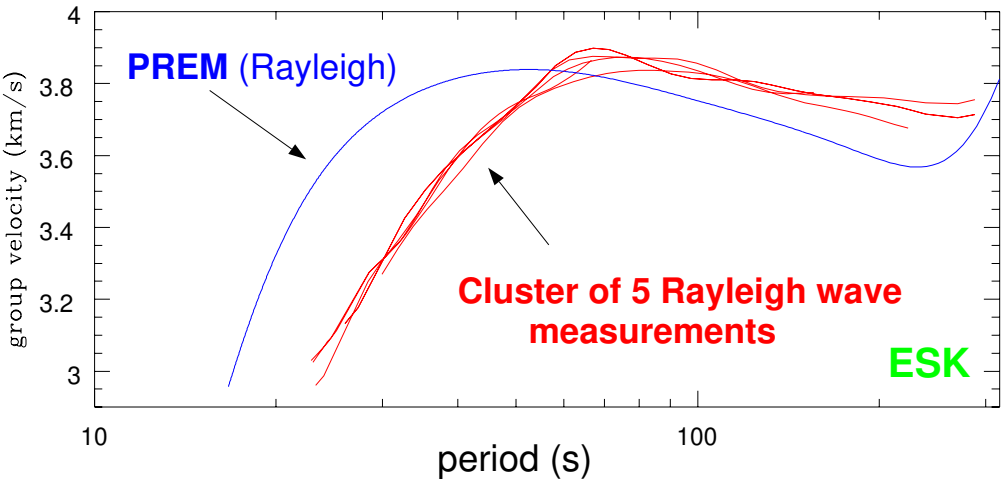
Cluster Analysis:

Redundancy, Reliability, & Uncertainty Estimates

Cluster: **A set of measurements that share approximately the same path.**

e.g., One Station & Several events (aftershocks)
Two or more nearby stations & One or more events
(MAJO/INU, TOL/PAB, KNET)

A cluster defines a unique measurement (path or ray).



Eskdalemuir, Scotland
Mindanao, Philippines

12/15/89	**Ms = 7.4**
2/8/90	**Ms = 6.6**
2/18/91	**Ms = 6.6**
11/13/91	**Ms = 6.4**
5/9/92	**Ms = 7.1**

Redundancy:

Reduce the number of source - receiver pairs to a set of unique paths -- one from each cluster -- with unique dispersion characteristics.

Reliability:

Rayleigh: 5200 → 3500
Love: 5800 → 3800

Discard inconsistent measurements.

Uncertainty Estimates:

Estimate the frequency-dependent standard deviation of measurements within each cluster. Weight the measurement from the cluster accordingly.

Estimate the average frequency-dependent standard deviation over the entire data set.

Problems:

1. **Clustering is rare at long periods -- very difficult beyond 200 s.**
2. **The sample contains certain paths preferentially -- those originating from regions with large earthquakes.**

(e.g., Aleutians, Kuriles, Sumbawa, W. China, Iran, etc.)