RETRIEVAL OF ATMOSPHERIC TEMPERATURE AND MOISTURE PROFILE USING PRINCIPAL COMPONENT ANALYSIS

A presentation from Abhishek Verma, AE04B026

Under the guidance of
Dr. C. Balaji, Associate Professor, Dept of Mechanical Engg.
Dr. Amit Kumar, Associate Professor, Dept. of Aerospace Engg.

ORGANIZATION

- Introduction
- Literature Survey
- KCARTA as forward model
- Statistical Tool for Retrieval
  - Linear Regression
  - Artificial Neural Network
  - Principal Component Analysis
- Methodology
- Result
- Conclusion

INTRODUCTION

- "Remote Sensing" → Implicit measurements.
- Atmospheric temperature and trace gases profile measured indirectly (*retrieval process*).
- Necessary for weather / climate prediction.
- Uses upwelling electromagnetic radiations emerging from top of the earth's atmosphere.
- Fast and accurate retrieval is the need of the hour.

INTRODUCTION

Physical Method

Retrieval Process

Statistical Method

- Physical Method: Accurate, Slow, “local inversion”, high computation budget.
- Statistical Model: Dependent Accuracy, Fast, “global inversion”, relatively low computation budget (*Blackwell*(2005))

INTRODUCTION

Atmospheric Infrared Sounder (AIRS)

- Onboard NASA's Aqua satellite.
- Hyper spectral Sounder.
- 2378 channels in infrared range.
- Resolution of $\frac{\nu}{\Delta \nu} \sim 1200$.
- Covers 90% of the globe in 24 hours.

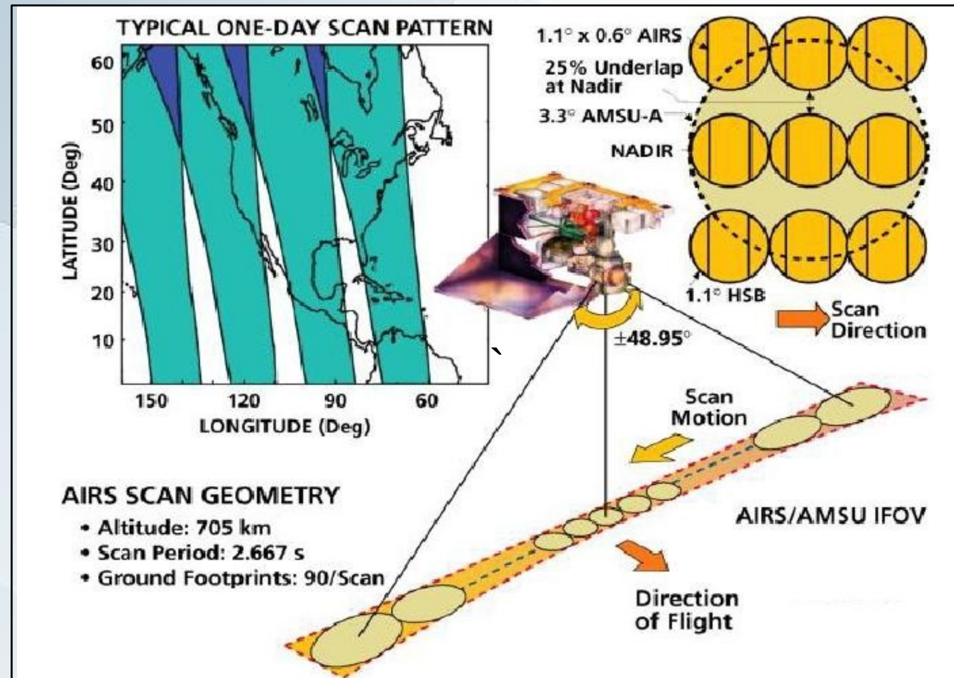


Fig. 1. AIRS viewing geometry and coverage

AIRS viewing geometry and coverage [1]

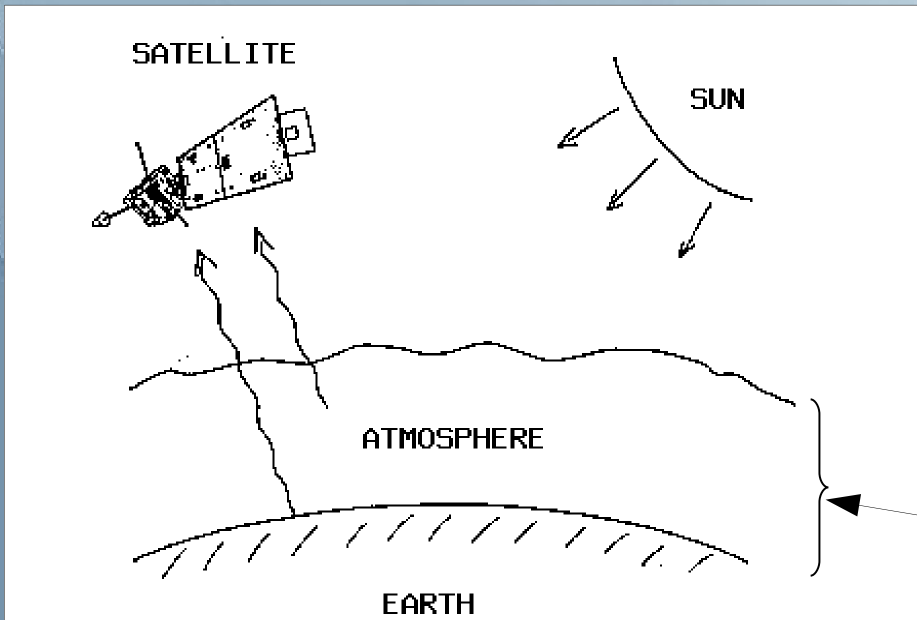
LITERATURE SURVEY

- *Huang and Antonelli* (2000) have investigated the application of PCA with regression in retrieval of temperature / humidity distribution. Besides obtaining an optimum compression ratio, a desirable subset of principal components (PCs) were obtained which gave high retrieval accuracy.
- An advanced version of PCA, known as Projected PC Transform (PPC) was used by *Blackwell* (2005) along with ANN for retrieval purposes.

kCARTA AS FORWARD MODEL

- k Compressed Atmospheric Radiative Transfer Algorithm (kCARTA) is a radiative transfer code for a non-scattering earth's atmosphere.
- Used to output monochromatic gas optical depths, layer-to-space transmittances or radiances.
- Absorption coefficients(κ) used by the code are computed using a database of compressed look up tables.
- Point spacing of the output is 0.0025cm^{-1}, which is the average value calculated over 5 points spaced at 0.0005 cm^{-1}

kCARTA AS FORWARD MODEL

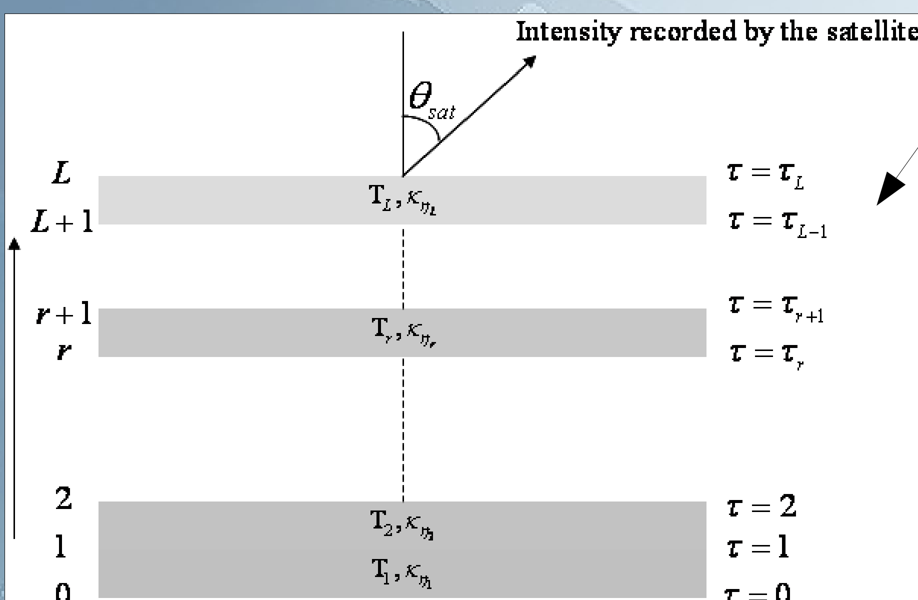


Courtesy: http://cimss.ssec.wisc.edu/goes/comet/2.html

$$R_\lambda^{sfc} = \varepsilon_\lambda^{sfc} B_\lambda(T_{sfc}) \tau_\lambda(sfc \rightarrow top)$$

$$\tau_\lambda = e^{-k_\lambda u}$$

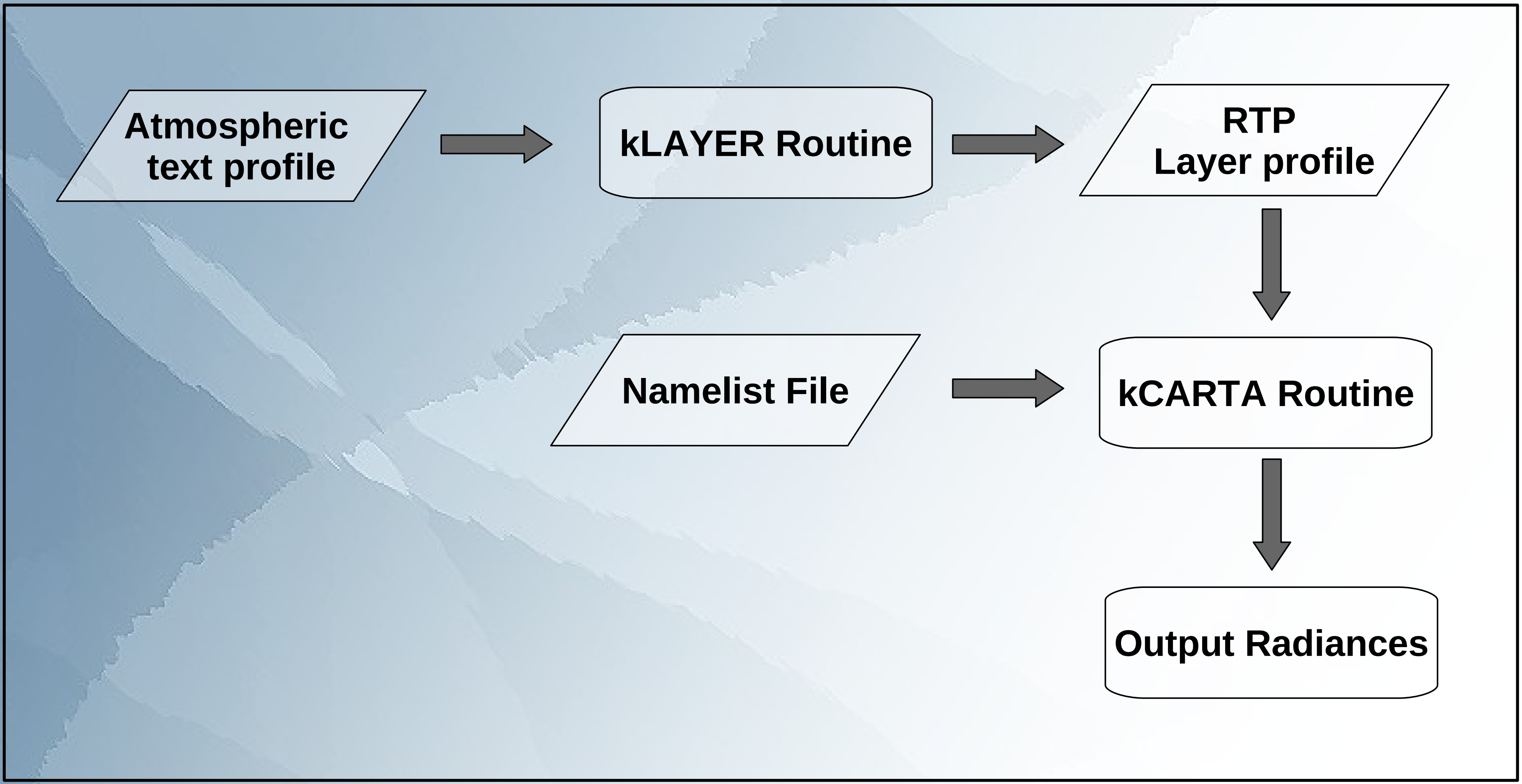
Divided into many layers



$$R_\lambda^{atm} = \sum_{layers} \varepsilon_\lambda^l B_\lambda(T_l) \tau_\lambda(layer \rightarrow top)$$

$$R_\lambda = \varepsilon_\lambda^{sfc} B_\lambda(T_{sfc}) \tau_\lambda(sfc) - \int_0^{P_{sfc}} B_\lambda(T(p)) \frac{d\tau_\lambda(p)}{dp} dp$$

kCARTA AS FORWARD MODEL



Flow chart of data processing steps in kCARTA

LINEAR REGRESSION

- Assumes the output variable as a linear function of predictor variables(input)
- Amongst the most widely used retrieval techniques
- Independent of any radiative transfer model
- Benchmark for other methods like ANN retrieval

LINEAR ADDITIVE REGRESSION

[Y] = the matrix of atmospheric parameters

[A] = the matrix of regression coefficients

[P] = the matrix of inputs (principal components)

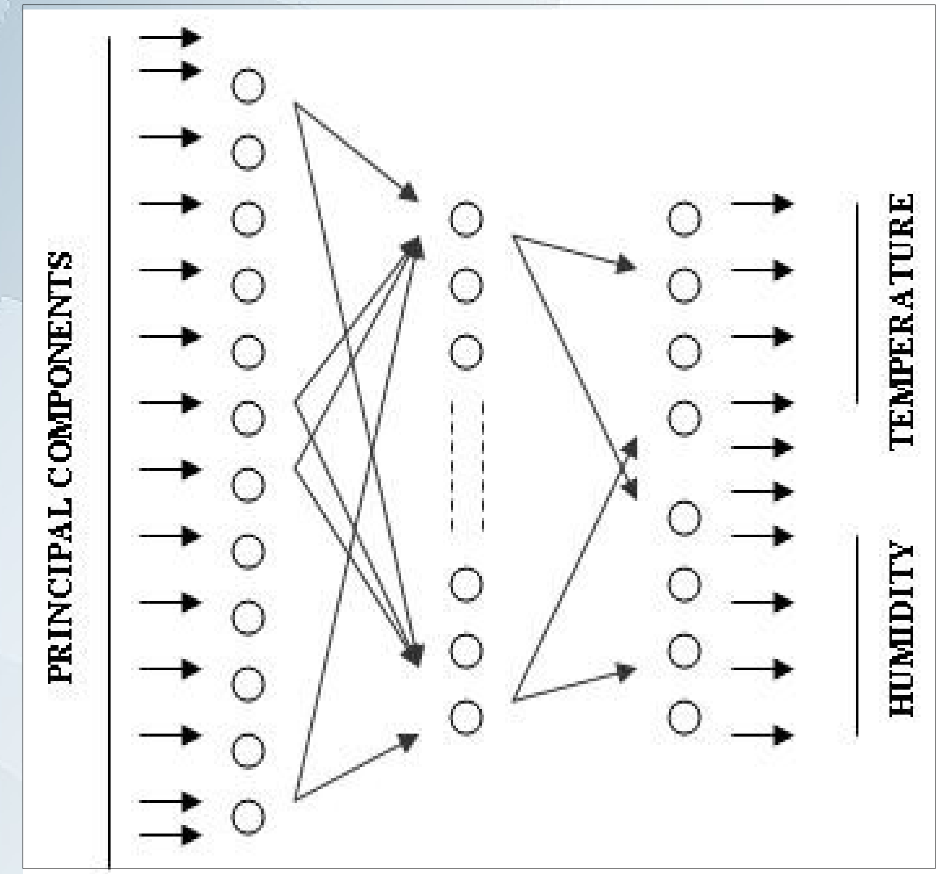
$$[Y] = [A]^T[P]$$

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

ARTIFICIAL NEURAL NETWORK

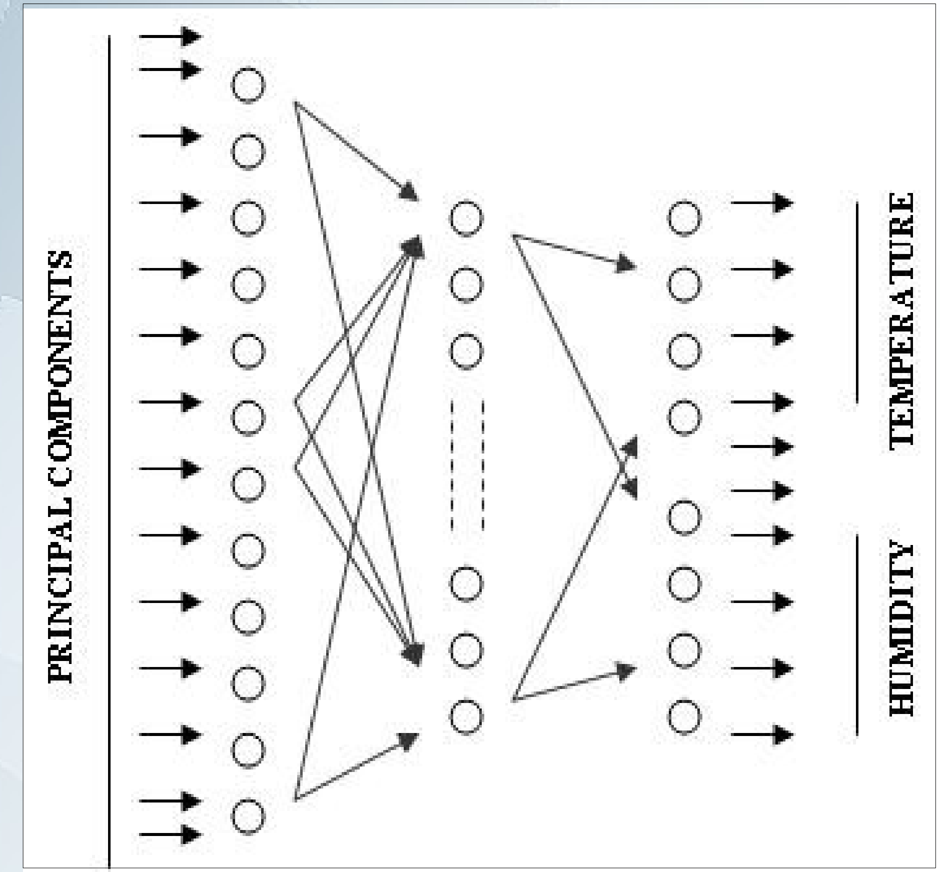
- Architecture linking various processing units(nodes)
- Processed by Activation function(linear, sigmoidal)
- Weighted linkage with bias



Neural network for retrieval process

ARTIFICIAL NEURAL NETWORK

- Used MATLAB® 7.0 neural network toolbox.
- Used TRAINSCG Routine.
- No hidden layer with linear activation function at output nodes.



Neural network for retrieval process

PRINCIPAL COMPONENT ANALYSIS__________

- Transformation of the data into entirely new dimensionality based on variance
- *Criteria*: To choose those features which represents most of the linear variation within the data (observation)
- Effectively used in Image Compression, Dimensionality Reduction and Data Assimilation.

PRINCIPAL COMPONENT ANALYSIS

[R] = Matrix of AIRS observation at all instants

{R} = Observation of AIRS at an instant

$\mathbf{R_m}$ = Radiance measured by mth channel

$\mathbf{<R_m,R_n>}$ = Covariance of mth and nth channel

$$\mathbf{Cov(R) = [<R_m,R_n>]_{\ m = 1:2378\ ,\ n = 1:2378}}$$

gives the covariance matrix for all the AIRS channels

- Solving for the Eigen Vectors of **Cov(R)** yields the required Principal Components

$$\mathbf{Cov(R) - [\lambda] \times I = 0}$$

PRINCIPAL COMPONENT ANALYSIS

[E] = Matrix of Eigen Vectors of **Cov(R)**

[P] = Matrix of Principal Components

[P_s] = Matrix of Selected Eigen Vectors

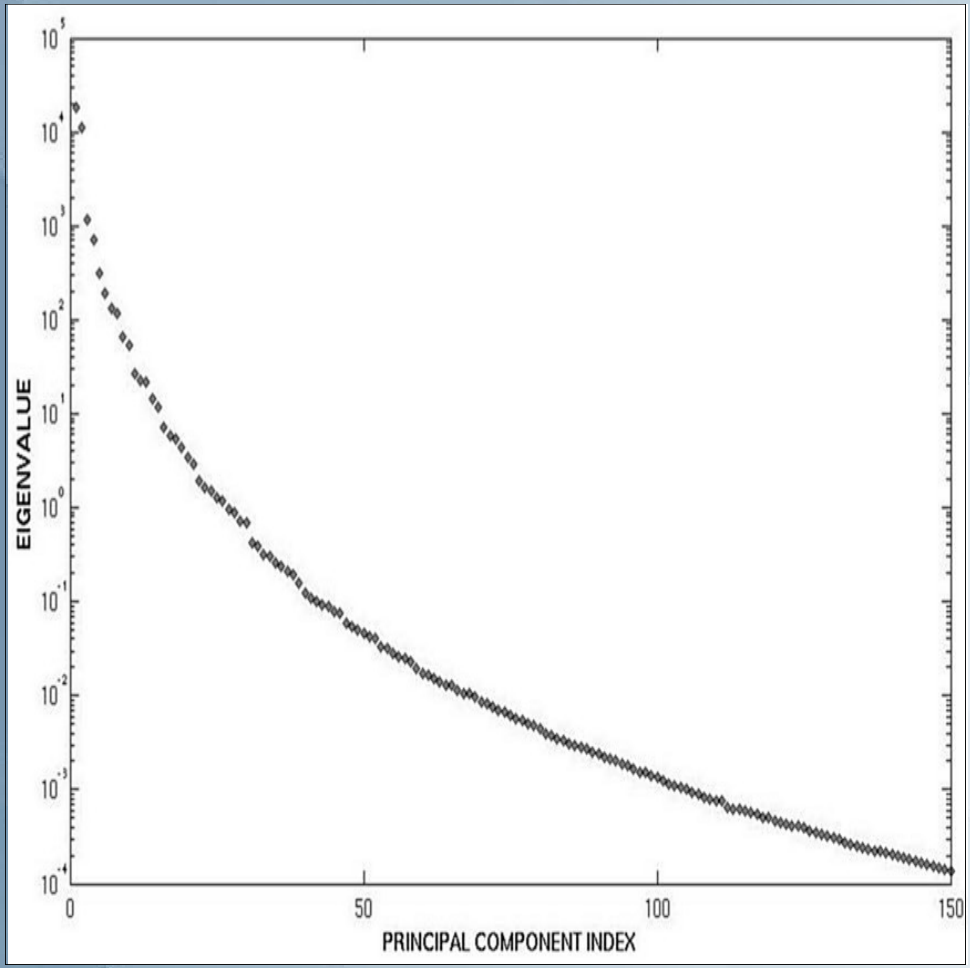
<R_m,R_n> = Covariance of **m**th and **n**th channel

$$[P_s] = [\ \{E\}_i \mid i = 1,2,3\ldots\ldots,k\ ]$$

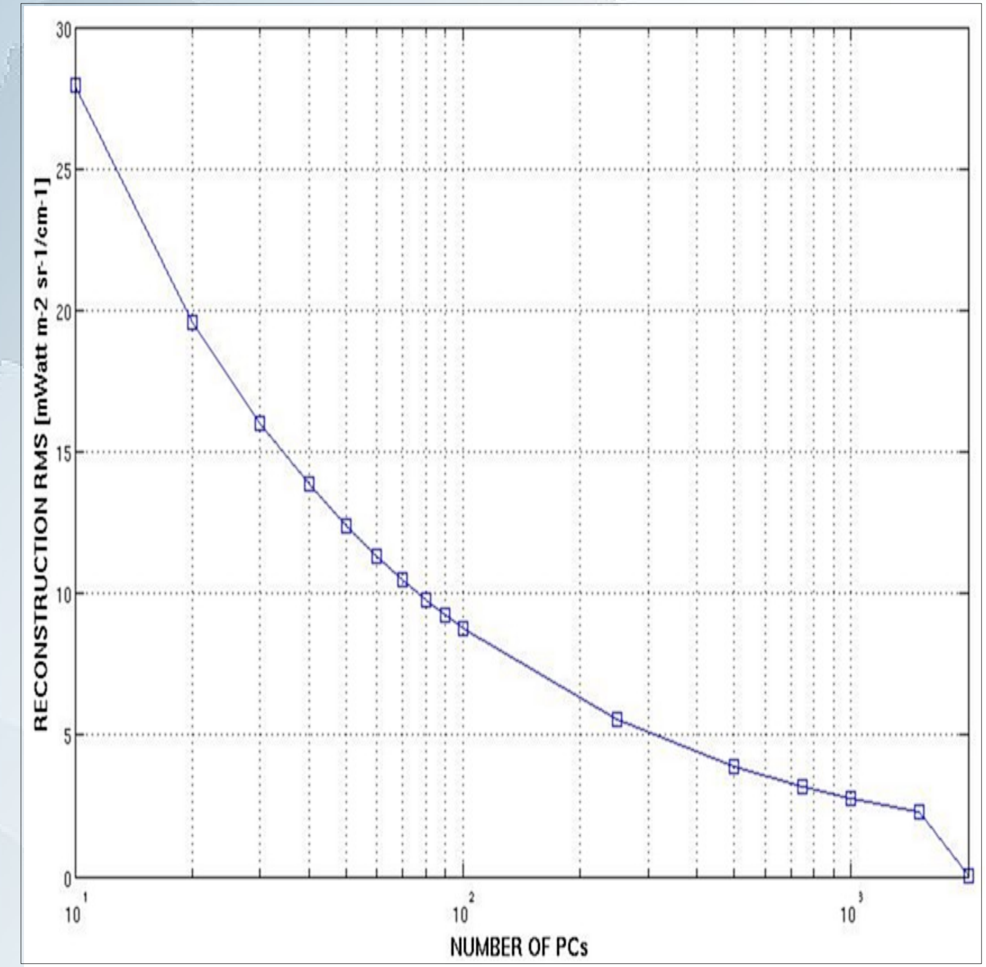
where **{E}**$_i$ are the eigen vectors sorted by their eigen values and k is the number of principal components to be used

$$[P] = [R]^T . [P_s]$$

PRINCIPAL COMPONENT ANALYSIS



Eigen Values of each the Principal Components



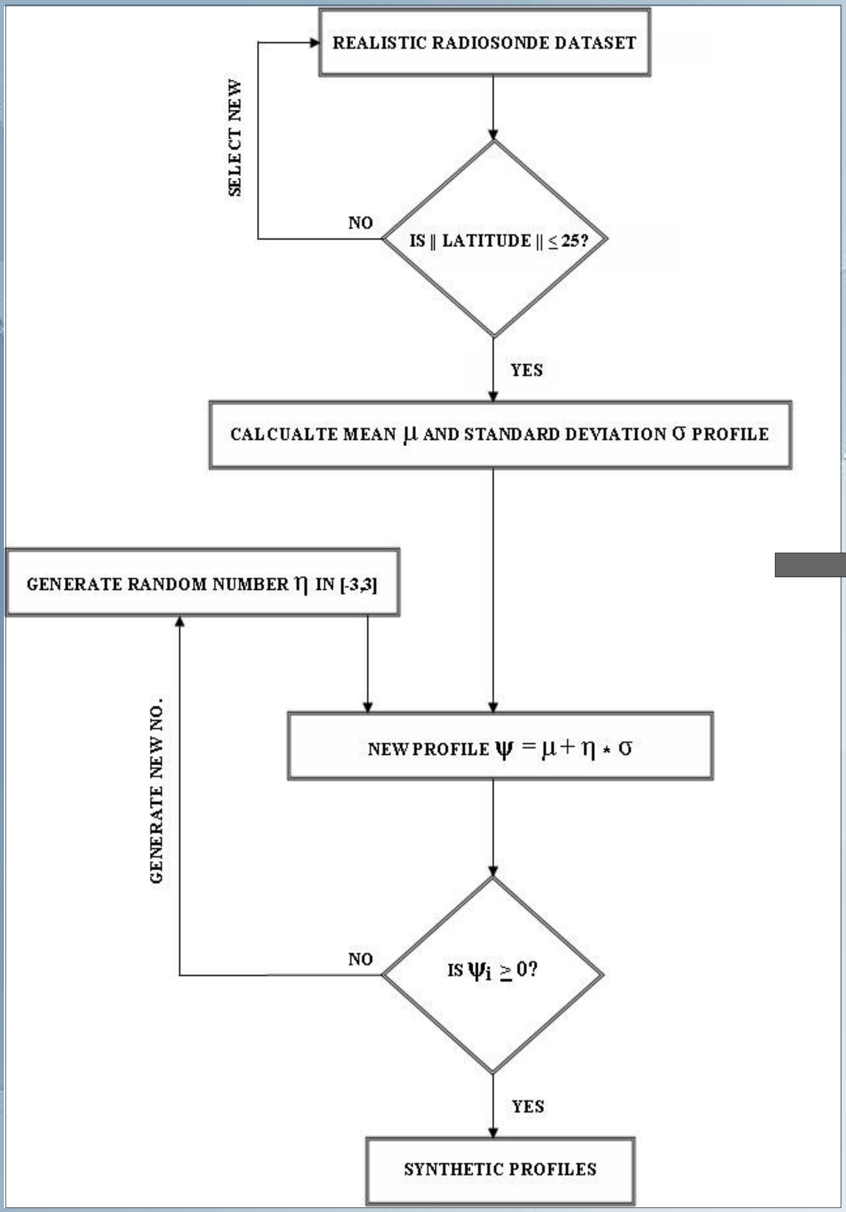
Performance of Reconstruction using Principal Components

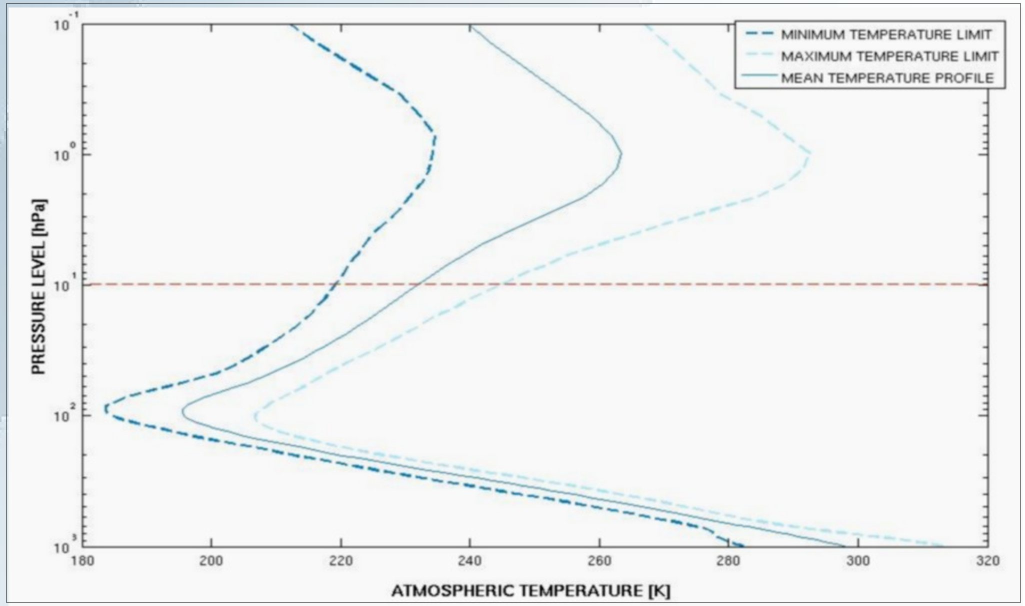
PRINCIPAL COMPONENT ANALYSIS__________

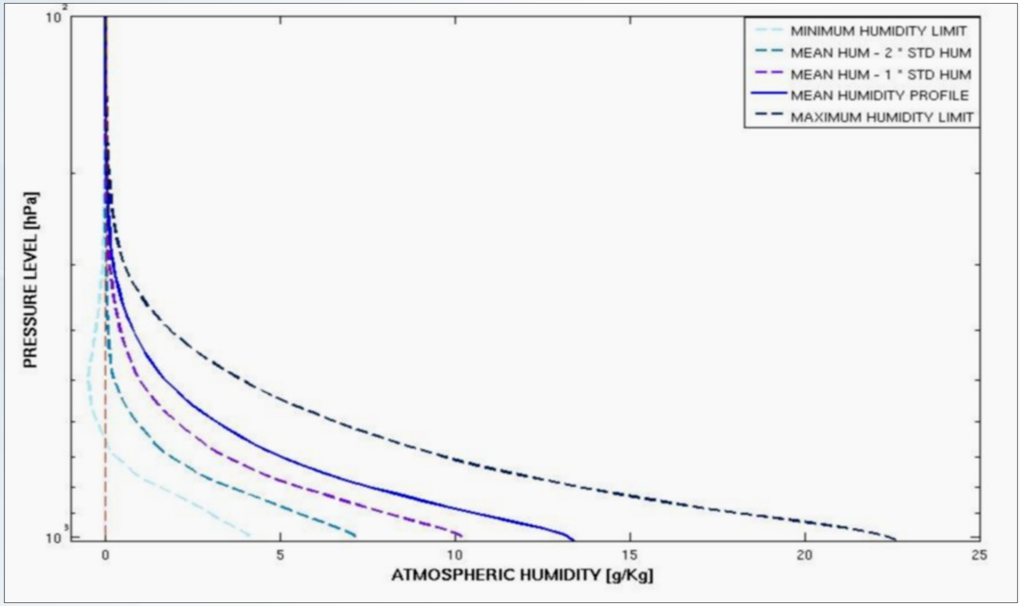
- Applying PCA on the ANN model for input reduction
- Uses the most prominent feature of the dataset in retrieving
- Results in the reduction of hidden layer / hidden nodes
- Requires Linear Activation Function explicitly
- Evidently leads to efficient computation

METHODOLOGY

SYNTHETIC DATABASE GENERATION







METHODOLOGY

Database used for Statistical Model

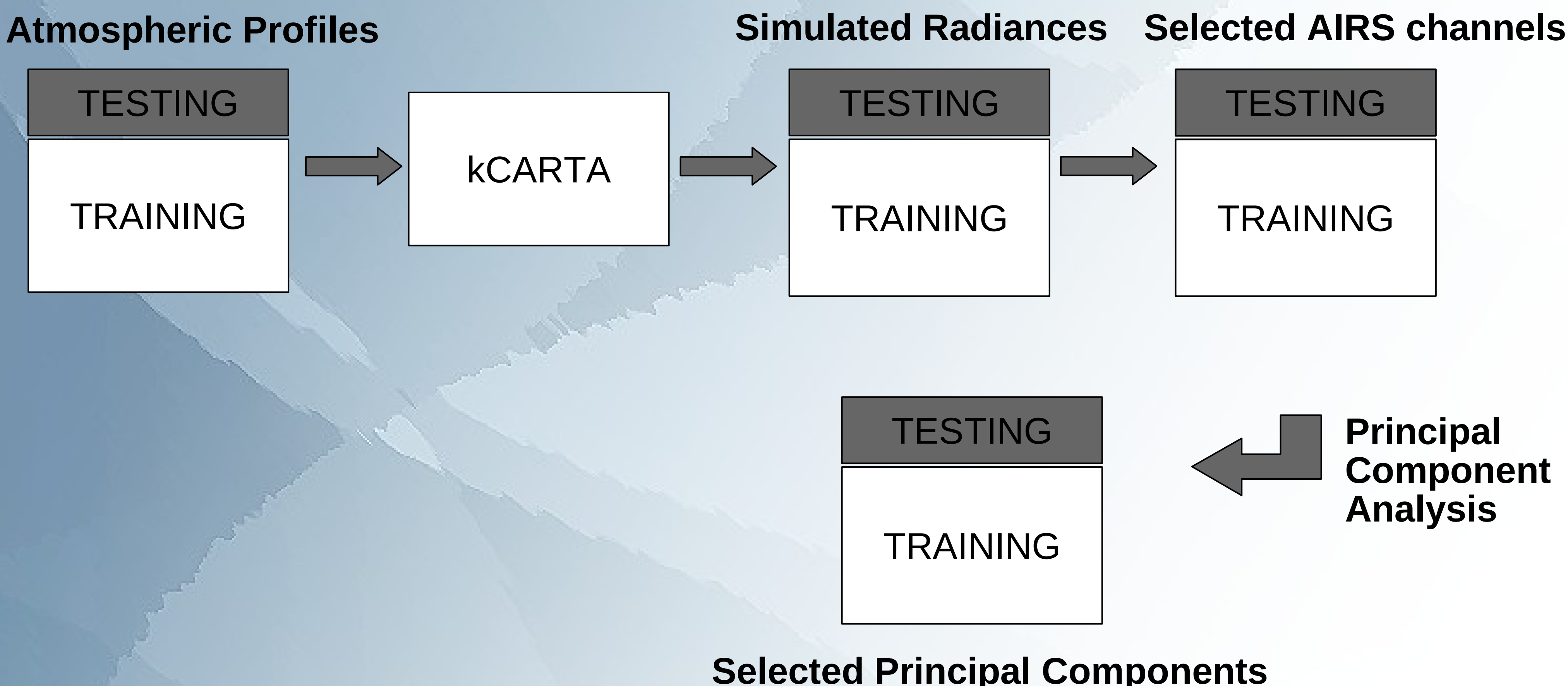
TYPE1

- Training set of 10000 profile (includes all 3876 tropical profile and synthetic profiles)
- Testing set of 3000 synthetic profiles

TYPE2

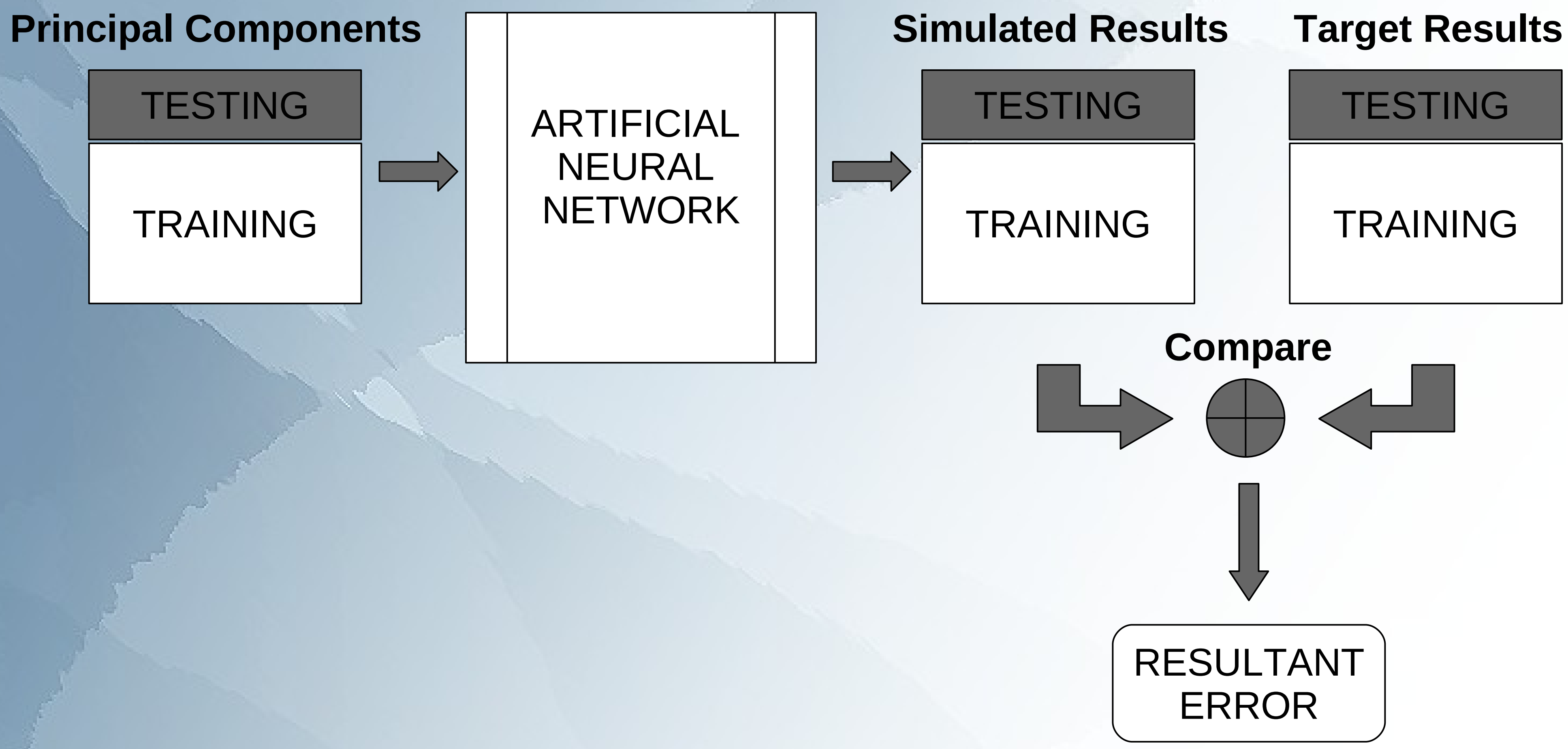
- Training set of ~12000 profile (same as Type 1 but excluding the RS profile)
- Testing set of 545 RS profile

METHODOLOGY



Schematic of the initial step to create the testing and training dataset

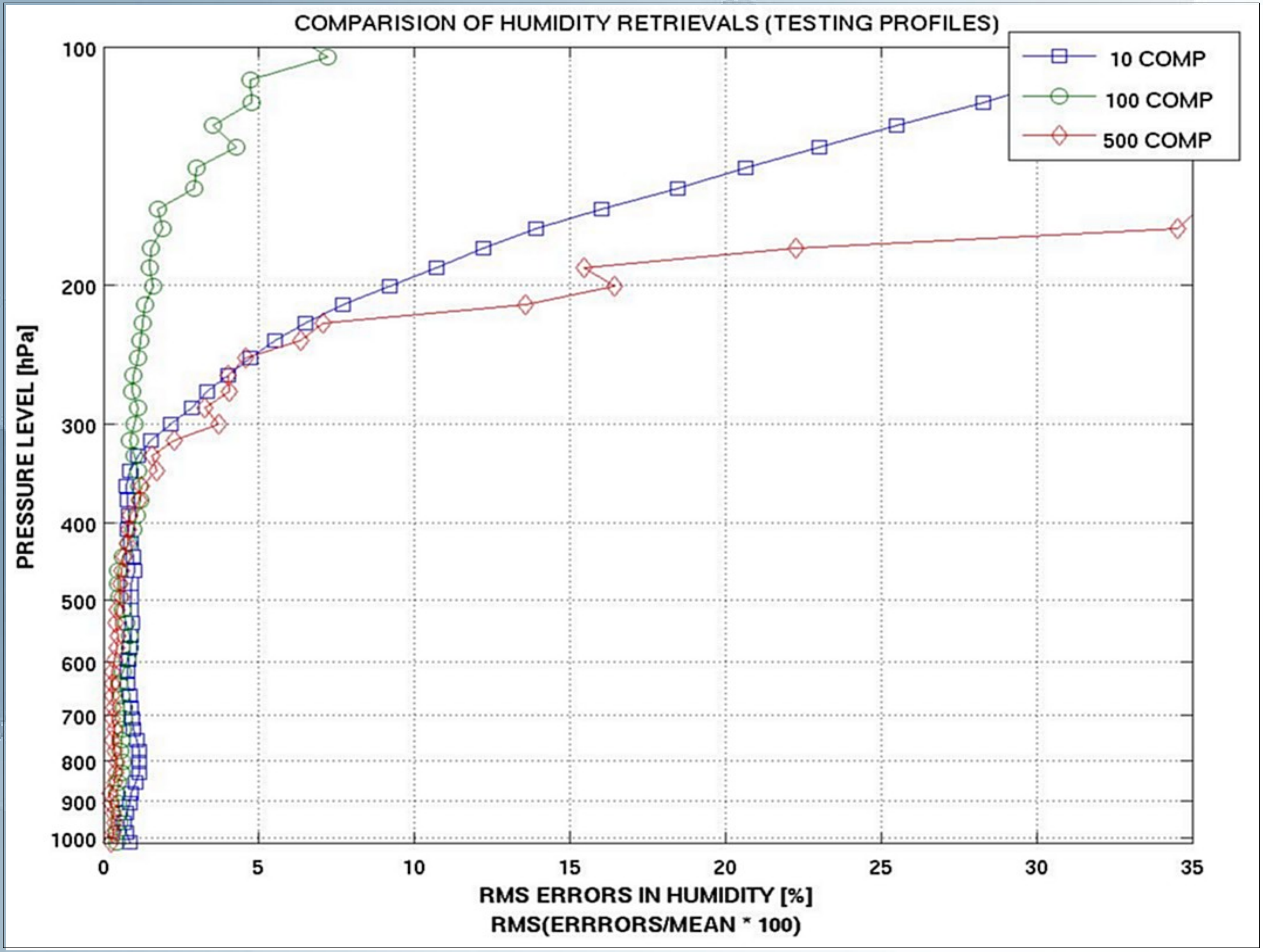
METHODOLOGY



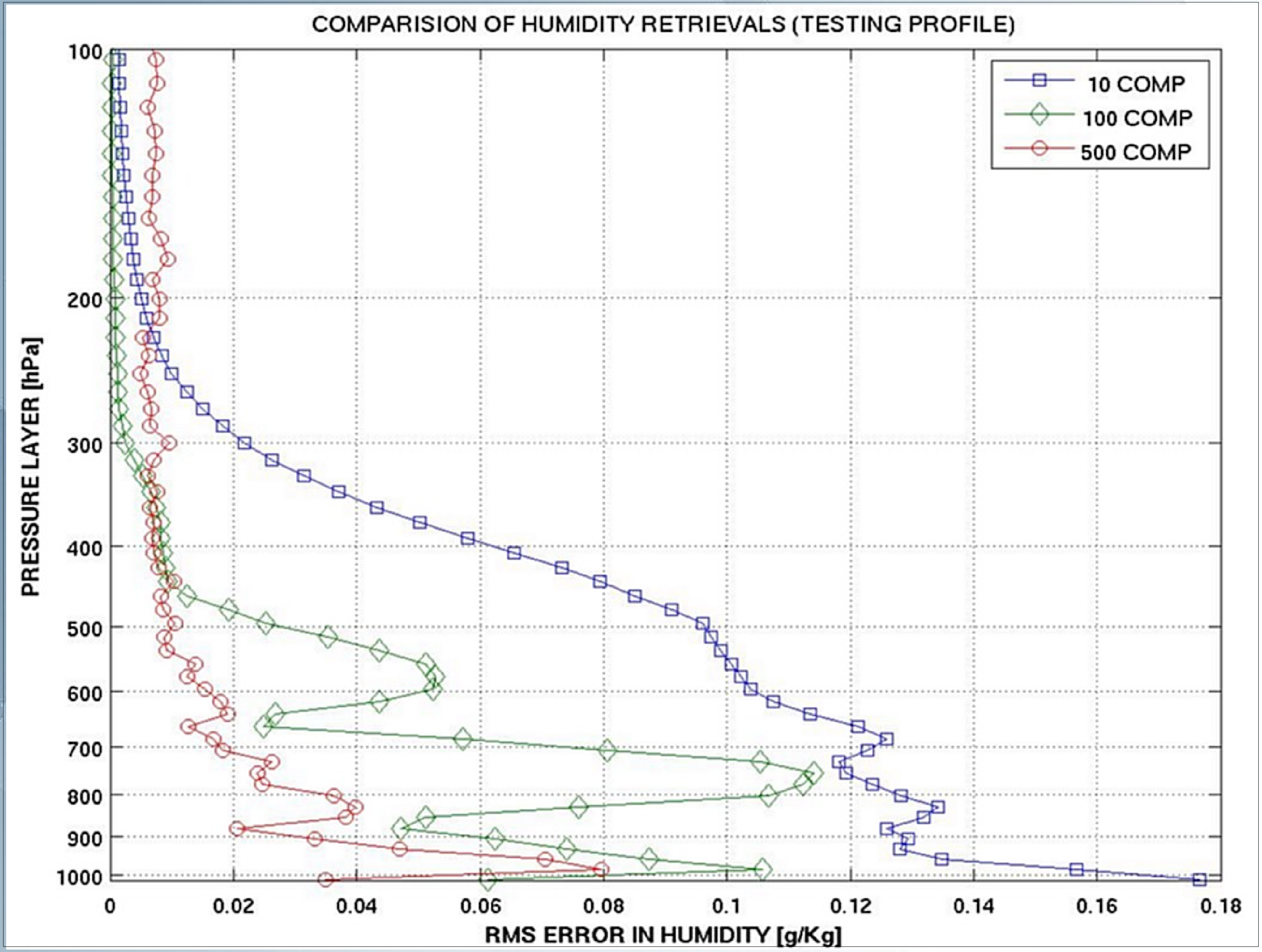
Schematic of ANN model used in retrieval process

RESULTS

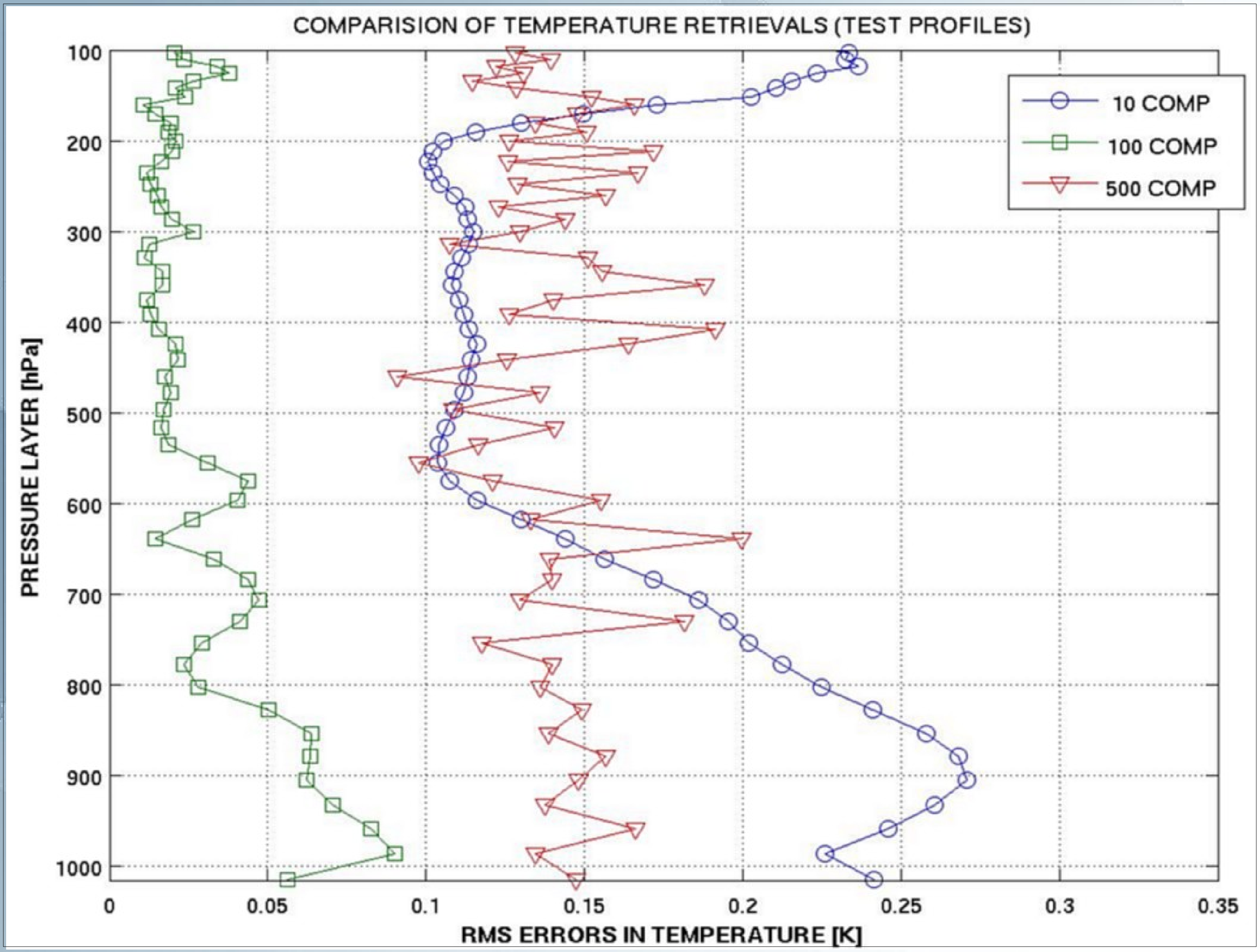
Results on PCA Based ANN (Type 1)



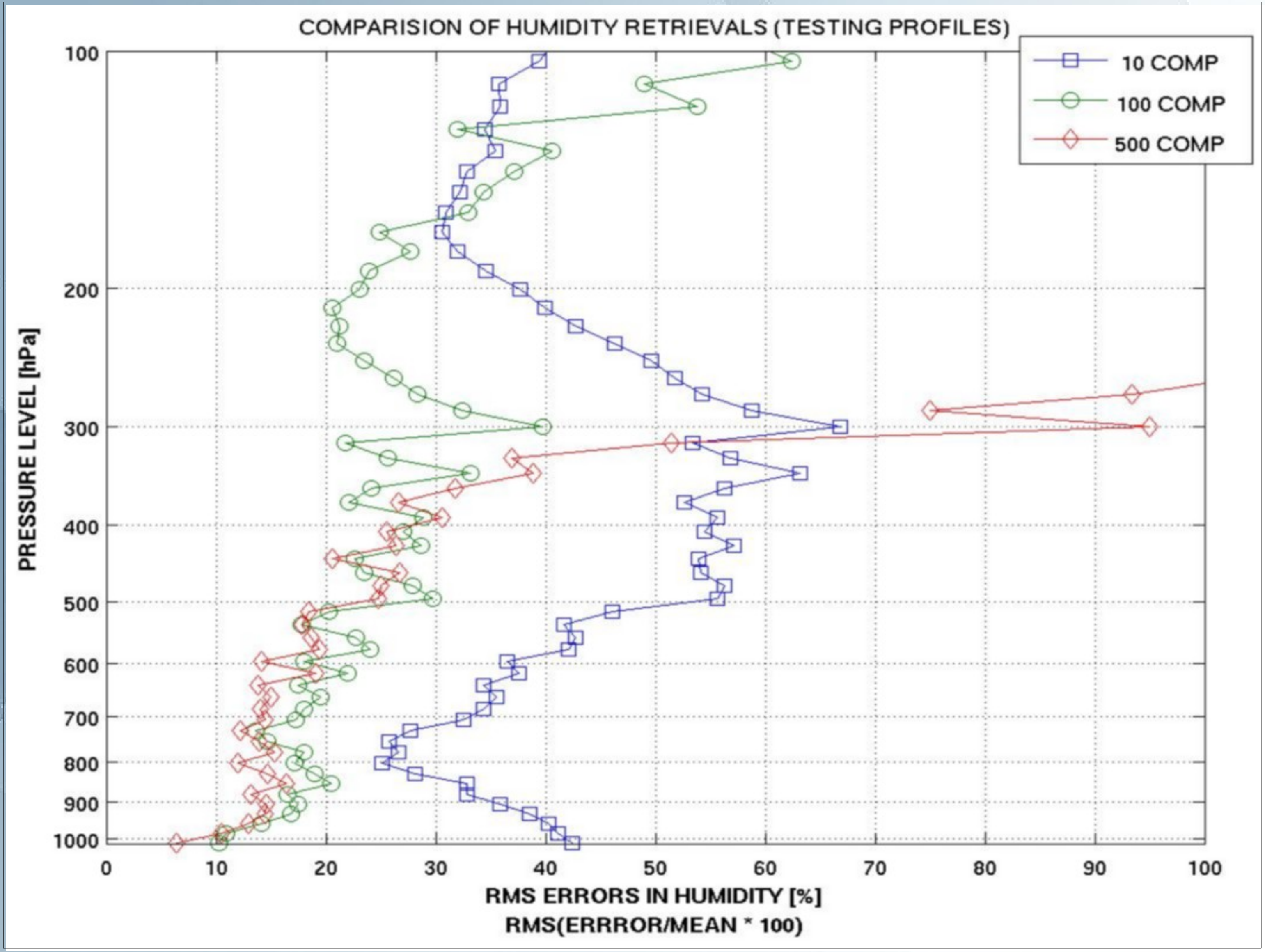
Results on PCA Based ANN (Type 1)



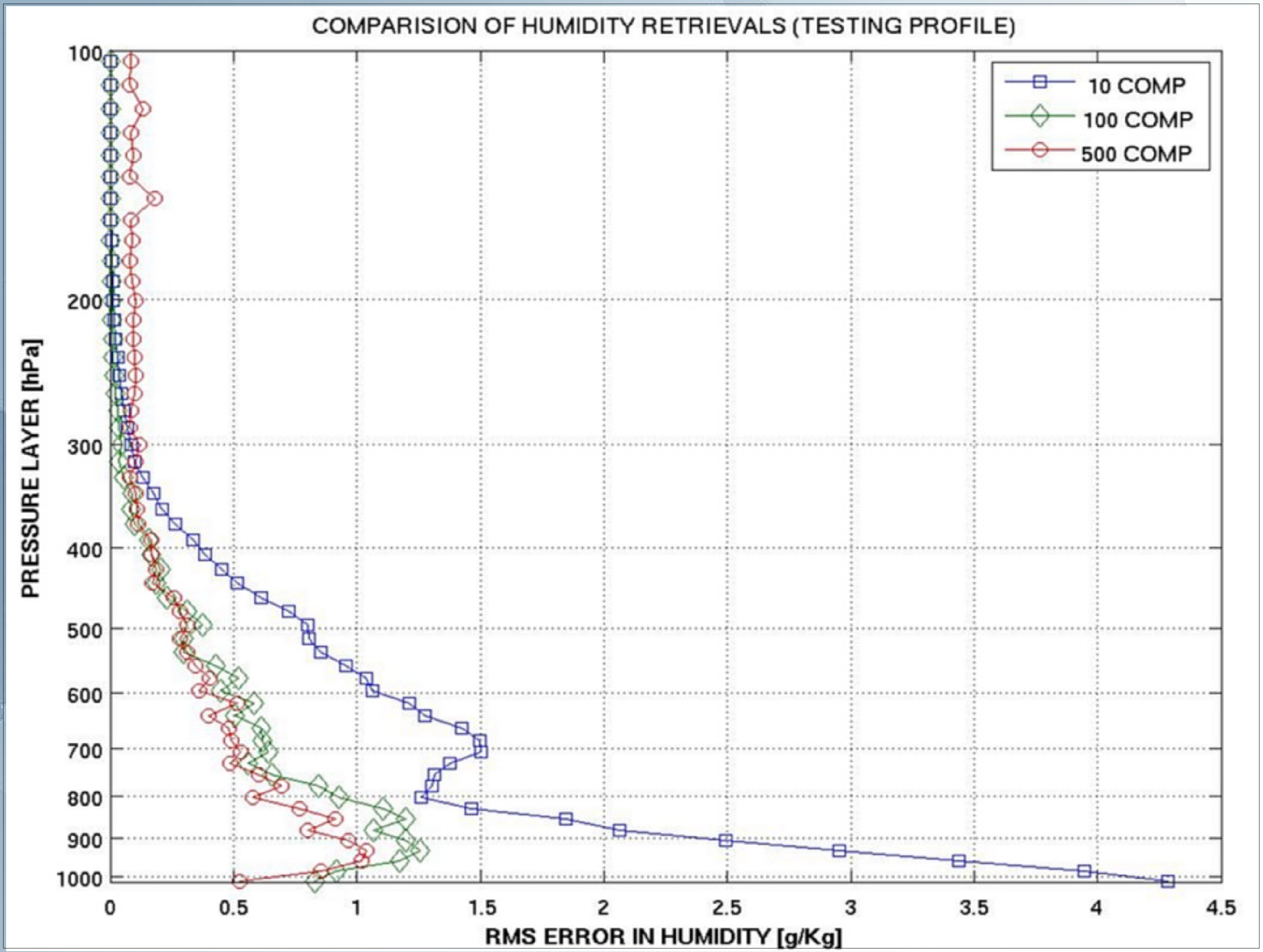
Results on PCA Based ANN (Type 1)



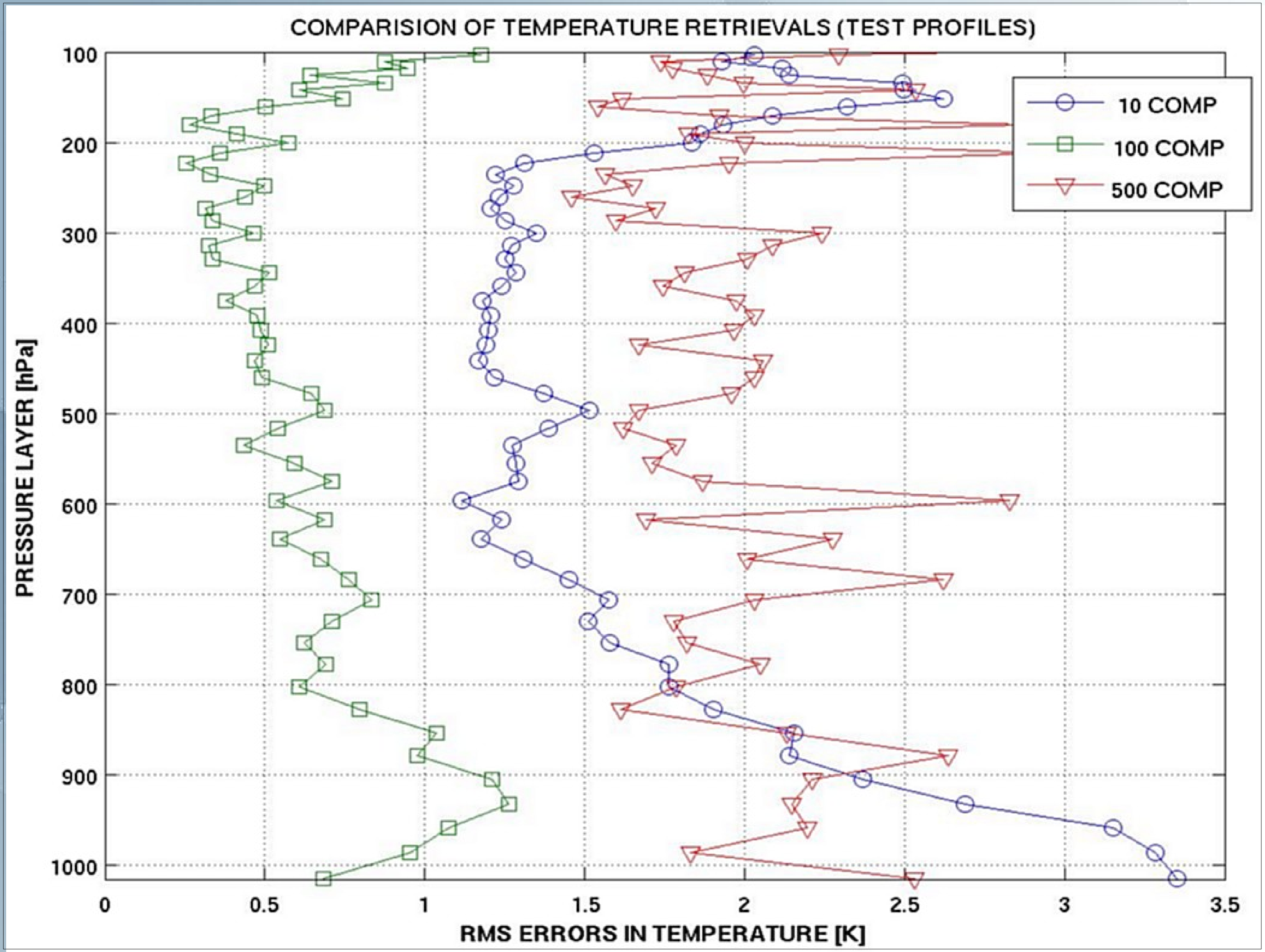
Results on PCA Based ANN (Type 2)



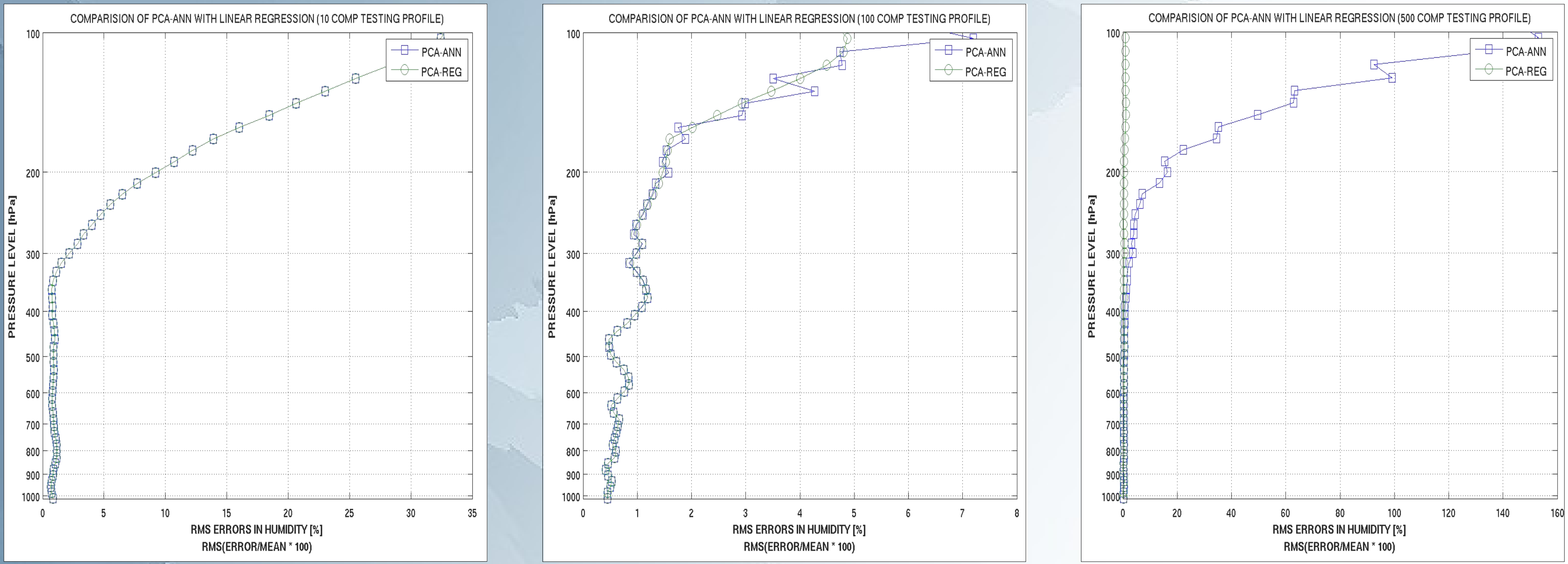
Results on PCA Based ANN (Type 2)



Results on PCA Based ANN (Type 2)

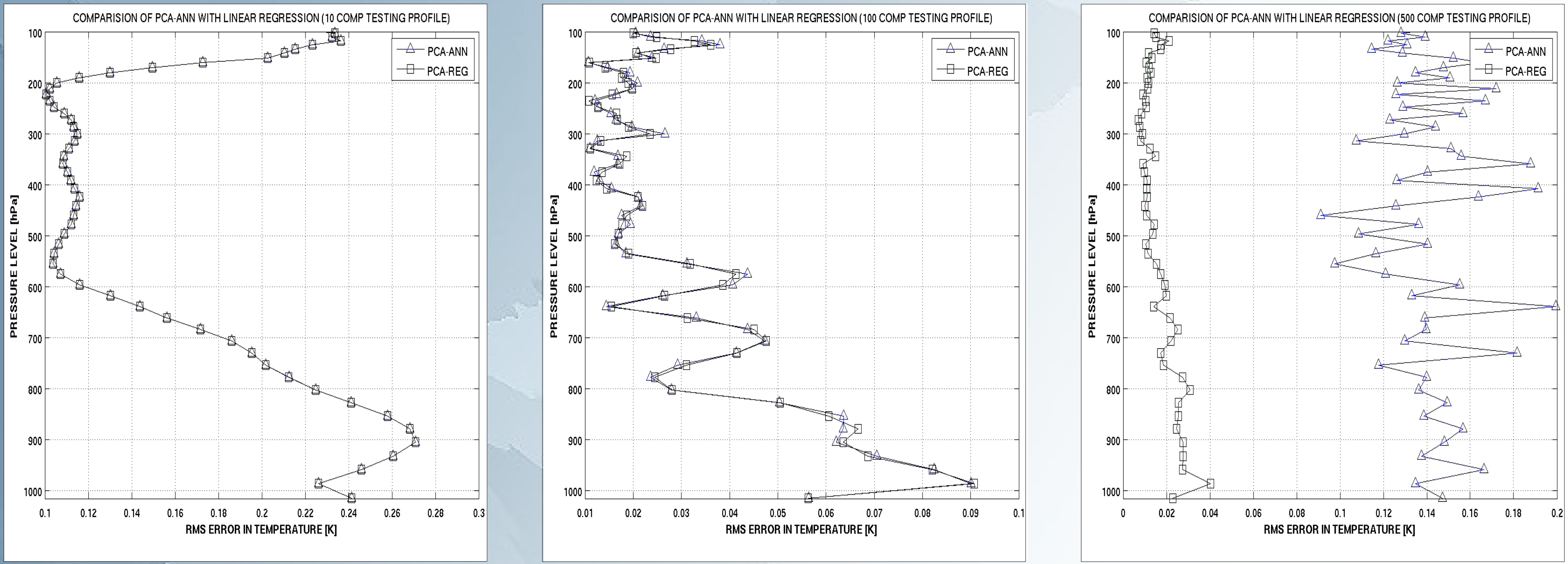


PCA Based ANN (Type 1) versus PCA Linear Regression



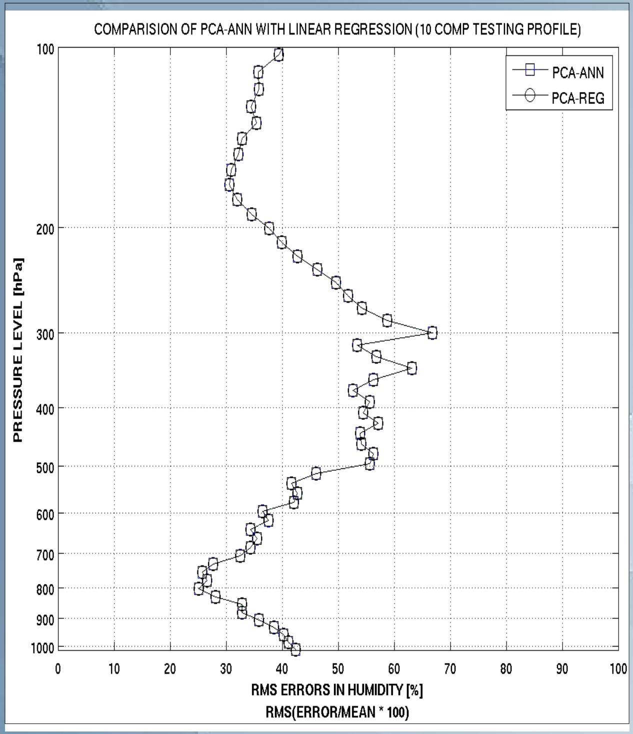
Comparison of PCA-ANN and PCA-Regression Humidity Retrieval using 10, 100, 500 Principal Components

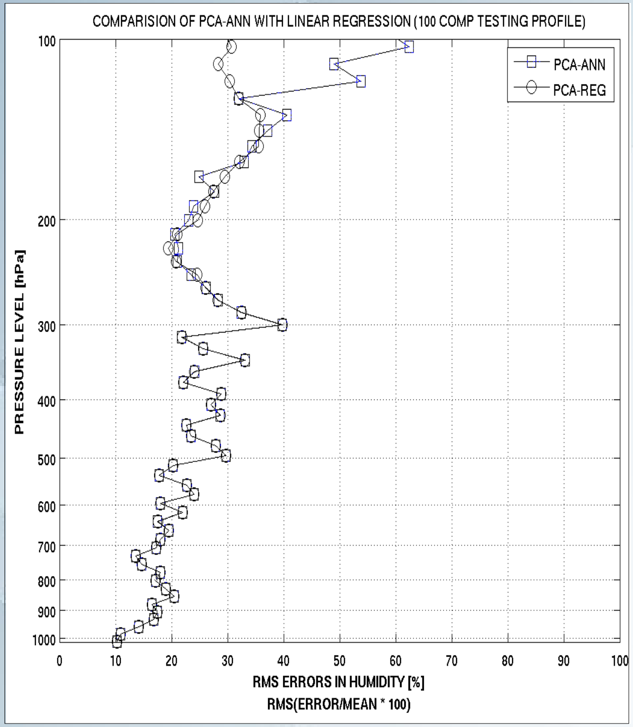
PCA Based ANN (Type 1) versus PCA Linear Regression

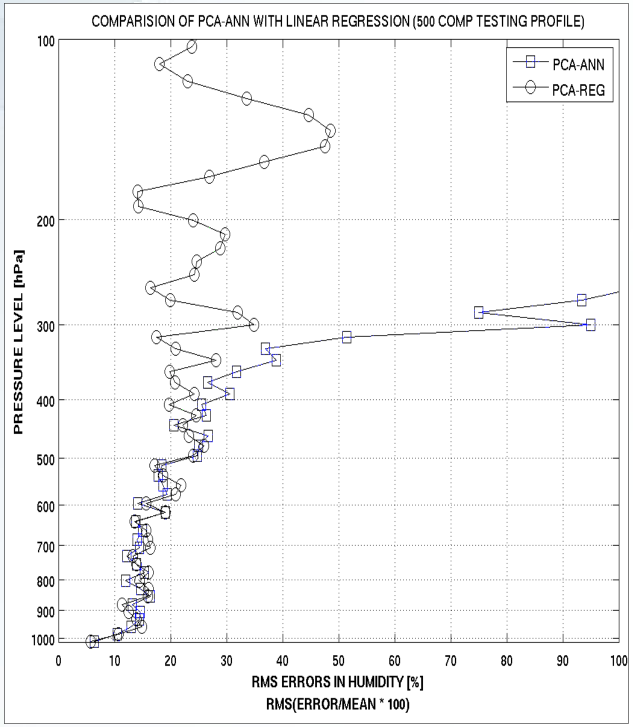


Comparison of PCA-ANN and PCA-Regression Humidity Retrieval using 10, 100, 500 Principal Components

PCA Based ANN (Type 2) versus PCA Linear Regression

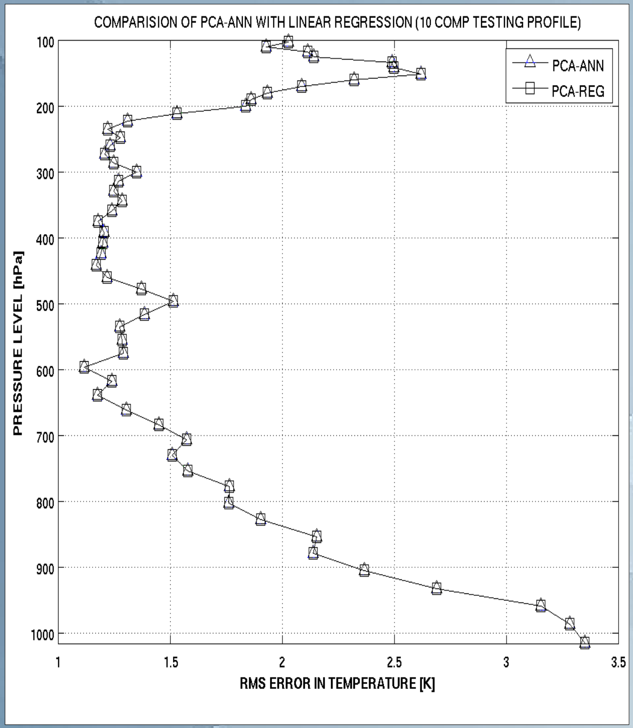


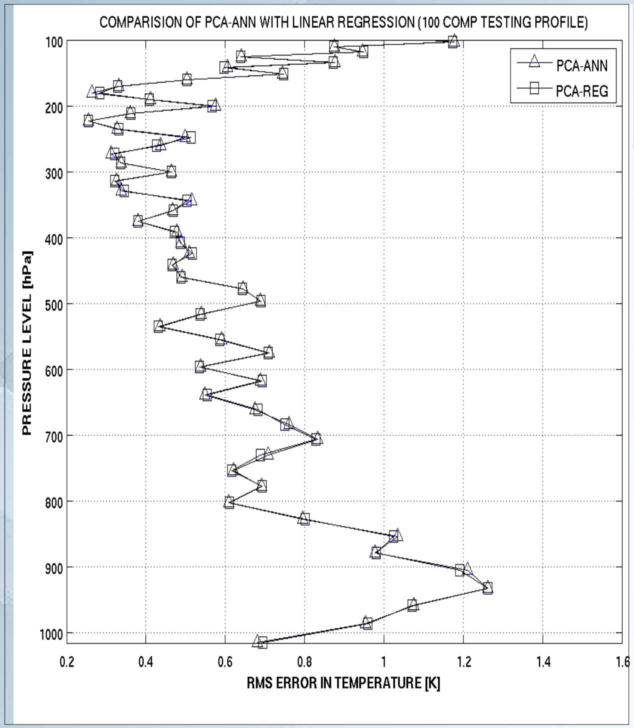


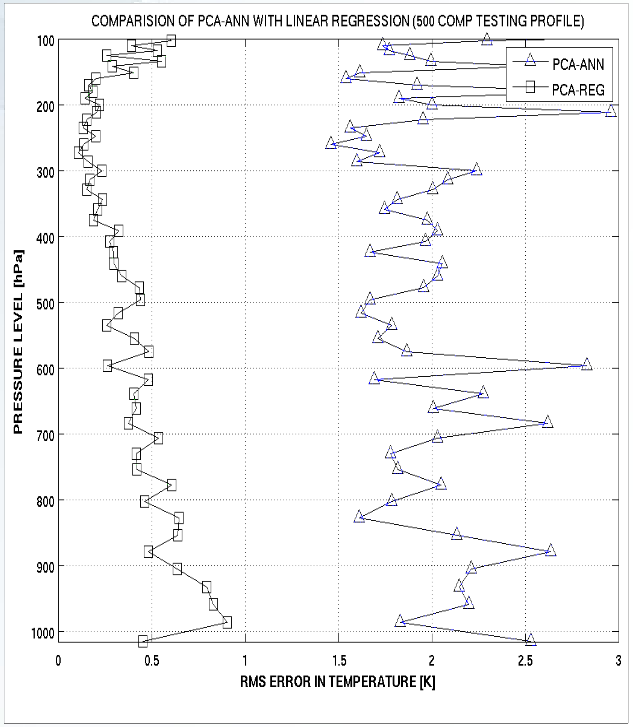


Comparison of PCA-ANN and PCA-Regression Humidity Retrieval using 10, 100, 500 Principal Components

PCA Based ANN (Type 2) versus PCA Linear Regression







Comparison of PCA-ANN and PCA-Regression Humidity Retrieval using 10, 100, 500 Principal Components

CONCLUSION

- Humidity and Temperature were retrieval within the limits set by World Meteorological Organization
- Optimal number of components exist for the most accurate prediction in PCA based ANN
- PCA based ANN and PCA based Regression converges to almost the same accuracy

REFERENCE

- **Lambrigtsen, B. , E. Fetzer, E. Fishbein, S.Y. Lee and T. Pagano**, (2004): AIRS-The Atmospheric Infrared Sounder, *IEEE*
- **Blackwell, W. , J.** , A Neural-Network Technique for the Retrieval of Atmospheric Temperature and Moisture Profiles From High Spectral Resolution Sounding Data, (2005):*IEEE Transaction on Geoscience and Remote Sensing*, (**43**),No. 11
- **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