# A Statistical Glance at the L3 Data

*Alexander Ruzmaikin*

*in collaboration with*

George Aumann & *Evan Manning*

# *AIRS Level 3 Data Products*

The v5  L3 monthly products are on a 1° x 1° grid

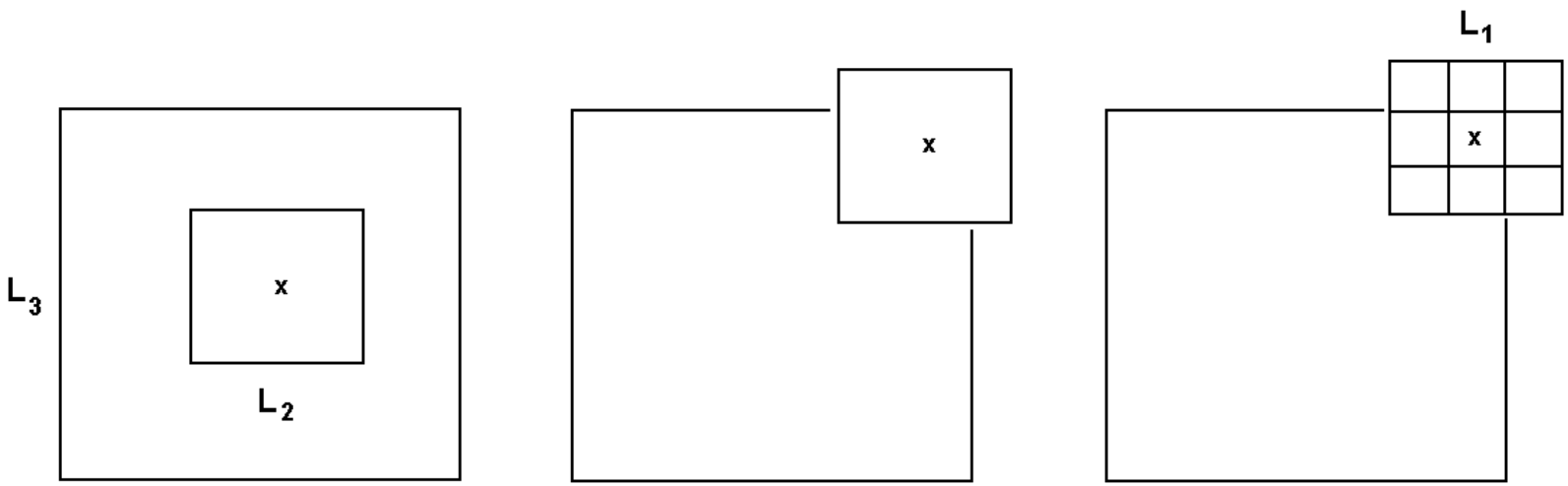
The L3 prescription uses L2 quality flags

The L3 product  could be an excellent resource for climate research

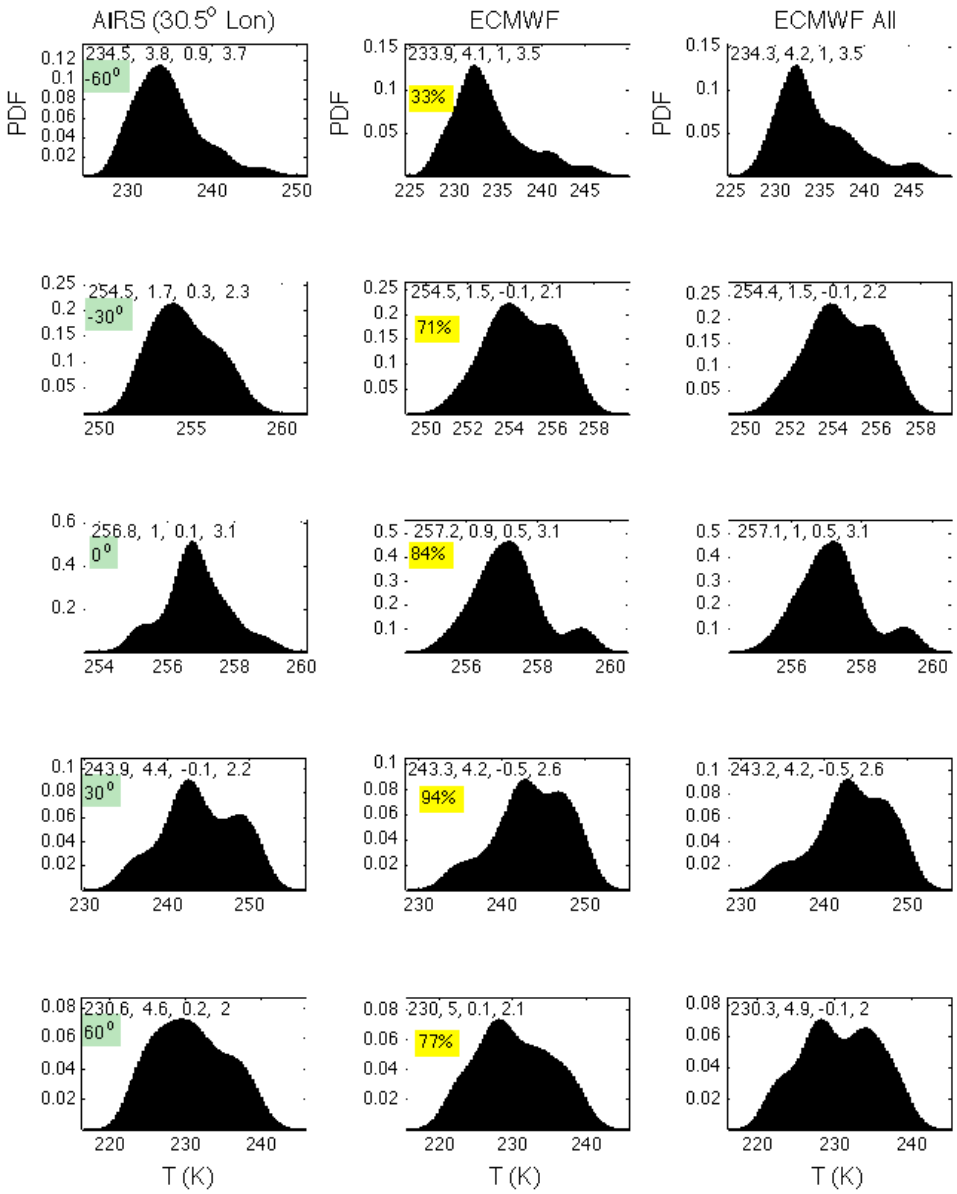
# Objective and Data

- ✧ Compare Airs L3 temperature distributions with temperatures from ECMWF
- ✧ Daily L2 data for January 2005 on 1° x 1° grid
- ✧ We preserve the latitude-longitude information on AIRS grid (13.5 km x 13.5 km), i.e. there are 9 pixels in the 1° x 1° box
- ✧ ECMWF Forecast data at the same time and place

# Location, Location

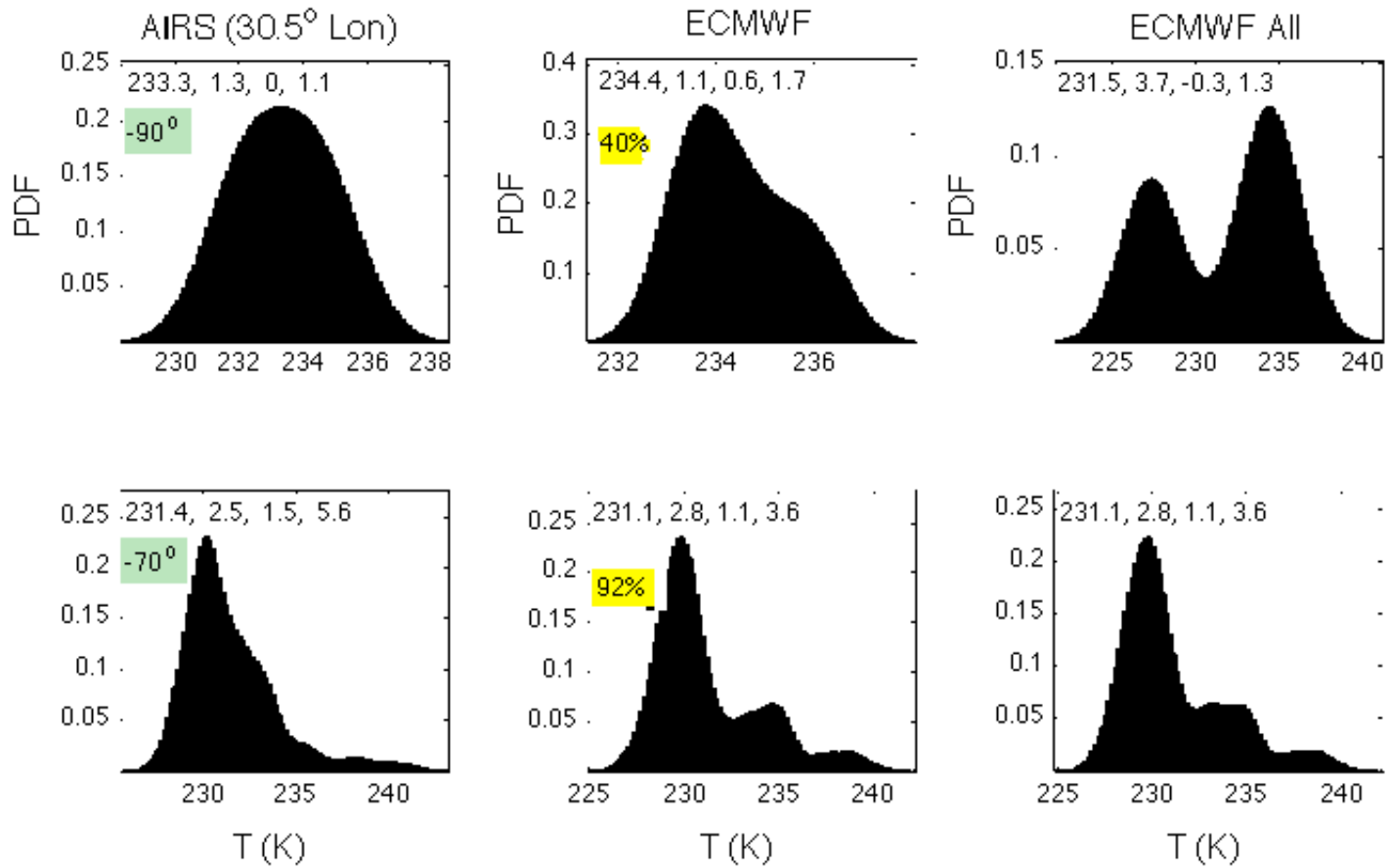


# Distributions: T at 400 hPa AIRS (Pgood) and ECMWF




Selected lat regions along 30° longitude (green), % of total data (yellow). Numbers on top: mean, std, skewness and kurtosis.

# More Distributions: T at 400 hPa

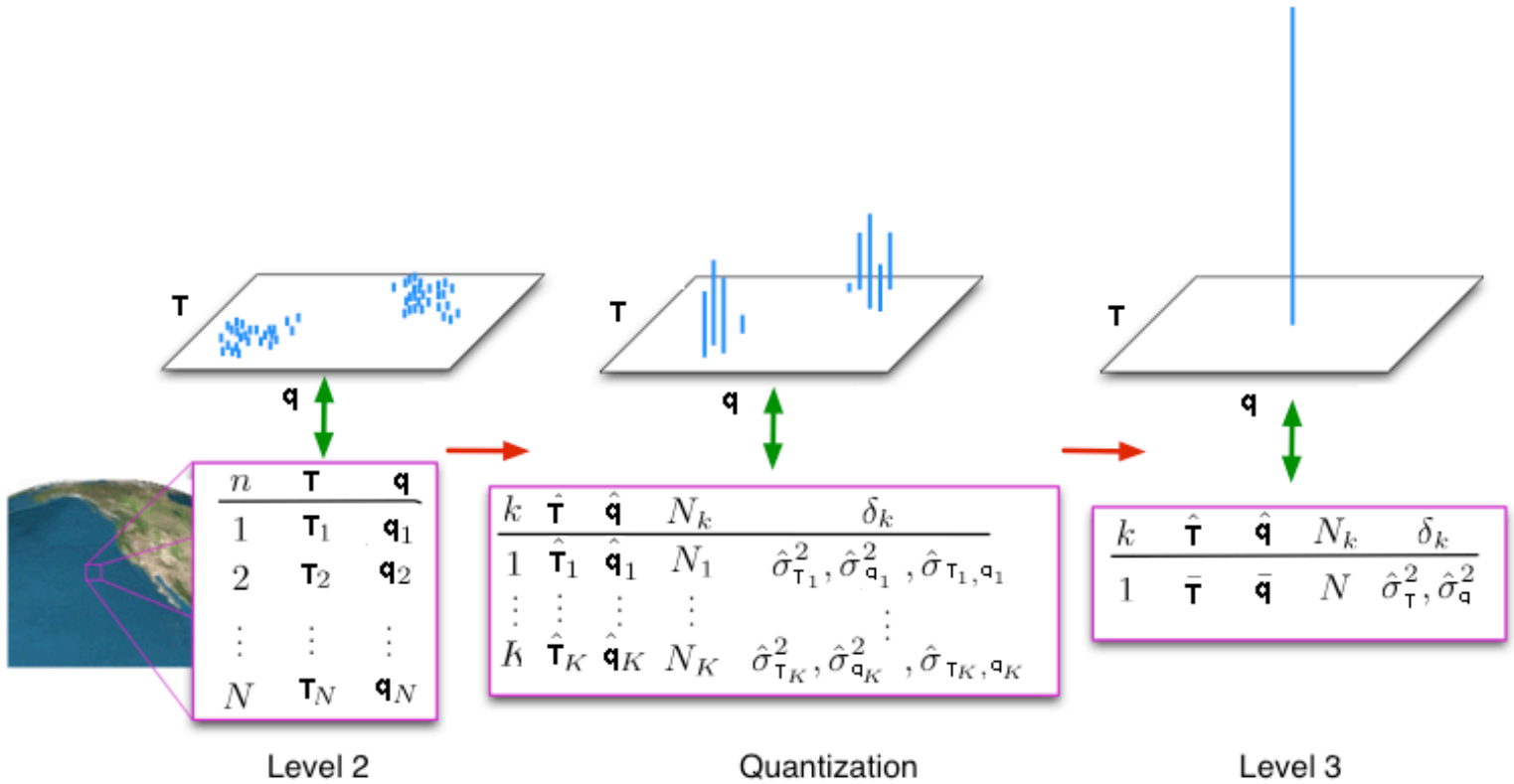


# Observation

- ✧ In some L3 pixels good quality L2 data fill less than 40% and may not be statistically representative
- ✧AIRS and ECMWF have close distributions
- ✧ Distributions are non-Gaussian, i.e. are not characterized by means and standard deviations
- ✧ Adding skewness and kurtosis is not sufficient due to double peaks

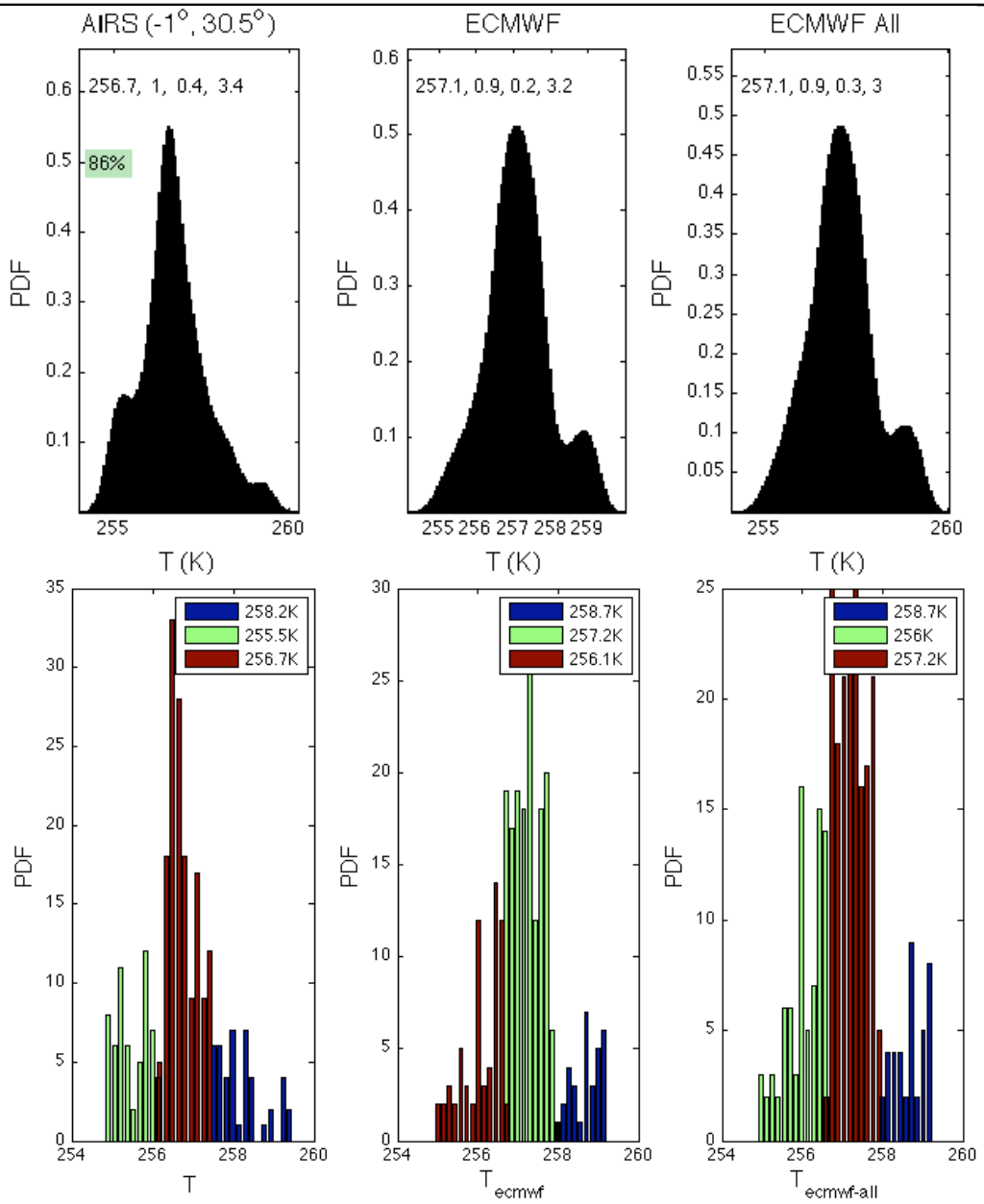
# Can We Preserve the Distributions Without Vastly Increasing the Data Volume?

# Distribution Estimates




Entropy-Constrained Vector Quantization (Chou et al., 1989; Braverman, 2002) reduces the distribution to a few clusters. Outputs: centroid, number of data points assigned to a cluster, and elements of cluster covariance matrix.

# Example of Cluster Reduction for AIRS Data



# Conclusions

- Mean and STD -- insufficient output and may bring errors
- Adding skewness and kurtosis --helps but not adequate for bi-modal distributions
- Distribution Estimates (ECVQ) -- a possible solution