

# Appendix A- Great Smokey Mountains Supplement

Trajectory analysis results at  
Great Smokey Mountains National  
Park.

# Equations for Different Metrics

## Everyday Residence-time Probability

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

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

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

## Incremental Probability

$$IP = HP - EP$$

## High Day Residence-time Probability

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

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

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

## Cluster-Weighted Probability

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

$L$  = total number of clusters calculated

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

$\bar{C}$  = 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.

# GRSM All Trajectories 00-04, Top 10 Clusters

Modes defined at: R=12, 120hr BT, 500m Start ht, 9827 Valid Trajectories, 4552 Invalid  
 Reassigned Trajectories Based on 72hr BT, 500m Start Ht, 12557 Valid Trajectories

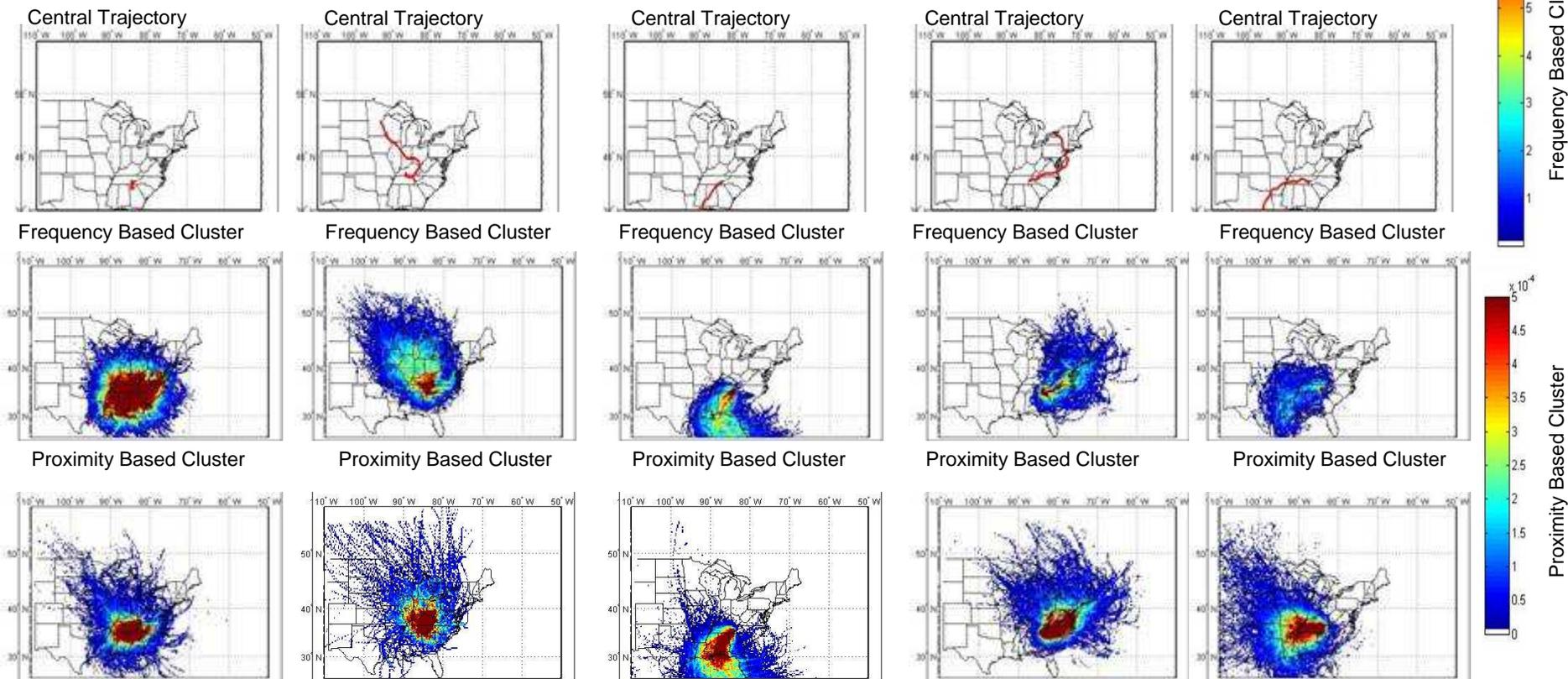
Cluster 1

Cluster 2

Cluster 3

Cluster 4

Cluster 5



	Frequency	Proximity
Sulfate	5.45	6.08
Bext	101.65	112.58
PM	13.62	15.08
OC	3.23	3.38
# Trajs	7141	1648
# Trajs w. Poll	2099	494

	Frequency	Proximity
Sulfate	4.47	6.37
Bext	89.30	113.89
PM	10.78	14.14
OC	2.58	2.96
# Trajs	1330	392
# Trajs w. Poll	396	130

	Frequency	Proximity
Sulfate	4.26	4.17
Bext	77.11	78.79
PM	10.85	10.96
OC	2.44	2.38
# Trajs	787	1515
# Trajs w. Poll	251	476

	Frequency	Proximity
Sulfate	4.23	4.95
Bext	75.88	93.00
PM	10.39	12.19
OC	2.33	2.70
# Trajs	483	1032
# Trajs w. Poll	141	260

	Frequency	Proximity
Sulfate	3.25	4.28
Bext	70.48	91.09
PM	8.00	11.78
OC	2.59	3.22
# Trajs	427	1239
# Trajs w. Poll	97	367

# GRSM All Trajectories 00-04, Top 10 Clusters

Modes defined at: R=12, 120hr BT, 500m Start ht, 9827 Valid Trajectories, 4552 Invalid  
 Reassigned Trajectories Based on 72hr BT, 500m Start Ht, 12557 Valid Trajectories

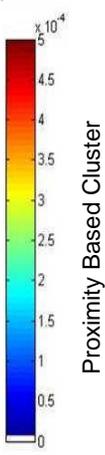
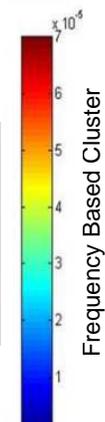
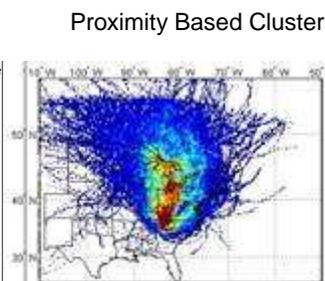
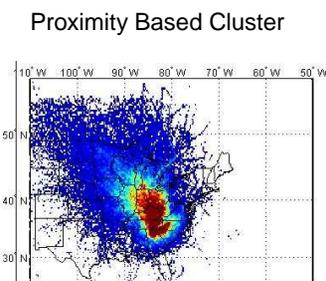
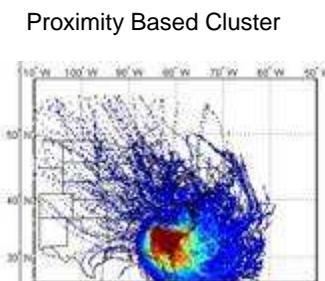
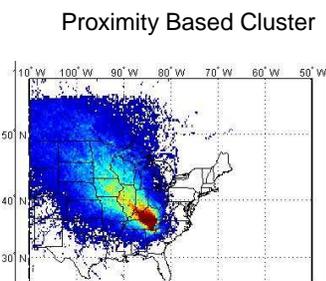
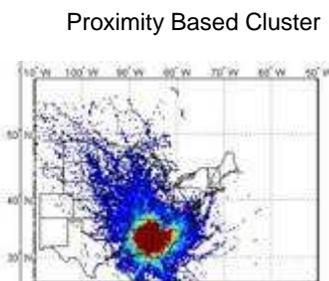
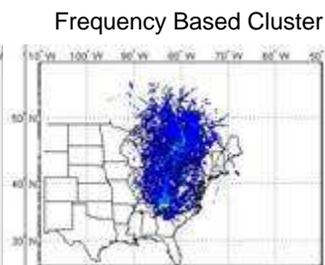
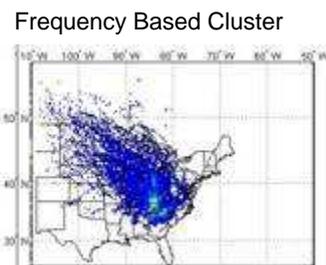
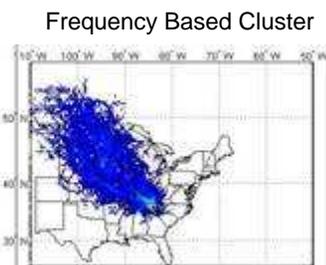
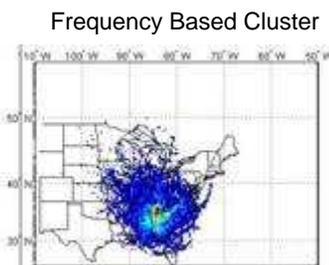
## Cluster 6

## Cluster 7

## Cluster 8

## Cluster 9

## Cluster 10



	Frequency	Proximity
Sulfate	3.05	4.71
Bext	71.00	89.15
PM	8.74	12.70
OC	2.41	3.25
# Trajs	396	1168
# Trajs w. Poll	122	336

	Frequency	Proximity
Sulfate	1.94	2.70
Bext	50.52	62.65
PM	6.02	7.48
OC	1.56	2.14
# Trajs	292	2103
# Trajs w. Poll	97	648

	Frequency	Proximity
Sulfate	6.52	4.52
Bext	116.58	80.59
PM	15.83	11.60
OC	3.06	2.85
# Trajs	221	944
# Trajs w. Poll	61	259

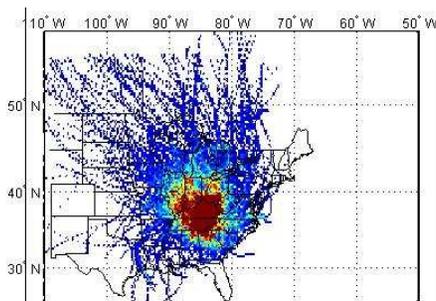
	Frequency	Proximity
Sulfate	1.78	6.17
Bext	54.67	114.09
PM	6.13	14.21
OC	1.99	3.08
# Trajs	144	1448
# Trajs w. Poll	42	362

	Frequency	Proximity
Sulfate	4.00	4.55
Bext	70.34	81.73
PM	9.69	10.64
OC	2.21	2.21
# Trajs	130	841
# Trajs w. Poll	27	267

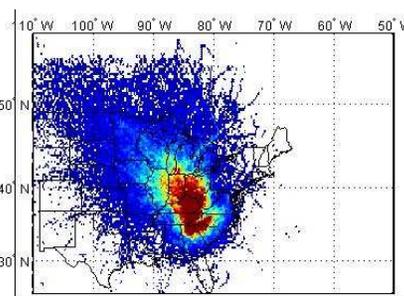
# GRSM All Trajectories 00-04, Best/Worst Sulfate

Modes defined at: R=12, 120hr BT, 500m Start ht, 9827 Valid Trajectories, 4552 Invalid Reassigned Trajectories Based on 72hr BT, 500m Start Ht, 12557 Valid Trajectories

Highest Sulfate  
(Proximity)

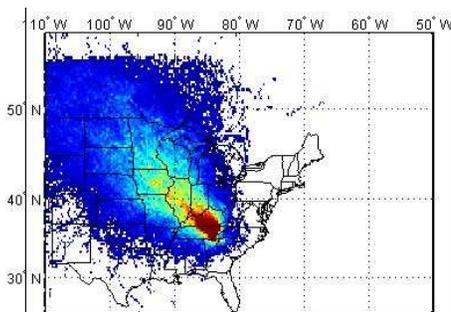


	Frequency	Proximity
Sulfate	4.47	6.37
Bext	89.30	113.89
PM	10.78	14.14
OC	2.58	2.96
# Trajs	1330	392
# Trajs w. Poll	396	130

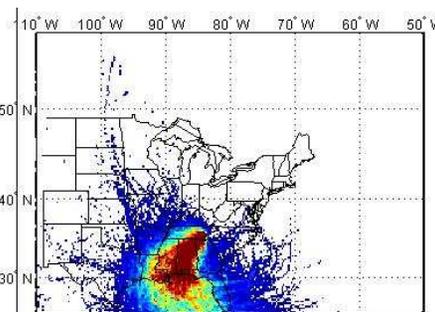


	Frequency	Proximity
Sulfate	1.78	6.17
Bext	54.67	114.09
PM	6.13	14.21
OC	1.99	3.08
# Trajs	144	1448
# Trajs w. Poll	42	362

Lowest Sulfate  
(Proximity)



	Frequency	Proximity
Sulfate	1.94	2.70
Bext	50.52	62.65
PM	6.02	7.48
OC	1.56	2.14
# Trajs	292	2103
# Trajs w. Poll	97	648

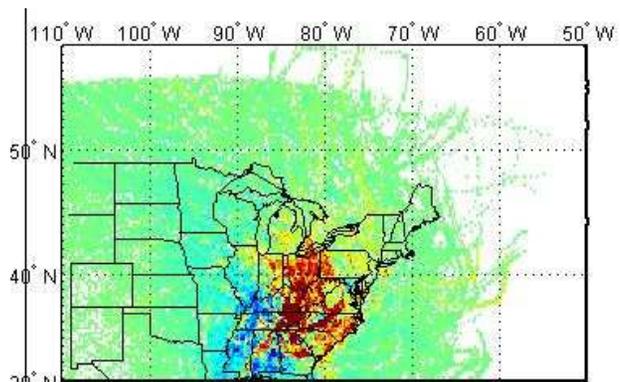


	Frequency	Proximity
Sulfate	4.26	4.17
Bext	77.11	78.79
PM	10.85	10.96
OC	2.44	2.38
# Trajs	787	1515
# Trajs w. Poll	251	476

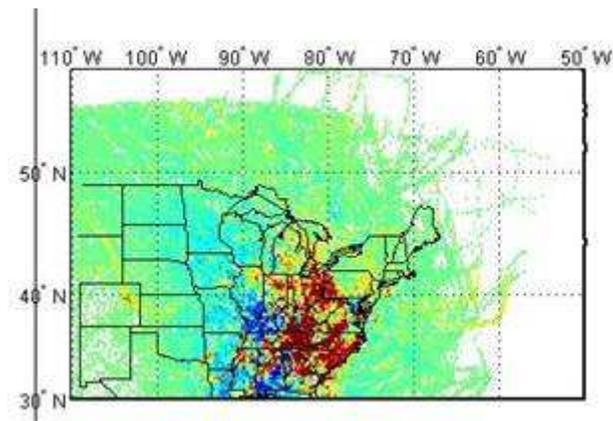
*Sulfate*- Sulfate ion Conc. (ug/m3)  
*Bext*- Extinction (Mm-1)  
*PM*- Particulate Matter Conc. (ug/m3)  
*OC*- Organic Carbon Conc. (ug/m3)  
*Num Trajs*- Number of trajectories in cluster  
*Num Trajs w. Poll*- Number of trajectories in cluster with associated pollution measurement (Based on number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

# GRSM All Trajectories 00-04, Incremental Probability

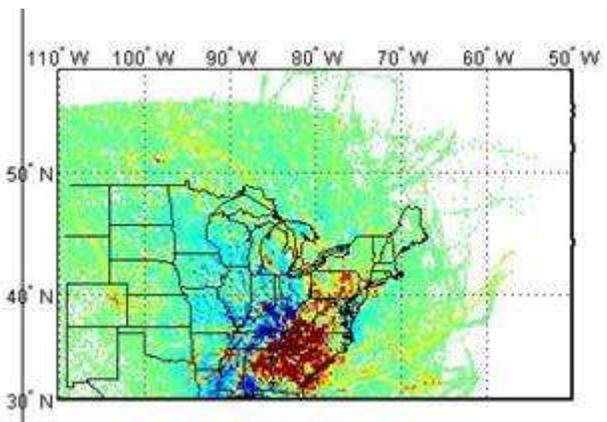
IP Based on Top10%, 500m



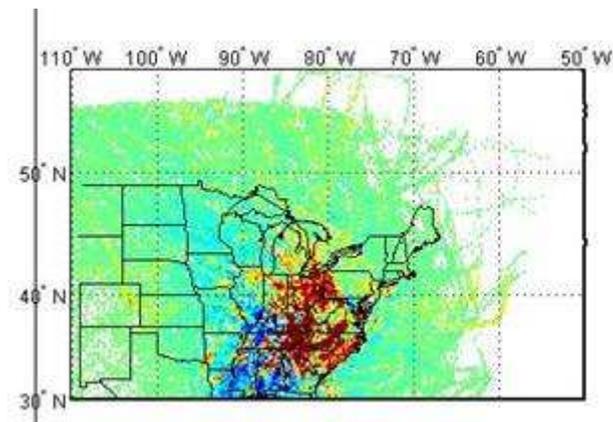
Sulfate



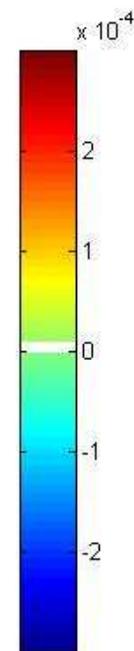
PM



OC

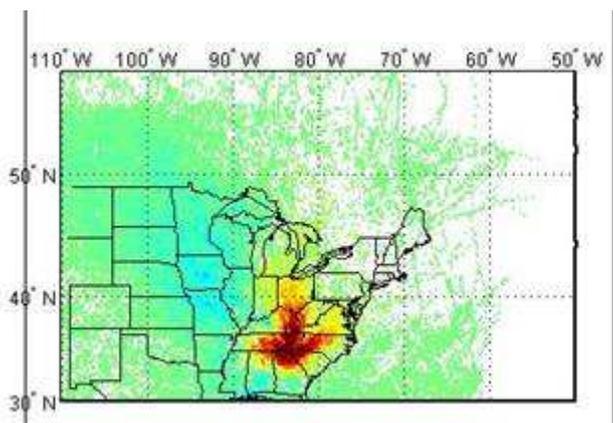


B-ext

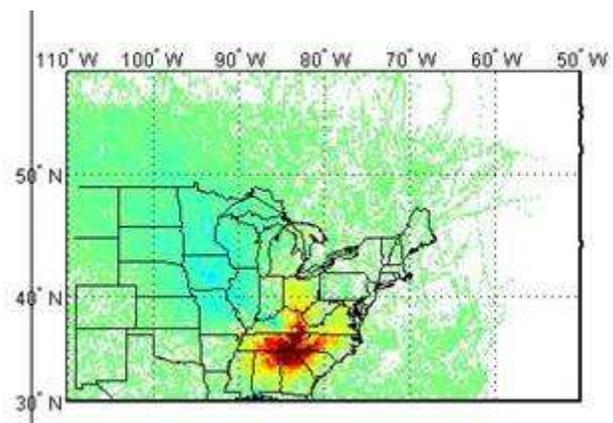


# GRSM All Trajectories 00-04, Cluster Weighted Probability

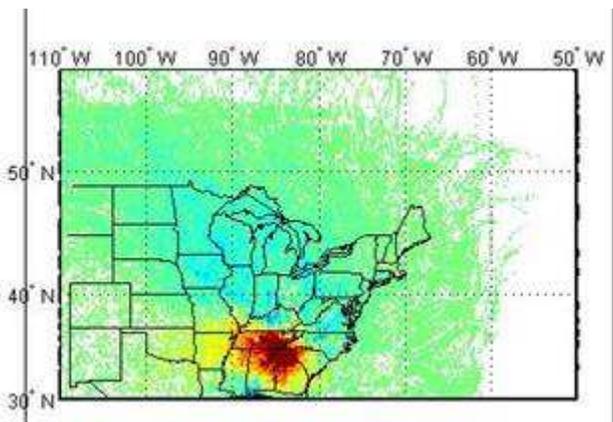
CWP calculated using Proximity Based Clusters, 500m



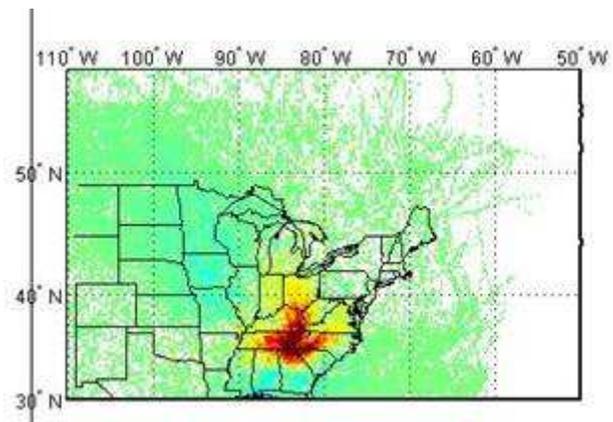
Sulfate



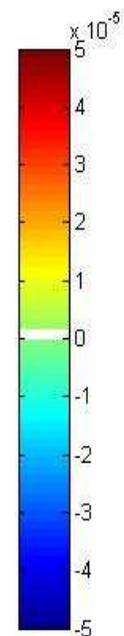
PM



OC

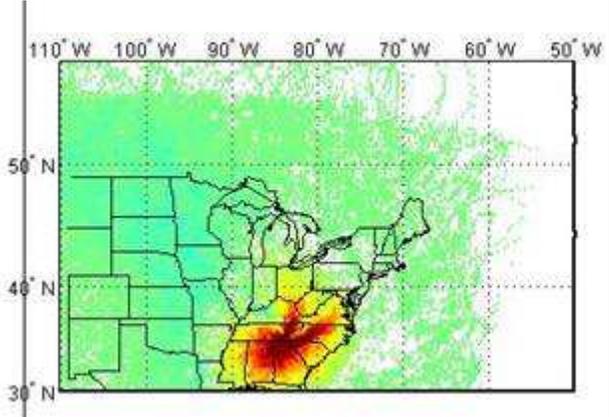


B-ext

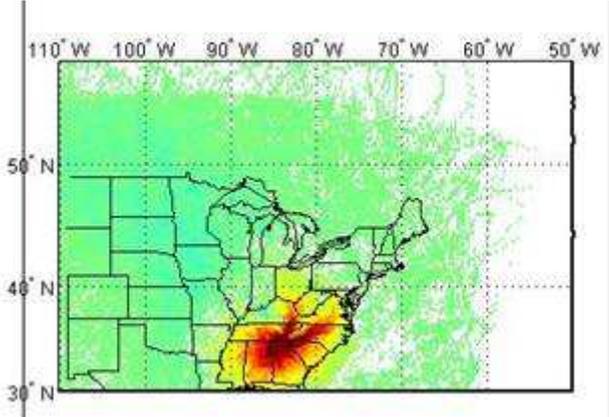


# GRSM All Trajectories 00-04, Cluster Weighted Probability

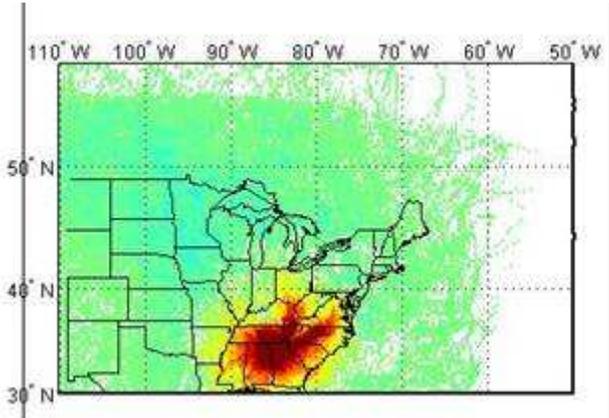
Calculated using Frequency Based Clusters, 500m



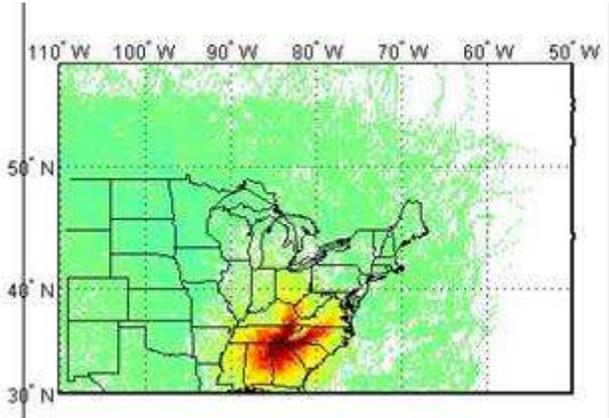
Sulfate



PM



OC



B-ext

