
Neural Network Applications in High-Resolution Atmospheric Remote Sensing

William J. Blackwell and Frederick W. Chen

■ Estimation techniques based on neural networks are becoming more common in high-resolution atmospheric remote sensing largely because of the simplicity, flexibility, and ability of the neural network techniques to accurately represent complex multidimensional statistical relationships. Spaceborne atmospheric sounders with increasingly finer spatial and spectral resolution are generating formidable amounts of radiance data. This abundance of data presents two major challenges in the development of algorithms that retrieve geophysical information from the radiance measurements. The first challenge 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 challenge is the implementation of the robust algorithm within a given computational budget. The neural network estimation techniques described in this article allow both of these challenges to be overcome. Sample results are presented for retrievals of (1) atmospheric temperature and moisture profiles and (2) precipitation rates.

MODERN SPACEBORNE ATMOSPHERIC sounders measure radiance with unprecedented resolution and accuracy in spatial, spectral, and temporal dimensions. For example, the Atmospheric Infrared Sounder (AIRS), operational on the NASA Earth Observing System (EOS) Aqua satellite since 2002, provides a spatial resolution of 15 km, a spectral resolution of $\nu/\Delta\nu \approx 1200$ (with 2378 channels from 650 to 2675 cm^{-1}), and a radiometric accuracy on the order of ± 0.2 K. Typical polar-orbiting atmospheric sounders measure approximately 90% of the earth's atmosphere (in the horizontal dimension) approximately every twelve hours. Retrieval algorithms estimate the geophysical state of the atmosphere as a function of space and time from upwelling spectral radiances measured by the sensor.

In this article, we present two examples of neural-network-based atmospheric retrieval algorithms

being developed and implemented at Lincoln Laboratory. In the first example, we consider the retrieval of atmospheric temperature and moisture profiles (quantity as a function of altitude) from hyperspectral radiance measurements in the thermal infrared. A projected principal component (PPC) transform is used to reduce the dimensionality of the spectral-radiance data and optimally extract geophysical information. A multilayer feed-forward neural network (NN) is subsequently used to estimate the desired geophysical profiles. This algorithm is known as the PPC/NN algorithm. The PPC/NN algorithm offers the numerical stability and efficiency of statistical methods while achieving accuracies comparable to those of physical, model-based methods.

In the second example, we consider the retrieval of precipitation rates from passive microwave radiance measurements at frequencies near the oxygen and

water-vapor resonances at 50 to 60 GHz and 183.31 GHz, respectively. In this case, the models relating the precipitation rate to the radiance intensities measured by the sensor are extremely complicated and difficult to validate. The use of neural networks to empirically learn the statistical relationship between rain rate and spectral intensity obviates the need for complex and often inaccurate models in the retrieval algorithm.

In this article, we first review the physics of spaceborne atmospheric remote sensing. Then we present the application of principal-component transforms to hyperspectral sounding data and introduce a new approach, in which the sensor radiances are projected into a subspace that reduces spectral redundancy and maximizes the resulting correlation to a given parameter. This method is 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.

Next, we review the use of multilayer feed-forward neural networks for geophysical parameter retrieval from hyperspectral measurements, first proposed in 1993 [2], and we give an overview of the network parameters used in this work. We discuss the combination of the projected principal components (PPC) radiance compression operator with a neural network for estimating temperature and water-vapor profiles, and present performance analyses comparing the PPC/NN algorithm to traditional retrieval methods. We then explain the novel application of neural network techniques to the retrieval of precipitation rates from spaceborne microwave data.

Next we present the microphysical relationships between precipitation intensity and the electromagnetic perturbations measured by the sensor, and discuss the neural network precipitation-rate retrieval technique. Then we show examples of precipitation-rate retrievals, including rain rates observed in and around Hurricane Isabel in September 2003. This topic leads to a study of global diurnal variation of precipitation. We close with a summary and suggestions for potential areas of future research.

Spaceborne Atmospheric Remote Sensing

Figure 1 shows the typical measurement scenario for spaceborne atmospheric remote sensing. A sensor

measures upwelling spectral radiance (intensity as a function of frequency) at various incidence angles. The sensor data are usually calibrated to remove measurement artifacts such as gain drift, nonlinearities, and noise. The spectral radiances measured by the sensor are related to geophysical quantities, such as the vertical temperature profile of the atmosphere. An appropriate retrieval algorithm is necessary to convert these radiances into a geophysical quantity of interest.

The radiative transfer equation describing the radiation intensity observed at altitude L , viewing angle θ , and frequency ν can be formulated by including the emitted atmospheric contribution, the reflected atmospheric and cosmic contributions, and the radiance emitted by the surface as follows [3, 4]:

$$\begin{aligned} \mathbf{R}_\nu(L) = & \int_0^L \kappa_\nu(z) \mathcal{J}_\nu[T(z)] e^{\alpha(z)} \sec \theta dz \\ & + \rho_\nu e^{-\tau^* \sec \theta} \int_0^L \kappa_\nu(z) \mathcal{J}_\nu[T(z)] e^{\beta(z)} \sec \theta dz \\ & + \rho_\nu e^{-2\tau^* \sec \theta} \mathcal{J}_\nu(T_c) \\ & + \varepsilon_\nu e^{-\tau^* \sec \theta} \mathcal{J}_\nu(T_s), \end{aligned} \quad (1)$$

where ε_ν is the surface emissivity, ρ_ν is the surface reflectivity, $\kappa_\nu(z)$ is the atmospheric absorption coefficient, τ^* is the atmospheric zenith opacity, $T(z)$ is the temperature profile, T_s is the surface temperature, T_c is the cosmic background temperature (2.736 ± 0.017 K), and

$$\alpha(z) = - \int_z^L \sec \theta \kappa_\nu(z') dz'$$

and

$$\beta(z) = - \int_0^z \sec \theta \kappa_\nu(z') dz'.$$

$\mathcal{J}_\nu(T)$ is the radiance intensity emitted by a blackbody at temperature T , which is given by the Planck equation:

$$\mathcal{J}_\nu(T) = \frac{h\nu^3}{c^2} \frac{1}{e^{h\nu/kT} - 1} \quad \text{W} \cdot \text{m}^{-2} \cdot \text{ster}^{-1} \cdot \text{Hz}^{-1}.$$

The first term in Equation 1 can be recast in terms of a transmittance function $T_v(z)$:

$$R_v(L) = \int_0^L \mathcal{J}_v[T(z)] \left(\frac{dT_v(z)}{dz} \right) dz.$$

The derivative of the transmittance function with respect to altitude is often called the temperature weighting function

$$W_v(z) \triangleq \frac{dT_v(z)}{dz},$$

and gives the relative contribution of the radiance emanating from each altitude. Figure 2 shows the temperature and water-vapor profile weighting functions for the Advanced Microwave Sounding Unit (AMSU).

Geophysical Parameter Retrieval

The objective of the geophysical parameter retrieval algorithm is to estimate the state of the atmosphere (represented by a parameter matrix X) given observations of spectral radiance (represented by a radiance matrix R). Note that the inverse model typically does not exist, as there are generally an infinite number of atmospheric states that could give rise to a particular radiance measurement.

There are generally two approaches to this retrieval problem, as shown in Figure 3. The first approach, called the variational approach, uses a forward model (for example, the transmittance and radiative transfer models previously discussed) to calculate the sensor radiance that would be measured, given a specific atmo-

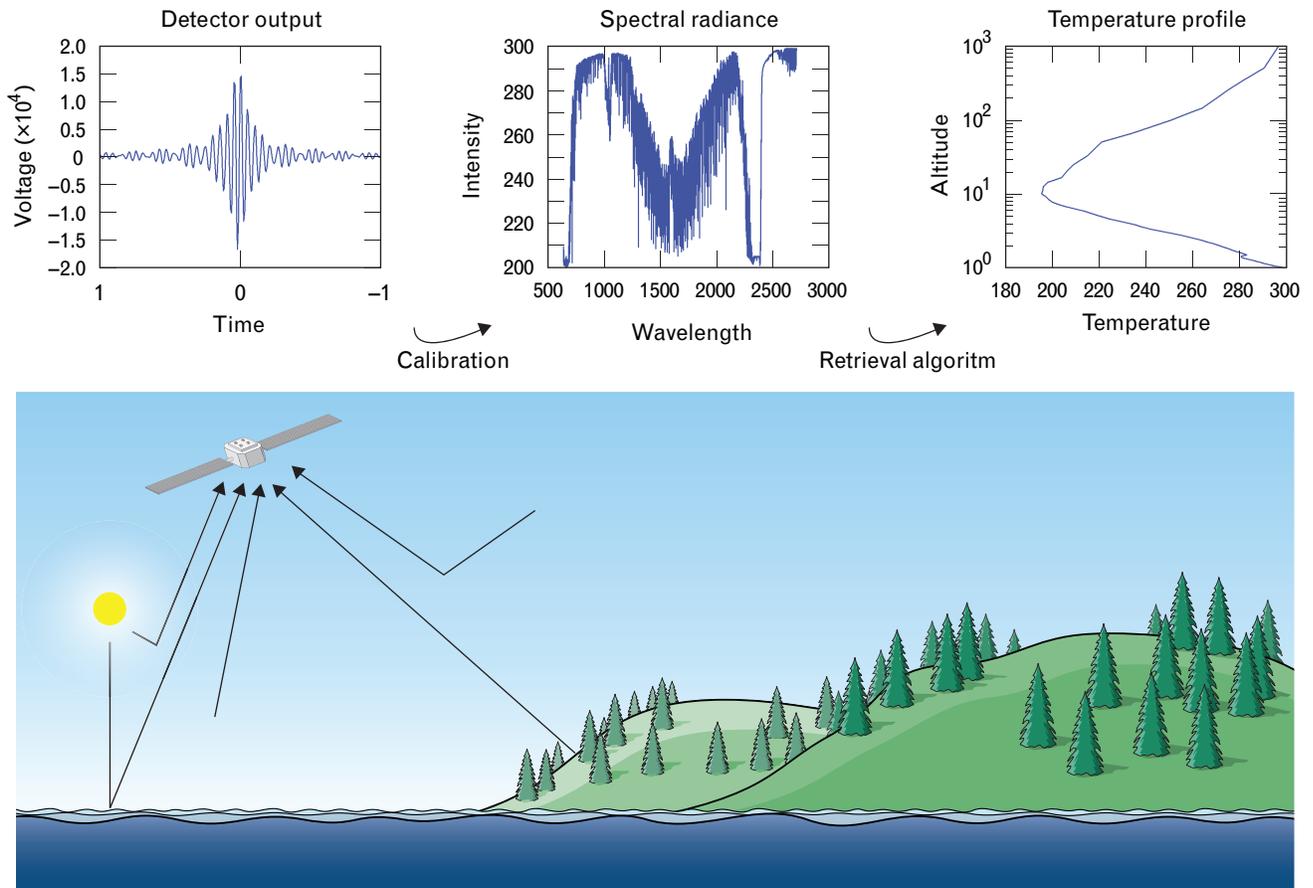


FIGURE 1. Typical measurement scenario for spaceborne atmospheric remote sensing. Electromagnetic radiation that reaches the sensor is emitted by the sun, atmosphere, surface, clouds, and cosmic background. This radiation can also be reflected or scattered by the surface, atmosphere, or clouds. The spectral radiances measured by the sensor are related to geophysical quantities such as the vertical temperature profile of the atmosphere. An appropriate retrieval algorithm is necessary to convert these radiances into a geophysical quantity of interest.

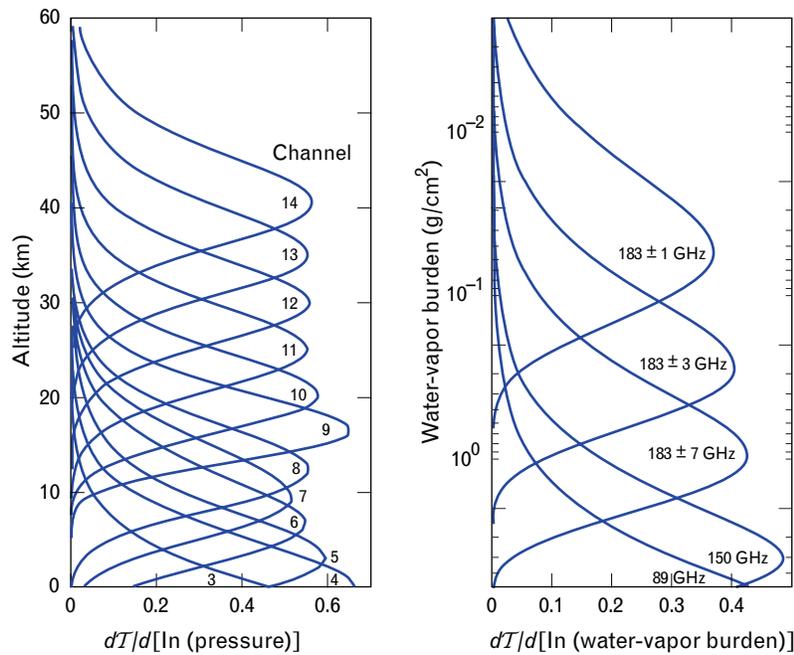


FIGURE 2. The Advanced Microwave Sounding Unit (AMSU) temperature profile (left) and AMSU water-vapor profile (right) weighting functions. The water-vapor burden at an altitude z is the integral of water-vapor profile from z to the top of the atmosphere. Note that channels 7 through 14 show marked ground-level insensitivities when compared to channels 3 through 6. Later in the article, the specific channel frequencies are tabulated and the reasons for their selection are detailed.

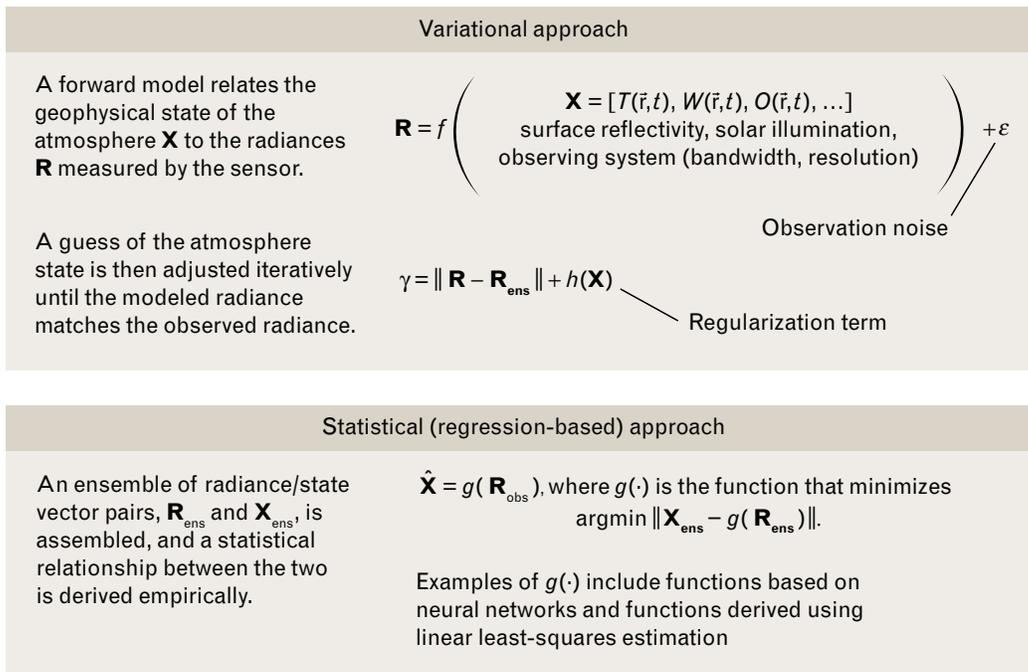


FIGURE 3. Variational and statistical approaches to geophysical parameter retrieval. In the variational approach, a forward model is used to predict at-sensor radiances based on atmospheric state. In the statistical approach, an empirical relationship between at-sensor radiances and atmospheric state is derived using an ensemble of radiance/state vectors.

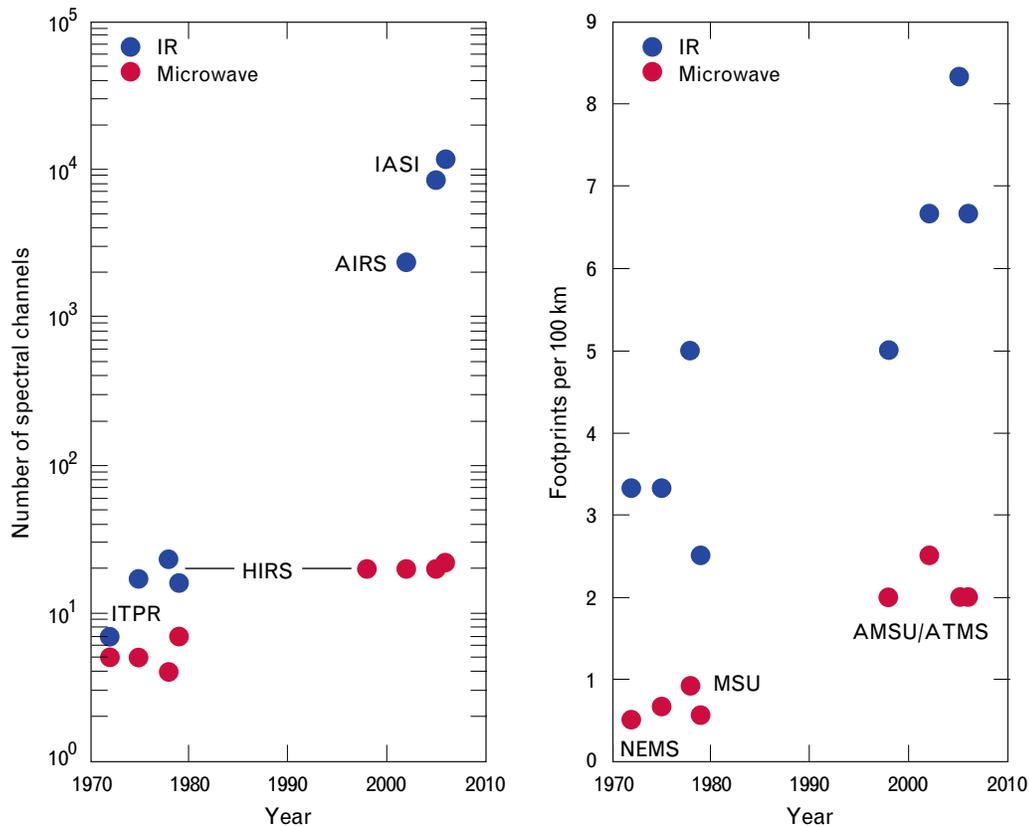


FIGURE 4. Improvements in spectral and spatial resolution of satellite-based remote sensors over the last thirty-five years. The recent increases in the spectral resolutions afforded by infrared sensors have far surpassed those available from microwave sensors. The trends in spatial resolution are similar for infrared and microwave sensors.

spheric state. The variational approach makes a guess of the atmospheric state (usually obtained through a forecast model or historical statistics) and propagates through the forward models, thereby producing an estimate of the at-sensor radiance. The measured radiance is compared with this estimated radiance, and the state vector is adjusted so as to reduce the difference between the measured and estimated radiance vectors. C.D. Rodgers provides a more thorough treatment of the methodology and implementation of variational retrieval methods [5]. The second approach, called the statistical, or regression-based, approach, does not explicitly use the forward model to derive the estimate of the atmospheric state vector. Instead, an ensemble of radiance/state vector pairs is selected, and a statistical characterization [$p(\mathbf{X})$, $p(\mathbf{R})$, and $p(\mathbf{X}, \mathbf{R})$] is sought. In practice, it is difficult to obtain these probability density functions directly from the data, and alterna-

tive methods are often used. Two of these methods are linear least-squares estimation (LLSE), or linear regression, and nonlinear least-squares estimation (NLLSE). Neural networks are a special class of NLLSE, and are discussed later.

The Motivation for Computationally Efficient Algorithms

The principle advantage of regression-based methods is their simplicity—once the coefficients are derived from training data, the calculation of atmospheric state vectors is relatively easy. The variational approaches require multiple calls to the forward models, which can be computationally prohibitive. The computational complexity of the forward models is usually nonlinearly related—often $O(n^2)$ or more—to the number of spectral channels. As shown in Figure 4, the spectral and spatial resolution of infrared sounders has

increased dramatically over the last thirty-five years, and the computation required for real-time operation with variational algorithms has outpaced Moore's Law. There is, therefore, a motivation to reduce the computational burden of current and next-generation retrieval algorithms to allow real-time ingestion of satellite-derived geophysical products into numerical weather forecast models.

Dimensionality Reduction with Principal Component Transforms

Principal component transforms can be used to represent radiance measurements in a statistically compact form, enabling subsequent retrieval operators to be substantially more efficient and robust [6]. Furthermore, measurement noise can be dramatically reduced through the use of principal component filtering [7, 8] and it has also been shown [9] that principal component transforms can be used to represent variability in high-spectral-resolution radiances perturbed by clouds. In the following sections, we briefly discuss several variants of the principal component transform and focus on the ability of each to extract geophysical information from noisy radiance data.

The Principal Component Transform

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

$$c_1(\cdot) = E \left[(\hat{\tilde{\mathbf{R}}}_r - \tilde{\mathbf{R}})^T (\hat{\tilde{\mathbf{R}}}_r - \tilde{\mathbf{R}}) \right], \quad (2)$$

where $\hat{\tilde{\mathbf{R}}}_r \triangleq \mathbf{G}_r \tilde{\mathbf{R}}$ 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

* The following mathematical notation is used here: $(\cdot)^T$ denotes the transpose, $\hat{(\cdot)}$ denotes an estimate of a random vector, and $\tilde{(\cdot)}$ denotes a noisy random vector. Matrices are indicated by bold upper case, vectors by upper case, and scalars by lower case.

\mathbf{Q}_r^T contain the r most significant eigenvectors (ordered by descending eigenvalue) of the noisy-data covariance matrix $\mathbf{C}_{\tilde{\mathbf{R}}\tilde{\mathbf{R}}} = \mathbf{C}_{\mathbf{R}\mathbf{R}} + \mathbf{C}_{\Psi\Psi}$.

The Noise-Adjusted Principal Component Transform

Cost criteria in the form of error functions other than Equation 2 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 error equation thus becomes

$$c_2(\cdot) = E \left[(\hat{\mathbf{R}}_r - \mathbf{R})^T (\hat{\mathbf{R}}_r - \mathbf{R}) \right], \quad (3)$$

where $\hat{\mathbf{R}}_r \triangleq \mathbf{H}_r \tilde{\mathbf{R}}$ for some linear operator \mathbf{H}_r with rank r . The noise-adjusted principal component (NAPC) transform [10], 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

$$\mathbf{C}_{\tilde{\mathbf{W}}\tilde{\mathbf{W}}} = \mathbf{C}_{\Psi\Psi}^{1/2} (\mathbf{C}_{\mathbf{R}\mathbf{R}} + \mathbf{C}_{\Psi\Psi}) \mathbf{C}_{\Psi\Psi}^{-1/2},$$

maximizes the signal-to-noise ratio of each component, and is superior to the PC transform for most noise-filtering applications in which the noise statistics are known *a priori*.

The Projected Principal Component 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 Equation 3. We can show [11] that the optimal linear operator with rank r that minimizes Equation 3 is

$$\mathbf{L}_r = \mathbf{E}_r \mathbf{E}_r^T \mathbf{C}_{\mathbf{R}\mathbf{R}} (\mathbf{C}_{\mathbf{R}\mathbf{R}} + \mathbf{C}_{\Psi\Psi})^{-1}, \quad (4)$$

where $\mathbf{E}_r = [\mathbf{E}_1 | \mathbf{E}_2 | \dots | \mathbf{E}_r]$ are the r most significant eigenvectors of

$$\mathbf{C}_{\mathbf{R}\mathbf{R}} (\mathbf{C}_{\mathbf{R}\mathbf{R}} + \mathbf{C}_{\Psi\Psi})^{-1} \mathbf{C}_{\mathbf{R}\mathbf{R}}.$$

Examination of Equation 4 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 matrix containing the r most significant right eigenvectors, \mathbf{V}_r , of the reduced-rank linear-regression matrix \mathbf{L}_r , given in Equation 4. We then define the projected principal components of $\tilde{\mathbf{R}}$ as

$$\tilde{\mathbf{P}} = \mathbf{V}_r^T \tilde{\mathbf{R}}.$$

Note that the elements of $\tilde{\mathbf{P}}$ are correlated, since

$$\mathbf{V}_r^T (\mathbf{C}_{\text{RR}} + \mathbf{C}_{\Psi\Psi}) \mathbf{V}_r$$

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 correlated to the temperature profile is similar to Equation 4 and is given as

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

where $\mathbf{E}_r = [\mathbf{E}_1 | \mathbf{E}_2 | \dots | \mathbf{E}_r]$ are the r most significant eigenvectors of

$$\mathbf{C}_{\text{TR}} (\mathbf{C}_{\text{RR}} + \mathbf{C}_{\Psi\Psi})^{-1} \mathbf{C}_{\text{TR}}^T,$$

and \mathbf{C}_{TR} is the cross-covariance of the temperature profile and the spectral radiance.

Evaluation of Compression Performance Using Two Different Metrics

The compression performance of each of the PC transforms discussed previously was evaluated by 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. We first compress a radiance spectrum by using each of the PC transforms for a given number of coefficients. Then we use the resulting coefficients in a linear regression to estimate the temperature profile. The results were obtained with simulated clear-air radiance intensity spectra from an AIRS-like sounder. We generated approximately 7500 1750-channel radiance vectors with spectral coverage from approxi-

mately $4 \mu\text{m}$ to $15 \mu\text{m}$ by using a National Oceanic and Atmospheric Administration radiosonde set. The simulated intensities were expressed in spectral radiance units ($\text{mW m}^{-2} \text{sr}^{-1} (\text{cm}^{-1})^{-1}$).

PC Filtering. Figure 5(a) shows the sum-squared radiance distortion (from Equation 2) 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, in which the first variant (the PC curve) uses eigenvectors of $\mathbf{C}_{\tilde{\text{RR}}}$ 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. We show in Figure 5(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 . Note 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.

PC Regression. The PC coefficients derived in the previous example are now used in a linear regression to estimate the temperature profile. The temperature profile error (integrated over all altitude levels) for each of the PC transforms is shown in Figure 5(b) as a function of the number of coefficients used in the linear regression. To reach the theoretically optimal value achieved by linear regression with all channels requires approximately twenty PPC coefficients, 200 NAPC coefficients, and 1000 PC coefficients. Thus the PPC transform results in a factor of 10 improvement over the NAPC transform when we compress 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 sixty of the most significant NAPC coefficients of the 2378-channel

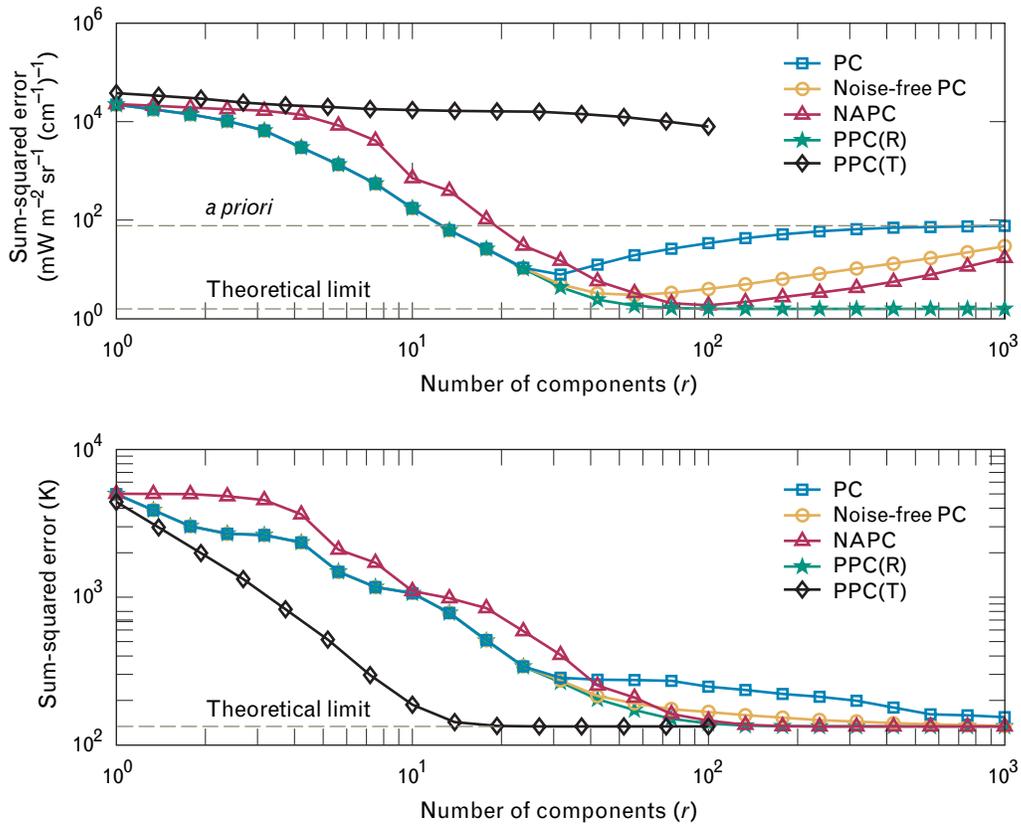


FIGURE 5. Performance comparisons of the principal components, in which the components are derived from both noisy and noise-free radiances, noise-adjusted principal components (NAPC), and projected principal components (PPC) transforms for a hypothetical 1750-channel infrared ($4\ \mu\text{m}$ to $15\ \mu\text{m}$) 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 upper plot shows the sum-squared error of the reduced-rank reconstruction of the noise-free spectral radiances. The lower plot shows the temperature-profile retrieval error (trace of the error covariance matrix) obtained by using linear regression with r components. The appropriate PPC transforms are more efficient, using fewer components to reach the theoretical limits.

AIRS spectrum (in units of brightness temperature) [6]. Results for the moisture profile are similar, although more coefficients (typically 35 versus 25 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 Figure 5, namely, PPC(R), when the basis for the noise-free radiance subspace

is desired, and PPC(T), when the basis for only the temperature profile information is desired. As shown in Figure 5(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, shown in Figure 5(b), offers a significantly less compact representation of temperature-profile information because the transform is representing information that is not correlated with temperature and thus superfluous when retrieving the temperature profile.

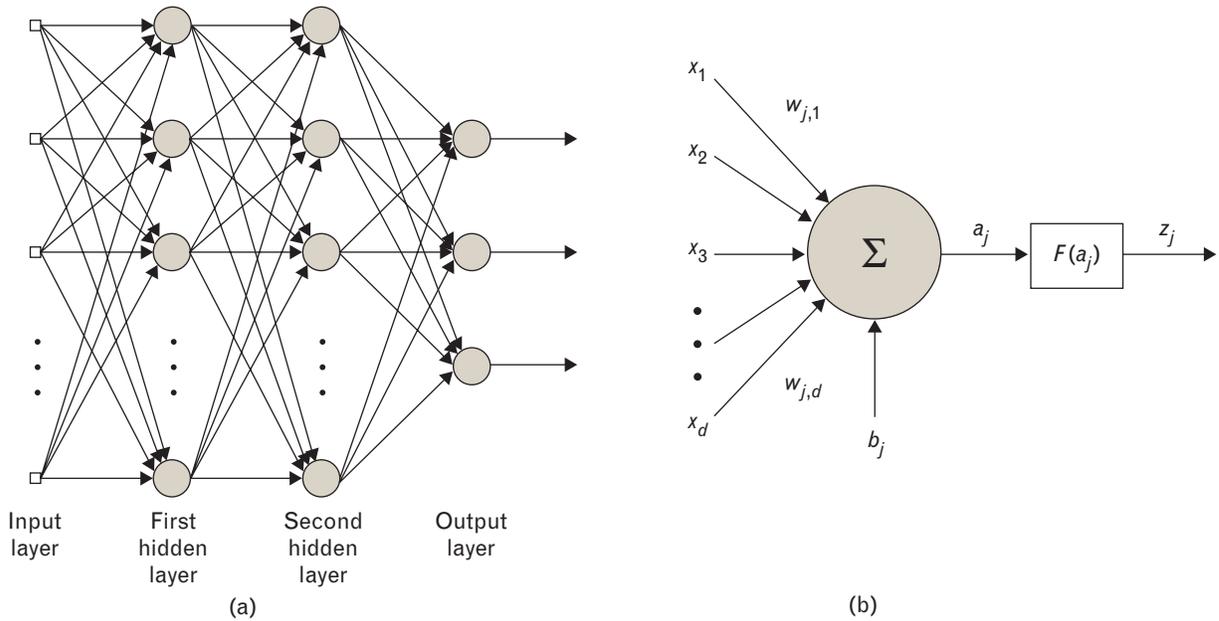


FIGURE 6. Neural network structure. (a) Interconnection of the multilayer feed-forward neural network, specifically, the multilayer perceptron, (b) the perceptron or node.

Neural Network Retrieval of Temperature and Moisture Profiles From High-Resolution Infrared and Microwave Sounding Data

A neural network is an interconnection of simple computational elements, or nodes, with activation functions that are usually nonlinear, monotonically increasing, and differentiable. Neural networks 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 [12, 13].

Introduction to Multilayer Neural Networks

Consider a multilayer feed-forward neural network, as illustrated in Figure 6, that consists 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 j th 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

$$E(\mathbf{w}) = \frac{1}{2} \sum_p \sum_k (t_k^{(p)} - y_k^{(p)})^2,$$

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, in which the local approximation of the cost function by a quadratic form is given by

$$E(\mathbf{w} + d\mathbf{w}) \approx E(\mathbf{w}) + \nabla E(\mathbf{w})^T d\mathbf{w} + \frac{1}{2} d\mathbf{w}^T \nabla^2 E(\mathbf{w}) d\mathbf{w}, \quad (5)$$

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 Equation 5 to zero and solving for the weight update vector $d\mathbf{w}$ yields

$$d\mathbf{w} = -\left[\nabla^2 E(\mathbf{w})\right]^{-1} \nabla E(\mathbf{w}). \quad (6)$$

Direct application of Equation 6 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, we can show that

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

and

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

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

$$\mathbf{S} = \sum_{p=1}^P \mathbf{e}_p \nabla^2 \mathbf{e}_p.$$

We can compute the Jacobian matrix by using a standard back-propagation technique [14] that is significantly more computationally efficient than direct calculation of the Hessian matrix [15]. 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, we assume that \mathbf{S} is zero (a reasonable assumption only near the solution), and Equation 6 becomes

$$d\mathbf{w} = -\left[\mathbf{J}^T \mathbf{J}\right]^{-1} \mathbf{J} \mathbf{e}.$$

The Levenberg-Marquardt modification [16] to the Gauss-Newton method is

$$d\mathbf{w} = -\left[\mathbf{J}^T \mathbf{J} + \mu \mathbf{I}\right]^{-1} \mathbf{J} \mathbf{e}.$$

As μ varies between zero and infinity, $d\mathbf{w}$ 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 [17]. The size of this region is governed by the value μ .

The use of multilayer feed-forward neural networks, such as the multilayer perceptron, to retrieve temperature profiles from hyperspectral radiance measurements has been addressed by several investigators [18, 19]. Neural network retrieval of moisture profiles from hyperspectral data is relatively new [20], but follows the same methodology used to retrieve temperature.

The PPC/INN Algorithm

A first attempt to combine the properties of both neural network estimators and PC transforms for the inversion of microwave radiometric data to retrieve atmospheric temperature and moisture profiles is reported in Reference 21, and a more recent study with hyperspectral data is presented in Reference 20. We take a conceptually similar approach in this work by combining the compression technique described in the section on the projected PC transform with the neural network estimator discussed in the previous section. PPC compression offers substantial performance advantages over traditional PC algorithms and is the cornerstone of the present work.

Network Topology. All multilayer perceptrons used in the PPC/INN 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, twenty-five PPC coefficients are input to six neural networks, each with a single hidden layer of fifteen nodes. We use separate neural networks for different vertical regions of the atmosphere; a total of six networks are used to estimate the temperature profile at sixty-five points from the surface to 50 mbar. For the water-vapor retrieval, we input thirty-five PPC coefficients to nine neural networks, each with a single hidden layer of twenty-five nodes. The water-vapor profile (mass-mixing ratio) is estimated at fifty-eight points from the surface to 75 mbar. We determined these network parameters largely through empirical analyses. Work is underway to dynamically optimize these parameters as the neural network is trained. We use separate training and testing datasets and discuss them in more detail later in this article.

Network Training. We initiated the weights and bi-

ases by using the Nguyen-Widrow method [22]. This method reduces the training time by initializing the weights so that each node is active (in the linear region of the activation function) over the input range of interest. We trained the neural network by using the Levenberg-Marquardt back-propagation algorithm discussed in the introductory section on multilayer neural networks. For each epoch, the μ parameter was initialized to 0.001. If a step succeeded, i.e., $E(\mathbf{w} + d\mathbf{w}) < E(\mathbf{w})$, then we decreased μ 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}). We stopped the network training when the error on a separate data set did not decrease for ten consecutive epochs. The sensor noise was changed on each training epoch to desensitize the network to radiance measurement errors.

Validation of the PPC/NN Algorithm

In this section, we evaluate the performance of the PPC/NN algorithm by using cloud-cleared AIRS data where the cloud clearing is performed with both AIRS/AMSU data and collocated European Centre for Medium-Range Weather Forecasts (ECMWF) forecast fields. We compare the PPC/NN retrieval performance with that obtained by using the AIRS Level-2 algorithm. We consider both ocean and land cases, including elevated surface terrain, and retrievals at all sensor scan angles (out to $\pm 48^\circ$) are derived. Finally, we present a sensitivity analysis of PPC/NN retrieval performance with respect to cloud amount.

Cloud Clearing of AIRS Radiances. We applied the cloud-clearing approach discussed in Reference 23 to the AIRS data by the AIRS Science Team. Several 3.x versions of the algorithm were used in this work. 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.

The AIRS/AMSU/ECMWF Data Set. We evaluated the performance of the PPC/NN algorithm by using 352,903 AIRS/AMSU observations and collocated ECMWF atmospheric fields collected on seven days from September 2002 to December 2003. We made various software version changes over the course of

this work, but these changes were primarily with regard to quality control and do not significantly affect the results presented here. However, the version 4.x release of the AIRS software, which was not available in time to be included in this work, should offer many enhancements over version 3.x, 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.

We randomly divided the 352,903 observations 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* root mean square (RMS) variation of the temperature and water-vapor (mass-mixing ratio) profiles contained in the validation set are shown in Figure 7. We matched the observations in the validation set with AIRS Level-2 retrievals obtained from the EOS Data Gateway (EDG). As advised in the AIRS Version 3.0 L2 Data Release Documentation, only retrievals that met certain quality standards (specifically, a Retrieval Quality Assurance Flag, RetQA-Flag, of zero for ocean and 256 for land) were included in the analyses. There were 17,856 AIRS Level-2 retrievals (all within $\pm 40^\circ$ latitude) that met these cri-

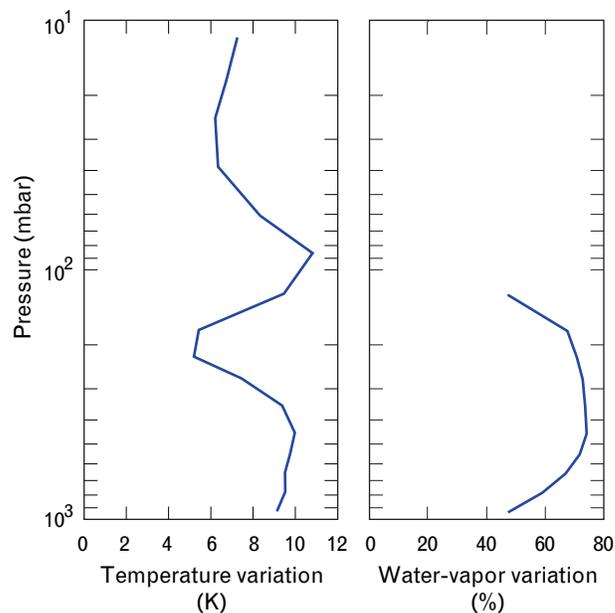


FIGURE 7. *A priori* temperature and water-vapor profile statistics for the validation dataset used in the analysis.

teria. Reanalysis with AIRS Level-2 version 4.x software is planned, since the version 4.x products have been validated over both ocean and land at near-polar latitudes.

To facilitate a comparison with results published in the AIRS v3.0 Validation Report, we calculated twenty-five layer error statistics as follows. 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, we calculated weighted water-vapor errors in each layer 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).

AIRS/AMSU Channel Selection. We discarded 37% (888 of the 2378) of the AIRS channels for the analysis; the radiance values for these channels frequently were flagged as invalid by the AIRS calibration software. Figure 8 illustrates a simulated AIRS brightness-temperature spectrum, showing the original 2378 AIRS channels and the 1490 channels that were selected for use with the PPC/NN algorithm. All fifteen AMSU channels were used. The algorithm automatically discounts channels that are excessively corrupted by sensor noise (for example, AMSU channel 7 on EOS Aqua) or other interfering signals (for example, the effects of non-local thermodynamic equilibrium) because the corruptive signals are largely uncorrelated with the geophysical parameters to be estimated.

PPC/NN Retrieval Enhancements for Variable Sensor Scan Angle and Surface Pressure. When dealing with global AIRS/AMSU data, we must accommodate a variety of scan angles and surface pressures. Therefore, we added two additional inputs to the neural networks discussed previously: (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 the number of inputs to the neural networks increased, we increased the number of hidden nodes in the neural networks used for temperature retrievals from fifteen to twenty. For water-vapor retrievals, we

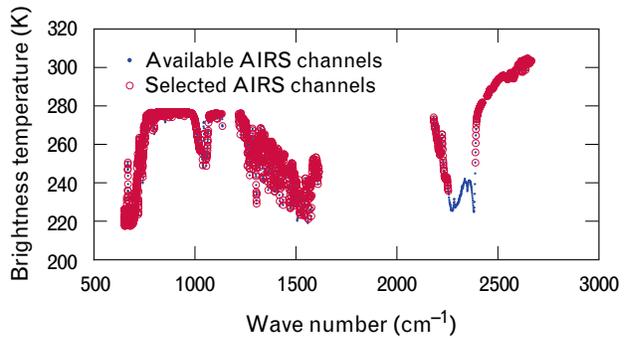


FIGURE 8. A typical AIRS spectrum (simulated) is shown. 1490 out of 2378 AIRS channels were selected.

maintained the number of hidden nodes in the first hidden layer at twenty-five, but we added a second layer of fifteen hidden nodes.

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, we derived the PPC/NN and linear-regression retrievals by using the same training set and the same validation set for all methods.

Figure 9 shows the temperature-profile retrieval performance over ocean for the linear regression retrieval, the PPC/NN retrieval, and the AIRS Level-2 retrieval. Figure 10 shows the corresponding water-vapor retrieval performance. We calculated the error statistics by using the 13,156 (out of 40,000) AIRS Level-2 retrievals that converged successfully. We found a bias of approximately 1K near 100 mbar between the AIRS Level-2 temperature retrievals and the ECMWF data (ECMWF was colder). We removed this bias prior to the computation of the AIRS Level-2 retrieval error statistics.

Figure 11 shows the temperature-profile retrieval performance over land for the linear regression retrieval, the PPC/NN retrieval, and the AIRS Level-2 retrieval. Figure 12 shows the corresponding water-vapor retrieval performance. We calculated the error statistics by using the 4700 (out of 10,000) AIRS Level-2 retrievals that converged successfully.

Several features in Figures 9 through 12 are noteworthy. First, for all retrieval methods, the performance over land is worse than that over ocean, as expected. The cloud-clearing problem is significantly more difficult over land, since variations in surface emissivity

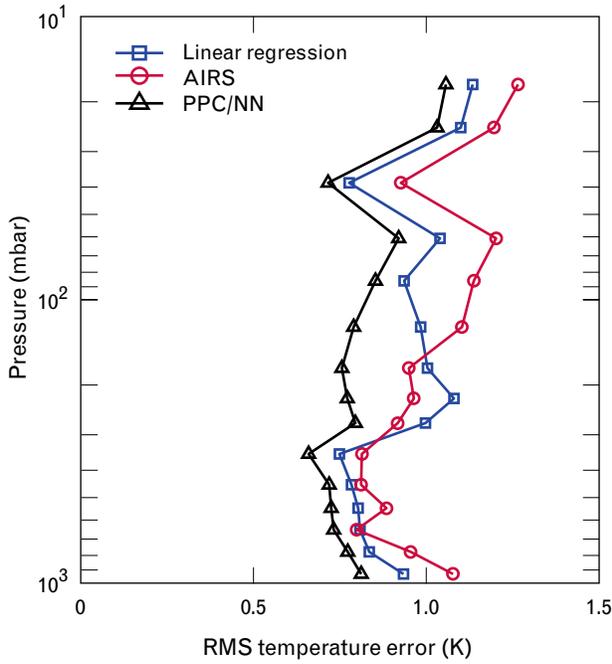


FIGURE 9. Temperature retrieval performance of the PPC/NN, linear regression, and AIRS Level-2 methods over ocean. Statistics were calculated over 13,156 fields of regard.

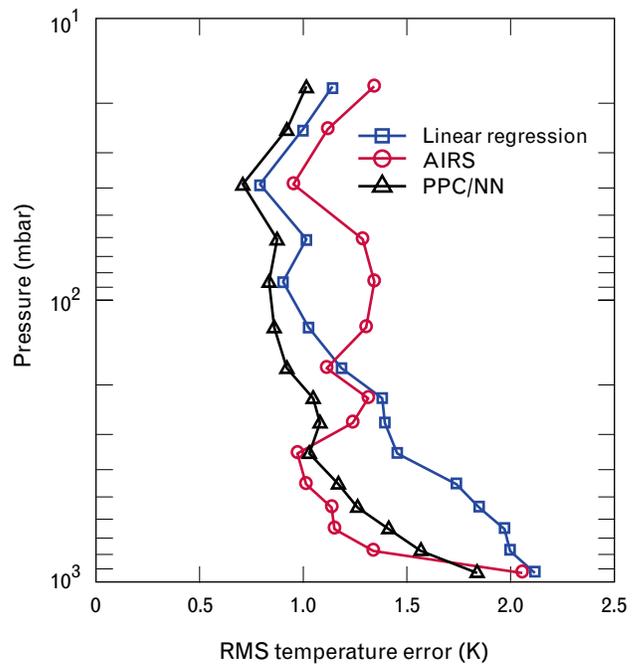


FIGURE 11. Temperature retrieval performance of the PPC/NN, linear regression, and AIRS Level-2 methods over land. Statistics were calculated over 4700 fields of regard.

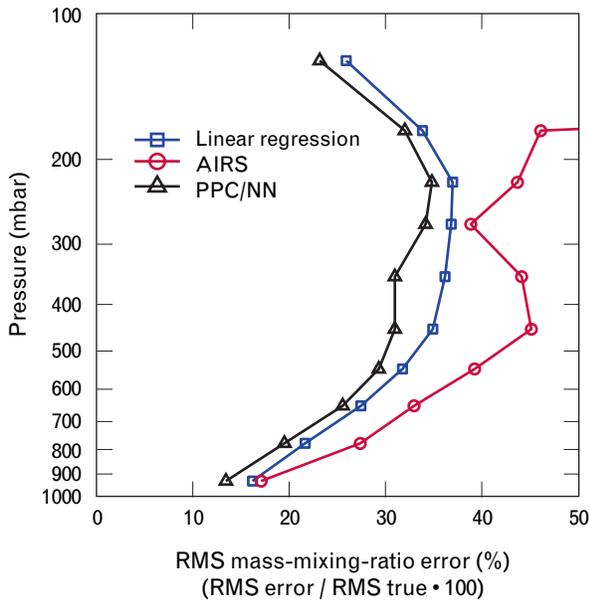


FIGURE 10. Water-vapor (mass-mixing ratio) retrieval performance of the PPC/NN, linear regression, and AIRS Level-2 methods over ocean. Statistics were calculated over 13,156 fields of regard. The anomalously high value for the AIRS Level-2 data below 200 mbar is an artifact of the errors in the atmospheric fields.

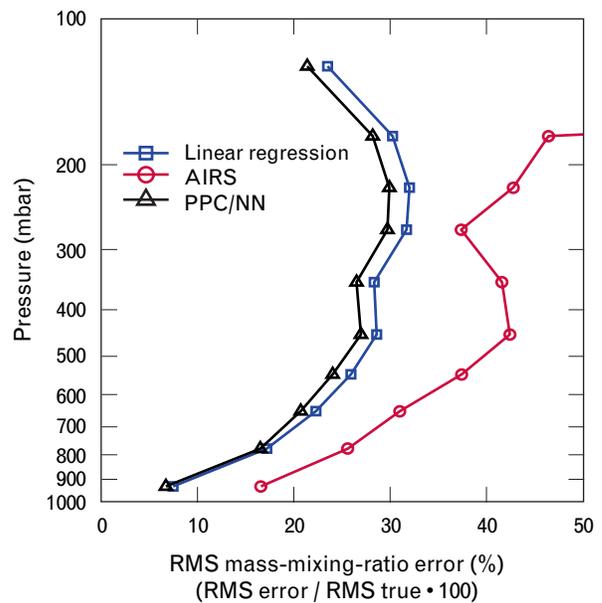


FIGURE 12. Water-vapor (mass-mixing ratio) retrieval performance of the PPC/NN, linear regression, and AIRS Level-2 methods over land. Statistics were calculated over 4700 fields of regard. The anomalously high value for the AIRS Level-2 data below 200 mbar is an artifact of the errors in the atmospheric fields.

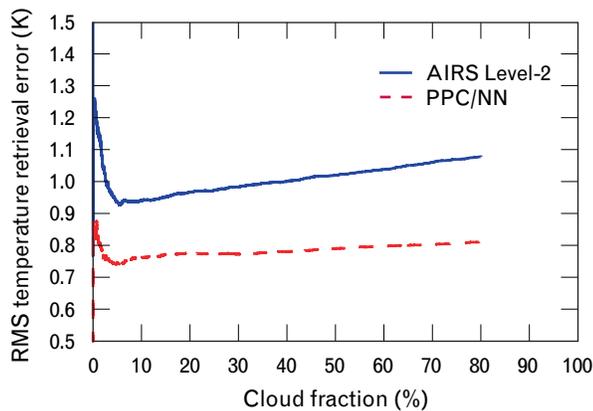


FIGURE 13. Cumulative RMS temperature error over ocean in the layer closest to the surface. Pixels were ranked in order of increasing cloudiness according to the retrieved cloud fraction from the AIRS Level-2 algorithm. No retrievals were attempted by the AIRS Level-2 algorithm if the retrieved cloud fraction exceeded 80%.

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 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. We are currently working on improving this performance.

Retrieval Sensitivity to Cloud Amount. Figure 13 shows the temperature retrieval error over ocean in the layer closest to the surface as a function of the cloud fraction retrieved by the AIRS Level-2 and PPC/NN algorithms. Similar curves for the water-vapor retrieval performance over ocean are shown in Figure 14. The large 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 retrievals are less sensitive than the AIRS Level-2 retrievals to cloud amount. Furthermore, it has been shown that the PPC/NN retrieval technique

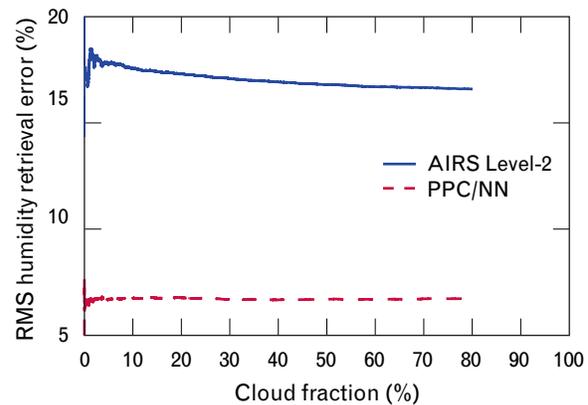


FIGURE 14. Cumulative RMS water-vapor error over ocean in the layer closest to the surface. Pixels were ranked in order of increasing cloudiness, according to the retrieved cloud fraction from the AIRS Level-2 algorithm. No retrievals were attempted by the AIRS Level-2 algorithm if the retrieved cloud fraction exceeded 80%.

is relatively insensitive to sensor scan angle, orbit type, and training set comprehensiveness [26].

Discussion and Future Work

While the PPC/NN performance results presented in the previous section are encouraging, we must mention several caveats. The ECMWF fields used for ground truth contain errors, and the neural network 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. Figures 10 and 12 show examples of these inaccuracies in the AIRS Level-2 retrieval of upper tropospheric water vapor. 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 in Figures 13 and 14. Work is currently under way 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.

Note that the PPC/NN algorithm as implemented in this work is currently not a stand-alone system, since 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, we must consider the steps required to implement the PPC/NN retrieval technique in an operational system. The required training methodology is most important. For reasons previously discussed, it is probably not feasible to derive the PPC/NN coefficients by using a forecast model. Two products from the EOS Aqua system should be invaluable sources of training data for future missions. The first is the generated database of radiances (both raw and cloud-cleared) and atmospheric parameter retrievals. Second, and more importantly, is the improved validation of surface, cloud, and transmittance models as a direct result of product validation efforts. The forecasting models, together with the database of atmospheric retrievals, will provide a ground-truth laboratory from which training datasets for future sensors can be derived.

Neural Network Retrieval of Precipitation from Microwave Sounding Data

In addition to estimating temperature and water-vapor profiles, neural networks have been shown to be useful for estimating precipitation. Because of the complexity of precipitation, existing physical models are not able to capture all of the physical variability of precipitation. An attractive alternative, therefore, is to derive a relationship between precipitation rate and observed brightness temperature by using neural networks that learn the dependencies through a training set of sample observations. This approach has been demonstrated to yield promising results. This section describes recent work in which we have used neural networks to estimate precipitation rates with data from opaque microwave frequency bands.

Precipitation Retrieval with AMSU

F.W. Chen and D.H. Staelin developed a neural-network-based precipitation-rate estimator for the Advanced Microwave Sounding Unit instruments AMSU-A and AMSU-B on the National Oceanic and Atmospheric Administration NOAA-15, NOAA-16, and NOAA-17 satellites, and the nearly identical AMSU/Humidity Sounder for Brazil (HSB) suite aboard the NASA EOS Aqua satellite [27–29]. A notable feature of these instruments is the presence of opaque microwave channels in the 54 GHz oxygen resonance band and in the 183 GHz water-vapor resonance band. Previous satellite-based passive microwave instruments operated only near frequencies where atmospheric absorption is relatively small (window channels). Some examples of these include the Tropical Rainfall Measurement Mission (TRMM) Microwave Imager and the Advanced Microwave Sounding Radiometer for the Earth Observing System on the NASA EOS Aqua satellite.

The Chen-Staelin algorithm is novel in that it relies primarily on opaque microwave channels; previous efforts relied exclusively on window channels such as those with frequency bands near 10, 19, 24, 31, 37, and 89 GHz. Window channels are useful for precipitation sensing over ocean because they are sensitive to the scattering signatures of ice particles and the emission signatures for water vapor. Over land, however, the emission signatures can be confused with surface variations.

Physical Basis for Precipitation Retrievals near 54 and 183 GHz. The opacity of the 54 GHz and 183 GHz channels on AMSU makes it possible to supplement the geophysical information found in window channels. A useful feature of the opaque channels is their sensitivity to specific layers of the atmosphere. Tables 1 and 2 show the channel frequencies used in this study. Figure 2 shows that AMSU-A channels above 53.6 GHz are largely surface blind and that they sample tropospheric temperatures and hydrometeors in broad layers at altitudes that increase with radio frequency up to around 57 GHz. This surface blindness is primarily due to oxygen attenuation.

All solid, liquid, or mixed-phase hydrometeors absorb and scatter electromagnetic radiation, typically

Table 1. AMSU-A Channels

<i>Channel</i>	<i>Channel Frequencies* (MHz)</i>	<i>Bandwidth (MHz)</i>
1	23,800 ± 72.5	2 × 125
2	31,400 ± 50	2 × 80
3	50,300 ± 50	2 × 80
4	52,800 ± 105	2 × 190
5	53,596 ± 115	2 × 168
6	54,400 ± 105	2 × 190
7	54,940 ± 105	2 × 190
8	55,500 ± 87.5	2 × 155
9	57,290.344 ± 87.5	2 × 155
10	57,290.344 ± 217	2 × 77
11	57,290.344 ± 322.2 ± 48	4 × 35
12	57,290.344 ± 322.2 ± 22	4 × 15
13	57,290.344 ± 322.2 ± 10	4 × 8
14	57,290.344 ± 322.2 ± 4.5	4 × 3
15	89,000 ± 900	2 × 1000

* The specific channel frequencies are determined from the table by adding and subtracting the values from the center frequency. For example, AMSU-A channel 1 frequencies are 23,727.5 and 23,872.5. There are four distinct frequencies for AMSU-A channels 11 through 14. They are calculated by adding and subtracting the primary frequency shift to the center and then adding and subtracting the secondary frequency shift to these two points. For example, the four AMSU-A channel 12 frequencies are 57634.544 (57290.344 + 322.2 + 22), 57590.544, 56990.144, and 56946.144.

introducing cold spots in the AMSU-A surface-blind radiance images. Both ice and liquid hydrometeors above 4 km are generally visible against the warmer opaque atmospheric background below. This visibility contrasts with the tendency of window channels over land to respond to both liquid hydrometeors and random surface variations [30]. Hydrometeors at very high altitudes produce a cold spot across all tropospheric AMSU-A channels, whereas low-altitude hydrometeors impact primarily those AMSU-A chan-

nels at the lower, more transparent radio frequencies that see down to those atmospheric depths. Because of this frequency-dependent penetration depth, 54 GHz spectra reveal the altitudes of precipitating cell tops even when they are hidden under thin cirrus or other light clouds [31]. In addition, the more transparent 54 GHz channels provide a measure of cloud albedo, which also depends on the hydrometeor particle-size spectrum [32] and type; ice clouds exhibit substantially higher albedos than the water clouds.

The observed cloud-top altitudes of convective cells generally correspond to the top of the graupel cloud thrust aloft. In this region, the average graupel sensed near 183 GHz is typically much smaller than that sensed near 54 GHz, and both populations are much larger and fall out much more rapidly than typical cloud particles sensed at infrared wavelengths. For such convective cells, the rain rate is closely related to the vertical velocity and the absolute humidity of the saturated air. This velocity is also directly related to the cell-top altitude, and we can estimate the humidity from nearby non-precipitating spots. Since higher vertical wind velocities are better able to support larger hydrometeors for long periods, they permit growth of water droplets and ice particles to sizes sufficient to perturb 54 GHz and lower frequencies.

This predicted relation between cell-top altitude and precipitation rate has been observed by others. For example, G.A. Vincent et al. compared 4 km resolution pairs of GOES-8 infrared images and instantaneous radar rainfall estimates obtained from operational 5 and 10 cm radars in the central Great

Table 2. AMSU-B Channels

<i>Channel</i>	<i>Channel Frequencies (MHz)</i>	<i>Bandwidth (MHz)</i>
1	89 ± 1	2 × 1
2	150 ± 0.9	2 × 1
3	183.31 ± 1	2 × 0.5
4	183.1 ± 3	2 × 1
5	183.31 ± 7	2 × 12

Plains and areas adjacent to the Gulf of Mexico [33]. In this work, only those rain systems showing clearly delineated convective cores in both the infrared and radar images were analyzed. A simple power law relationship between GOES-8 cloud-top temperature T (degrees K) and radar rainfall rate R was obtained for these special cases,

$$R \approx 1.12 \times 10^{11} e^{-0.0364T^{1.2}} \text{ mm/h},$$

yielding an RMS discrepancy between GOES-8 and radar rainfall rate estimates of 7.2 mm/h over a dynamic range of 0 to 150 mm/h.

The advantages of AMSU data relative to such GOES observations include an improved ability to retrieve temperature and humidity profiles near precipitation and an ability to see through overlying cirrus down to the more substantive graupel. Moreover, the differences in microwave responses near 183 and 54 GHz reveal information about the graupel size distribution. Larger hydrometeors strongly affect both 54 GHz and 183 GHz spectra, whereas smaller ones affect primarily 183 GHz spectra alone. Particle size is a particularly good indicator of vertical updraft velocities because only larger velocities can sustain large particles aloft, and these large velocities are also directly conveying saturated air into low-temperature zones where the humidity must condense and precipitate. Thus a simple relationship linking convective velocity and cell-top altitudes to precipitation rates might be expected. Although GOES can resolve smaller precipitation cells, those with diameters over 15 km produce the most rain. In addition to properties of precipitation particles, the 54 GHz and 183 GHz channels also provide information about the temperature and water-vapor profiles, which determine the amount of water that can be precipitated.

Overview of the Neural-Network-Based Precipitation-Rate Retrieval Algorithm. The key to the method of Chen and Staelin was to process the sensor data in a way that extracted the most useful information related to precipitation. On the basis of the preceding discussion and experimental analysis, the following were selected as inputs to the neural network: (1) cloud-induced brightness-temperature perturbations from five AMSU-A channels in the 54 GHz band; (2) three temperature-profile PPC computed by using

cloud-cleared brightness temperatures in five AMSU-A channels in the 54 GHz band; (3) two water-vapor profile principal components; (4) brightness temperatures from three AMSU-B channels in the 183 GHz band; and (5) the secant of the satellite zenith angle.

Precipitation rates should vary monotonically with these inputs. Therefore, a simple neural network with a structure similar to that shown in Figure 6(a) would be appropriate. We trained our Chen and Staelin method by using data from the NEXRAD ground-based radar network, which has good coverage over the eastern U.S. This method also has shown plausible results for precipitation globally.

Brightness-Temperature Perturbation Analysis. Figure 15 shows images of intermediate variables based on data over a storm front on 13 September 2000, around 0130 UTC. Figure 15(a) shows brightness temperatures from the 183 ± 7 GHz channel. In this image, the storm front is evident as a series of cold perturbations that span an area from southern Louisiana to northern Alabama. In the algorithm, the 183 ± 7 GHz channel is the primary channel for detecting precipitation, since it is sensitive to most precipitation without being sensitive to the surface. In some cold and dry regions, the 183 ± 7 GHz channel can be sensitive to the surface. In such cases, the more opaque 183 ± 3 GHz channel, shown in Figure 15(b), is used. Of the 183 GHz channels available on AMSU-B, the 183 ± 1 GHz channels is the most opaque. However, in extremely rare cases, even this channel can be sensitive to surface variations.

Figure 15(c) shows cloud-induced brightness-temperature perturbations near 52.8 GHz. This image lacks the fine structure evident in Figure 15(a) because AMSU-A has coarser resolution than AMSU-B. In order to retrieve precipitation rate at 15 km resolution, we had to sharpen this image from 50 km to 15 km resolution by using 183 ± 7 GHz data; Figure 15(d) shows the result.

Sample Retrieval Results. Figure 16 shows a comparison of 15 km AMSU rain rates that is based on the Chen and Staelin method with results from the NEXRAD ground-based radar network. The AMSU image agrees with the NEXRAD image in morphology. The AMSU method also tends to place heavy precipitation where NEXRAD places heavy precipitation.

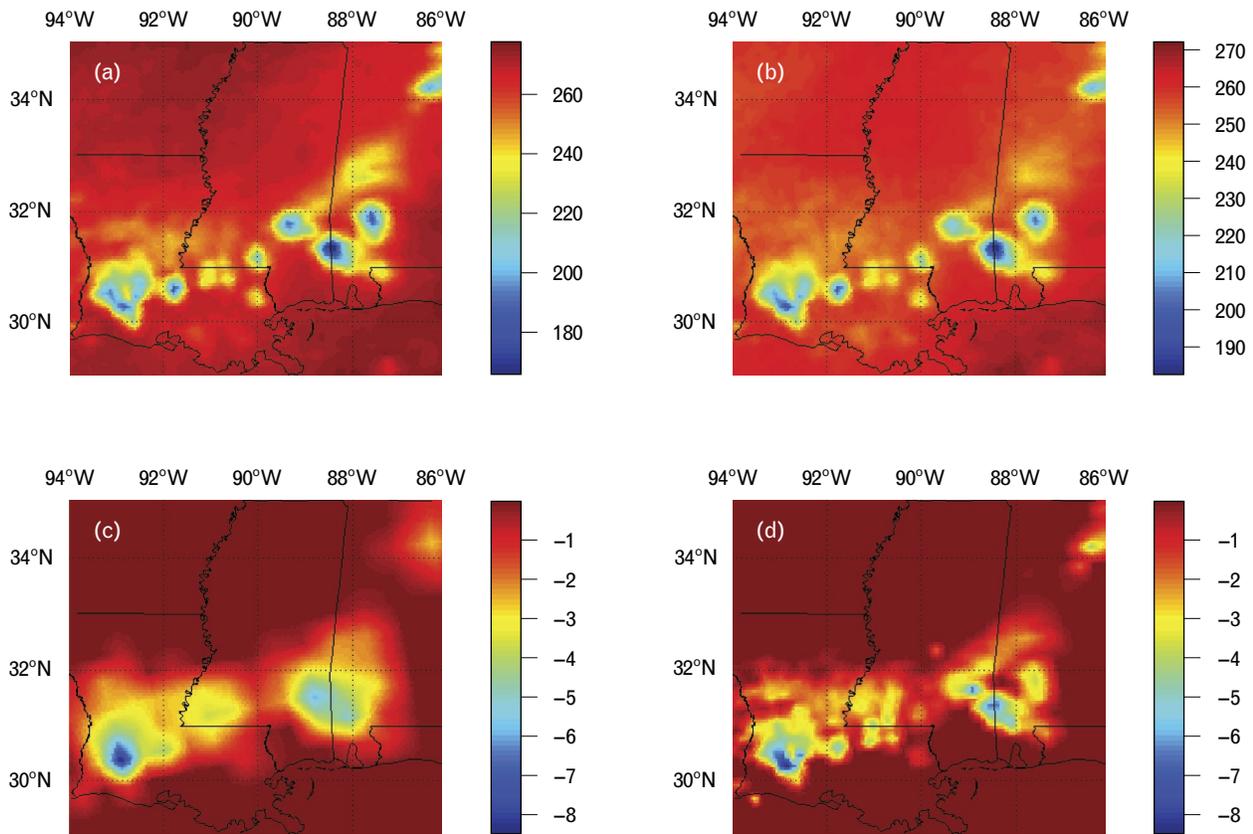


FIGURE 15. Frontal system 13 September 2000, 0130 UTC: (a) Brightness temperatures near 183 ± 7 GHz. (b) Brightness temperatures near 183 ± 3 GHz. (c) Brightness temperature perturbations near 52.8 GHz. (d) Inferred 15-km-resolution brightness temperature perturbations near 52.8 GHz [28].

Figure 17 shows AMSU retrievals over Hurricane Isabel (2003) while it was a Category 5 hurricane. The eye is clearly visible in the precipitation image, and the precipitation rates appear plausible on the basis of the 183 ± 7 GHz brightness temperatures.

This method is also useful for detecting snow. Figure 18 shows the retrieval results for a snowstorm on 17 February 2003. The image of 183 ± 3 GHz brightness temperatures shows a dark-blue cold band moving northward through New England. The peak precipita-

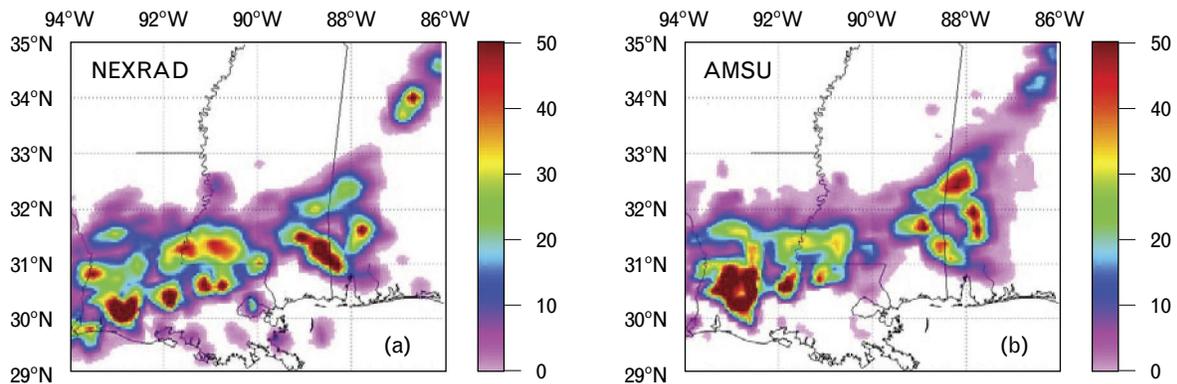


FIGURE 16. Precipitation rates (millimeters per hour) above 0.5 mm/h observed 13 September 2000, 0130 UTC. (a) Fifteen-kilometer-resolution NEXRAD retrieval and (b) 15 km resolution AMSU retrieval.

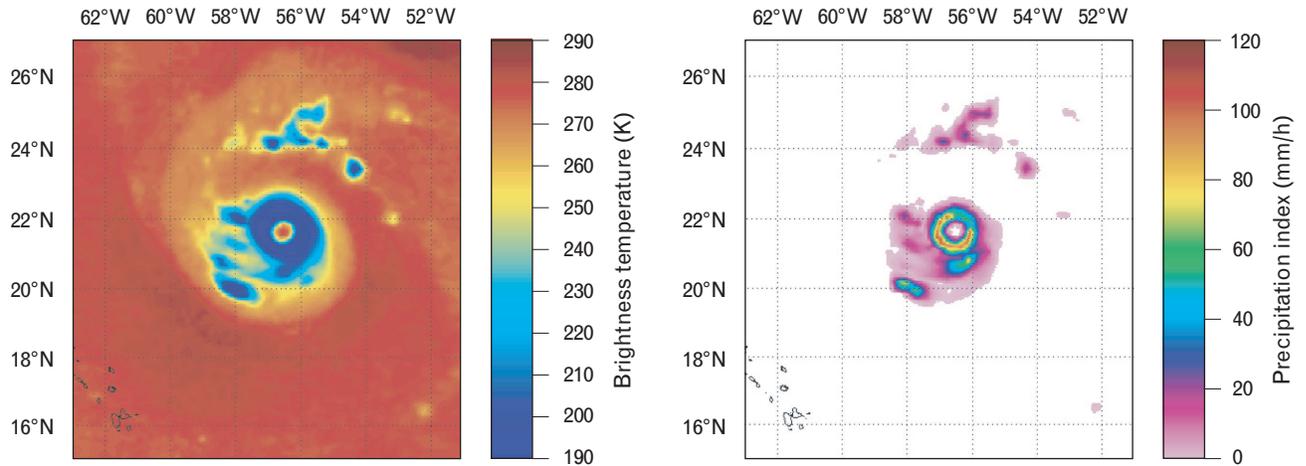


FIGURE 17. Brightness temperatures near 183 ± 7 GHz (left) and retrieved 15 km precipitation (right) over Hurricane Isabel (category 5) observed 12 September 2003, 0610 UTC.

tion index over New England is about 1 mm/h, which was not able to produce the 27.5 inches reported for Boston on that day [34]. Although the algorithm has yet to be tuned to accurately estimate snowfall rate, it still has demonstrated the ability to detect snowfall, which is not surprising; the algorithm is sensitive to ice particles that proceed upward from the tops of convective precipitation. Whether the particles end up as rain or snow at the surface depends largely on the surface temperature.

Precipitation Climatology

The presence of AMSU-A/B on the NOAA-15, -16, and -17 satellites, whose orbital paths are shown in Figure 19, have local equatorial crossing times at near-

ly regular intervals throughout the day, which facilitated a study of the diurnal variations of precipitation. The diurnal variation of precipitation is a critical issue that researchers have hoped the Tropical Rainfall Measurement Mission (TRMM) would help to address. Knowledge of the diurnal variation of precipitation will contribute to a better understanding of the daily variation of atmospheric heat transport.

The studies in References 35 and 36 present results based on data from the NOAA-15, -16, and -17 satellites extending toward 65° N and 65° S. We used the retrievals based on the method of Chen and Staelin to compute parameters of diurnal variations. Figure 20 shows the diurnal variations of precipitation frequency for retrievals from July 2002 to June 2003

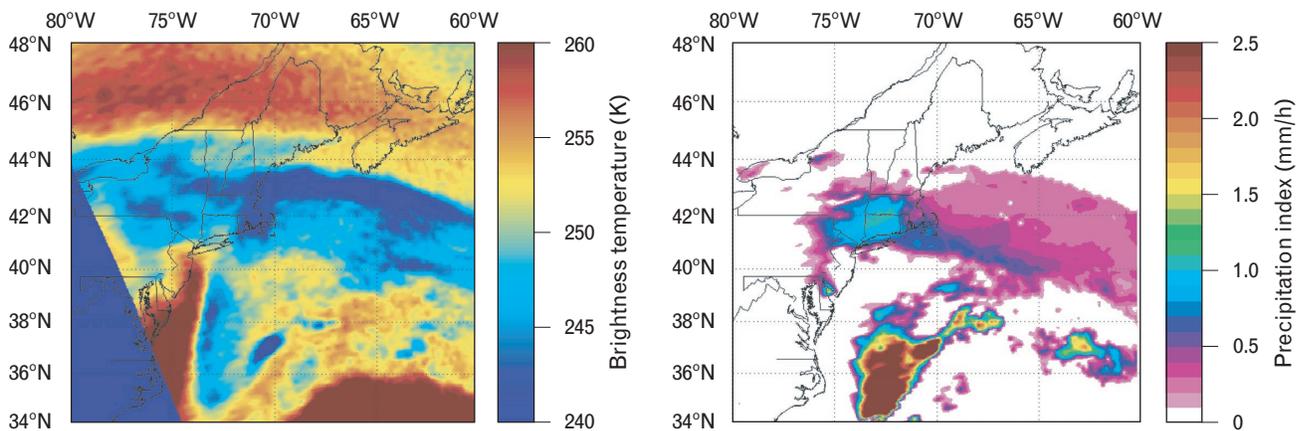


FIGURE 18. Brightness temperatures near 183 ± 3 GHz (left) and retrieved 15 km precipitation (right) over a snowstorm observed on 17 February 2003, 1742 to 1747 UTC.

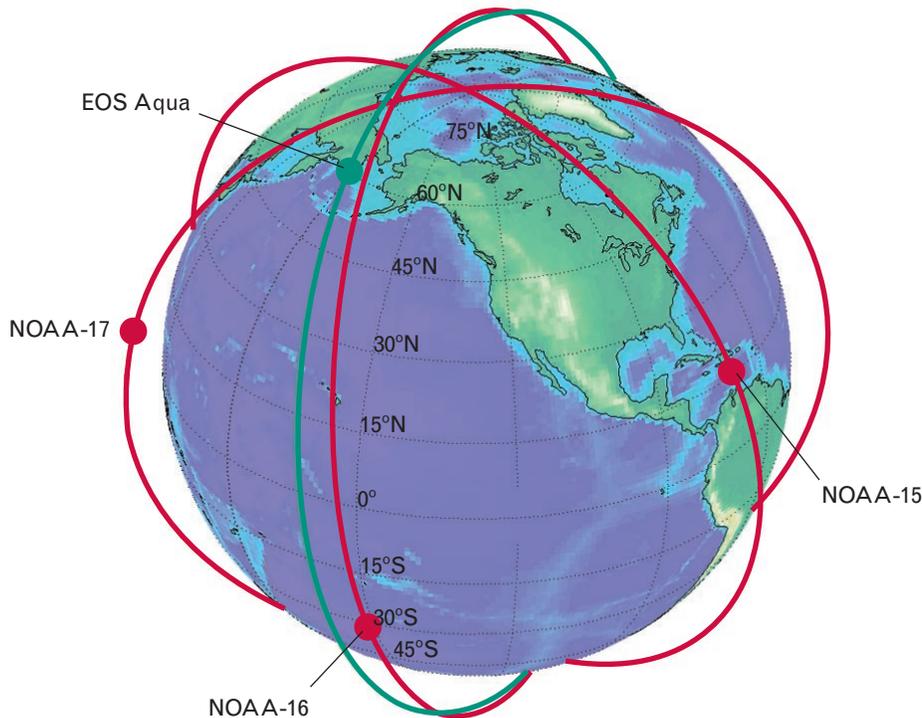


FIGURE 19. Orbital patterns for the National Oceanographic and Atmospheric Administration NOAA-15, NOAA-16, and NOAA-17 satellites (red) and the NASA EOS Aqua satellite (green).

[35, 36]. For this image, precipitation frequency for a location is defined as the number of events in which the retrieved precipitation rate is greater than 1 mm/h divided by the number of observations. The precipitation frequency image shows large areas where the precipitation frequency is less than 1%; in the Pacific Ocean west of South America and west of the United States, in the northern Atlantic Ocean west of northwestern Africa, and in the southern Atlantic Ocean west of southern Africa. Most of the area where the precipitation frequency is higher than 10% is in the tropical latitudes. This morphology is consistent with other global precipitation averages such as those produced by the Global Precipitation Climatology Project and the Climate Prediction Center Merged Analysis of Precipitation, both of which combine satellite-based retrievals and rain gauge measurements [37]. In Figure 20, the arrows representing preferred local solar time of precipitation and the circles representing the percentage diurnal variation (relative to the mean precipitation frequency) show large areas of smooth variations. For example, over most of Brazil the preferred

local time of precipitation is about 2200, and east of the United States, it is about 1300. The largest percentage of diurnal fluctuations are found in Australia, in southern Africa, and along the Rocky Mountains. Note that the largest percentage diurnal fluctuations are found in areas where the mean precipitation frequency is low.

Discussion and Future Work

The algorithm of Chen and Staelin has demonstrated the utility of opaque microwave channels for estimating precipitation. The results obtained with this neural-network-based method show that neural networks can be used in place of or in addition to physics-based methods. We hope that this neural-network-based method can be improved by using more sophisticated methods to process the observations before they are presented to the neural network. Some possibilities include the use of other signal processing techniques such as independent component analysis and the use of scattering and emission signatures seen in window-channel measurements.

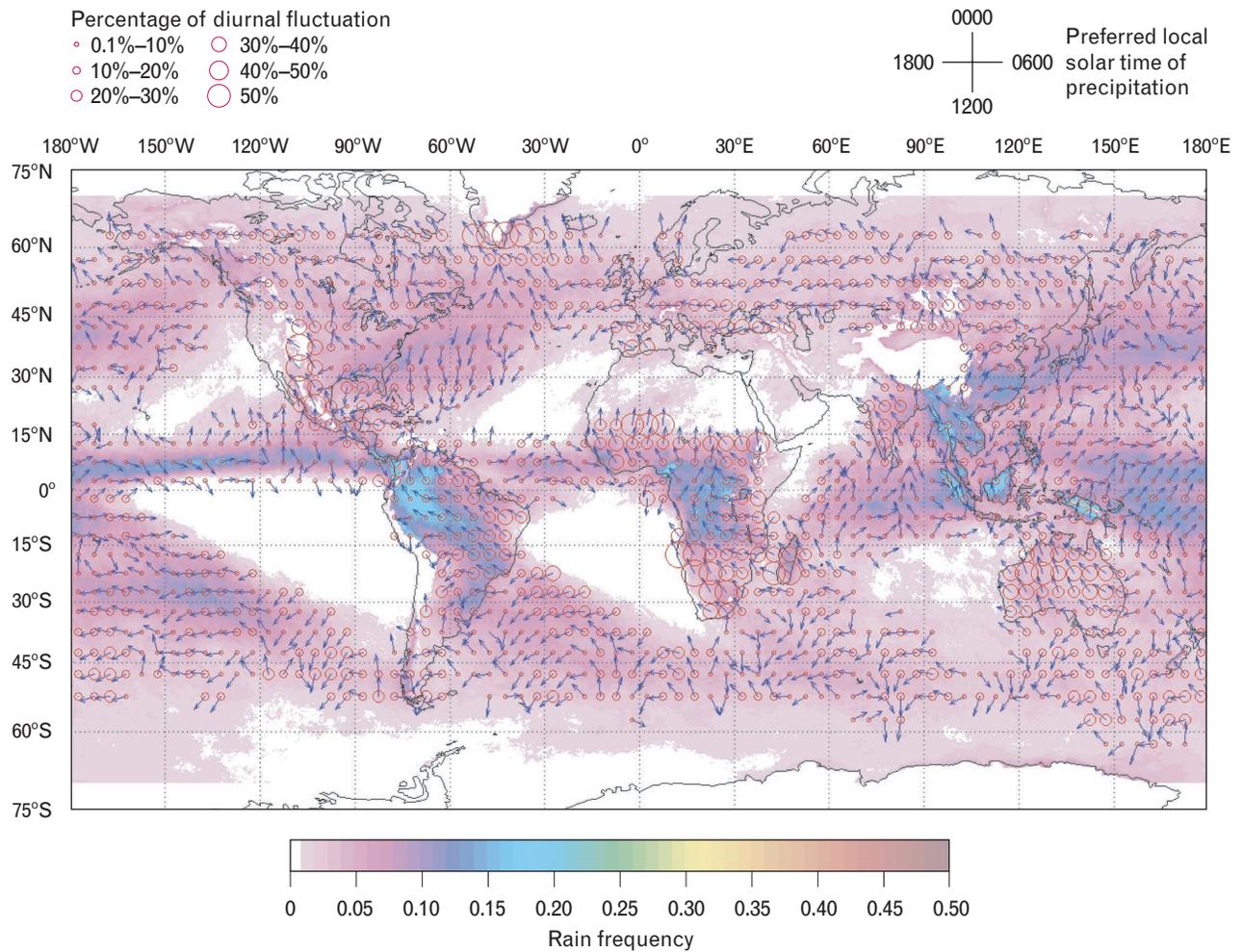


FIGURE 20. Diurnal variations of precipitation frequency computed for AMSU observations between July 2002 and June 2003 from the NOAA-15, NOAA-16, and NOAA-17 satellites. Precipitation frequency is defined as the fraction of observations in which the precipitation retrieval is greater than 1 mm/h. Precipitation frequency is represented by the background image, the arrows indicate the preferred local solar time of precipitation, and the circles indicate a range of percentage of diurnal fluctuation.

The work of Chen and Staelin can easily be adapted and improved for future instruments like the Advanced Technology Microwave Sounder (ATMS) on the National Polar-Orbiting Operational Environmental Satellite System (NPOESS) and the NPOESS Preparatory Project. ATMS has most of the channels found on AMSU-A/B and a few additional channels. ATMS has several advantages over AMSU-A/B and AMSU/HSB: (1) ATMS samples brightness temperatures from all channels at a fixed sampling rate that is finer than that of AMSU-A and AMSU; (2) the 54 GHz channels have finer resolution on ATMS than on AMSU-A and AMSU; and (3) ATMS observes

at a wider range of scan angles than AMSU-A/B and AMSU/HSB

The 23.8 GHz and 31.4 GHz channels on ATMS have significantly coarser resolution on ATMS than on AMSU-A and AMSU. However, the advantages of ATMS are likely to lead to retrievals with better accuracy and resolution.

Summary

We presented two neural-network-based parameter retrieval methods: a temperature- and moisture-profile retrieval method using infrared and microwave data and a precipitation-rate retrieval method us-

ing primarily opaque microwave data. The PPC/NN temperature- and moisture-profile retrieval technique combines a linear-radiance-compression operator with a neural network estimator. We showed that the projected principal component (PPC) transform is 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 neural networks to be used, thereby improving both the stability and computational efficiency of the algorithm. Test cases with 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. We obtained 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 by using the PPC/NN retrieval method with AIRS/AMSU data in partially cloudy areas. We also performed PPC/NN retrievals with partially cloudy AIRS/AMSU data over land. The PPC/NN retrieval technique is 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. Our 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.

We also presented the precipitation-rate retrieval method of Chen and Staelin. Examples illustrate that AMSU and similar operational instruments flying on NOAA-15, -16, -17, NASA EOS Aqua, and other meteorological satellites will provide an important new resource for routine precipitation monitoring on

a global scale. Evaluations of rain rate with 15 km and 50 km nominal resolution suggest that AIRS/AMSU rain rate retrievals will usefully supplement other global precipitation data sets over both land and sea at rates approaching 100 mm/h. The examples presented show that 50 km resolution instantaneous precipitation rate retrievals agree quite well with NEXRAD results when both 54 GHz and 183 GHz spectral data are employed, and most 15 km spots precipitating more than 1 mm/h should be readily identifiable.

Acknowledgments

The authors would like to thank the National Oceanic and Atmospheric Administration National Environmental Satellite, Data, and Information Service (NOAA NESDIS), Office of Systems Development, and the National Aeronautics and Space Administration for financial support of this work. The collocated AIRS/AMSU/ECMWF data sets were provided by the AIRS Science Team, via Mitch Goldberg (NOAA NESDIS Office of Research Applications). The author would also like to thank Lihang Zhou for help obtaining and interpreting the AIRS/AMSU/ECMWF data sets, and David Staelin and Philip Rosenkranz for many helpful discussions on all facets of this work.

REFERENCES

1. H. Hotelling, "The Most Predictable Criterion," *J. Educational Psychology* **26**, 1935, pp. 139–142.
2. J. Escobar-Munoz, A. Chédin, F. Cheruy, and N. Scott, "Réseaux de Neurones Multi-Couches pour la Restitution de Variables Thermodynamiques Atmosphériques à l'Aide de Sondes Verticaux Satellitaires," *Comptes Rendus de L'Académie des Sciences, série 2*, **317** (7), 1993, pp. 911–918.
3. D.H. Staelin, "Passive Remote Sensing at Microwave Wavelengths," *Proc. IEEE* **57** (4), 1969, pp. 427–439.
4. M.A. Janssen, *Atmospheric Remote Sensing by Microwave Radiometry* (Wiley, New York, 1993).
5. C.D. Rodgers, "Retrieval of Atmospheric Temperature and Composition from Remote Measurements of Thermal Radiation," *J. Geophys. Res.* **14** (7), 1976, pp. 609–624.
6. 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 Sens.* **41** (2), 2003, pp. 379–389.
7. H.-L. Huang and P. Antonelli, "Application of Principal Component Analysis to High-Resolution Infrared Measurement

- Compression and Retrieval," *J. Appl. Meteorol.* **40** (3), 2001, pp. 365–388.
8. 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.* **107** (D22), 2002.
 9. W.J. Blackwell and D.H. Staelin, "Cloud Flagging and Clearing Using High-Resolution Infrared and Microwave Sounding Data," *Proc. IGARSS 3, Toronto, 24–28 June 2002*, pp. 1860–1862.
 10. 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 Sens.* **28** (3), 1990, pp. 295–304.
 11. W.J. Blackwell, "Retrieval of Cloud-Cleared Atmospheric Temperature Profiles from Hyperspectral Infrared and Microwave Observations," Ph.D. Thesis, Dept. of Electrical Engineering and Computer Science, MIT, Cambridge, Mass., 2002.
 12. K.M. Hornik, M. Stinchcombe, and H. White, "Multilayer Feedforward Networks Are Universal Approximators," *Neural Netw.* **4** (5), 1989, pp. 359–366.
 13. S.S. Haykin, *Neural Networks: A Comprehensive Foundation* (Macmillan College Publishing Co., New York, 1994).
 14. D.E. Rumelhart, J.L. McClelland, and the PDP Research Group, *Parallel Distributed Processing: Explorations in the Microstructure of Cognition, Vol. 1: Foundations* (Bradford Books/MIT Press, Cambridge, Mass., 1986).
 15. M.T. Hagan and M.B. Menhaj, "Training Feedforward Networks with the Marquardt Algorithm," *IEEE Trans. Neural Netw.* **5** (6), 1994, pp. 989–993.
 16. P.E. Gill, W. Murray, and M.H. Wright, "The Levenberg-Marquardt Method," in *Practical Optimization* (Academic Press, London, 1981), pp. 136–137, 140, 259.
 17. C.M. Bishop, *Neural Networks for Pattern Recognition* (Oxford University Press, New York, 1995).
 18. 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.* **34** (24), 1995, pp. 5390–5397.
 19. 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.* **41** (2), 2002, pp. 144–159.
 20. 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.* **107** (D22), 2002.
 21. F. Del Frate and G. Schiavon, "A Combined Natural Orthogonal Functions/Neural Network Technique for the Radiometric Estimation of Atmospheric Profiles," *Radio Sci.* **33** (2), 1998, pp. 405–410.
 22. D. Nguyen and B. Widrow, "Improving the Learning Speed of Two-Layer Neural Networks by Choosing Initial Values of the Adaptive Weights," *Proc. Int. Joint Conf. on Neural Networks 3, San Diego, 17–21 June 1990*, pp. 21–26.
 23. J. Susskind, C.D. Barnett, and J.M. Blaisdell, "Retrieval of Atmospheric and Surface Parameters from AIRS/AMSU/HSB Data in the Presence of Clouds," *IEEE Trans. Geosci. Remote Sens.* **41** (2), 2003, pp. 390–409.
 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. E. Fetzer, ed., "Validation of AIRS/AMSU/HSB Core Products for Data Release Version 3.0," NASA, JPL Technical Report D-26538 (13 Aug 2003).
 26. W.J. Blackwell, "A Neural-Network Technique for the Retrieval of Atmospheric Temperature and Moisture Profiles from High Spectral Resolution Sounding Data," *IEEE Trans. Geosci. Remote Sens.*, 2005, accepted for publication.
 27. D.H. Staelin and F.W. Chen, "Precipitation Observations near 54 and 183 GHz Using the NOAA-15 Satellite," *IEEE Trans. Geosci. Remote Sens.* **38** (5), 2000, pp. 2322–2332.
 28. F.W. Chen and D.H. Staelin, "AIRS/AMSU/HSB Precipitation Estimates," *IEEE Trans. Geosci. Remote Sens.* **41** (2), 2003, pp. 410–417.
 29. F.W. Chen, "Global Estimation of Precipitation Using Opaque Microwave Bands," Ph.D. Thesis, Dept. of Electrical Engineering, MIT, Cambridge, Mass., 2004.
 30. M.D. Conner and G.W. Petty, "Validation and Intercomparison of SSM/I Rain-Rate Retrieval Methods over the Continental United States," *J. Appl. Meteorol.* **75** (7), 1998, pp. 679–700.
 31. M.J. Schwartz, J.W. Barrett, P.W. Fieguth, P.W. Rosenkranz, M.S. Spina, and D.H. Staelin, "Observations of Thermal and Precipitation Structure in a Tropical Cyclone by Means of Passive Microwave Imagery near 118 GHz," *J. Appl. Meteorol.* **35** (5), 1996, pp. 671–678.
 32. A.J. Gasiewski and D.H. Staelin, "Numerical Modeling of Passive Microwave O₂ Observations over Precipitation," *Radio Sci.* **25** (3), 1990, pp. 217–235.
 33. G.A. Vincente, R.A. Scofield, and W.B. Menzel, "The Operational GOES Infrared Rainfall Estimation Technique," *Bull. AMS* **99** (9), 1998, pp. 1883–1898.
 34. S. Potter, "Name That Storm," *Weatherwise* **58** (1), 2005, pp. 24–29.
 35. F.W. Chen, D.H. Staelin, and C. Surussavadee, "Global Monthly and Diurnal Precipitation Statistics Based on Passive Microwave Observations from AMSU," *AMS 13th Conf. on Satellite Meteorology and Climatology, Norfolk, Va., 20–23 Sept. 2004*.
 36. F.W. Chen and D.H. Staelin, "Diurnal Variations of Precipitation Observed Using Opaque Microwave Frequency Bands," *85th AMS Mgt: 19th Conf. on Hydrology, San Diego, 9–13 Jan. 2005*.



WILLIAM J. BLACKWELL is a staff member in the Sensor Technology and System Applications group. He received a B.E.E. degree from the Georgia Institute of Technology and S.M. and Sc.D. degrees from MIT in 1994, 1995, and 2002, respectively, all in electrical engineering. 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.



FREDERICK W. CHEN is a staff member in the Sensor Technology and System Applications group. He is currently working on atmospheric remote sensing problems including satellite-based estimation of precipitation and atmospheric trace gases using microwave and infrared data. His research interests include signal separation, transform coding, data fusion, data compression, and other areas in signal processing and information theory. He has previously worked at Argonne National Laboratory (Argonne, IL) in the Reactor Analysis Division, and at the MIT Research Laboratory of Electronics in the Remote Sensing and Estimation Group. He received S.B., M.Eng., and Ph.D. degrees in electrical engineering from MIT in 1998, 1998, and 2004, respectively.