PCA enabled retrieval of temperature and humidity with hyperspectral sounder radiances

Abhishek Verma[a] and Dr.C.Balaji[b]
[a] IIT Madras, Chennai 600036, India
verma.iitm@gmail.com
[b]Associate Professor, Department of Mechanical Engineering, IIT Madras, Chennai 600036, India
balaji@iitm.ac.in

Abstract: A simulations study is used to demonstrate the application of principal component analysis(PCA) in retrieving geophysical parameters using high spectral infrared radiances from the Atmospheric Infrared Sounder (AIRS) as inputs. Realistic and synthetic atmospheric profiles were used to simulate upwelling long wave radiances at TOA by kCompressed Atmospheric Radiative Transfer Algorithm (kCARTA) model. These combinations of information profiles with their respective infrared radiances were utilized to derive their statistical relationships in two stages. First, principal component transform was applied to infrared radiances recorded at each channels of the sounder. Second, a feed forward back propagation type artificial neural network (ANN) was used to retrieve temperature and humidity profile using reduced number of components. The number of components in input was varied in orders of 10^3 to 10^1 and performance of the algorithm was recorded. Similarly, principal component based linear regressions were carried out with same datasets and its performance was compared with ANN performance. Overall, the study concludes that principal component analysis optimizes statistical prediction of geophysical parameters.

Keywords: high spectral resolution, retrieval algorithm, principal component, artificial neural network

NOMENCLATURE

Alphabetic Symbols

a_{im}	regression coefficient of i^{th} pressure layer and m^{th} term
b	bias of the node
p_l	l^{th} principal coefficient
q	atmospheric humidity, $\frac{g}{Kg}$
t	atmospheric parameter measured (K, $\frac{g}{Kg}$)
w	weight of the linkage between nodes
w	weight matrix of the Artificial Neural Network
y	atmospheric parameter predicted (K, $\frac{g}{Kg}$)
	response of node
z	height from the earth's surface, m
	activation function output
A	design matrix for regression
C	covariance matrix
E	cost function
P	matrix of principal coefficients

Q	matrix of eigen vectors
R	matrix of radiances
Ŕ	matrix of mean subtracted radiances
Ř	matrix of noisy mean subtracted radiances
T	atmospheric temperature, K
Y	matrix of atmospheric parameters

Greek Symbols

μ	mean atmospheric profile
η	randomly generated variable in the range [-3,3]
σ	standard deviation of atmospheric profile
Ψ	synthetic atmospheric profile

Subscripts

i	for i^{th} pressure level
	i^{th} node in previous layer
j	j^{th} node in current layer
nor	normalized
η	at a given wavenumber, cm^{-1}
r	in a given layer

Superscript

p	for p^{th} step of learning
T	transpose of matrix

1 INTRODUCTION

In the past few decades, monitoring of the earth's atmosphere has improved drastically, leading to accurate observation of atmospheric temperature, water vapor and other constituents present in minute quantities. Modern atmospheric sounders equipped with high resolution interferometer are providing highly dense scanning of the atmosphere with all-round temporal coverage. The fabrication of new sensor to measure radiances with high resolution are providing with an increased information of physical properties related to the earth's radiation budget, which has improved our understanding of various physical phenomenon, vaguely known previously. This advancement has come at the cost of generating voluminous correlated data, which are testing the usage of conventional retrieval techniques.

Many retrieval techniques (both physical and statistical) have successfully been implement in the prediction of temperature, moisture profiles although the new sounder instruments are setting new standards of retrieval accuracy, unprecedented by any other sounder. In specific, the operational Atmospheric InfraRed Sounder(AIRS) onboard the NASA's Aqua Satellite, has improved this aspect with surpassing accuracy. The sounder provides a spatial resolution of ≈ 15km, a spectral resolution of $\nu / \Delta\nu \approx 1200$ and a radiometric accuracy on the order of $\pm$ 0.2 K. With 2378 channels ranging from 650 cm^{-1} to 2675 cm^{-1}, we intend to exploit high dimensional measurements concept for retrieval algorithm development.

In this paper, a simple nonlinear retrieval algorithm is developed using Principal Component Analysis (PCA) technique on an artificial neural network (ANN), trained for solving the inverse problem related to

the transference of earth's radiative energy. The goal of this study is to reveal the performances of PCA based ANN with coupled temperature and humidity retrievals. The network architecture is designed such that it predicts the temperature and humidity profiles together in one processing. The performance of the network is studied based on the robustness, accuracy of retrieval. This study is further extended to analyze the effect of interferences from instrument noise on the predictions. In addition, comparison were drawn between the performances of PCA based ANN and PCA based regression, which represents the currently used regression model, for the temperature and humidity retrieval respectively.

The following explains the organization of this paper: In section 2, literatures relevant to the present work are explained. Section 3 gives a context description of the mathematical formulation involved with a brief review toward the forward radiative transfer model used. This section also includes the inverse model retrieval using PCA to extract, reduce and denoise the input space. Section 4 provides the results of the retrieval for various cases discussed. Finally, conclusions are presented in section 5.

2 REVIEW OF LITERATURE

Inverse solution to Radiative Transfer Equation (RTE) mainly deals with the estimation of atmospheric properties from known or measured quantities like radiation intensity or brightness temperatures from the sounder instruments. The algorithms implemented to solve them are generally categorized as physical retrievals and statistical approaches depending on their formulation and operation. The physical retrievals, basically, work by solving the physics of RTE, an integral-differential equation and validating the solution to satellite measured data. A statistical method, on other hand, derives a numerical relationship between the measured radiances and atmospheric parameters.

Regardless of its type, the basis of any retrieval algorithm in meteorological application, is the forward model or RTE. Various methods like 'line by line' code, fast forward models have been developed to accurately simulate electromagnetic radiances for specified atmospheric conditions. The Atmospheric Infrared Sounder or AIRS fast forward algorithm (*Hannon et al* (1996)) is one of the advancements made to provide radiances measured by the AIRS in infrared range. In this algorithm, the effective monochromatic layer to space transmittance of the atmosphere is calculated by interpolating among a dataset of specifically chosen 100 atmospheric profiles, representing various possible atmospheric conditions. Models like k Compressed Radiative Transfer Algorithm (kCARTA) (*DeSouza-Machado* (1997)) has also been developed to simulate the infrared radiances using a compressed monochromatic absorption coefficient database. kCARTA, though relatively slow in computation, provides precise infrared radiances when compared with AIRS fast forward model and is therefore, used in this work for generating simulated radiances.

The use of multilayer feed forward Artificial Neural Network (ANN) for retrieval of atmospheric parameters using hyper spectral measurements was first proposed by *Escobar-Munozetal* (1993). Thereafter, many improved retrieval techniques of ANN have been documented using high spectral measurements. The advantage of ANN inversion method is that it provides a model of inverse radiative transfer function parameterized once and for all conditions, in contrast to conventional inversion techniques.

Development of an efficient data compression procedure is crucial for data handling from high spectral resolution measurements and subsequent remote sensing. The application of PCA to data compression is one of the advanced statistical methods used in reducing high resolution spectral reading from sounders. *Huang and Antonelli* (2000) have investigated the application of PCA with regression in retrieval of temperature and humidity distribution. Besides obtaining an optimum compression ratio, a desirable subset of principal components (PC) was obtained which gave high retrieval accuracy. They have also discussed the correlation between the PCs and the atmospheric temperature / humidity distribution and noise removal from the spectral data achieved by PCA.

Combining both the ANN model with PCA gives an enhanced performance of remote sensing. *Aires* [1] *et al* (2002) applied the concept of PCs for compressing, denoising and first guess retrieval from the reading of Infrared Atmospheric Sounding Interferometer (IASI), a high resolution sounder. This analysis used 8461 channel of IASI as against 2378 channels of AIRS. The first guess retrieval was then applied to a PC based ANN (*Aires*[2] *et al* (2002)) to retrieve simultaneously the temperature, water vapor and ozone profile. The result obtained showed an overall improvement in accuracy of the retrievals using ANN.

An advanced version of PCA, known as Projected PC Transform (PPC) was used by *Blackwell* (2005) along with ANN for retrieval purposes. The performance of PPC based ANN algorithm in retrieving temperature / humidity profiles was very good when compared with PPC regression and Iterated Minimum Variance Techniques (IMV).

The literature review shows that retrieval of vertical atmospheric profiles from observed infrared radiances were done by various techniques ranging from statistical to non-linear physical methods. Uncertainty in the retrievals is more pronounced for humidity profile then for temperature profile. In the present work, only clear sky situations were taken into account. Increase in number of spectral channels brings in the problem of channel selection too, but reduces the ill-posed nature of the retrieval. The inverse problem dealing with the retrieval of vertical atmospheric profiles from the observed infrared satellite radiances needs a retrieval algorithm that is robust so that real time retrievals can be performed with less uncertainty, whenever they are required. The present work aims at studying the performance of PC based ANN algorithm in retrieving the temperature / humidity profile using high spectral resolution measurements. PC based regression algorithm is also developed for the retrieval and its performance is compared. The algorithms are trained using mixed sets of synthetic and realistic profiles and the retrieval accuracies are tested extensively for both synthetic and realistic profile independently.

3 MATHEMATICAL FORMULATION AND SOLUTION PROCEDURE

3.1 Forward model

Lines by line radiative transfer models are very essential in the retrieval of atmospheric parameter using satellite measurements. They effectively utilize the physics of radiative transfer for any participating medium and successively calculate the physical attributes for each spectral line i.e. by 'line by line' radiative transfer. These models perform the computation of monochromatic layer transmittance for a specified atmospheric condition, which are further evaluated to compute spectral radiances.

3.2 Solution(kCARTA)

kCARTA is an infrared, monochromatic radiative transfer algorithms used here for non scattering condition in the earth's atmosphere. The radiance simulated in present work is pertaining to nadir viewing of the AIRS on board NASA's Aqua Satellite. The radiative transfer equation solved by the model uses absorption coefficients computed by k-compressed lookup tables, compressed using singular value decomposition techniques *Strow* (1996). These absorption coefficients are strongly dependent on absorbing gas amount and show a slow variation with temperature. Therefore, the absorption coefficients are computed by interpolated with respect to the temperature within the k-compressed database for any arbitrary atmospheric profile. These coefficients are then used in the radiative transfer equation to give accurate radiances in the infrared region of electromagnetic wave. The whole coverage of 2378 AIRS channels lying between 650 cm-1 to 2700-1, has the capability to capture the absorption behavior of almost all constituents of the atmosphere and therefore is simulated by kCARTA for various atmospheric conditions.

3.3 Database creation

The ground truths for the studies were radiosonde profiles from many sources. The profile chosen were irrespective of attributes considered for operational mode retrieval such as day-night classification, ascending-descending orbit and climatic seasons to demonstrate a generic retrieval procedure. For present work only tropical region profiles were used. With sufficient number of realistic profile available, various synthetic profiles feasible in nature were generated and added to enlarge the datasets.

Assuming that each parameter follows a Gaussian distribution about its mean value, an artificial value of the parameter Ψ is generated by $\Psi = \mu + \eta\sigma$, where $\eta \in R$ is a random number in the range [-3,3], σ is its standard deviation, giving a 99 percent probability of occurrence in nature. By applying the criterion of non negative values of Ψ and using a single random number for each profile generation, it is ensured that the profiles made are physically realistic. A mixed dataset of around 3500 realistic profiles and 7000 synthetic profiles was used for setting up the relationship between radiance parameters and their corresponding profiles. Including synthetic profiles in the training process increases the robustness of prediction at the same time amplifies statistical features of information used. Testing the generalization were two dataset, hereafter refered as TYPE1 and TYPE2, shown in the TABLE 1

3.3 PCA

It is imperative that retrieval procedures involve maximal usage of geophysical content of hyper spectral radiance data for its prediction. Principal Component Analysis or PCA is one of the powerful statistical tool which has been successfully utilized for improving data usage. It involves categorizing data on the basis of their covariance with each other and effectively reduces the data dimensionality by discarding less relevant dimensions. Besides being a valuable tool for compressing spectral data, PCA is also used for smoothening instrument noise that is contained in radiance measurements by neglecting higher order variances.

3.3.1 PCA without noise interference

In the computation, a single noise free infrared reading of the AIRS channels was considered as a radiance vector {R} in m dimensional subspace (m =2378). With p readings from the AIRS channels, the radiance vector was arranged column wise in a matrix $[R]_{mxn}$ of mxn dimensions. If Q_k^T is the PC operator, a linear, orthonormal operator which projects a m dimensional radiance vector into an k dimensional (k < m) subspace then principal components $[P]_{kxn}$ of the radiances are $[P] = Q_k^T[\acute{R}]$. $[\acute{R}] = [R] - R_\mu$ are reduced radiances with R_μ mean of radiance vector subtracted from each observation. The rows of Q_k^T contains the k most significant eigenvectors (sorted by descending eigenvalues) of data covariance matrix C_{RR} , where $C_{RR} = \langle \acute{R}.\acute{R} \rangle$.

3.3.2 PCA with noise interference

For computational convenience, the brightness temperature spectrum of AIRS channels was considered as input {B} in m = 2378 dimensional subspace. The noise from instrument was simulated using the noise equivalent temperature difference (NEΔT) of the sensors from the L2 IR Channel Properties IN_FLIGHT, VERSION v6.8.1, available at the AIRS website airs.jpl.nasa.gov. This noise was assumed to be Gaussian-distributed with zero mean and standard deviation σ_{noise} dependent on the measured radiances {R} as

$$\sigma_{noise} = NE\Delta T \left(\frac{\partial R}{\partial B} \Big/ \frac{\partial R_{250K}}{\partial B_{250K}} \right) \tag{1}$$

where R_{250K} , B_{250K} represents the value of the measured radiance and brightness temperature at a scene temperature of 250K, same as that of NEΔT provided. These σ_{noise} parameters for each channel are used to generate the noisy spectrum of brightness temperatures {B} corresponding to{R} radiances. For PCA on noisy data, the covariance matrix is modified as $C_{\acute{R}\acute{R}} = C_{RR} + C_{noise}$ and the corresponding eigen vectors are

then used for PCA reduction. C_{noise} is the covariance of channels noise, which is taken to be for the mean radiances observed for clear sky condition.

The evaluation of PCA retrievals were done on the basis of two important outcomes; (i) accuracy of prediction in terms of mean RMS error, which should be within the standards set by World Meteorological Organization (i.e. within 1K for temperature retrieval and 20-30% relative error for humidity retrieval) and (ii) best compression or component reduction possible to maintain the accuracy limits.

Each eigen vector of the covariance matrix C_{RR} represents a unique, linear and orthogonal variance contributor to the dataset. Projecting each radiance observation onto the eigen vector, will capture the variance feature represented by this vector. By ordering the vectors on the basis of their eigen values, we intend to classify the variance features of the data. The sorted eigen values decreases exponentially in magnitude, thereby limiting important features to only first few components, enabling the option of compression in the inputs. Another way of measuring the contribution of various components is by reconstructing the original data using first few principal components. The reconstructed data $[\acute{R}_R] = Q_k Q_k^T [\acute{R}]$ is obtained by the operator $Q_k Q_k^T$; comparison with the original data $[\acute{R}]$ will show an error deviation. FIGURE 1 depicts the reconstruction of radiances using variances contained in their principal components.

TABLE 5.2 shows typical eigen values of prominent vectors along with their cumulative contributions in variance. The first three eigen vector account to 95% variance contribution in dataset TYPE1 which makes them the most important features of the dataset. Yet, in the retrieval method an average of 100 to 300 principal components are used for the most optimal prediction (*Huang* (2001)) and up to 500 principal components are used for noisy radiances dataset (*Blackwell* (2005)). Shown in FIGURE 2, 3 are the first three prominent eigen vectors in TYPE1 and TYPE2 dataset respectively.

3.4 Retrieval technique

3.4.1 PCA based regression

Regression methods were the first retrieval algorithms used in sounding techniques. They are statistical or non-physical retrieval methods because they do not explicitly make use of the radiative transfer physics in computation. The advantage of such methods is that they are used independent of any radiative transfer model, which makes their performance much faster and efficient. Measured radiance data can be colocated with radiosonde measurements to build sets of training data directly (*Xuebao et al.*(2005)).

A linear additive (hyperplane) model is used in the regression of the atmospheric parameters y_i of i^{th} pressure level, with $y_i = a_{i0} + a_{i1}p_1 + a_{i2}p_2 + a_{i3}p_3 + a_{i4}p_4 + ... + a_{ik}p_k$ where p_l (l = 1,2,.....k) , a_{im} (m = 0,2,.....k) are the chosen k principal components or predictor variable with their respective regression coefficients. With n readings, the equation becomes a matrix equation of the form [Y] = [A][P] , $[Y]_{pxn}$ being the matrix of atmospheric parameters having temperature and humidity profiles arranged in columns, $[A]_{px(k+1)}$ being the matrix of regression coefficients arranged in rows for each parameter y_i and principal components matrix $[P]_{(k+1)xn}$ with the components arranged in column and $p_o = 1$ for all columns. The regression coefficients are calculated for the training sets, validated with the testing sets and the cost function E is evaluated to record the performance of the regression for each layer i.e.

$$E_i = \sum_{l=1}^{N} (t_{li} - y_{li})^2 \tag{2}$$

where t_i of i^{th} pressure level is target value of the parameter. This value of E or Mean Square Error (MSE) is utilized to analyze the overall accuracy of the retrieval, thereby assisting in searching for an optimal compression ratio for regression based retrievals.

The use of principal components as predictor vectors gives a large compression of radiance data and makes the regression more stable. Reduction of principal components becomes more relevant when number of channels involved is quite high, as in with the Atmospheric Infrared Sounder (AIRS) radiance data of 2378 channels used in present work. The linear additive model used here is a basic regression model, which could be improved by applying complex regression methods. However, in general, linear regression forms a sort of benchmark against which retrieval by other methods like ANN can be compared.

3.4.2 PCA based ANN

The architecture of any neural network comprises of an interconnection of basic processing units or nodes. The connectivity with in the network is such that, each node receives a weighted sum of responses from nodes present in previous layer. To create a following response, the incoming signal is processed by an activation function, with adjustments to the signal by adding a bias. A response when generated by the node is passed on to next nodes through connectivity, thereby repeating the same process till final output is reached at the outer layer (FIGURE 4) .

These activation function, linear or non-linear in nature, are the most integral part of neural network, as they determine the learning performance of the whole architecture. Most commonly used activation functions are linear, sigmoid, hyperbolic tangent and radial basis functions. The added advantage of ANN is its learning ability i.e. to deduce an input-output relationship from a training ensemble irrespective of the distribution of data.

For the present work, the neural network toolbox of MATLAB®7.0 (Demuth (2009)) was used. A feed forward back propagation neural network was chosen for an input layer of principal components with an output layer consisting of normalized temperature and water vapor content and no hidden layers. The first 96 outputs nodes of temperature were normalized by T_{nor} = 300 K while the water vapor content in the output were normalized by q_{nor} = 15 g/Kg . The output layer contains the linear activation function of the form $z_j = s_j$, where $s_j = \Sigma_i w_{ji} y_i + b_j$ and y_i is a response by i^{th} node in previous layer. The weights w_{ji} connecting i^{th} node with j^{th} node and bias b_j are iteratively calculated by the back propagation learning method. TRAINSCG® routine or scaled conjugate gradient method was typically used for the training, which can handle a large dataset (10000 to 12000 sets of input / output combination). TRAINSCG® routine uses the conjugate gradient search method, where the search direction is determined conjugate to the previous search direction (i.e. to combine new steepest descent direction with the previous search directions), along with the model-trust region approach (used in the Levenberg-Marquardt algorithm).

The overall performance of a trained ANN is given by the cost function defined in equation 1 . For analyzing training performance of an ANN, for a set of weights and bias w, a cost function is chosen E(w) such that

$$E^p(w) = \frac{1}{2}\sum_k (t_k^p - z_k^p)^2 \tag{3}$$

where $t_k(p)$ and $z_k(p)$ denotes the network target and output, respectively, of each output node k given at p^{th} iterative step of learning. Most of the back propagation algorithms uses this parameter explicitly to update weights / bias of the network. As many as 10000 epochs or generation were used for each case to ensure adequate training of the network and sufficient convergence of solution.

For noisy radiance data, the neural network was operated such that the noise in the training input were changed in every 10 epoch steps. The reason for such operation is to make the network more robust to any random noise of the instrument. Therefore, irrespective of any number in the noisy data, the neural network should give a prediction more or less equivalent to the prediction by the noise free component in the radiance spectrum.

There is no established method to determine the optimum number of hidden layers or the number of neurons in each layer and so they have to be determined iteratively. When there are too few neurons in the layers, a large bias exist between the output and target. Too many neurons in the network results in over fitting, where the network is trained excessively for training ensemble but performs badly when tested for generalization. Therefore, an optimal network is determined heuristically by operating for various configurations. Convergence and computation time are also some attributes which are considered when evaluating the network. TABLE 3 & 4 shows the performance of various network architecture using different number of PC typically 100, 250 and 500 in number as the optimal number of principal component is supposed to lie in this range. Typically, linear activation based neural network with no hidden layer have shown high performance in the retrievals. Principal component classification simplifies the variance distribution within the radiances which reduces the need for a complex architecture for simulation. It is quite likely that the relative performance for neural network TYPE2 be similar to TYPE1 with variations only in magnitude. Therefore, its definitive that TYPE2 generalization will produce accurate results with no hidden layer, linear activation architecture.

4 RESULTS AND DISCUSSION

For verification of retrievals by ANN model, results from linear regression were studied and compared with ANN results. For our analysis, principal components were used as predictor variables and the atmospheric temperature and humidity profile were retrieved simultaneously. As described by Aires[2] (2002), the coupled retrieval of temperature and humidity makes the problem more constrained, thereby resulting in an optimal retrieval of water vapor content. The trend in the retrievals by using different number of principal components is noted to give a conclusive idea on its performance. For overall comparison, the layer averaged mse, defined in equation (2), of various network were calculated and compared; individual retrieval of temperature and humidity were also taken into account. Temperature root mean square error (RMSE) evaluated in Kelvin represents the error in temperature retrieval where as humidity RMSE were evaluated in g/Kg of dry air and percentage with respect to mean humidity distribution for judging the humidity predictions.

Theoretically, by increasing the number of components, the result converges to most accurate prediction and this is evident in Principal Component (PC) based regression. For dataset TYPE 2, which include only realistic radiosonde profile, the PC based ANN and PC based regression retrievals were tested. In our case, PC based regression represent the most accurate prediction possible with the principal components and is plotted along with the PCA based ANN for both the cases i.e. (1) without noise interference and (2) with noise interference.

(a) *PCA based ANN without noise interference*: The optimal retrieval performance for PCA based ANN is shown in FIGURE 6 with the water vapor retrieval performance in FIGURE 7 and FIGURE 8. As shown in TABLE 5 the bias for lower PCs decrease to an optimal and increases for higher PCs retrievals. TABLE 6 depicts the percentage performance of humidity upto certain thresholds of pressure, upto 100hPa and upto 300hPa. Furthermore, the profile performance of water vapor retrieval for higher components is good only up to 300 hPa for most of the cases (FIGURE 9). This implies that an optimal number of components exist, for both, temperature and humidity prediction, which leads to the most accurate prediction by ANN. Overall, 400 PCs were found to give the least mse for the retrievals (FIGURE 10, 11,12).

(b) *PCA based ANN with noise interference*: The addition of instrument noise disturbs the retrieval performance leading to a decrease in accuracy inspite of PCA reduction. The number of principal component required to achieve optimal retrieval increases and therefore increase the dimensionality and computation required in the process. FIGURE 13, 14 shows the retrieval performance of temperature and humidity with noise respectively, with FIGURE 15 depicting the percentage humidity error. This being suggested in the TABLE 7 and 8, where performances for various numbers of PCs are classified and their respective mse mentioned. Overal, 500 PCs were found to give the least mse for retrieval with noise.

(c) Noise Effect on Retrievals: Noise interference has clearly increased the optimum number of PCs required for the most accurate predictions. Overall the magnitude of RMSE for both temperature and humidity profile, have increased. Predictions for higher pressure layers have deteriorated with noise, although upper layer retrievals have improved for humidity distribution.

5 CONCLUSION

The present work, which deals with the applications of Principal Components (PC) in solving inverse Radiative Transfer Equation (RTE) using Artificial Neural Network (ANN) on high spectral radiance measurement, was divided into two main parts. (i) TYPE1 dataset, which contains only synthetic profiles, was used for testing the generalization of various network architecture, which concluded the accuracy of linear activation based neural network in the retrieval. (ii) TYPE2 data was used for testing the generalization of ANN to realistic profiles. Comparative study of PCA based linear regression with PCA based ANN were performed and the results from each part were recorded, analyzed and compared with the norms set by World Meteorological Organization. Results from PC based ANN were promising and the performance when varying the number of PCs were in accordance to predictions. The estimates of atmospheric temperature and humidity profiles were less that 1K mean RMSE and in 20% - 30% mean RMSE for humidity retrieval.

PC based regression were also used for comparing its performance with PC based ANN. It was found that the retrieval of the two matched perfectly for smaller number of PCs, where as it deviated drastically as number of PCs were increased. For many cases, it has been shown that simple ANN model performs better in comparison to linear regression. In contrast, PC based regression retrieval was better than ANN retrieval in this case, even when tested with realist dataset TYPE2. The ANN predictions for higher altitudes, typically 300mbar–100mbar, were more deviated than regression prediction. The retrieval with instrument noise in the radiances were also performed which gave erroneous retrieval for lower layers.

Evaluation of PC based ANN for the retrieval of atmospheric temperature / humidity profile form the crux of the present work. Results of the analysis are encouraging for studying the application of advance PC methods on ANN model. Hence, PC based ANN promises to be an effective tool in compressing high dimensional radiance measurement and retrieving temperature and humidity profile with precision.

REFERENCE

[1] **Aires, F.[1] , W. B. Rossow, N. A. Scott, and A. Chédin**, 2002: Remote sensing from the infrared atmospheric sounding interferometer instrument: 1. Compression, denoising, first-guess retrieval inversion algorithms, *Journal of Geophysical Research*, **107(D22)**, 4619.

[2] **Aires, F.[2] , W. B. Rossow, N. A. Scott, and A. Chédin**, 2002: Remote sensing from the infrared atmospheric sounding interferometer instrument: 2. Simultaneous retrieval of temperature, water vapor, and ozone atmospheric profiles, *Journal of Geophysical Research*, **107(D22)**, 4620.

[3] **Blackwell, W. , J.** , 2005: A neural-network technique for the retrieval of atmospheric temperature and moisture profiles from high spectral resolution sounding data. *IEEE Transaction on Geoscience and Remote Sensing*, **43, No. 11**.

[4] **De Souza-Machado, S., L. Strow, and S. Hannon**, 1997: kCompressed atmospheric radiative transfer algorithm (kCARTA). *Satellite Remote Sensing Wavenumber (cm^{-1})of Clouds and the Atmosphere, Proceedings of the European Symposium on Aerospace Remote Sensing 3220, Europto Series, Institute of Electrical Engineers, London, Great Britain.*

[5] **Demuth, H. , M. Beale, M. Hagan**, 2009: Neural Network Toolbox™ 6, User's Guide, MATLAB® , © *COPYRIGHT 1992–2009 by The Math Works™, Inc.*

[6] **Escobar-Munoz, J. , A. Chedin, F. Cheruy, and N. Scott**, 1993: 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, 317(7)**, 911–918.

[7] **Hannon, S. E. , L. L. Strow, and W. W. McMillan,** 1996: Atmospheric infrared fast transmittance models: A comparison of two approaches. *In Proceedings of SPIE Conference 2830, Optical Spectroscopic Techniques and Instrumentation for Atmospheric and Space Research II.*

[8] **Huang, H. L. , and P. Antonelli**, 2001: Application of principal component analysis to high-resolution infrared measurement compression and retrieval, *Journal of Climate and Applied Meteorology*, **40**, 365–388.

[9] **Strow, L. L. , H. E. Motteler, R. G. Benson, S. E. Hannon and S. De Souza-Machado,** 1998: Fast Computation of Monochromatic Infrared Atmospheric Transmittances Using Compressed Look-Up Tables, *Journal of Quantitative Spectroscopy and Radiative Transfer*, **59(3-5)**, 481–493

[10] **Xuebao, W. , L. Jun, Z. Wenjian and W. Fang**, 2005: Atmospheric Profile Retrieval with AIRS Data and Validation at the ARM CART Site, *Advances in Atmospheric Sciences*, **22(5)**, 647-654.

LIST OF TABLES

Table 1 Specification of testing datasets

DATASET	PROFILES
TYPE1	Synthetic profile (3000)
TYPE2	Radiosonse profile (540)

TABLE 2 : Eigen values of first few eigen vectors (TYPE1)

PRINCIPAL INDEX	EIGEN VALUE	CUMULATIVE %
1	17993.3093	56.2658
2	11087.8259	90.9379
3	1185.3443	94.6446
4	699.7194	96.8326
5	315.0443	97.8178
6	188.2016	98.4063
7	130.9496	98.8158
8	117.8997	99.1844
9	65.0963	99.3880
10	54.6386	99.5589

TABLE 3: Performance of feed forward back propagation neural network with different architecture (TYPE1)

Network	Activation Function	MSE Test	RMS Temp Error (K)	RMS Hum Error (g/Kg)	RMS Hum Error (%)
100x192	Linear	0.0000	0.0877	0.0346	8.2735
100x150x192	Linear, Linear	0.0010	9.3657	0.4898	14730.2410
250x192	Linear	0.0000	0.0510	0.0190	46.2904
250x225x192	Linear, Linear	0.0100	30.1343	1.4957	44644.1486
	Log, Linear	0.0008	8.4933	0.4179	13088.9266
	Tan, Linear	0.0008	8.7440	0.4360	12811.7134
500x192	Linear	0.0000	0.1470	0.0173	213.7873
500x350x192	Linear, Linear	0.0185	40.5154	2.0587	64220.7132
	Log, Linear	0.0012	10.2830	0.5237	14938.5408
	Tan, Linear	0.0485	68.1975	3.1920	104890.4190

TABLE 4: Performance of feed forward back propagation neural network with different architecture with noise interference (TYPE1)

Network	Activation Function	MSE Test	RMS Temp Error (K)	RMS Hum Error (g/Kg)	RMS Hum Error (%)
100x192	Linear	0.00001	0.2510	0.0775	8.8555
100x150x192	Log, Linear	0.00420	19.3159	0.9698	31038.2023
250x192	Linear	0.00005	0.3137	0.1446	12.6457
250x225x192	Log, Linear	0.00260	15.2799	0.7581	23927.1812
500x192	Linear	0.00006	0.9163	0.1628	1308.7781
500x350x192	Log, Linear	0.00410	18.8589	0.9814	30391.1171

TABLE 5 : Results of retrieval using various principal components for a feed forward back propagation, linear activation function network without noise component (TYPE2)

Network	MSE Test	RMS Temp Error (K)	RMS Hum Error (g/Kg)	RMS Hum Error (%)
100x192	3.56E-04	1.1315	0.3962	1.10E+03
200x192	3.24E-04	0.8769	0.3795	4.71E+03
250x192	2.87E-04	0.9537	0.3562	7.72E+03
300x192	2.77E-04	0.9734	0.3497	1.19E+04
350x192	2.59E-04	1.1530	0.3366	2.06E+04
400x192	2.54E-04	1.5423	0.3294	3.70E+04
500x192	2.65E-04	2.1764	0.3276	5.20E+04

TABLE 6: Results of humidity retrieval using various principal components for a feed forward back propagation, linear activation function network without noise component (TYPE2)

Network	RMS Hum Error (%) (upto 100hpa)	RMS Hum Error (%) (upto 300hpa)
100x192	27.9568	22.5141
200x192	60.4996	20.8862
250x192	95.6786	21.1120
300x192	141.1524	20.7588
350x192	269.4383	21.6402
400x192	542.2269	25.2082
500x192	705.8045	28.6662

TABLE 7: Results of retrieval using various principal components for a feed forward back propagation, linear activation function network with noise component (TYPE2)

Network	MSE Test	RMS Temp Error (K)	RMS Hum Error (g/Kg)	RMS Hum Error (%)
100x192	0.002	1.6439	0.9513	451.2316
200x192	0.0017	1.5301	0.8752	423.5800
250x192	0.0016	1.5257	0.8546	458.8355
300x192	0.0014	1.5078	0.7785	508.0846
350x192	0.0011	1.5216	0.7074	605.3759
400x192	0.0011	1.5029	0.6956	788.1561
500x192	0.001	1.4905	0.6672	1468.0259

TABLE8: Results of humidity retrieval using various principal components for a feed forward back propagation, linear activation function network without noise component (TYPE2)

Network	**RMS Hum Error (%) (upto 100hpa)**	**RMS Hum Error (%) (upto 300hpa)**
100x192	45.2990	44.1362
200x192	40.3559	35.0699
250x192	40.4870	34.3497
300x192	39.1497	33.2881
350x192	38.0276	32.3651
400x192	37.9381	32.2723
500x192	39.8648	31.4074

LIST OF FIGURES

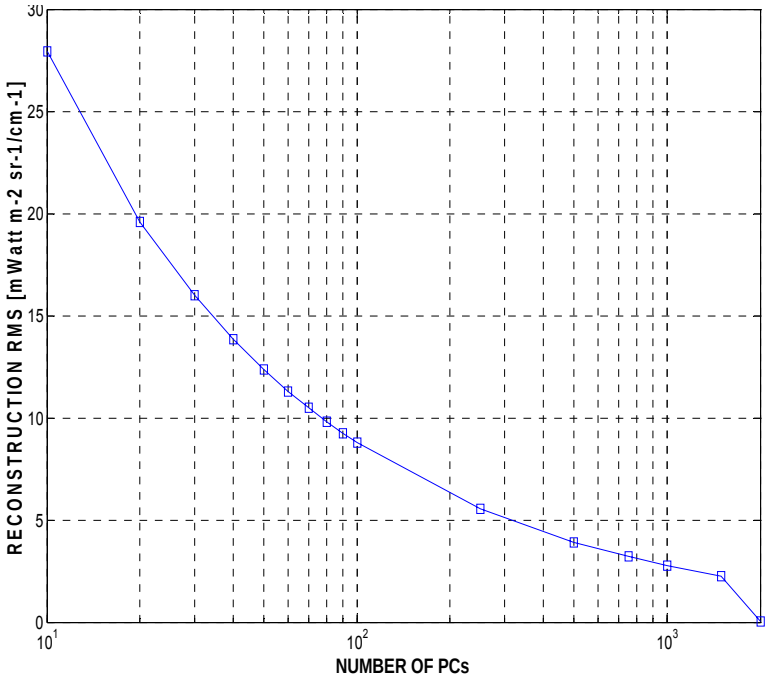
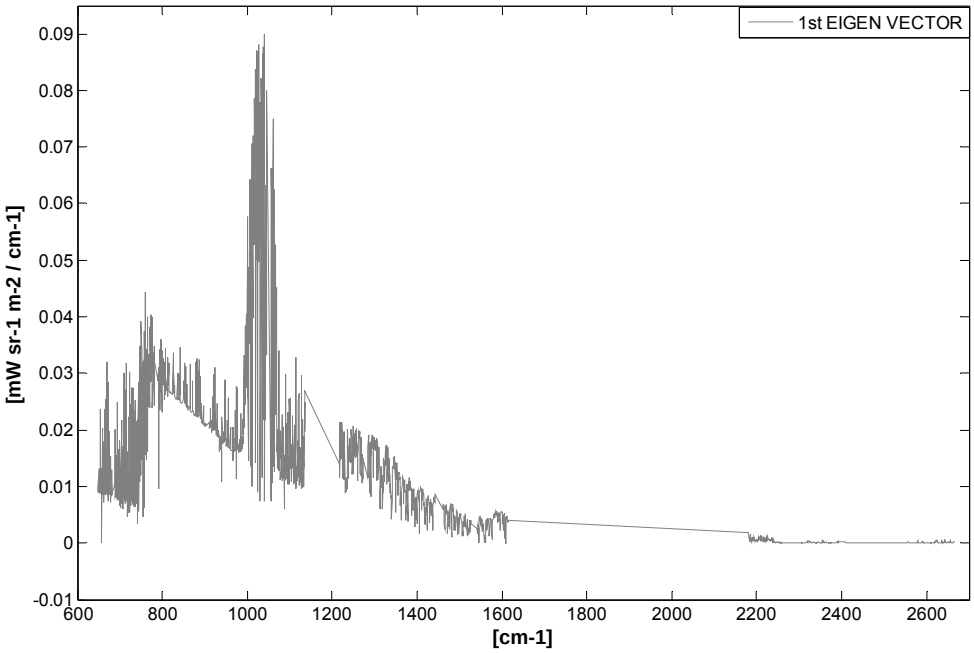
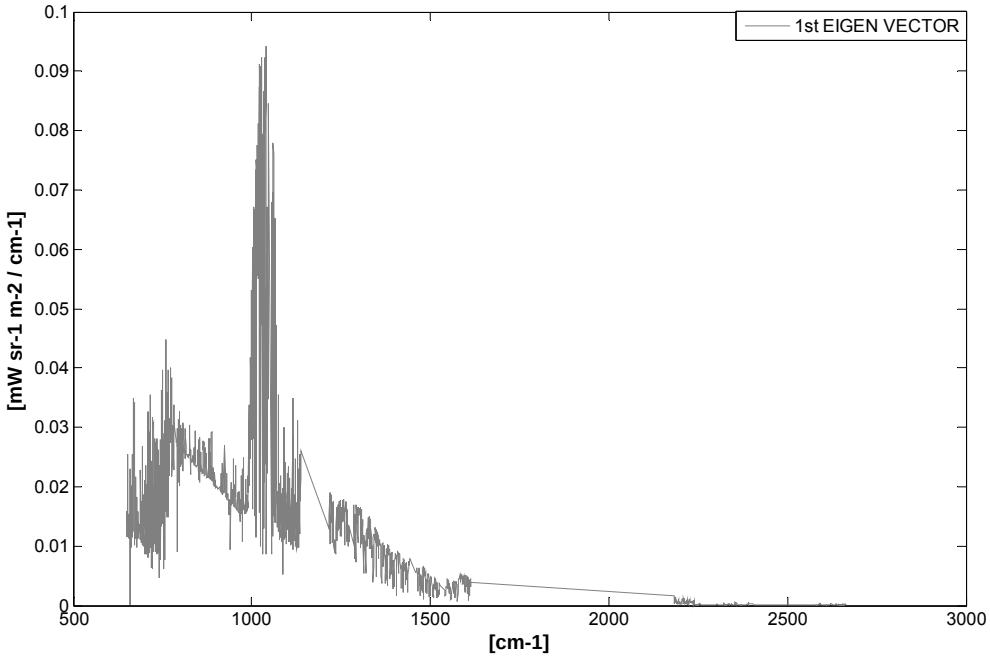


FIGURE 1: Reconstruction error using principal components (TYPE1)

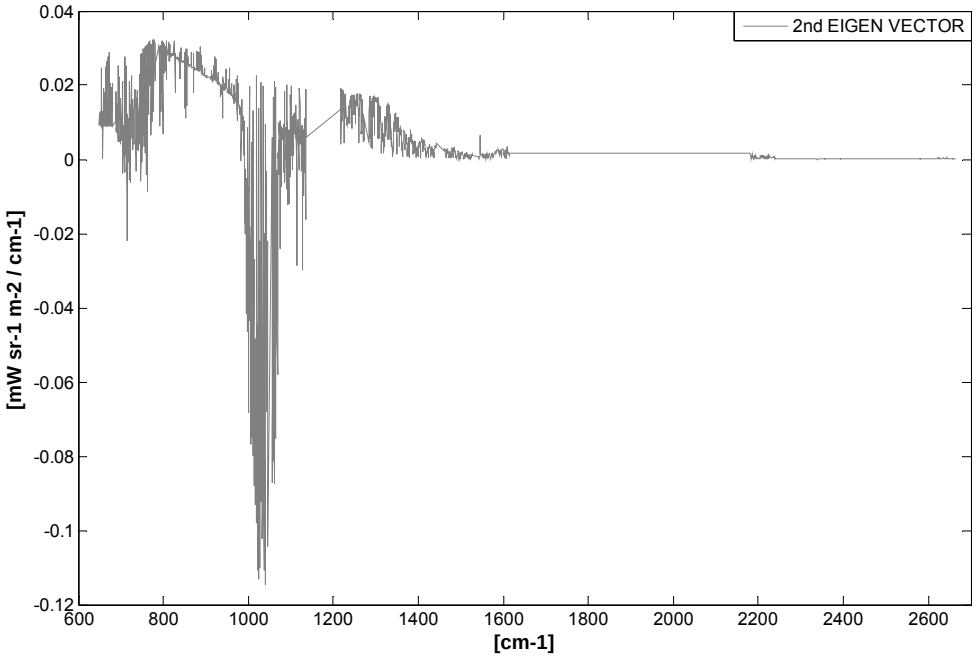


(a)

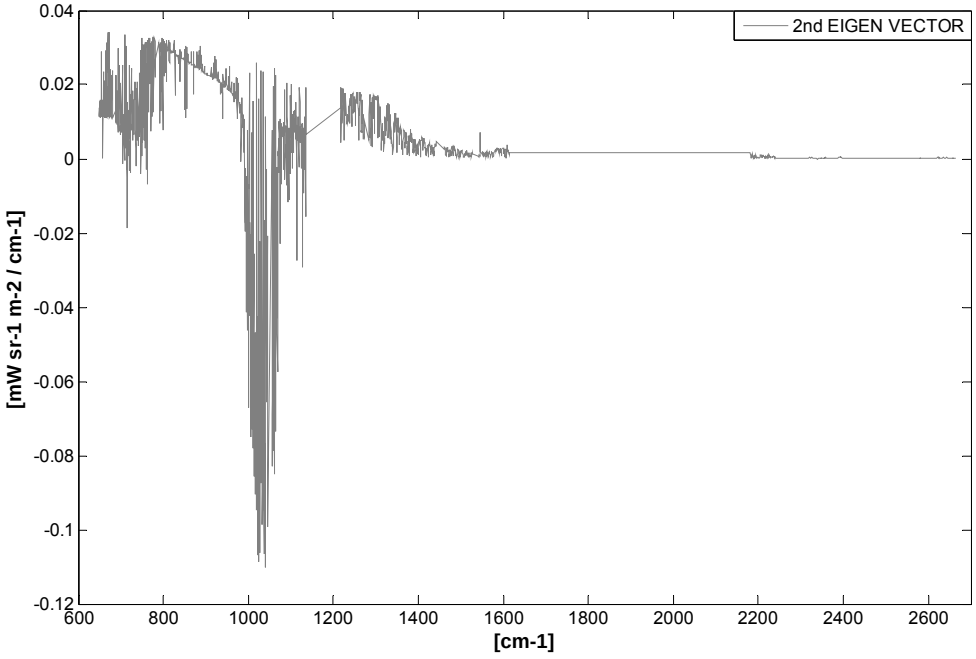


(b)

FIGURE 2 : First eigen vector derived from principal component analysis for TYPE1 and TYPE2

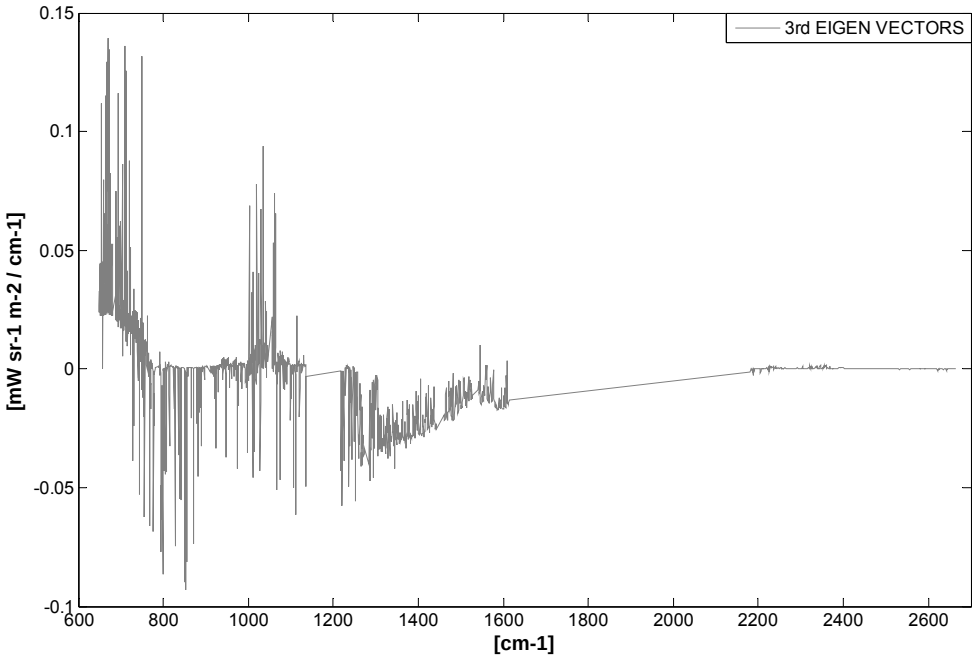


(a)

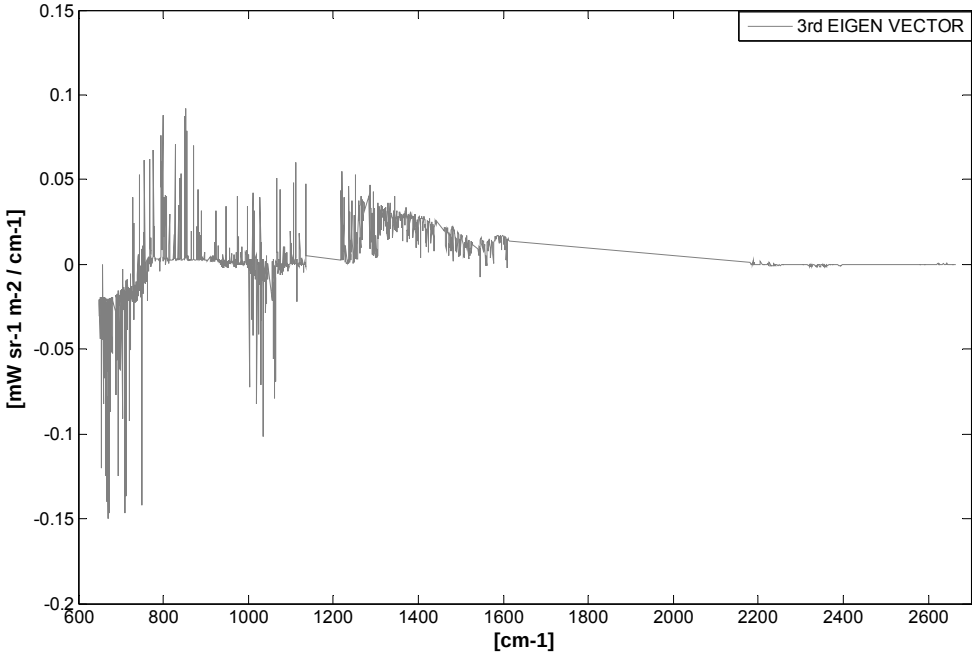


(b)

FIGURE 3 : Second eigen vector derived from principal component analysis for TYPE1 and TYPE2



(a)



(b)

FIGURE 4 : Third eigen vector derived from principal component analysis for TYPE1 and TYPE2

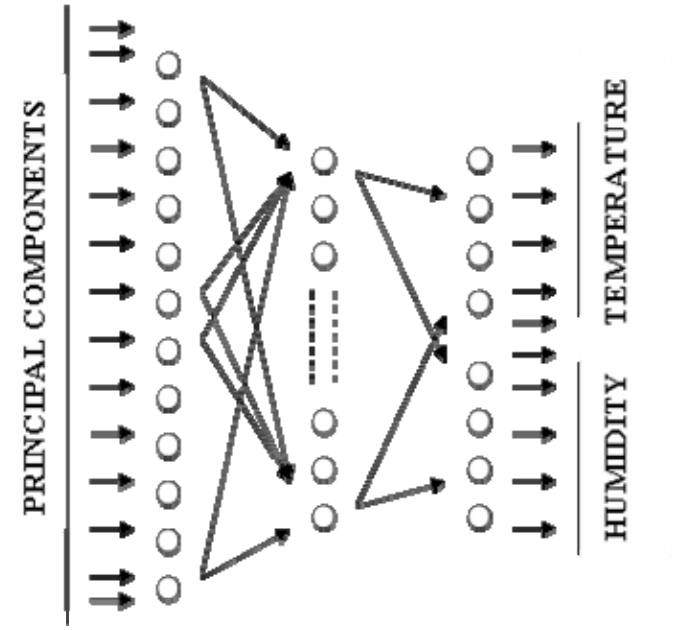


FIGURE 5 : An artificial neural network architecture used for retrieval

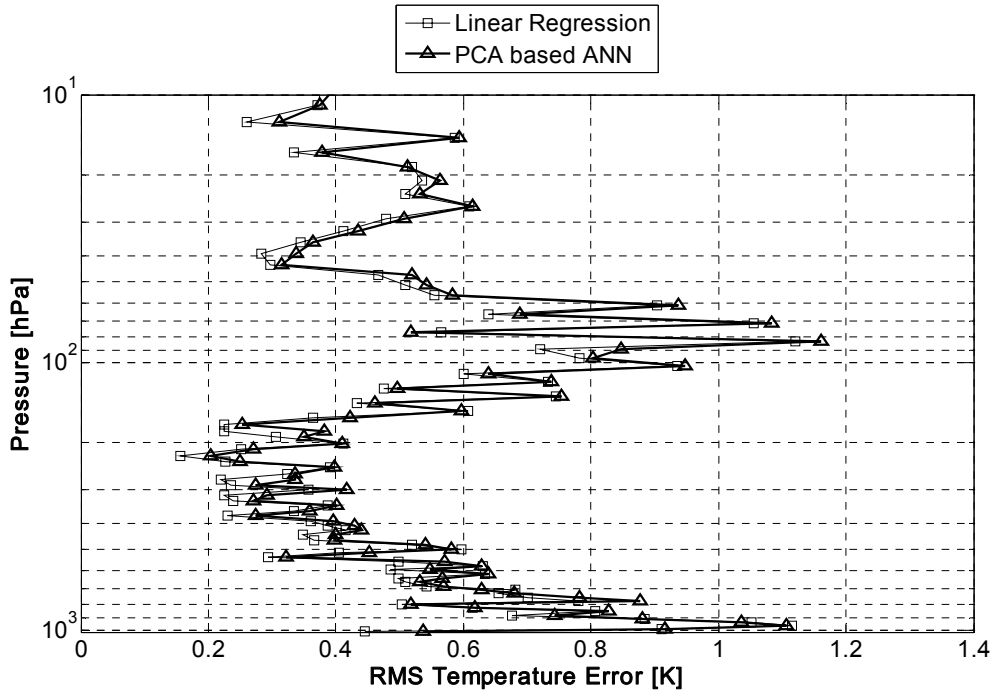


FIGURE 6 : Temperature retrieval performance using 200 PCs without noise interference (TYPE2) having least mean temperature error (K)

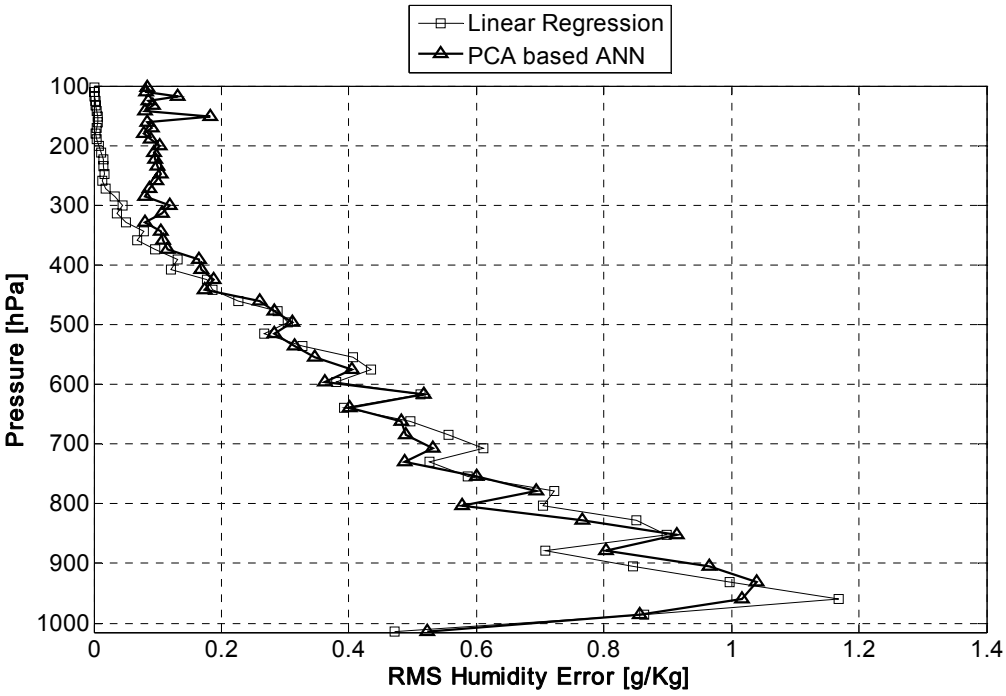


FIGURE 7 : Humidity retrieval performance using 500 PCs without noise interference (TYPE2) having least mean humidity error (g/Kg)

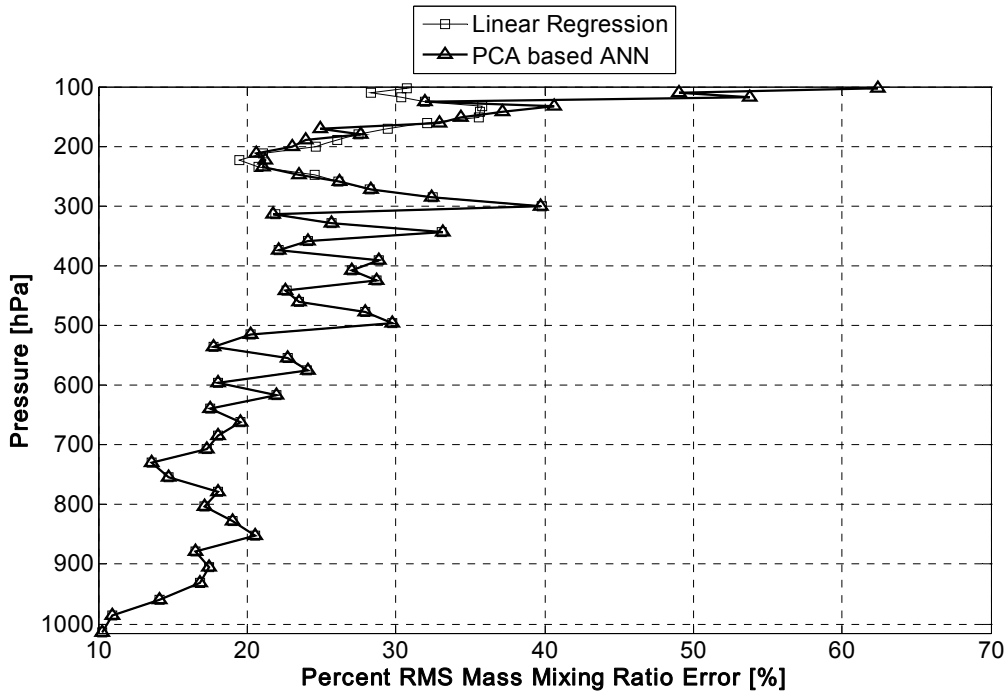


FIGURE 8 : Humidity retrieval performance using 100 PCs without noise interference (TYPE2) having least mean humidity error (%) upto 100 hPa

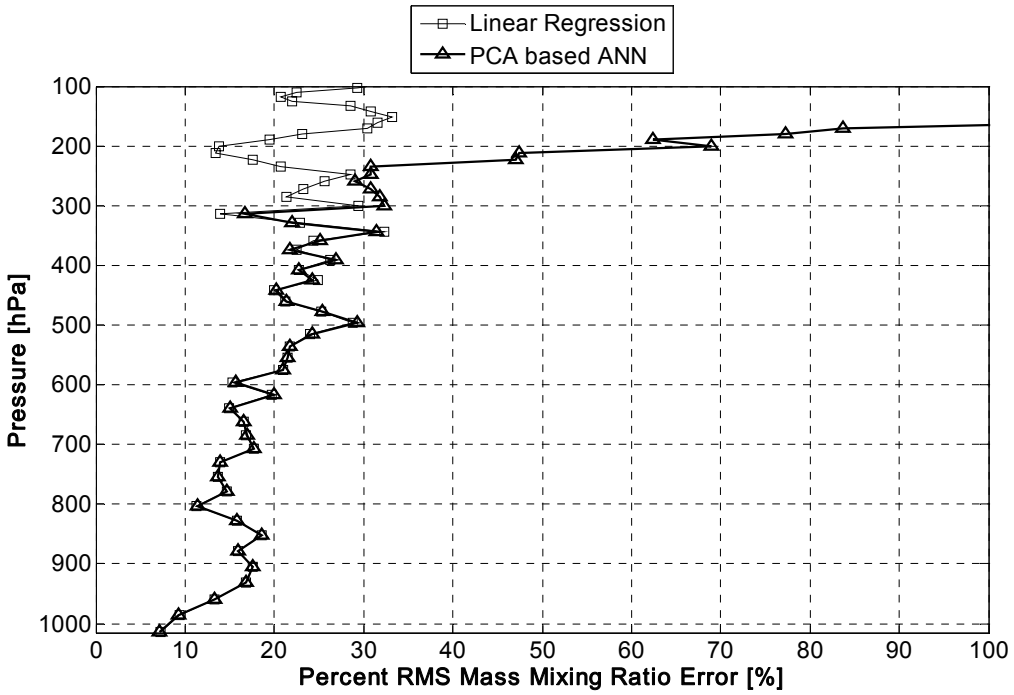


FIGURE 9 : Humidity retrieval performance using 300 PCs without noise interference (TYPE2) having least mean humidity error (%) upto 300 hPa

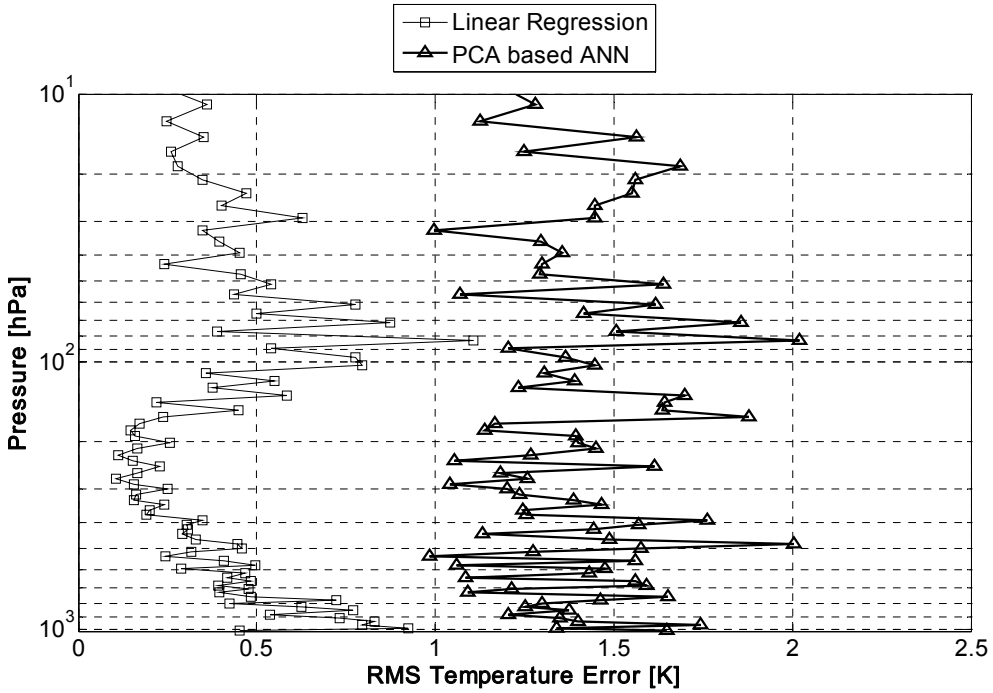


FIGURE 10 : Temperature retrieval performance using 400 PCs without noise interference (TYPE2) having least overall mse value

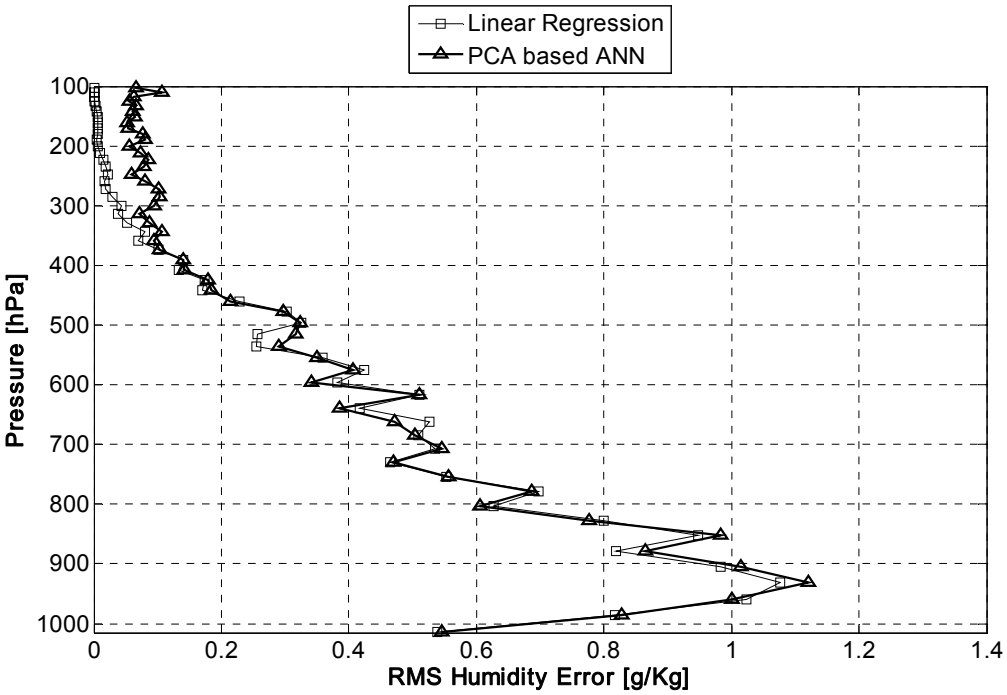


FIGURE 11 : Humidity retrieval performance using 400 PCs without noise interference (TYPE2) having least overall mse value

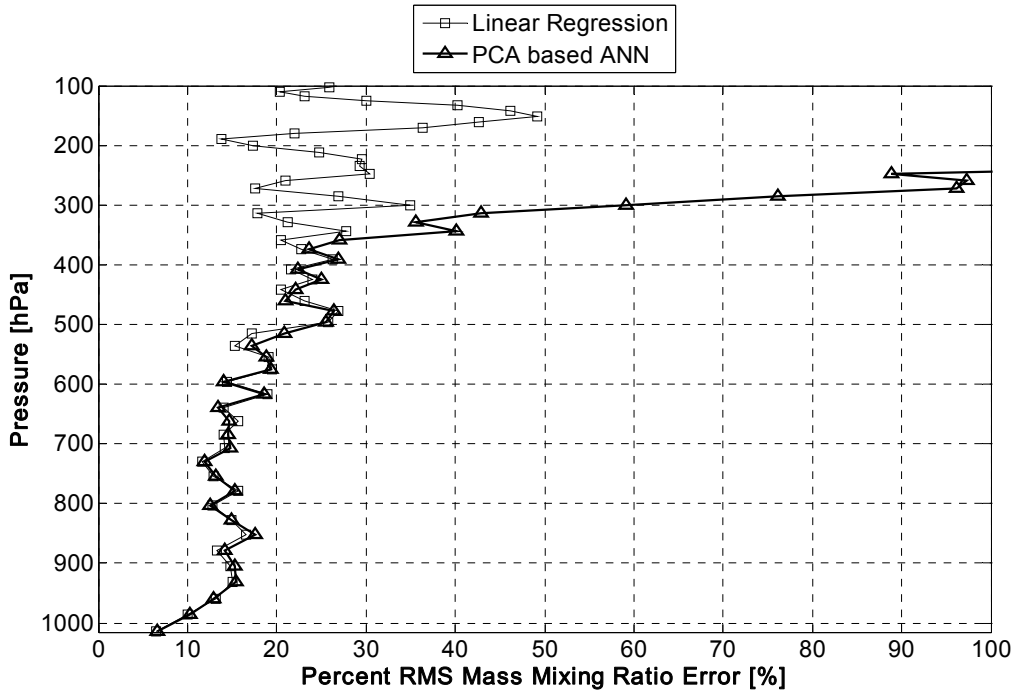


FIGURE 12 : Humidity retrieval performance using 400 PCs without noise interference (TYPE2) having least overall mse value

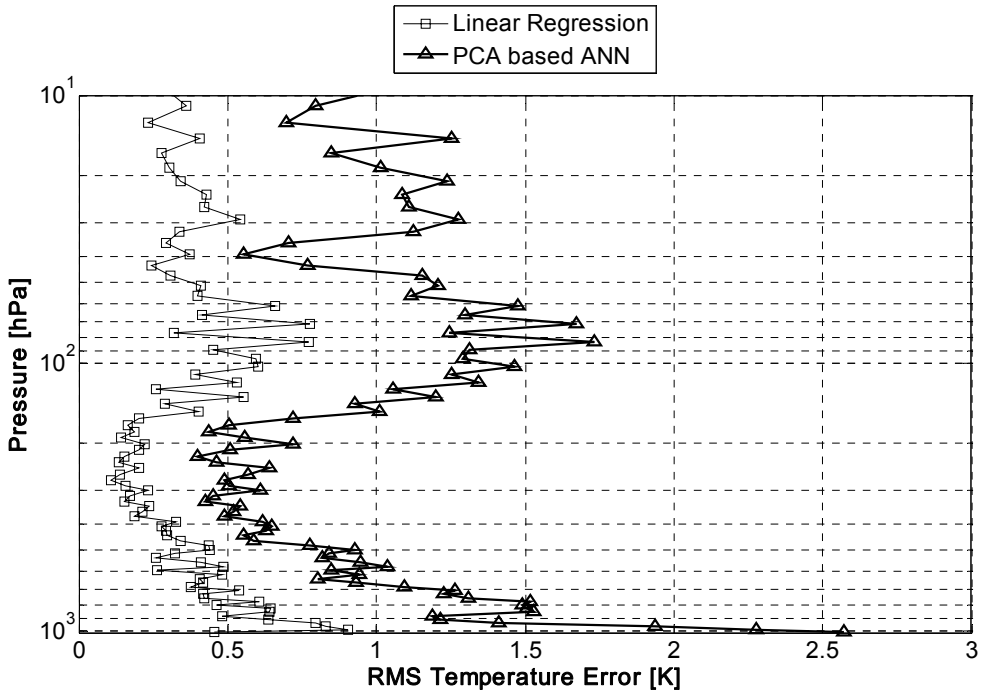


FIGURE 13 : Temperature retrieval performance using 500 PCs with noise interference (TYPE2) having least mean temperature error (K)

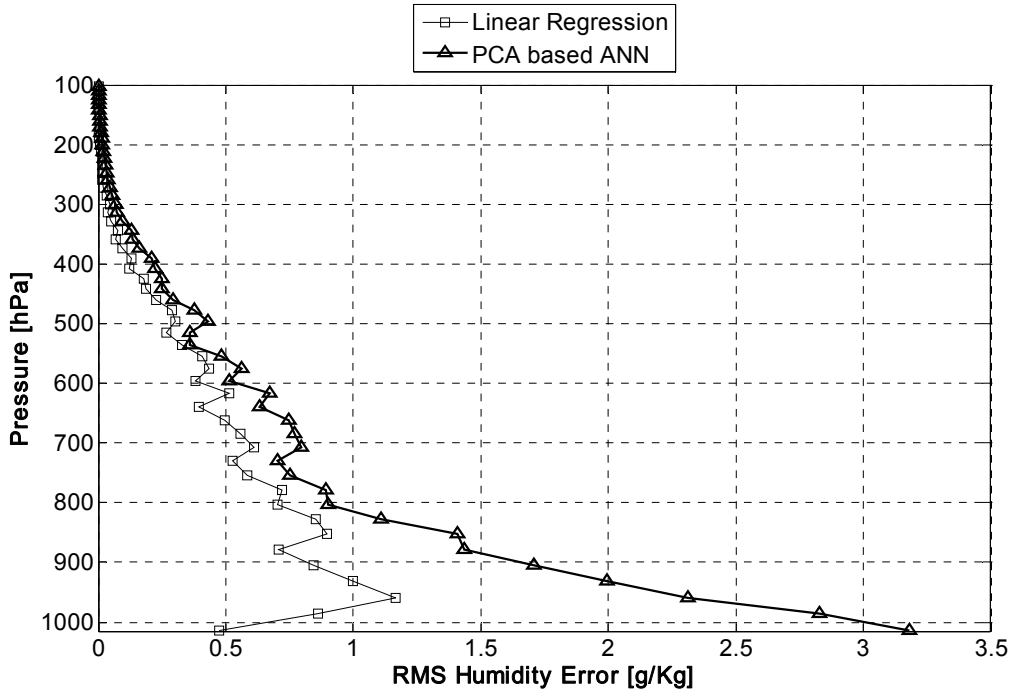


FIGURE 14 : Humidity retrieval performance using 500 PCs with noise interference (TYPE2) having least mean humidity error (g/kg)

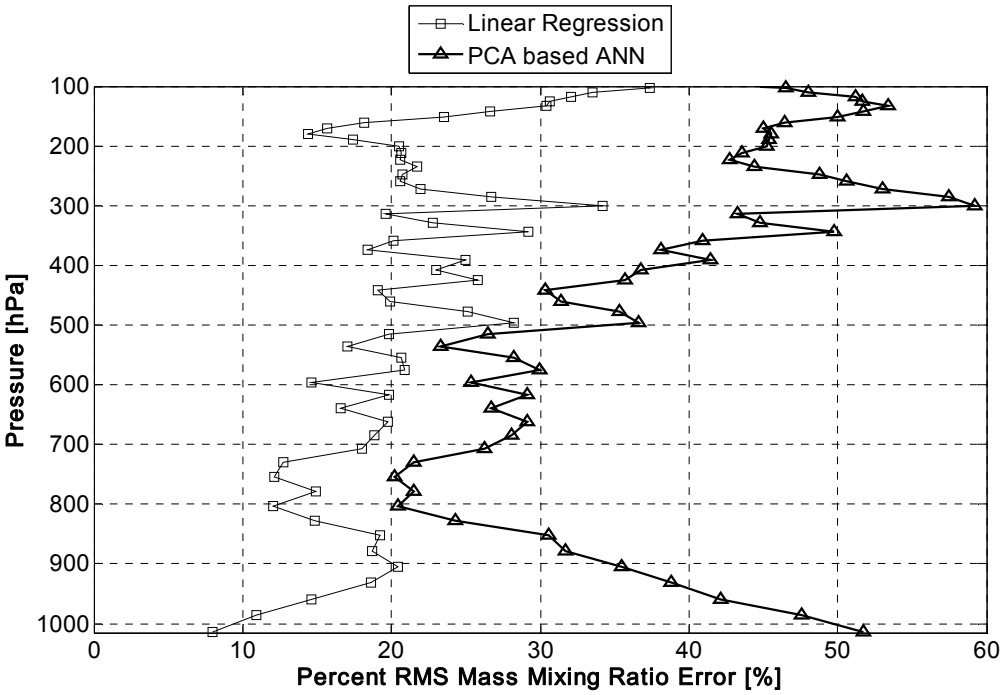


FIGURE 15 : Humidity retrieval performance using 200 PCs with noise interference (TYPE2) having least mean humidity error (%)

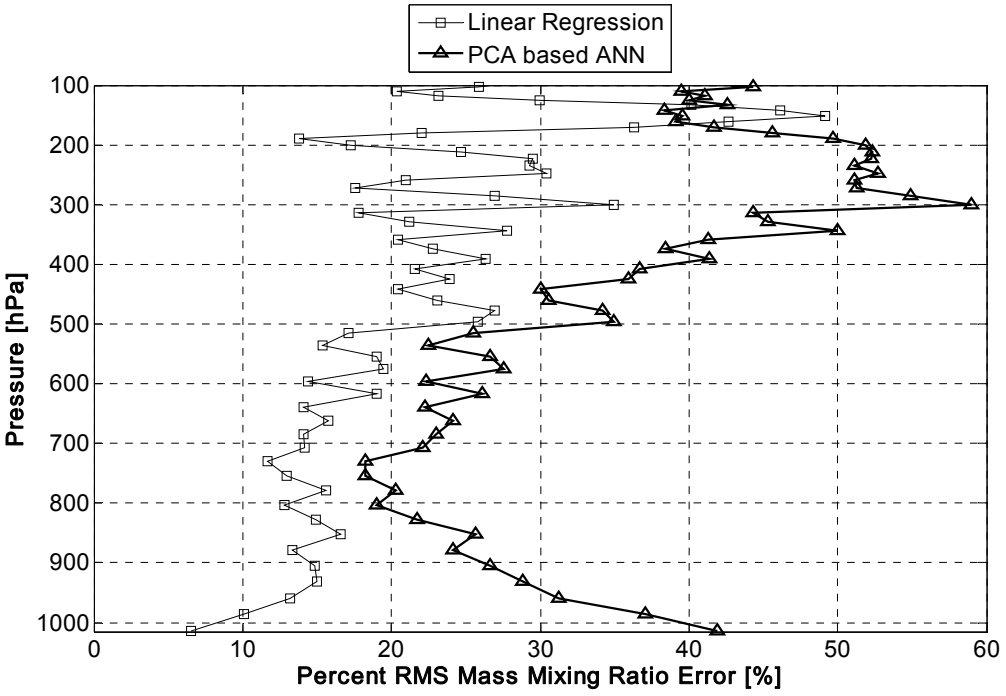


FIGURE 16 : Humidity retrieval performance using 400 PCs with noise interference (TYPE2) having least mean humidity error (%) upto 100hPa

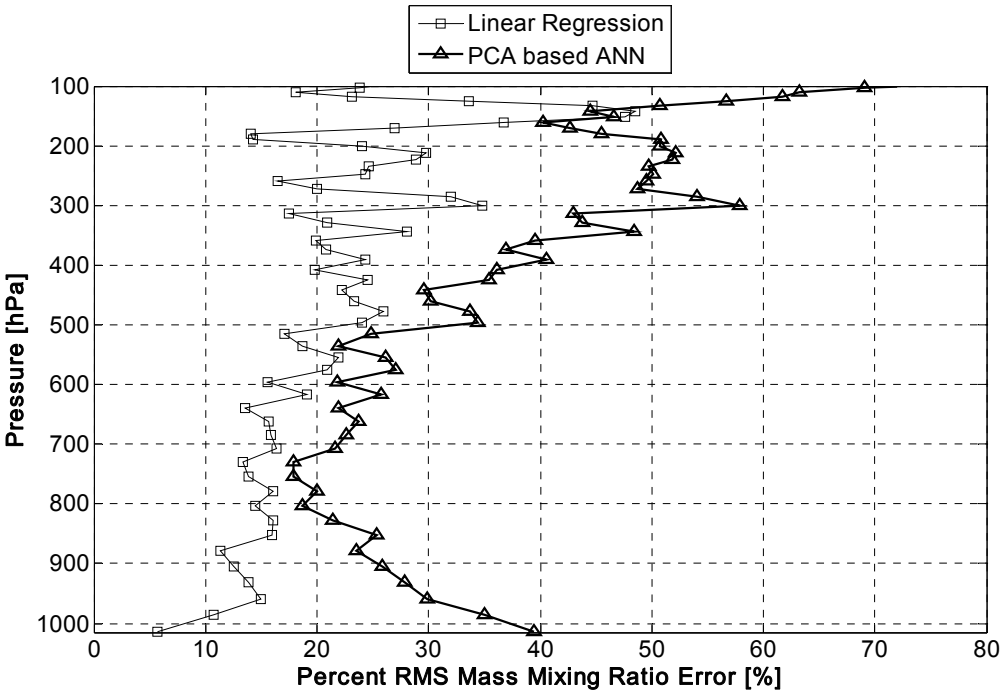


FIGURE 17 : Humidity retrieval performance using 500 PCs with noise interference (TYPE2) having least mean humidity error (%) upto 300hPa

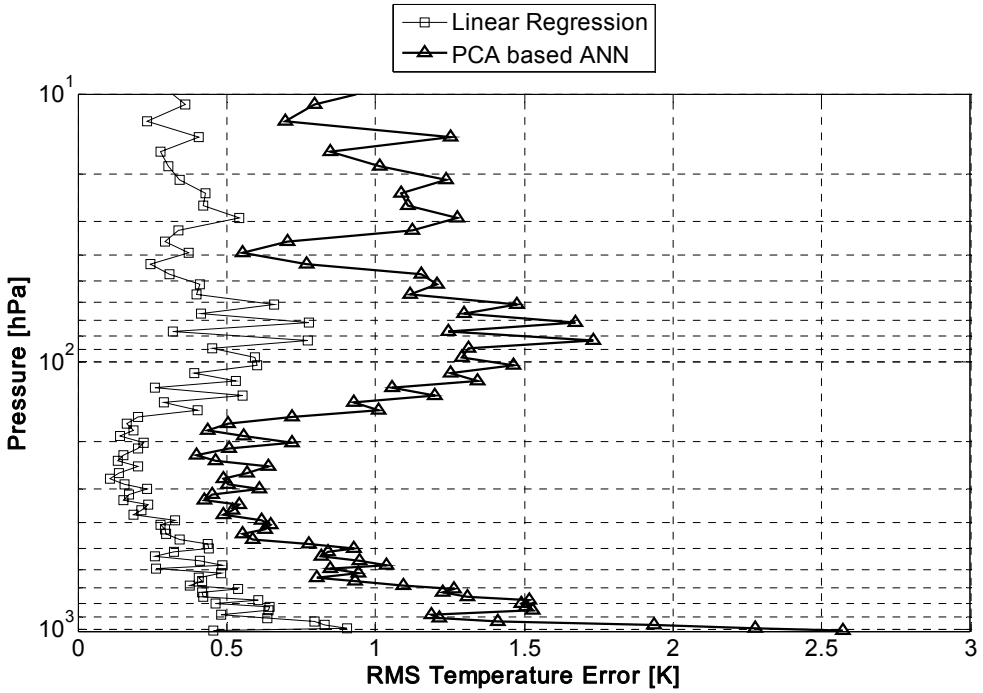


FIGURE 18 : Temperature retrieval performance using 500 PCs with noise interference (TYPE2) having least overall mse value

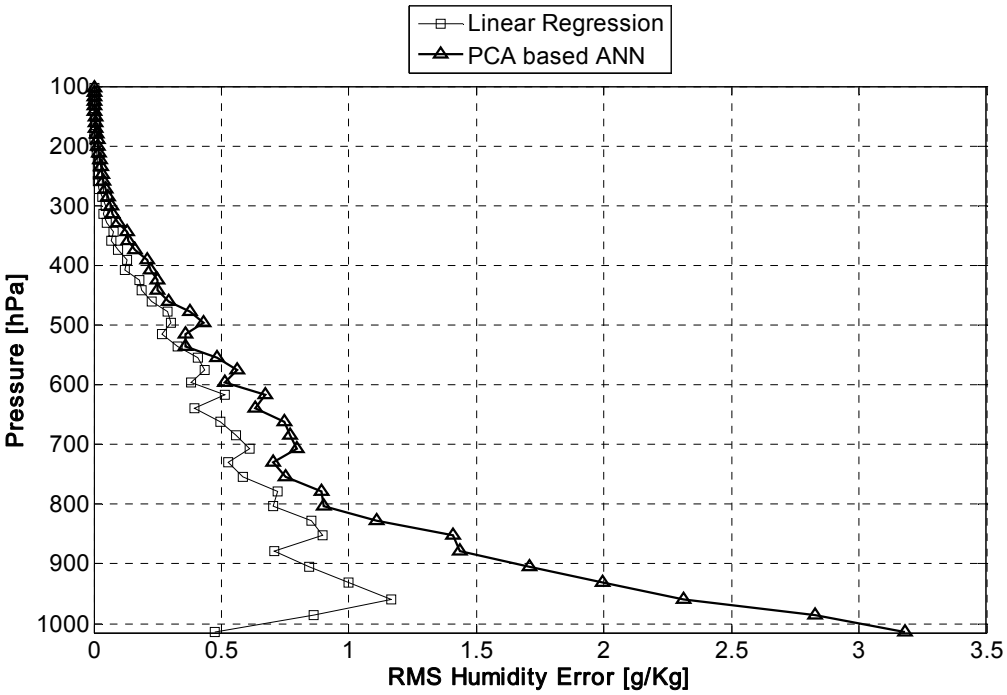


FIGURE 19 : Humidity retrieval performance using 500 PCs with noise interference (TYPE2) having least overall mse value

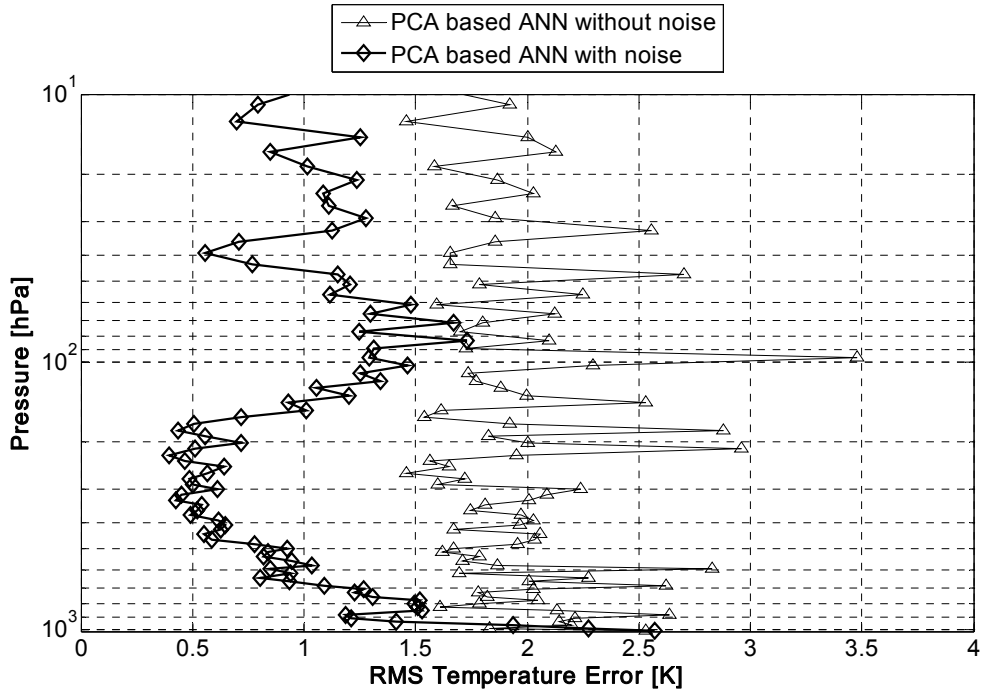


FIGURE 20: Comparison of temperature retrieval performance using 500PCs both for with and without noise interference PCA based ANN

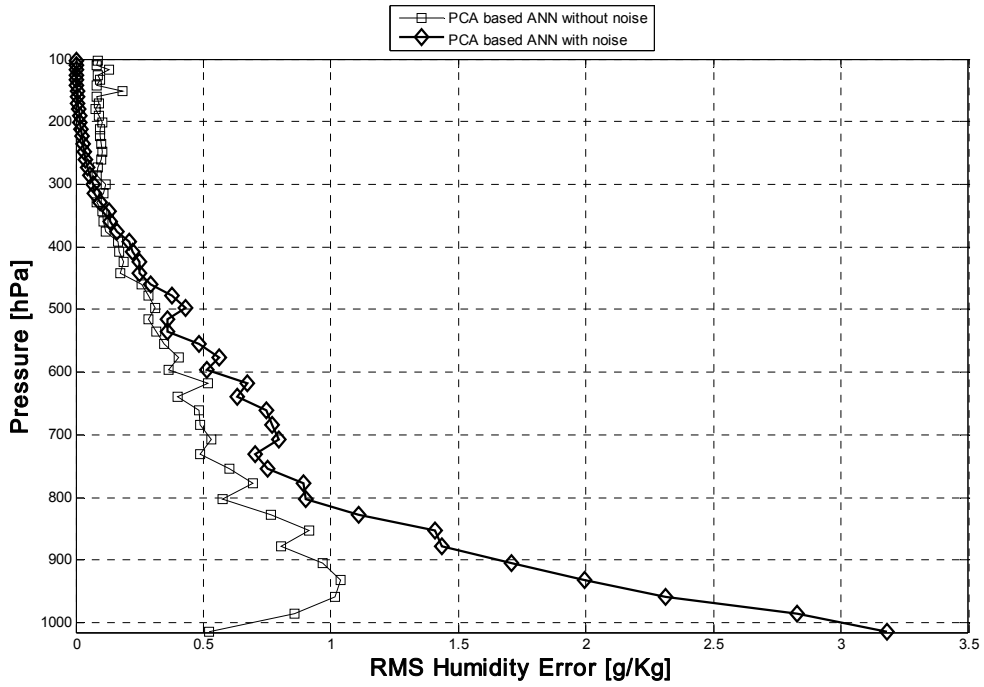


FIGURE 21: Comparison of humidity retrieval performance using 500PCs both for with and without noise interference PCA based ANN

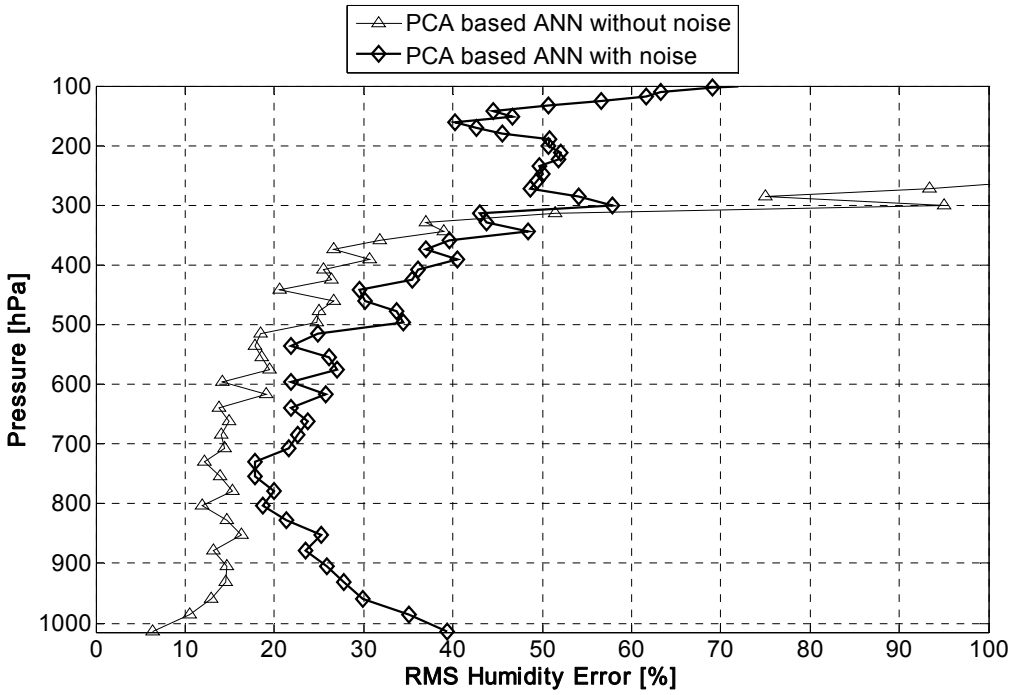


FIGURE 22: Comparison of humidity retrieval performance using 500PCs both for with and without noise interference PCA based ANN