

DEPARTMENT OF STATISTICS COLLOQUIUM SERIES

Texas A&M University

**TAO SHI**

Department of Statistics  
Ohio State University

**STATISTICAL MODELING OF AIRS LEVEL 3  
QUANTIZATION DATA**

**ABSTRACT:**

Atmospheric Infrared Sounder (AIRS) has been collecting temperatures, water vapor mass-mixing ratios, and cloud fraction at various atmosphere pressure levels. It generates 35 dimensional vectors at each  $45km$  ground footprint in each satellite path in its level-2 data. The level 3 quantization data (L3Q) summarize valid level-2 data in each  $5^\circ \times 5^\circ$  latitude-longitude grid box during a time period by a set of representative vectors and their associated weights. The specialty of the data set is that the observations are empirical distributions. Most statistical methods are mainly developed for handling datasets whose observations are in  $\mathbb{R}^d$ . Statistical inference for this type of data is an open problem. We start with the commonly used Mallows distance as a measure of distance between two distributions and build a mixture model on empirical distributions with each component being a Gaussian type distribution. We further fit the model using Data Spectroscopic type of methods for AIRS L3Q data. Finally, we will address some statistical questions such as classification and prediction on AIRS L3Q data.

**DATE:** Thursday, September 9, 2010

**TIME:** 3:45 p.m. – 4:45 p.m.

**PLACE:** Room 457, Blocker

Refreshments will be served in the Statistics Conference Room at 3:30 pm for those attending the seminar.