in the AIRS v3.0 Validation Report [25], layer error statistics are calculated slightly differently in this section than they were in Section V. First, layer averages are calculated in layers of approximately (but not exactly) 1 km width—the exact layer widths can be found in Appendix III in the AIRS v3.0 Validation Report. Second, weighted water vapor errors in each layer are calculated by dividing the RMS mass mixing ratio error by the RMS variation of the true mass mixing ratio (as opposed to dividing the mass mixing ratio error of each profile by the true mass mixing ratio for that profile and computing the RMS of the resulting ensemble).

C. AIRS/AMSU Channel Selection

Thirty-seven percent (888 of the 2378) of the AIRS channels were discarded for the analysis, as the radiance values for these channels frequently were flagged as invalid by the AIRS calibration software. A simulated AIRS brightness temperature spectrum is shown in Fig. 5, which shows the original 2378 AIRS channels and the 1490 channels that were selected for use with the PPC/NN algorithm. All 15 AMSU channels were used. The algorithm automatically discounts channels that are excessively corrupted by sensor noise (for example, AMSU channel 7) or other interfering signals (for example, the effects of nonlocal thermodynamic equilibrium) because the corruptive signals are largely uncorrelated with the geophysical parameters that are to be estimated.

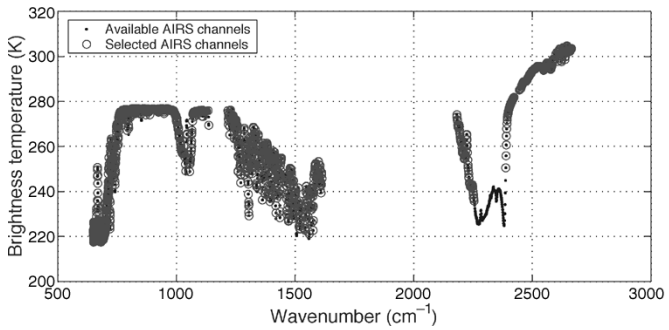

Fig. 5. Typical AIRS spectrum (simulated) is shown. A total of 1490 out of 2378 AIRS channels were selected.

D. PPC/NN Retrieval Enhancements for Variable Sensor Scan Angle and Surface Pressure

The PPC/NN retrieval results presented in Section V corresponded to simulated measurements at a fixed scan angle (nadir) and a fixed surface pressure (1013.25 mbar). When dealing with real AIRS/AMSU data, a variety of scan angles and surface pressures must be accommodated. Therefore, two additional inputs were added to the NNs discussed in Section IV: 1) the secant of the scan angle and 2) the forecast surface pressure (in mbar) divided by 1013.25. The resulting temperature and water vapor profile estimates were reported on a variable pressure grid anchored by the forecast surface pressure.

Because of the number of inputs to the NNs increased, the number of hidden nodes in NNs used for temperature retrievals was increased from 15 to 20. For water vapor retrievals, the number of hidden nodes in the first hidden layer was maintained at 25, but a second hidden layer of 15 hidden nodes was added.

E. Retrieval Performance

We now compare the retrieval performance of the PPC/NN, linear regression, and AIRS Level-2 methods. For both the ocean and land cases, the PPC/NN and linear regression retrievals were derived using the same training set, and the same validation set was used for all methods.

1) Ocean Cases: The temperature profile retrieval performance over ocean for the linear regression retrieval, the PPC/NN retrieval, and the AIRS Level-2 retrieval is shown in Fig. 6, and the water vapor retrieval performance is shown in Fig. 7. The error statistics were calculated using the 13 156 (out of 40 000) AIRS Level-2 retrievals that converged successfully. A bias of approximately 1 K near 100 mbar was found between the AIRS Level-2 temperature retrievals and the ECMWF data (ECMWF was colder). This bias was removed prior to computation of the AIRS Level-2 retrieval error statistics, which are shown in Fig. 6.

2) Land Cases: The temperature profile retrieval performance over land for the linear regression retrieval, the PPC/NN retrieval, and the AIRS Level-2 retrieval is shown in Fig. 8, and the water vapor retrieval performance is shown in Fig. 9. The error statistics were calculated using the 4700 (out of 10 000) AIRS Level-2 retrievals that converged successfully.

3) Discussion: There are several features in Figs. 6–9 that are worthy of note. First, for all retrieval methods, the performance over land is worse than that over ocean, as expected. The

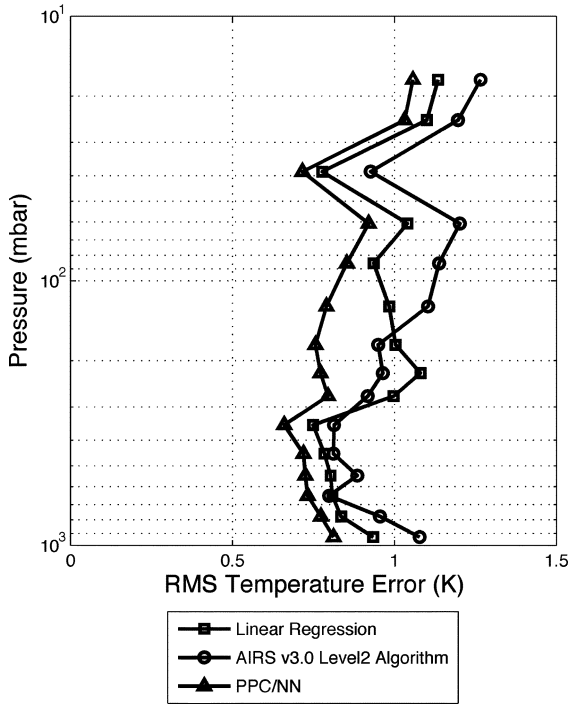

Fig. 6. Temperature retrieval performance of the PPC/NN, linear regression, and AIRS Level-2 methods over ocean. Statistics were calculated over 13 156 fields of regard.

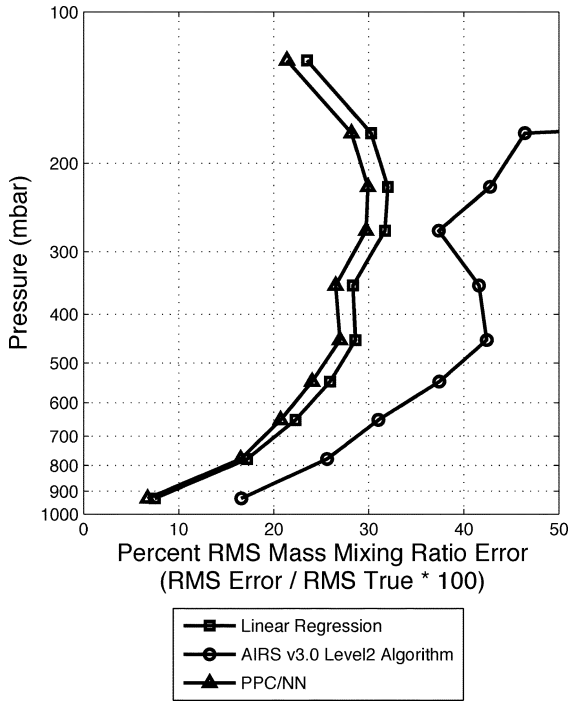

Fig. 7. Water vapor (mass mixing ratio) retrieval performance of the PPC/NN, linear regression, and AIRS Level-2 methods over ocean. Statistics were calculated over 13 156 fields of regard.

cloud-clearing problem is significantly more difficult over land, as variations in surface emissivity can be mistaken for cloud perturbations, thus resulting in improper radiance corrections. Second, the magnitude of the temperature profile error degradation for land versus ocean is larger for the PPC/NN algorithm than for the AIRS Level-2 algorithm. In fact, the temperature

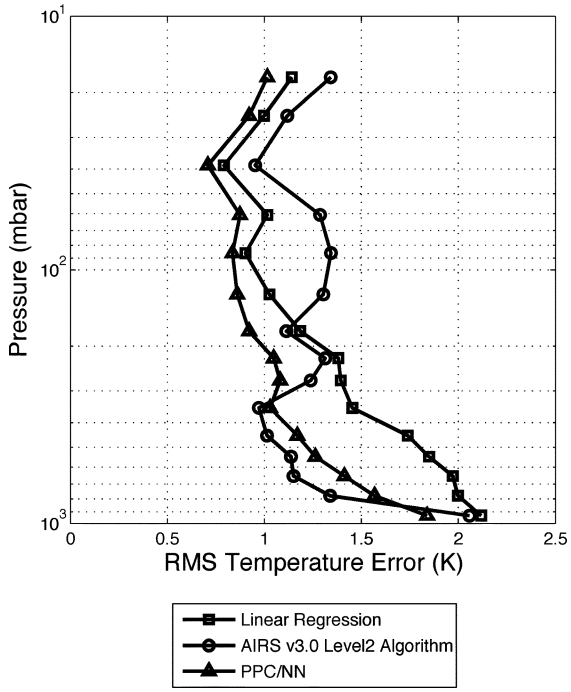

Fig. 8. Temperature retrieval performance of the PPC/NN, linear regression, and AIRS Level-2 methods over land. Statistics were calculated over 4700 fields of regard.

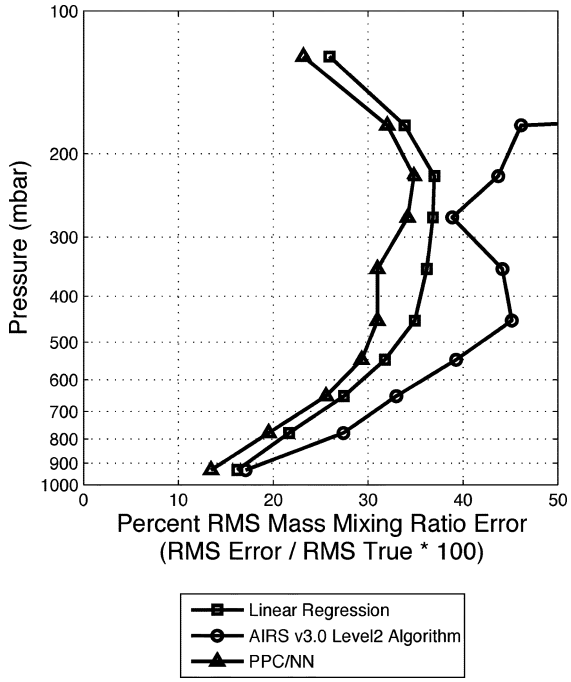

Fig. 9. Water vapor (mass mixing ratio) retrieval performance of the PPC/NN, linear regression, and AIRS Level-2 methods over land. Statistics were calculated over 4700 fields of regard.

profile retrieval performance of the AIRS Level-2 algorithm is superior to that of the PPC/NN algorithm throughout most of the lower troposphere over land. Further analyses of this discrepancy suggest that the performance of the PPC/NN method over elevated terrain is suboptimal, and could be improved. This work is currently underway.

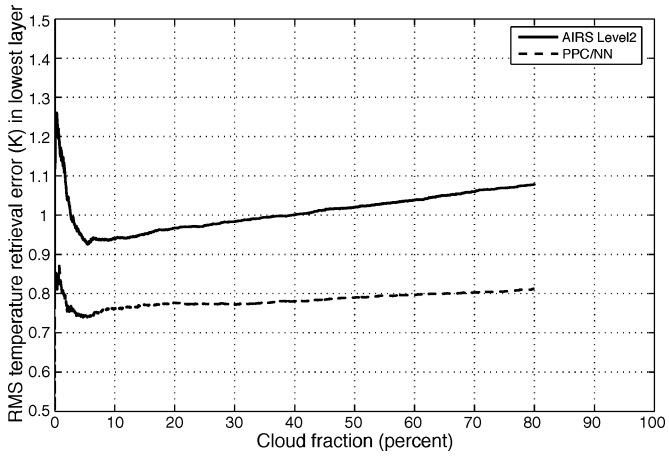

Fig. 10. Cumulative RMS temperature error over ocean in the layer closest to the surface. Pixels were ranked in order of increasing cloudiness according to the retrieved cloud fraction from the AIRS Level-2 algorithm. No retrievals were attempted by the AIRS Level-2 algorithm if the retrieved cloud fraction exceeded 80%.

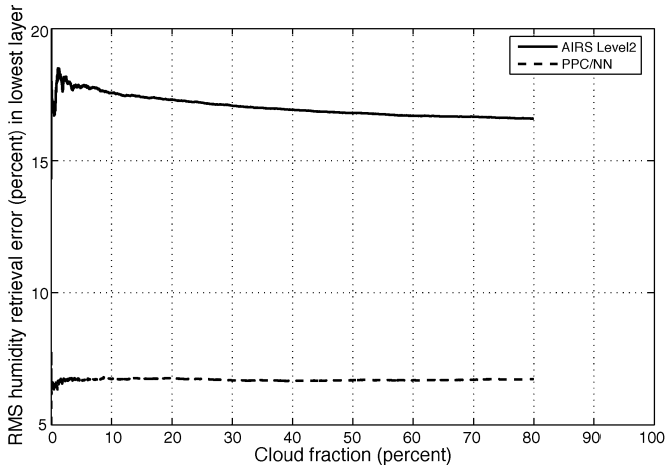

Fig. 11. Cumulative RMS water vapor error over ocean in the layer closest to the surface. Pixels were ranked in order of increasing cloudiness, according to the retrieved cloud fraction from the AIRS Level-2 algorithm. No retrievals were attempted by the AIRS Level-2 algorithm if the retrieved cloud fraction exceeded 80%.

F. Retrieval Performance Sensitivity Analyses

In this section, the PPC/NN retrieval performance over ocean is stratified by cloud amount, sensor scan angle, orbit type, and training set comprehensiveness, and it is shown that the PPC/NN retrieval is relatively insensitive to each of these parameters. The sensitivity of the PPC/NN retrieval to cloud fraction is compared with that exhibited by the AIRS Level-2 retrieval.

1) Sensitivity to Cloud Amount: A plot of the temperature retrieval error in the layer closest to the surface as a function of the cloud fraction retrieved by the AIRS Level-2 algorithm is shown in Fig. 10. Similar curves for the water vapor retrieval performance are shown in Fig. 11. The larger errors evident in regions with small retrieved cloud fractions are due to the AIRS Level-2 algorithm misclassifying some "completely cloudy" scenes as "completely clear." Both methods produce temperature and moisture retrievals with RMS errors near 1 K and 15%, respectively, even in cases with large cloud fractions. The figures show that the PPC/NN temperature and moisture

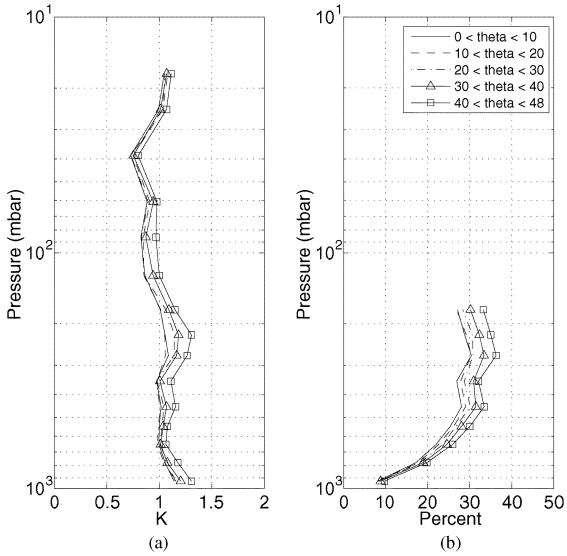

Fig. 12. PPC/NN retrieval performance over ocean stratified by sensor scan angle. Statistics were calculated over 40 000 fields of regard. (a) Temperature error. (b) Water vapor error.

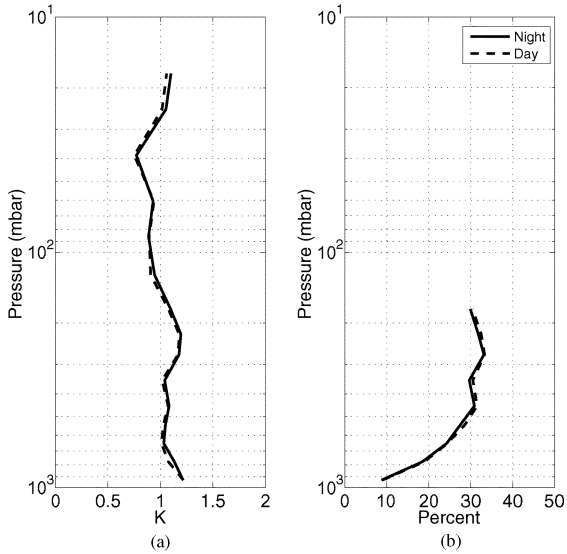

Fig. 13. PPC/NN retrieval performance over ocean stratified by satellite orbit type: ascending (daytime) and descending (nighttime). Statistics were calculated over 40 000 fields of regard. (a) Temperature error. (b) Water vapor error.

retrievals are less sensitive than the AIRS Level-2 retrievals to cloud amount.

2) Sensitivity to Sensor Scan Angle: Fig. 12 shows the PPC/NN temperature and moisture retrieval errors stratified by sensor scan angle. At the extreme scan angles, the temperature retrieval errors increase by approximately 0.2 K and the moisture retrieval errors increase by approximately 4%.

3) Sensitivity to Satellite Orbit Type: Fig. 13 shows the PPC/NN temperature and moisture retrieval errors stratified by satellite orbit type (ascending versus descending orbits). There is almost no dependence evident, which indicates the retrieval is robust to a variety of solar illumination conditions.

4) Sensitivity to Training Set Comprehensiveness: As a final test of the stability of the PPC/NN algorithm, the ability of

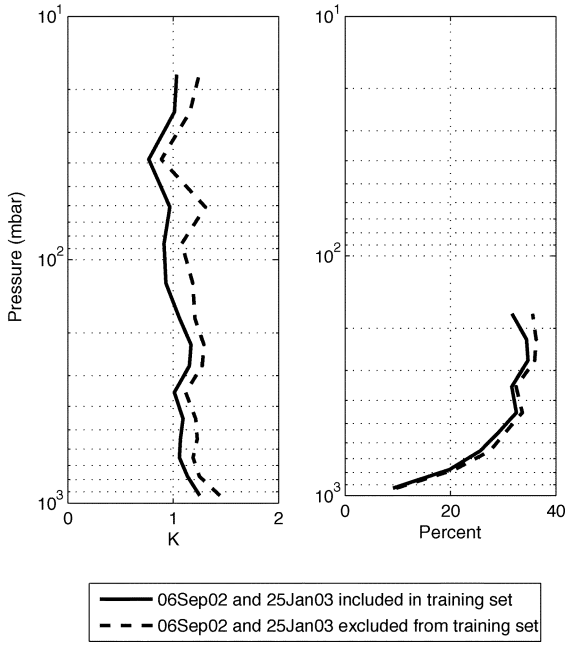

Fig. 14. PPC/NN retrieval performance over ocean for two training sets. The same validation set was used in both cases: 11 877 profiles observed on September 6, 2002 and June 8, 2003. The first training set includes profiles from all seven days. For the second training set, all profiles observed on September 6, 2002 and June 8, 2003 were removed.

the NN to generalize to time periods not represented in the training set is examined. For this experiment, the first two days (September 6, 2002 and January 25, 2003) of the seven-day training set were removed. A new validation set was assembled, consisting only of the two days not represented in the training set. This new validation set was applied to the original PPC/NN algorithm, which had been trained using all seven days, and the PPC/NN algorithm which was trained with the training set containing only five days. Temperature and moisture retrieval results for this experiment are shown in Fig. 14. The temperature and moisture retrieval errors increase by approximately 0.2 K and 2%, respectively, as a result of the limited training set. This suggests that the PPC/NN is relatively insensitive to seasonal and year-to-year changes, and any residual error due to this effect could be reduced by stratifying the training set by season and, if necessary, updating the PPC/NN coefficients periodically.

G. Discussion and Future Work

While the PPC/NN performance results presented in the previous section are very encouraging, several caveats must be mentioned. The ECMWF fields used for "ground truth" contain errors, and the NN will tune to these errors as part of its training process. Therefore, the PPC/NN RMS errors shown in the previous section may be underestimated, and the AIRS Level-2 RMS errors may be overestimated, as the ECMWF data are not an accurate representation of the true state of the atmosphere. This is especially true over ocean and at high altitudes where radiosonde data are largely absent or of poor quality. An example of this is evident in the AIRS Level-2 retrieval of upper tropospheric water vapor shown in Figs. 7 and 9. The marked

increase in AIRS Level-2 retrieval error near 200 mbar is probably due to errors in the ECMWF atmospheric fields, not the retrieval. Therefore, the "true" spread between the performance of the PPC/NN and AIRS Level-2 algorithms is almost certainly smaller than that shown here. Work is currently underway to test the performance of both the PPC/NN and AIRS Level-2 algorithms with additional ground truth data, including radiosonde data, and ground- and aircraft-based measurements. It should be noted that the PPC/NN algorithm as implemented in this work is currently not a standalone system, as both AIRS cloud-cleared radiances and quality flags produced by the AIRS Level-2 algorithm are required. Future work is planned to adapt the PPC/NN algorithm for use directly on cloudy AIRS/AMSU radiances and to produce quality assessments of the retrieved products. Finally, assimilation of PPC/NN-derived atmospheric parameters into numerical weather prediction models is planned, and the resulting impact on forecast accuracy will be an excellent indicator of retrieval quality.

In light of the previous comments, it is necessary to consider the steps that would be required to implement the PPC/NN retrieval technique in an operational system. Most importantly, perhaps, is the training methodology that will be needed. For reasons previously discussed, it will probably not be feasible to derive the PPC/NN coefficients using a forecast model. Two products from the EOS Aqua system should be an invaluable source of training data for future missions. The first is the database of radiances (both raw and cloud-cleared) and atmospheric parameter retrievals that is being generated. Second, and more importantly, is the improved validation of surface, cloud, and transmittance models as a direct result of product validation efforts. These models, together with the database of atmospheric retrievals being generated, will provide a "ground-truth laboratory" from which training datasets for future sensors can be derived.

VII. Summary and Conclusion

A novel statistical retrieval technique was introduced that combines a linear radiance compression operator with a neural-network estimator. The projected principal components (PPC) transform was shown to be well-suited for this application because information correlated to the geophysical quantity of interest is optimally represented with only a few dozen components. This substantial amount of radiance compression (approximately a factor of 100) allows relatively small NNs to be used thereby improving both the stability and computational efficiency of the algorithm. Test cases with both simulated clear-air and observed partially cloudy AIRS/AMSU data demonstrate that the PPC/NN temperature and moisture retrievals yield accuracies commensurate with those of physical methods at a substantially reduced computational burden. Retrieval accuracies (defined as agreement with ECMWF fields) near 1 K for temperature and 25% for water vapor mass mixing ratio in layers of approximately 1-km thickness were obtained using the PPC/NN retrieval method with AIRS/AMSU data in partially cloudy areas. PPC/NN retrievals with partially cloudy AIRS/AMSU data over land were also performed. The PPC/NN retrieval technique was shown to be relatively insensitive to cloud amount, sensor scan angle, orbit type, and training set comprehensiveness. These results further suggest the AIRS Level-2 algorithm that produced the cloud-cleared radiances and quality flags used by the PPC/NN retrieval is performing well.

The high level of performance achieved by the PPC/NN algorithm suggests it would be a suitable candidate for the retrieval of geophysical parameters other than temperature and moisture from high resolution spectral data. Potential applications include the retrieval of ozone profiles and trace gas amounts. Future work will involve further evaluation of the algorithm with simulated and observed partially cloudy data, including global radiosonde data and ground- and aircraft-based observations.

Acknowledgment

The author would like to thank the NOAA NESDIS, Office of Systems Development for financial support for this work. The colocated AIRS/AMSU/ECMWF datasets were provided by the AIRS Science Team, via M. Goldberg (NOAA NESDIS, Office of Research Applications). The author would also like to thank L. Zhou for help obtaining and interpreting the AIRS/AMSU/ECMWF datasets, D. Staelin and P. Rosenkranz for many helpful discussions on all facets of this work, and the two anonymous reviewers, whose suggestions substantially improved the quality of the final manuscript.

References

[1] H. Hotelling, "The most predictable criterion," *J. Educ. Psychol.*, vol. 26, pp. 139–142, 1935.

[2] J. Escobar-Munoz, A. Chedin, F. Cheruy, and N. Scott, "Reseaux de neurons multicouches pour la restitution de variables thermodynamiques atmosphériques à l'aide de sondeurs verticauz satellitairs," *Comptes-Rendus de L'Academie Des Sciences; Série II*, vol. 317, no. 7, pp. 911–918, 1993.

[3] M. D. Goldberg, Y. Qu, L. M. McMillin, W. Wolff, L. Zhou, and M. Divakarla, "AIRS near-real-time products and algorithms in support of operational numerical weather prediction," *IEEE Trans. Geosci. Remote Sensing*, vol. 41, no. 2, pp. 379–389, Feb. 2003.

[4] H. Huang and P. Antonelli, "Application of principal component analysis to high-resolution infrared measurement compression and retrieval," *J. Appl. Meteorol.*, vol. 40, pp. 365–388, Mar. 2001.

[5] F. Aires, W. B. Rossow, N. A. Scott, and A. Chédin, "Remote sensing from the infrared atmospheric sounding interferometer instrument: 1. compression, denoising, first-guess retrieval inversion algorithms," *J. Geophys. Res.*, vol. 107, Nov. 2002.

[6] W. J. Blackwell and D. H. Staelin, "Cloud flagging and clearing using high-resolution infrared and microwave sounding data," in *Proc. IGARSS*, vol. 3, Jun. 2002, pp. 1860–1862.

[7] J. B. Lee, A. S. Woodyatt, and M. Berman, "Enhancement of high spectral resolution remote-sensing data by a noise-adjusted principal components transform," *IEEE Trans. Geosci. Remote Sensing*, vol. 28, no. 3, pp. 295–304, May 1990.

[8] W. J. Blackwell, "Retrieval of cloud-cleared atmospheric temperature profiles from hyperspectral infrared and microwave observations,", Dept. Elect. Eng. Comput. Sci., Massachusetts Inst. Technol., Cambridge, Jun. 2002.

[9] K. M. Hornik, M. Stinchcombe, and H. White, "Multilayer feedforward networks are universal approximators," *Neural Netw.*, vol. 4, no. 5, pp. 359–366, 1989.

[10] D. E. Rumelhart, G. Hinton, and R. Williams, *Parallel Distributed Processing: Explorations in the Microstructure of Cognition*, D. E. Rumelhart and J. L. McClelland, Eds. Cambridge, MA: MIT Press, 1986, vol. 1, Foundations.

[11] M. T. Hagan and M. B. Menhaj, "Training feedforward networks with the Marquardt algorithm," *IEEE Trans. Neural Netw.*, vol. 5, no. 6, pp. 989–993, Nov. 1994.

[12] P. E. Gill, W. Murray, and M. H. Wright, "The Levenberg–Marquardt method," in *Practical Optimization*. London, U.K.: Academic, 1981.
[13] C. M. Bishop, *Neural Networks for Pattern Recognition*. Oxford, U.K.: Oxford Univ. Press, 1995.
[14] H. E. Motteler, L. L. Strow, L. McMillin, and J. A. Gualtieri, "Comparison of neural networks and regression based methods for temperature retrievals," *Appl. Opt.*, vol. 34, no. 24, pp. 5390–5397, Aug. 1995.
[15] F. Aires, A. Chédin, N. A. Scott, and W. B. Rossow, "A regularized neural net approach for retrieval of atmospheric and surface temperatures with the IASI instrument," *J. Appl. Meteorol.*, vol. 41, pp. 144–159, Feb. 2002.
[16] F. Aires, W. B. Rossow, N. A. Scott, and A. Chédin, "Remote sensing from the infrared atmospheric sounding interferometer instrument: 2. Simultaneous retrieval of temperature, water vapor, and ozone atmospheric profiles," *J. Geophys. Res.*, vol. 107, Nov. 2002.
[17] F. D. Frate and G. Schiavon, "A combined natural orthogonal functions/neural network technique for the radiometric estimation of atmospheric profiles," *Radio Sci.*, vol. 33, no. 2, pp. 405–410, Mar. 1998.
[18] D. Nguyen and B. Widrow, "Improving the learning speed of two-layer neural networks by choosing initial values of the adaptive weights," in *Int. Joint Conf. Neural Networks*, vol. 3, 1990, pp. 21–26.
[19] H. H. Aumann *et al.*, "AIRS/AMSU/HSB on the Aqua mission: Design, science objectives, data products, and processing systems," *IEEE Trans. Geosci. Remote Sens.*, vol. 41, no. 2, pp. 253–264, Feb. 2003.
[20] L. Strow, S. Hannon, and S. Desouza-Machado, "An overview of the AIRS radiative transfer model," *IEEE Trans. Geosci. Remote Sens.*, vol. 41, no. 2, pp. 303–313, Feb. 2003.
[21] J. R. Eyre, "Inversion of cloudy satellite sounding radiances by nonlinear optimal estimation. I: Theory and simulation for TOVS," *Q. J. R. Meteorol. Soc.*, vol. 115, pp. 1001–1026, Jul. 1989.
[22] C. D. Rodgers, "Retrieval of atmospheric temperature and composition from remote measurements of thermal radiation," *J. Geophys. Res.*, vol. 41, no. 7, pp. 609–624, Jul. 1976.
[23] J. Susskind, C. D. Barnet, and J. M. Blaisdell, "Retrieval of atmospheric and surface parameters from AIRS/AMSU/HSB data in the presence of clouds," *IEEE Trans. Geosci. Remote Sensing*, vol. 41, no. 2, pp. 390–409, Feb. 2003.
[24] J. Susskind *et al.*, "Accuracy of geophysical parameters derived from AIRS/AMSU as a function of fractional cloud cover," *J. Geophys. Res.*, 2005, submitted for publication.
[25] H. H. Aumann *et al.*, "Validation of AIRS/AMSU/HSB core products for data release version 3.0," Goddard Space Flight Center, Greenbelt, MD, NASA JPL Tech. Rep. D-26538, Aug. 2003.

William J. Blackwell (S'92–M'02) received the B.E.E. degree from the Georgia Institute of Technology, Atlanta, and the S.M. and Sc.D. degrees from the Massachusetts Institute of Technology (MIT), Cambridge, in 1994, 1995, and 2002, respectively, all in electrical engineering.

He has been a member of the technical staff at the MIT Lincoln Laboratory, Lexington, MA, since 2002. His primary research interests are in the area of atmospheric remote sensing, including the development and calibration of airborne and spaceborne microwave and hyperspectral infrared sensors, the retrieval of geophysical products from remote radiance measurements, and the application of electromagnetic, signal processing, and estimation theory.

Dr. Blackwell held a National Science Foundation Graduate Research Fellowship from 1994 to 1997 and is a member of Tau Beta Pi, Eta Kappa Nu, Phi Kappa Phi, and Sigma Xi. He is currently serving as Chair of the Boston Section of the IEEE Geoscience and Remote Sensing Society.