Appendix A- Mohawk Mountain Supplement

Trajectory analysis results at Mohawk Mountain.

Equations for Different Metrics

Everyday Residence-time Probability

$$EP = \left(n_{ij} \big/ N \right)$$

n_{ij} = total endpoints passing through grid cell i, j

N = total endpoints passing through all grid cells from all trajectories

High Day Residence-time Probability

$$HP = \left(m_{ij} \big/ M \right)$$

m_{ij} = total high day endpoints passing through grid cell i, j

M = total high day endpoints passing through all grid cells from high day trajectories

Incremental Probability

$$IP = HP - EP$$

Cluster-Weighted Probability

$$CWP = \frac{1}{\overline{C}} \left(\sum_{i=1}^{L} (\overline{C})_i \cdot RP_i - \overline{C} \cdot EP \right)$$

L = total number of clusters calculated

$(\overline{C})_i$ = Average pollutant concentration (based on observations associated with cluster i)

$\overline{C}$ = Average pollutant concentration (based on all days)

Description of Figures

- Central Trajectory (CT)- Trajectory with the largest number of nearest neighbors in the dataset.

- Frequency Based Clusters- These clusters are formed by finding the "central" trajectory which has the greatest number of neighboring trajectories within a subjectively selected radius of proximity (R). These trajectories are then removed from the dataset and the process is applied to the remaining trajectories.

- Proximity Based Clusters- Clustering relies on the frequency-based cluster groups, but forms trajectory groups based on proximity rather than frequency. In the first step, the frequency-based approach is used to identify the central trajectories that represent the most populated frequency-based clusters (approximately 10 clusters typically contain at least 98% of the trajectories in the dataset using R=12 and 120 hour back-trajectory (BT) time).   These 10 central trajectories are then used to develop 10 proximity-based clusters by assigning every trajectory in the dataset to its nearest central trajectories (calculated back to 72 hours).

- Incremental Probability- Difference between the everyday probability (probability derived from all the trajectories in the dataset) and high day probability (probability derived from trajectories arriving at the site on the subset of high pollution days).

- Cluster Weighted Probability- Each PATH-derived cluster's residence-time probability is weighted by the average sulfate (or other pollutant) value for any measurements corresponding to a trajectory which is a member of that cluster. The weighted residence-time probability is summed over *all* clusters calculated for a site. The everyday probability is subtracted from the sum of cluster-weighted probabilities to identify areas of increased (or in the case of negative values, decreased) probability of being associated with a meteorological pathway for pollutant transport.

Mohawk Mt All Trajectories 00-04, Top 9 Clusters

Modes defined at: R=12, 120hr BT, 500m Start ht, 6456 Valid Trajectories, 7743 Invalid
Reassigned Trajectories Based on 72hr BT, 500m Start Ht, 10161 Valid Trajectories
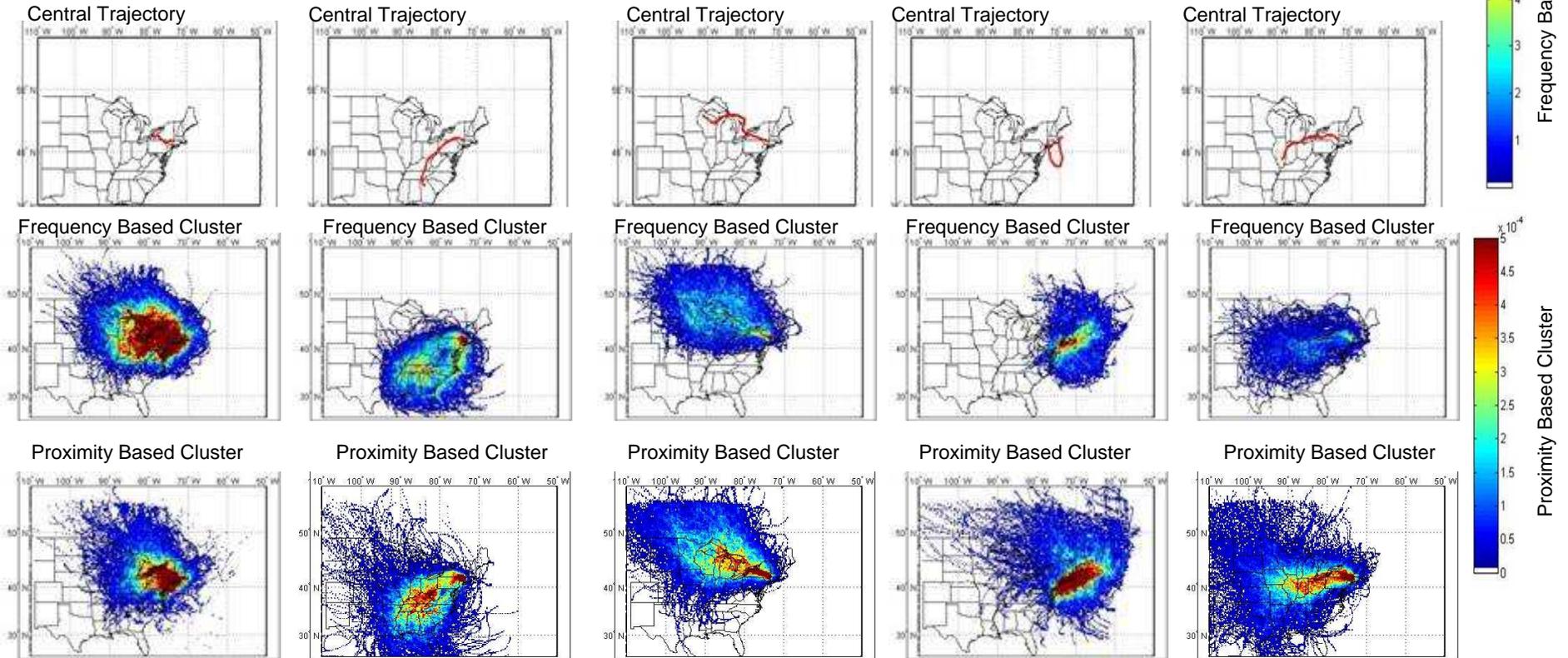

Cluster 1		Cluster 2		Cluster 3		Cluster 4		Cluster 5	
	Frequency	Proximity		Frequency	Proximity		Frequency	Proximity	

Cluster 1

	Frequency	Proximity
Sulfate	4.00	4.25
Bext	74.59	77.74
PM	9.83	10.12
OC	2.85	3.03
# Trajs	4562	1286
# Trajs w. Poll	931	293

Cluster 2

	Frequency	Proximity
Sulfate	1.62	5.10
Bext	35.40	91.95
PM	4.55	11.78
OC	1.41	3.19
# Trajs	1318	917
# Trajs w. Poll	254	205

Cluster 3

	Frequency	Proximity
Sulfate	4.85	2.01
Bext	89.55	41.43
PM	10.99	5.35
OC	2.87	1.50
# Trajs	925	1130
# Trajs w. Poll	196	200

Cluster 4

	Frequency	Proximity
Sulfate	1.75	2.65
Bext	36.74	49.81
PM	5.12	6.45
OC	1.66	1.73
# Trajs	812	698
# Trajs w. Poll	168	141

Cluster 5

	Frequency	Proximity
Sulfate	5.04	4.84
Bext	89.33	83.57
PM	12.38	10.98
OC	3.31	2.58
# Trajs	641	1135
# Trajs w. Poll	110	220

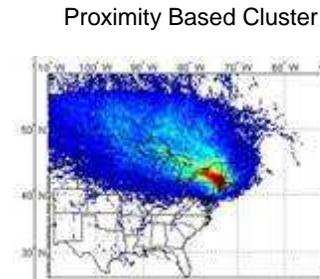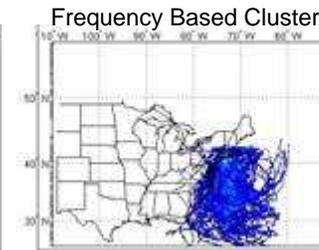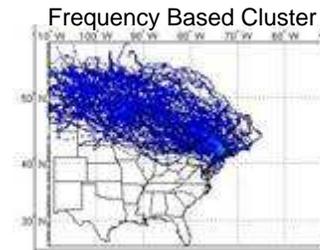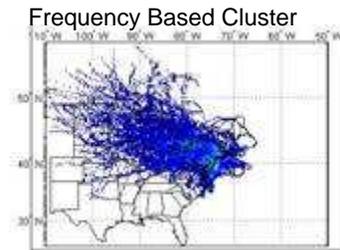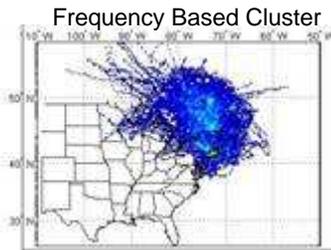# Mohawk Mt All Trajectories 00-04, Top 9 Clusters

Modes defined at: R=12, 120hr BT, 500m Start ht, 6456 Valid Trajectories, 7743 Invalid
Reassigned Trajectories Based on 72hr BT, 500m Start Ht, 10161 Valid Trajectories

Cluster 6

Central Trajectory
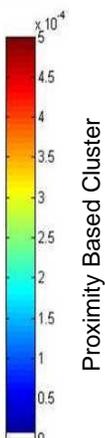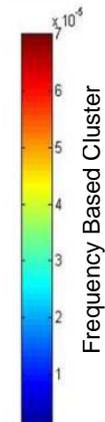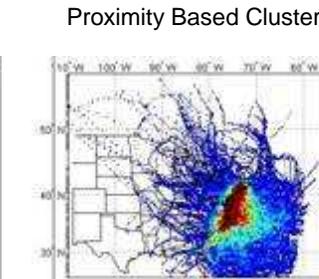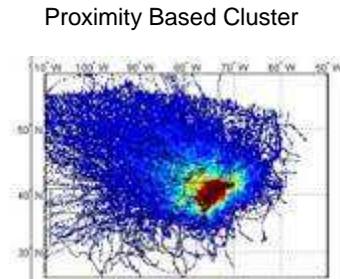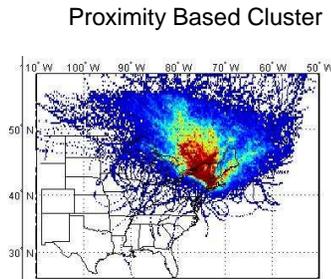
Frequency Based Cluster

Proximity Based Cluster

	Frequency	Proximity
Sulfate	1.96	1.53
Bext	38.83	40.71
PM	5.53	6.44
OC	1.62	2.95
# Trajs	301	1141
# Trajs w. Poll	41	239

Cluster 7

Central Trajectory

Frequency Based Cluster

Proximity Based Cluster

	Frequency	Proximity
Sulfate	2.38	4.62
Bext	47.70	83.40
PM	6.37	10.99
OC	1.69	2.80
# Trajs	270	1061
# Trajs w. Poll	48	217

Cluster 8

Central Trajectory

Frequency Based Cluster

Proximity Based Cluster

	Frequency	Proximity
Sulfate	2.51	2.26
Bext	49.09	47.14
PM	6.52	6.32
OC	1.60	1.87
# Trajs	253	2161
# Trajs w. Poll	58	369

Cluster 9

Central Trajectory

Frequency Based Cluster

Proximity Based Cluster

	Frequency	Proximity
Sulfate	2.34	4.08
Bext	45.33	74.60
PM	5.49	8.50
OC	1.63	2.10
# Trajs	188	641
# Trajs w. Poll	37	118

Frequency Based Cluster

Proximity Based Cluster

Mohawk Mt All Trajectories 00-04, Best/Worst Sulfate

Modes defined at: R=12, 120hr BT, 500m Start ht, 6456 Valid Trajectories, 7743 Invalid

Reassigned Trajectories Based on 72hr BT, 500m Start Ht, 10161 Valid Trajectories
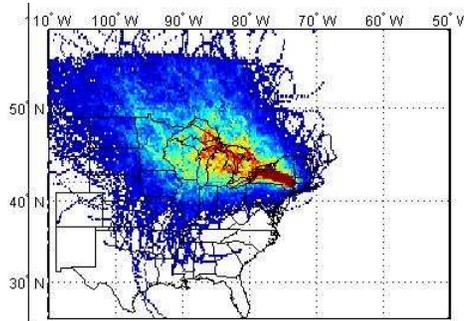
Highest Sulfate
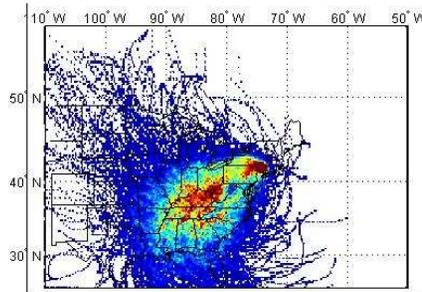
	Frequency	Proximity
Sulfate	1.62	5.10
Bext	35.40	91.95
PM	4.55	11.78
OC	1.41	3.19
# Trajs	1318	917
# Trajs w. Poll	254	205
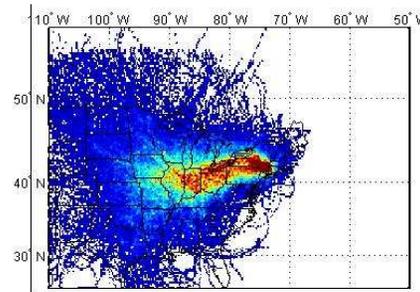
	Frequency	Proximity
Sulfate	5.04	4.84
Bext	89.33	83.57
PM	12.38	10.98
OC	3.31	2.58
# Trajs	641	1135
# Trajs w. Poll	110	220

Lowest Sulfate

	Frequency	Proximity
Sulfate	4.85	2.01
Bext	89.55	41.43
PM	10.99	5.35
OC	2.87	1.50
# Trajs	925	1130
# Trajs w. Poll	196	200
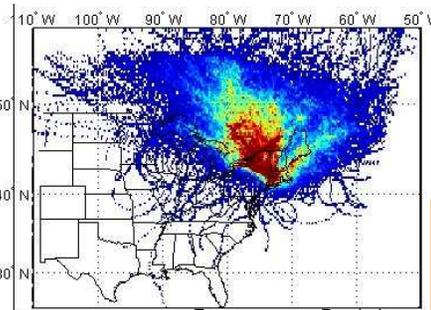
	Frequency	Proximity
Sulfate	1.96	1.53
Bext	38.83	40.71
PM	5.53	6.44
OC	1.62	2.95
# Trajs	301	1141
# Trajs w. Poll	41	239

Sulfate- Sulfate ion Conc. (ug/m3)
Bext- Extinction (Mm-1)
PM- Particulate Matter Conc. (ug/m3)
OC- Organic Carbon Conc. (ug/m3)
Num Trajs- Number of trajectories in cluster
Num Trajs w. Poll- Number of trajectories in cluster with associated pollution measurement (Based on number of IMPROVE samples taken during the 2000-2004 period).

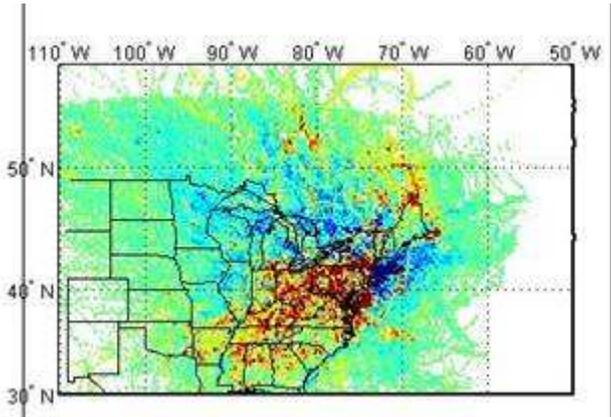# Mohawk Mt All Trajectories 00-04, Incremental Probability
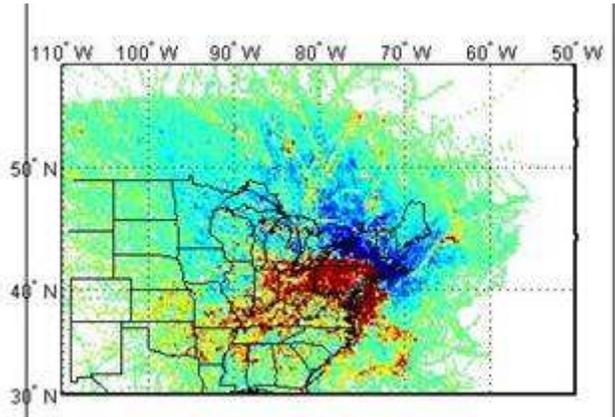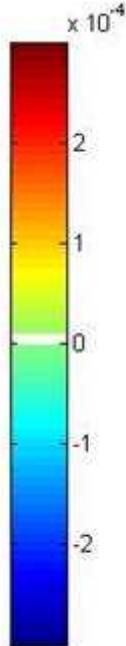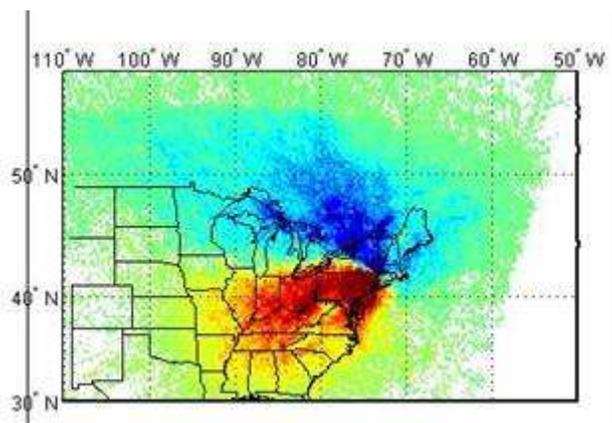## IP Based on Top10%, 500m
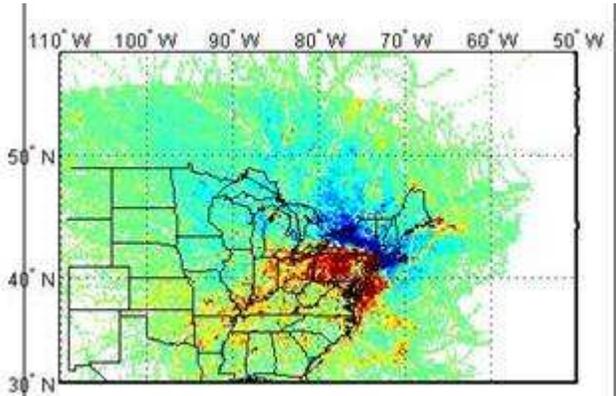
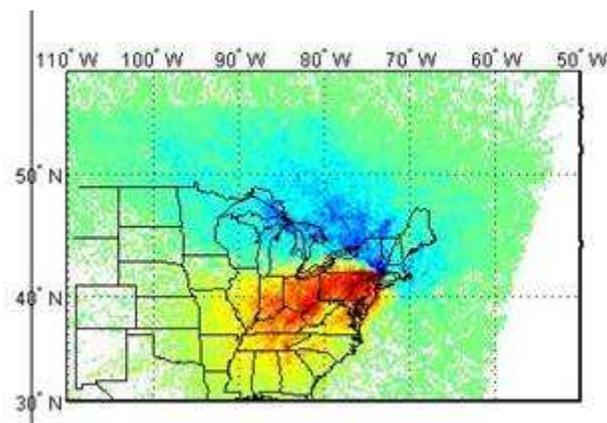Sulfate
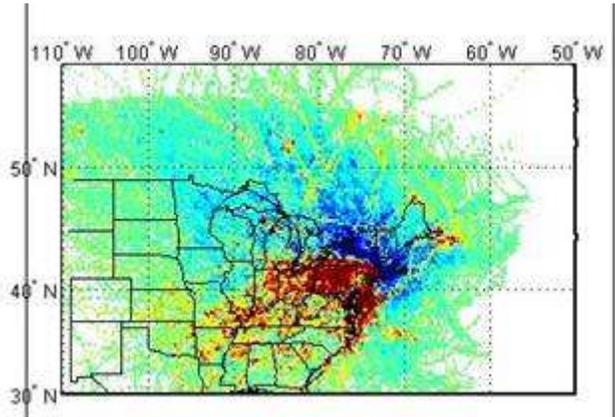
PM

OC

B-ext

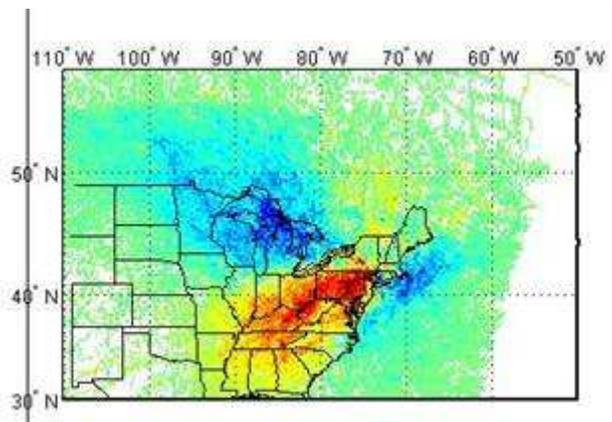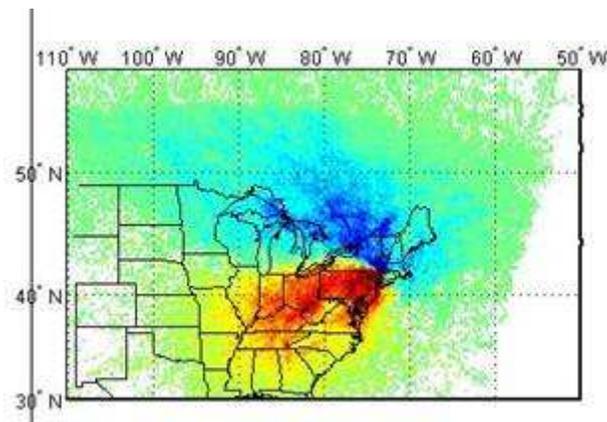# Mohawk Mt All Trajectories 00-04, Cluster Weighted Probability
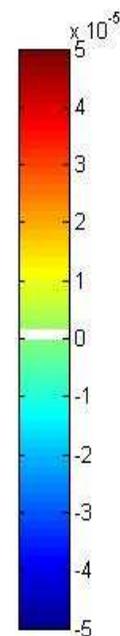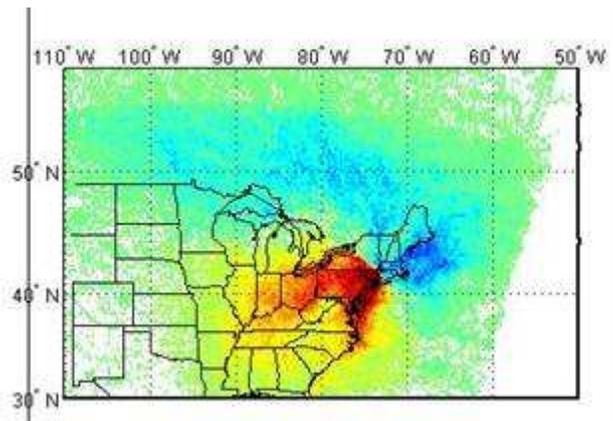## Calculated using Proximity Based Clusters, 500m

Sulfate

PM

OC

B-ext

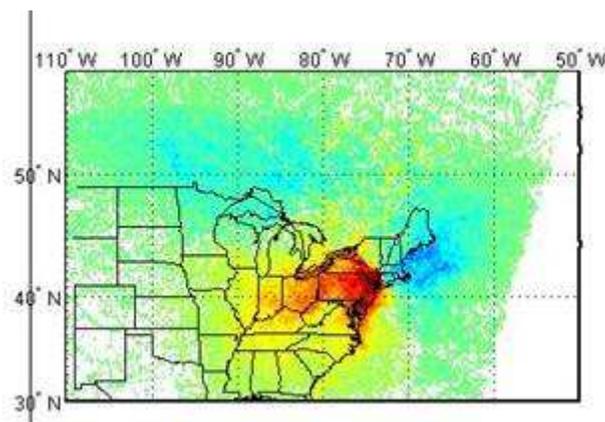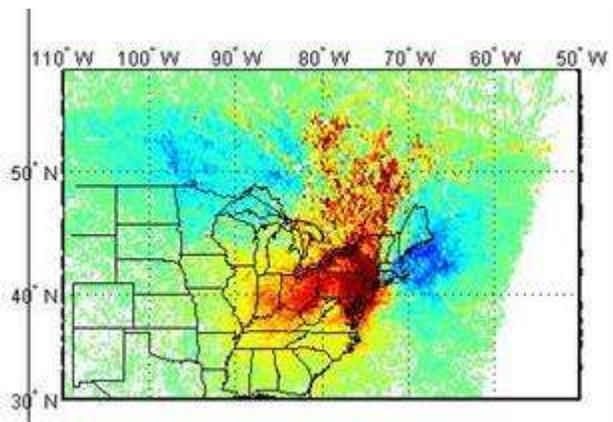# Mohawk Mt All Trajectories 00-04, Cluster Weighted Probability

Calculated using Frequency Based Clusters. 500m

Sulfate

PM

OC

B-ext