Appendix A- Washington D.C. Supplement

Trajectory analysis results at Washington D.C.

Equations for Different Metrics

Everyday Residence-time Probability

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

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

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

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.

Washington All Trajectories 00-04, Top 10 Clusters

Modes defined at: R=12, 120hr BT, 500m Start ht, 7438 Valid Trajectories, 6778 Invalid
Reassigned Trajectories Based on 72hr BT, 500m Start Ht, 11020 Valid Trajectories
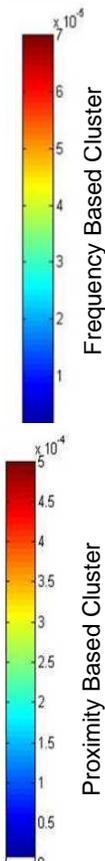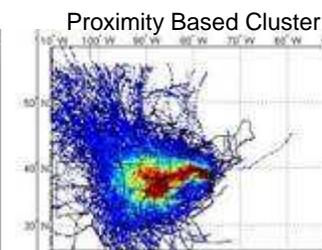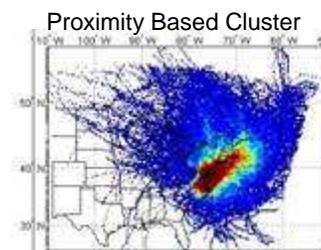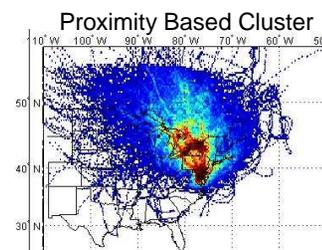
Cluster 1

Central Trajectory
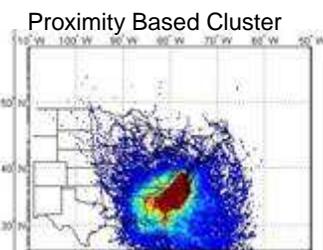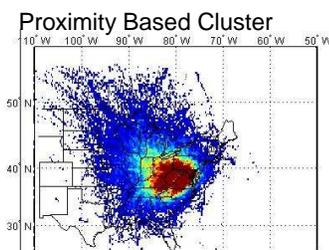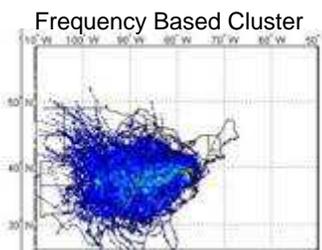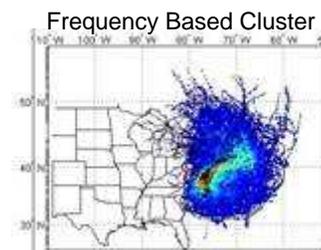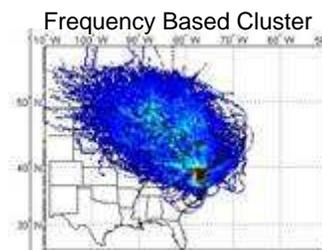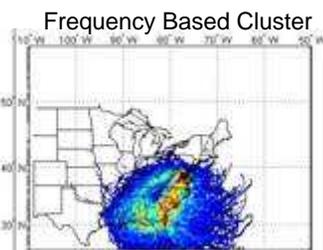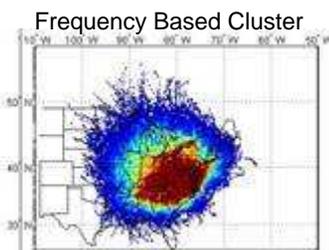
Frequency Based Cluster

Proximity Based Cluster

	Frequency	Proximity
Sulfate	6.75	7.94
Bext	120.54	134.78
PM	17.04	18.88
OC	4.12	4.47
# Trajs	5229	1297
# Trajs w. Poll	1584	374

Cluster 2

Central Trajectory

Frequency Based Cluster

Proximity Based Cluster

	Frequency	Proximity
Sulfate	4.09	5.70
Bext	87.13	105.95
PM	12.11	15.17
OC	3.17	3.82
# Trajs	1760	1344
# Trajs w. Poll	572	365

Cluster 3

Central Trajectory

Frequency Based Cluster

Proximity Based Cluster

	Frequency	Proximity
Sulfate	5.62	4.26
Bext	105.20	87.39
PM	16.08	12.06
OC	4.62	3.24
# Trajs	806	1090
# Trajs w. Poll	275	296

Cluster 4

Central Trajectory

Frequency Based Cluster

Proximity Based Cluster

	Frequency	Proximity
Sulfate	3.26	5.28
Bext	75.57	100.34
PM	10.13	13.76
OC	2.62	3.05
# Trajs	673	1081
# Trajs w. Poll	216	294

Cluster 5

Central Trajectory

Frequency Based Cluster

Proximity Based Cluster

	Frequency	Proximity
Sulfate	5.02	6.16
Bext	93.63	113.55
PM	13.95	16.11
OC	3.46	4.21
# Trajs	545	768
# Trajs w. Poll	172	234

Washington All Trajectories 00-04, Top 10 Clusters

Modes defined at: R=12, 120hr BT, 500m Start ht, 7438 Valid Trajectories, 6778 Invalid
Reassigned Trajectories Based on 72hr BT, 500m Start Ht, 11020 Valid Trajectories
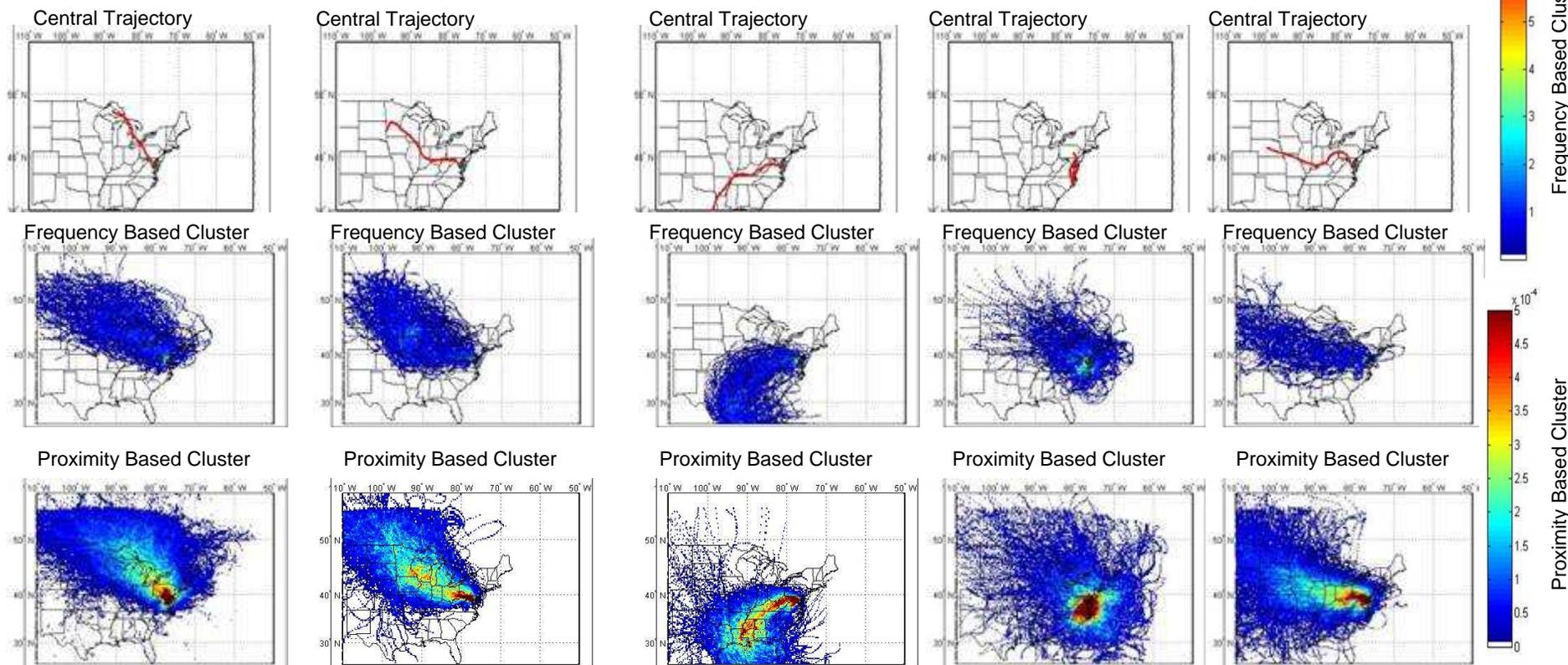

Cluster 6		
	Frequency	Proximity
Sulfate	3.82	4.41
Bext	72.29	87.88
PM	10.13	12.18
OC	2.43	3.18
# Trajs	423	1958
# Trajs w. Poll	106	527

Cluster 7		
	Frequency	Proximity
Sulfate	3.48	3.44
Bext	74.25	73.24
PM	10.39	10.06
OC	2.93	2.69
# Trajs	287	912
# Trajs w. Poll	92	259

Cluster 8		
	Frequency	Proximity
Sulfate	5.85	6.43
Bext	106.34	117.19
PM	16.18	17.74
OC	3.93	4.79
# Trajs	218	655
# Trajs w. Poll	71	214

Cluster 9		
	Frequency	Proximity
Sulfate	3.91	5.74
Bext	68.44	106.14
PM	10.90	15.66
OC	2.97	3.87
# Trajs	142	702
# Trajs w. Poll	36	191

Cluster 10		
	Frequency	Proximity
Sulfate	2.90	5.31
Bext	58.47	103.23
PM	8.60	15.43
OC	2.63	3.96
# Trajs	140	1223
# Trajs w. Poll	48	340

Washington All Trajectories 00-04, Top 10 Clusters

Modes defined at: R=12, 120hr BT, 500m Start ht, 7438 Valid Trajectories, 6778 Invalid
Reassigned Trajectories Based on 72hr BT, 500m Start Ht, 11020 Valid Trajectories
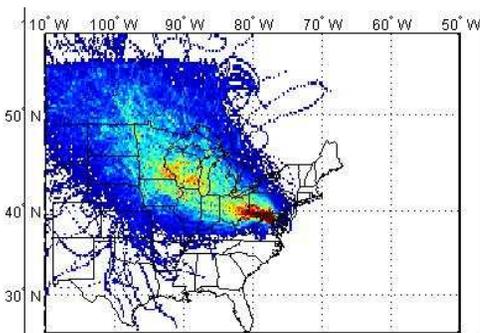Best and Worst Days

Highest Sulfate (Proximate)
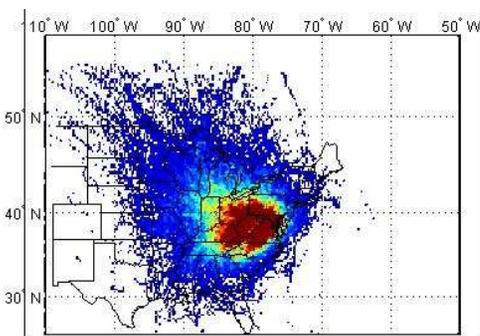
	Frequency	Proximity
Sulfate	6.75	7.94
Bext	120.54	134.78
PM	17.04	18.88
OC	4.12	4.47
# Trajs	5229	1297
# Trajs w. Poll	1584	374
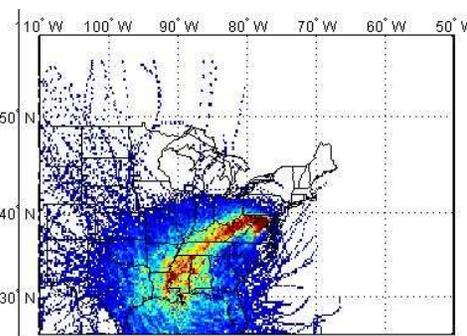
	Frequency	Proximity
Sulfate	5.85	6.43
Bext	106.34	117.19
PM	16.18	17.74
OC	3.93	4.79
# Trajs	218	655
# Trajs w. Poll	71	214

Lowest Sulfate (Proximate)

	Frequency	Proximity
Sulfate	3.48	3.44
Bext	74.25	73.24
PM	10.39	10.06
OC	2.93	2.69
# Trajs	287	912
# Trajs w. Poll	92	259
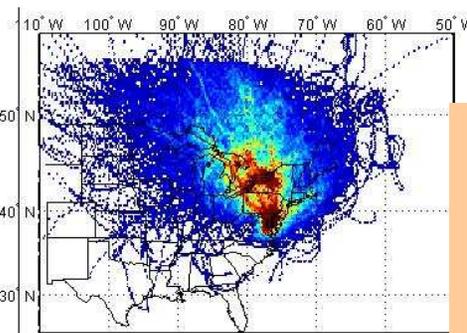
	Frequency	Proximity
Sulfate	5.62	4.26
Bext	105.20	87.39
PM	16.08	12.06
OC	4.62	3.24
# Trajs	806	1090
# Trajs w. Poll	275	296

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).
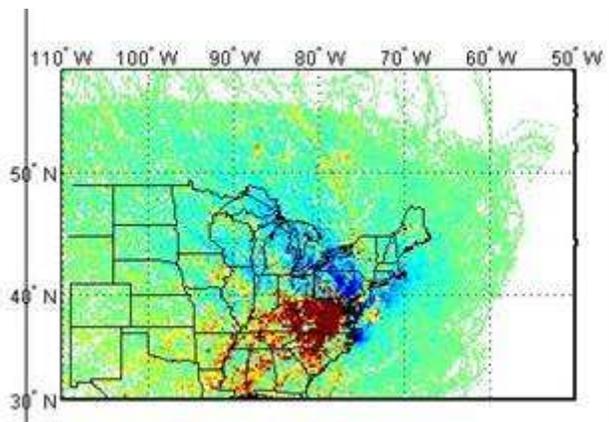
Washington All Trajectories 00-04, Incremental Probability
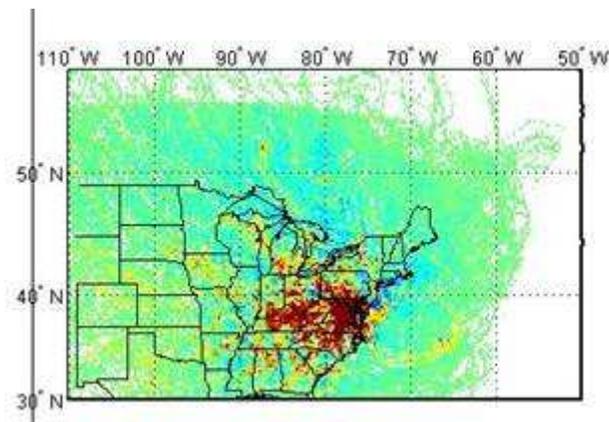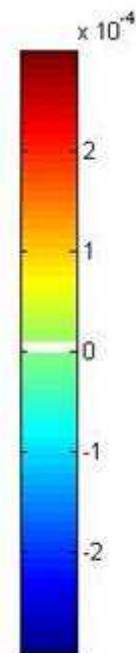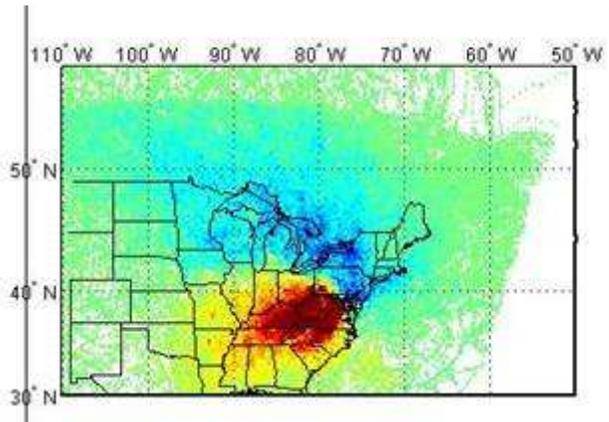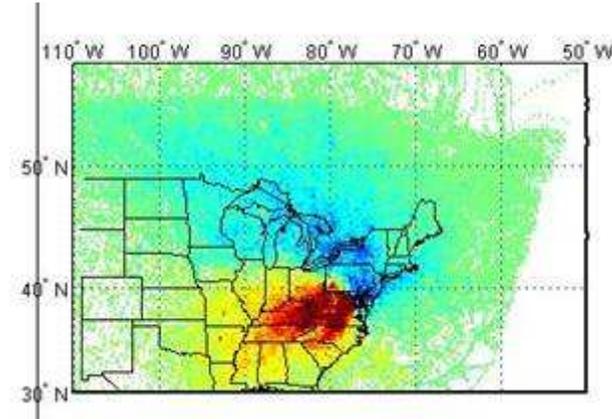## IP Based on Top10%, 500m
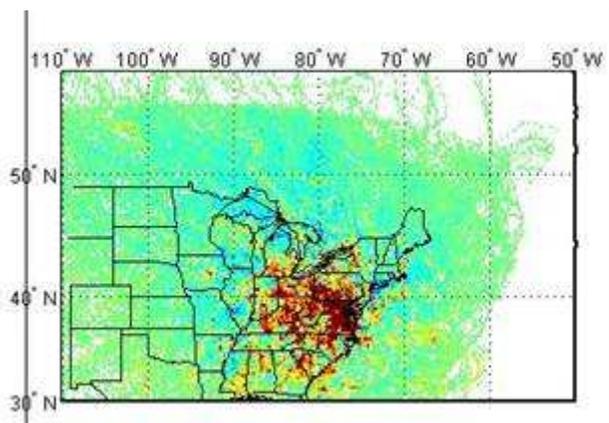
Sulfate
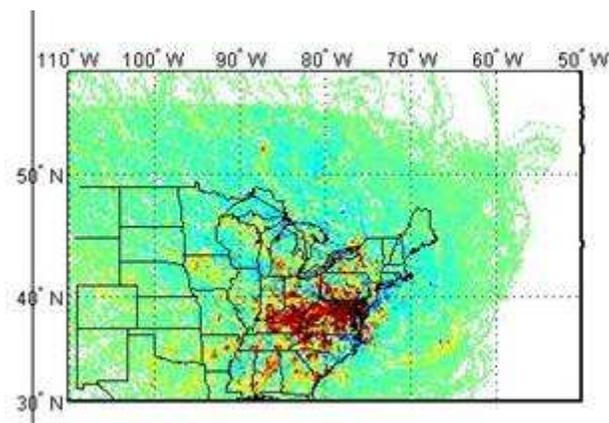
PM

OC

B-ext

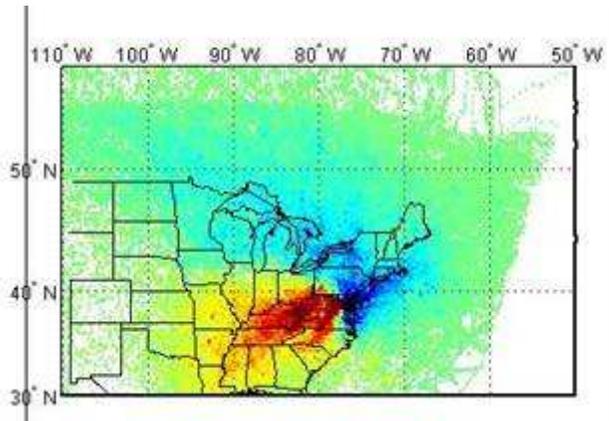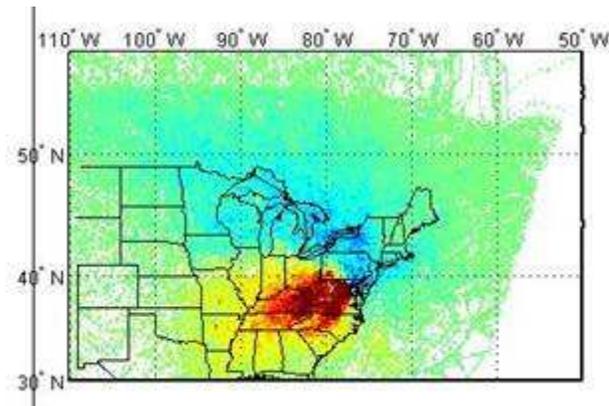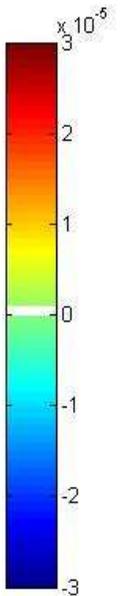# Washington All Trajectories 00-04, Cluster Weighted Probability
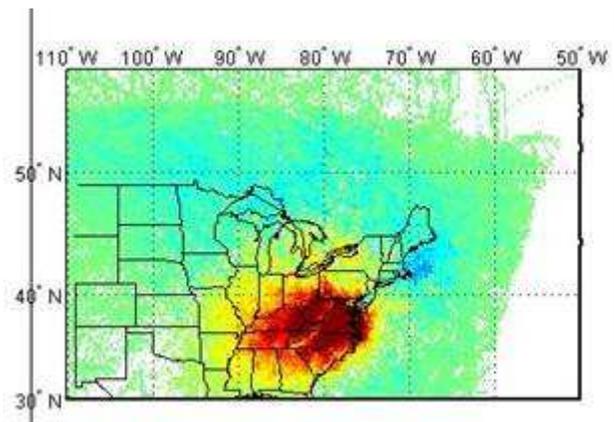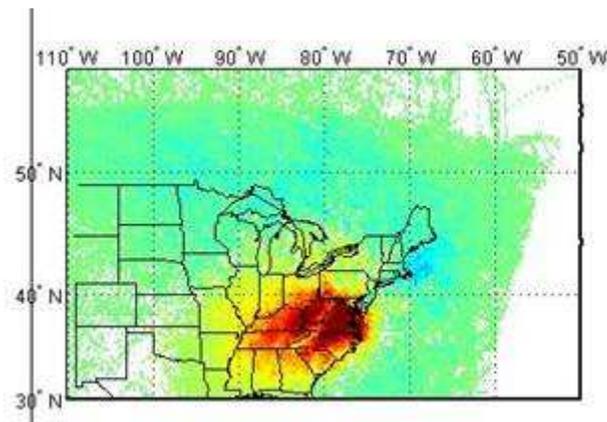## Calculated using Proximity Based Clusters, 500m

Sulfate

PM

OC

B-ext

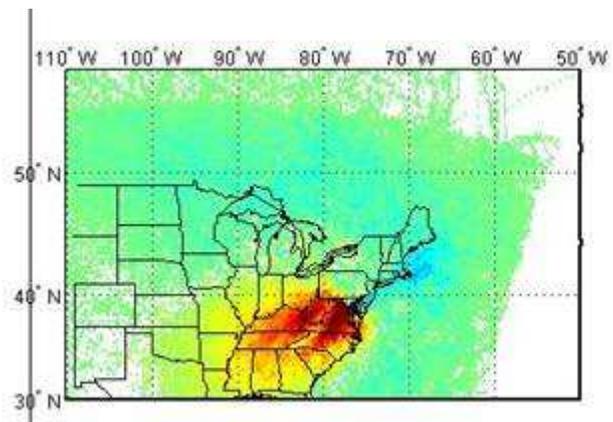# Washington All Trajectories 00-04, Cluster Weighted Probability
Calculated using Frequency Based Clusters, 500m

Sulfate

PM

OC

B-ext