

# Appendix A- New York City Supplement

Trajectory analysis results at  
Queens College, New York City.

# 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.

# NYC All Trajectories 00-04, Top 10 Clusters

Modes defined at: R= 12, 120hr BT, 500m start height, 6488 valid trajectories, 7698 invalid  
 Reassign Trajectories Based on 72hr BT, 500m start height, 10116 Valid Trajectories

Cluster 1

Cluster 2

Cluster 3

Cluster 4

Cluster 5

Central Trajectory



Central Trajectory



Central Trajectory



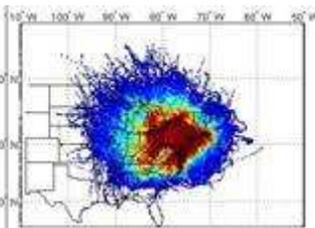
Central Trajectory



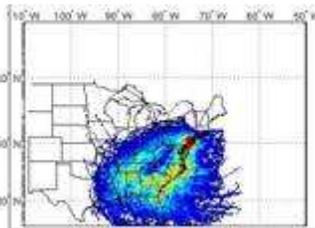
Central Trajectory



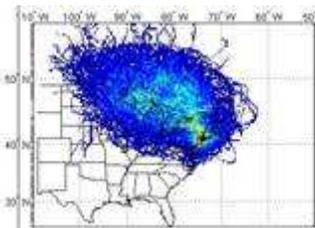
Frequency Based Cluster



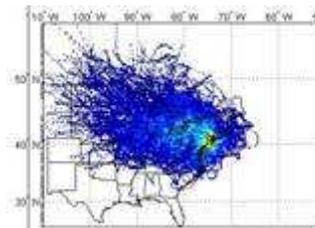
Frequency Based Cluster



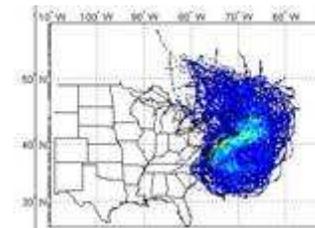
Frequency Based Cluster



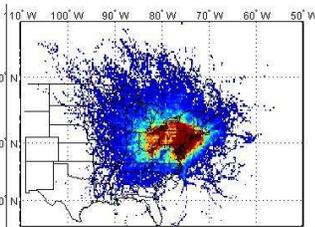
Frequency Based Cluster



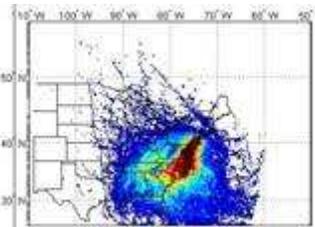
Frequency Based Cluster



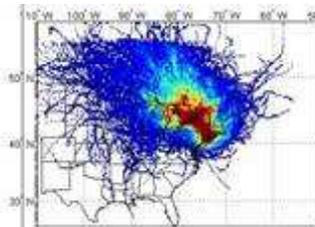
Proximity Based Cluster



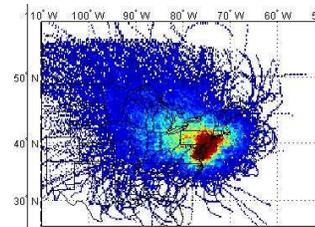
Proximity Based Cluster



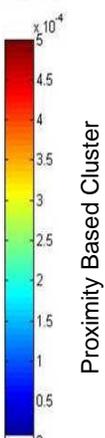
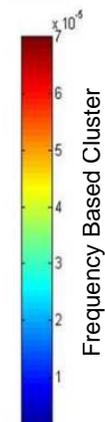
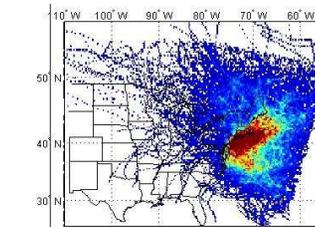
Proximity Based Cluster



Proximity Based Cluster



Proximity Based Cluster



	Frequency	Proximity
Sulfate	5.87	6.44
Bext		
PM	18.30	18.86
OC	3.02	3.14
# Trajs	4617	1180
# Trajs w. Poll	1105	284

	Frequency	Proximity
Sulfate	3.42	5.86
Bext		
PM	13.10	17.27
OC	2.30	2.44
# Trajs	1638	1134
# Trajs w. Poll	339	276

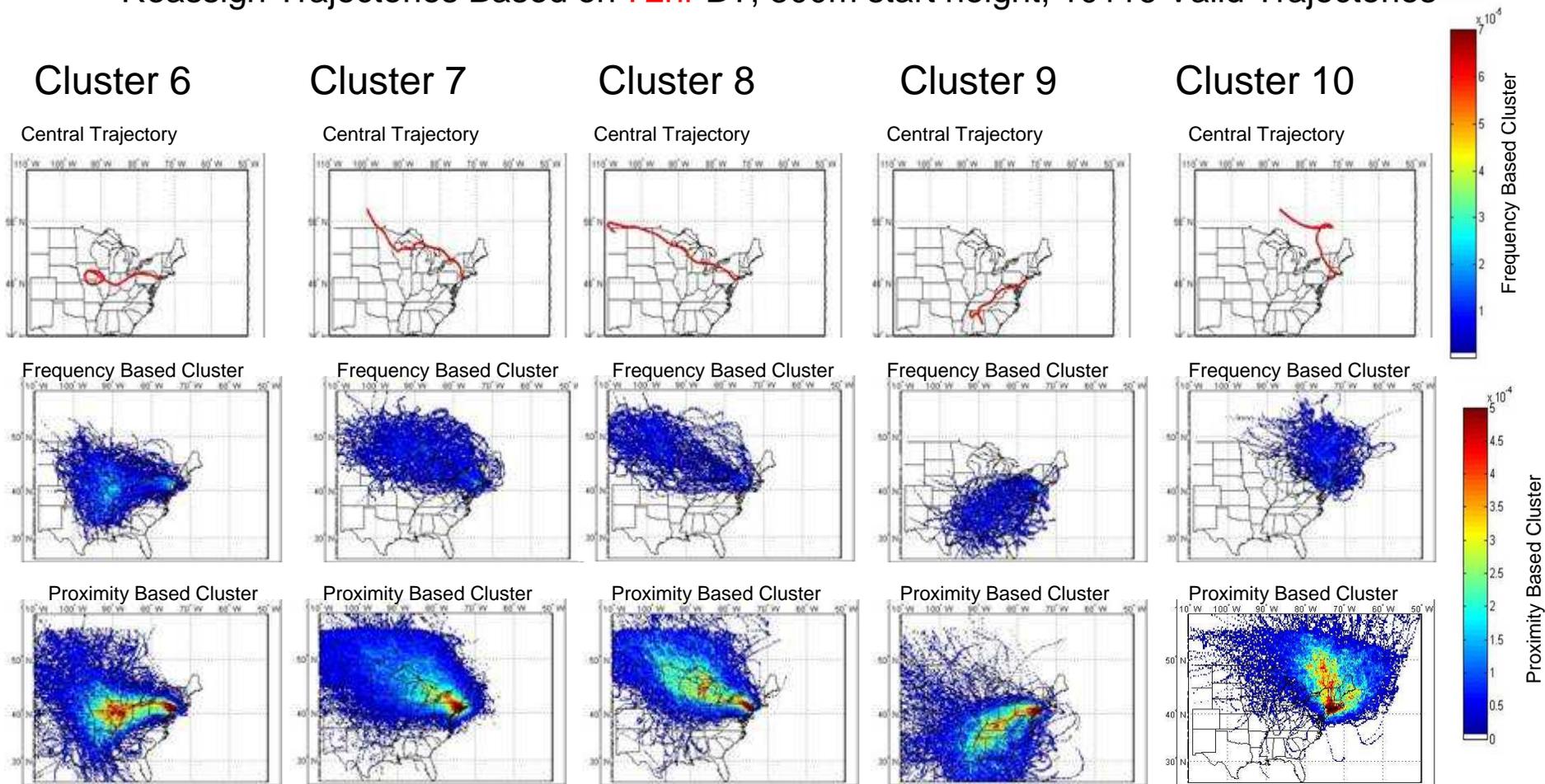
	Frequency	Proximity
Sulfate	4.41	3.32
Bext		
PM	12.98	13.48
OC	1.93	3.10
# Trajs	851	976
# Trajs w. Poll	182	226

	Frequency	Proximity
Sulfate	2.26	5.63
Bext		
PM	9.15	18.32
OC	2.42	3.07
# Trajs	687	927
# Trajs w. Poll	146	230

	Frequency	Proximity
Sulfate	6.42	3.48
Bext		
PM	19.36	10.73
OC	3.60	1.97
# Trajs	438	775
# Trajs w. Poll	85	190

# NYC All Trajectories 00-04, Top 10 Clusters

Modes defined at: R= 12, 120hr BT, 500m start height, 6488 valid trajectories, 7698 invalid  
 Reassign Trajectories Based on 72hr BT, 500m start height, 10116 Valid Trajectories



	Frequency	Proximity
Sulfate	4.53	6.05
Bext		
PM	15.47	19.79
OC	2.89	3.22
# Trajs	317	770
# Trajs w. Poll	72	167

	Frequency	Proximity
Sulfate	4.97	3.95
Bext		
PM	18.04	14.50
OC	2.56	2.33
# Trajs	305	2119
# Trajs w. Poll	69	470

	Frequency	Proximity
Sulfate	2.75	3.46
Bext		
PM	12.05	12.63
OC	2.73	2.39
# Trajs	278	775
# Trajs w. Poll	49	150

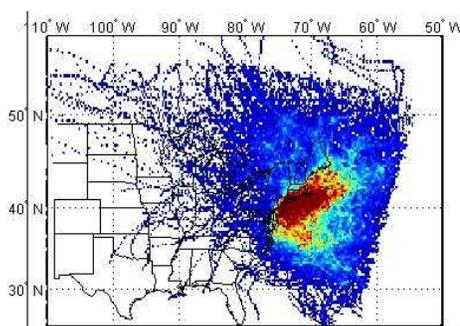
	Frequency	Proximity
Sulfate	3.01	7.20
Bext		
PM	9.26	20.90
OC	1.53	3.26
# Trajs	150	658
# Trajs w. Poll	37	123

	Frequency	Proximity
Sulfate	2.44	2.67
Bext		
PM	9.15	10.10
OC	1.21	2.21
# Trajs	110	812
# Trajs w. Poll	17	166

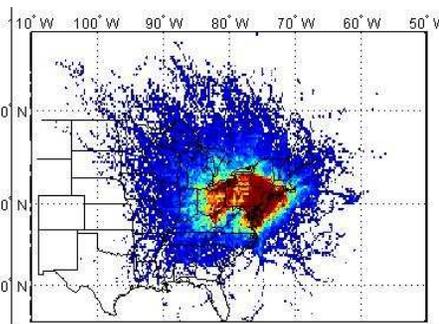
# NYC All Trajectories 00-04, Top 10 Clusters

Modes defined at: R= 12, 120hr BT, 500m start height, 6488 valid trajectories, 7698 invalid  
 Reassign Trajectories Based on 72hr BT, 500m start height, 10116 Valid Trajectories  
 Best and Worst Days

Highest Sulfate  
(Proximity)

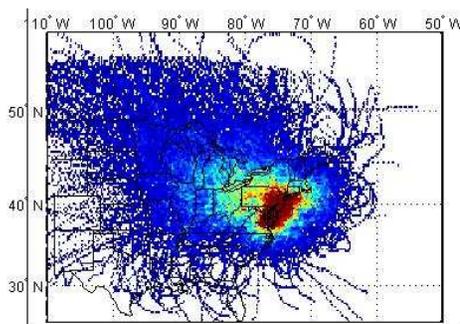


	Frequency	Proximity
Sulfate	3.48	6.42
Bext		
PM	10.73	19.36
OC	1.97	3.60
# Trajs	775	438
# Trajs w. Poll	190	85

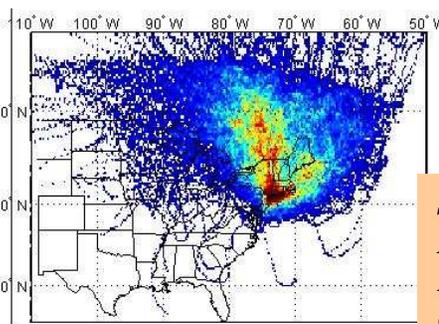


	Frequency	Proximity
Sulfate	6.44	5.87
Bext		
PM	18.86	18.30
OC	3.14	3.02
# Trajs	1180	4617
# Trajs w. Poll	284	1105

Lowest Sulfate  
(Proximity)



	Frequency	Proximity
Sulfate	5.63	2.26
Bext		
PM	18.32	9.15
OC	3.07	2.42
# Trajs	927	687
# Trajs w. Poll	230	146

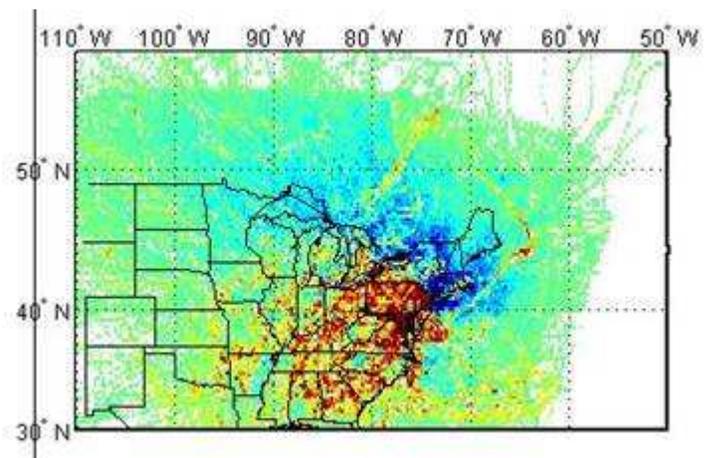


	Frequency	Proximity
Sulfate	2.67	2.44
Bext		
PM	10.10	9.15
OC	2.21	1.21
# Trajs	812	110
# Trajs w. Poll	166	17

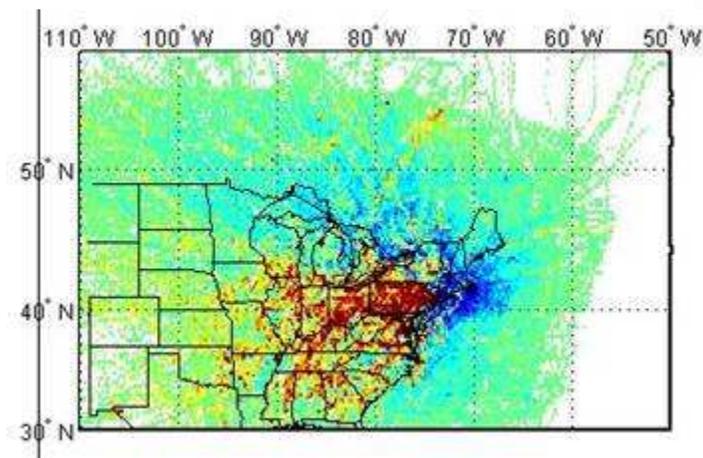
*Sulfate*- Sulfate ion Conc. (ug/m3)  
*Bext*- Extinction (Mm-1)  
*PM*- Particulate Matter Conc. (ug/m3)  
*OC*- Organic Carbon Conc. (ug/m3)  
*Num Trajs*- Number of trajectories in cluster  
*Num Trajs w. Poll*- Number of trajectories in cluster with associated pollution measurement (Based on number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

# NYC All Trajectories 00-04, Incremental Probability

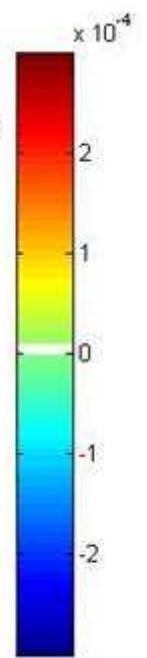
IP Based on Top10%, 500m



Sulfate

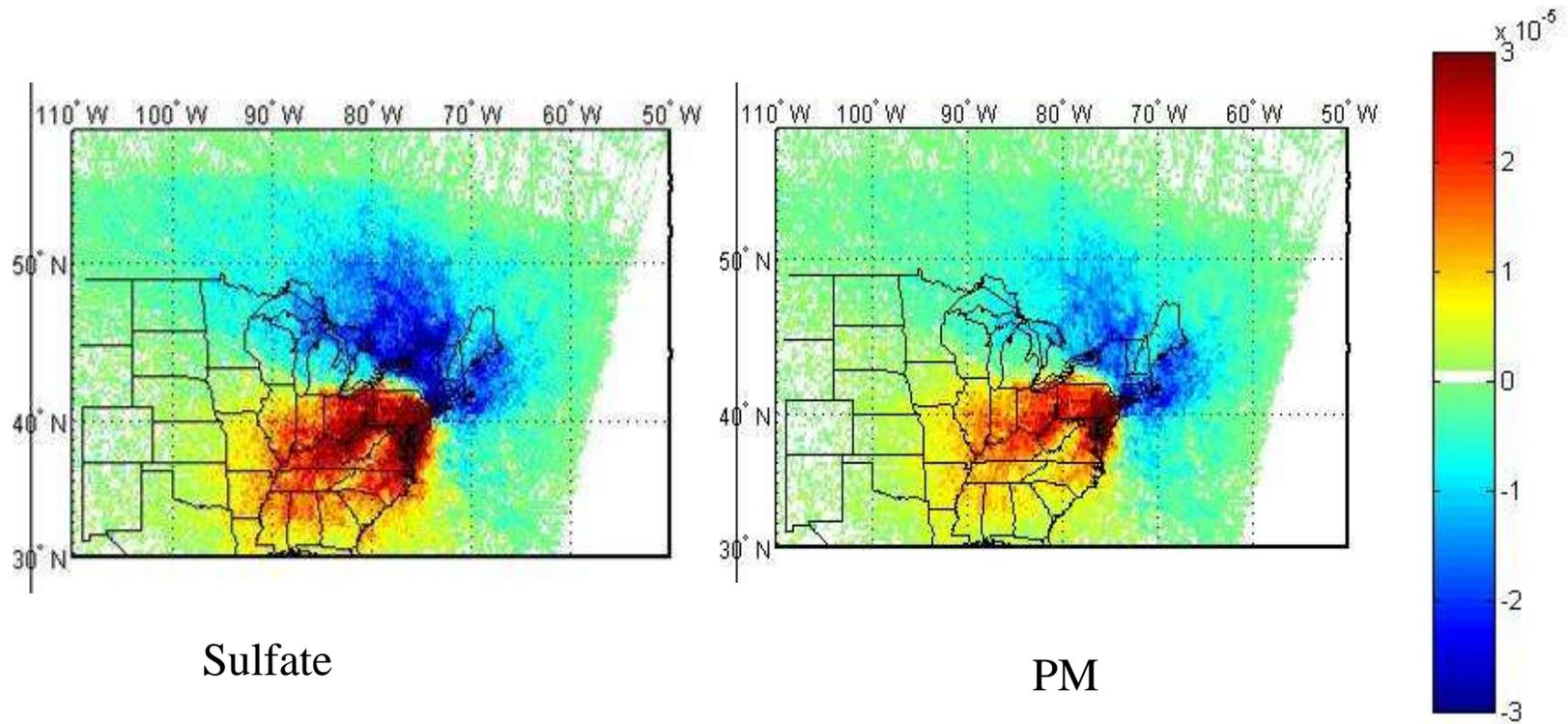


PM



# NYC All Trajectories 00-04, Cluster Weighted Probability

CWP calculated using Proximity Based Clusters, 500m



# NYC All Trajectories 00-04, Cluster Weighted Probability

Calculated using Frequency Based Clusters, 500m

