The Use of Principal Component Analysis (PCA) in Processing AIRS Data
Lihang Zhou1, Mitchell D. Goldberg1, Walter W. Wolf2, Chris Barnet1 and Yanni Qu3
1NOAA/NESDIS/ORA, Camp Springs, MD 20746; 2QSS Group, Inc, Lanham, MD 20706; 3ITT

1. INTRODUCTION

Atmospheric Infrared Sounder (AIRS) on the NASA EOS AQUA platform is providing much improved atmospheric temperature and moisture profiles when compared to soundings from current operational sounders. Principal Components Analysis (PCA) provides an effective way to approximate the AIRS observations. This poster presents the application of PCA to AIRS data, including the generation and application of eigenvectors, and the use of PCA for data compression, high noise channel detection and reconstruction, noise filtering and estimation, and regression retrieval.

2. METHOD

In our application the data vectors are radiances spectra that are divided by the expected instrumental noise, which we refer to as normalized radiances. The eigenvectors are related to the covariance matrix by:

\[ S = E \Theta T \]

where S is a N x N covariance matrix, E is the matrix of eigenvectors, and T is the diagonal matrix of eigenvalues. The principal component scores P are computed from:

\[ P = E^T R \]

where R is the vector of centered normalized radiances. An overall estimate of how well the principal component scores can reconstruct the original data is provided by Reconstruction Score (RS), which is defined as:

\[ RS = \left[ \frac{1}{N} (O_i - R_i)^2 \right]^{1/2} \]

with N = number of channels, Oi and Ri are noise scaled observed and reconstructed radiances for i-th channel.

Application 1: Reconstruct Radiance

Our studies indicate that AIRS brightness temperature can be reconstructed within the noise level, using about 10 eigenvectors.

Application 2: Channel Monitoring and Bad Channel Handling

In operational AIRS processing, when a channel has a bad radiance value, it has to be replaced by something useful since information from each channel is required to reconstruct the radiances.

Bad channel handling technique: when a bad channel is detected, use the neighboring channels to compute the pcs, and reconstruct the bad channel.

Application 3: Noise Estimate/Filtering

Application 4: Regression Retrieval

Principal component regression of near real time AIRS data provides a very computational efficient retrieval of atmospheric temperature, moisture, and ozone, etc.

Principal component scores have been a very effective tool for data compression, detector monitoring and bad channel handling, as well as the noise estimation.

Use the first 85 PCS as predictors, generate regression coefficients generated from June have been stable.

For more information, please visit http://orbit35i.nesdis.noaa.gov/crad2/airsgrid/AIRS_NRT