Appendix A- Pittsburgh Supplement

Trajectory analysis results at Pittsburgh, PA.

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

Incremental Probability

$$IP = HP - EP$$

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

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

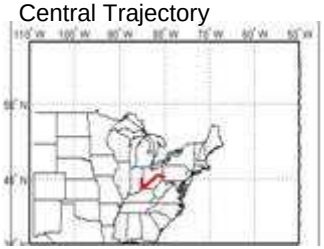
- Central Trajectory (CT)- Trajectory with the largest number of nearest neighbors in the dataset.
- Frequency Based Clusters- These clusters are formed by finding the “central” trajectory which has the greatest number of neighboring trajectories within a subjectively selected radius of proximity (R). 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.

Pittsburgh All Trajectories 00-04, Top 10 Clusters

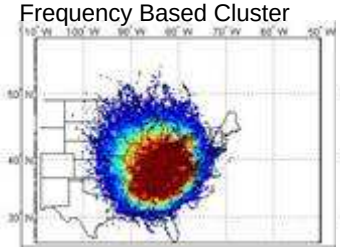
Modes defined at: R=12, 120hr BT, 500m Start ht, 7852 Valid Trajectories, 6394 Invalid
Reassigned Trajectories Based on 72hr BT, 500m Start Ht, 11330 Valid Trajectories

Cluster 1

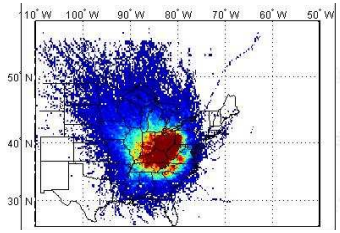
Central Trajectory



Frequency Based Cluster



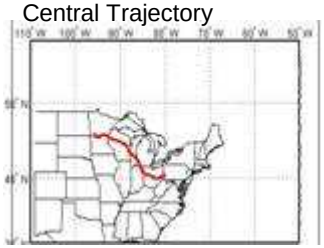
Proximity Based Cluster



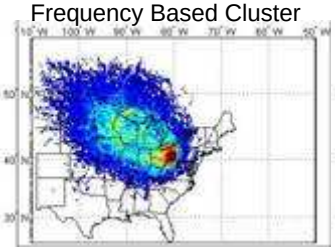
	Frequency	Proximity
Sulfate	7.64	8.12
PM	20.81	22.32
# Trajs	5383	1887
# Trajs w. Poll	1187	418

Cluster 2

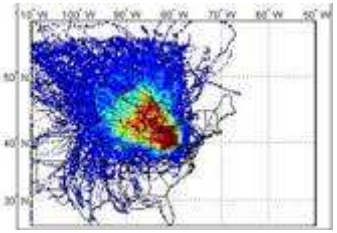
Central Trajectory



Frequency Based Cluster



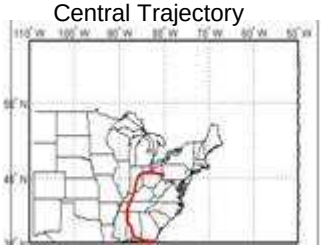
Proximity Based Cluster



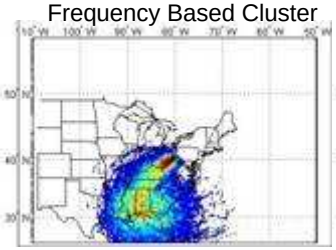
	Frequency	Proximity
Sulfate	5.19	6.02
PM	16.17	18.23
# Trajs	1506	761
# Trajs w. Poll	331	171

Cluster 3

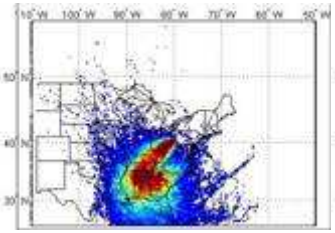
Central Trajectory



Frequency Based Cluster



Proximity Based Cluster

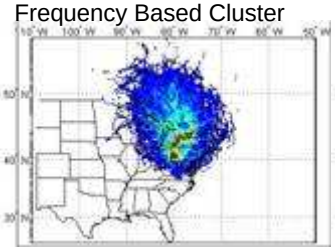


	Frequency	Proximity
Sulfate	5.71	6.18
PM	16.88	17.65
# Trajs	1048	1275
# Trajs w. Poll	228	237

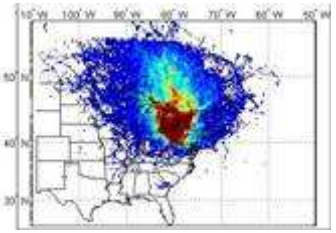
Cluster 4

Central Trajectory

Frequency Based Cluster



Proximity Based Cluster

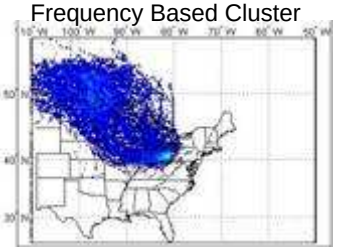


	Frequency	Proximity
Sulfate	4.25	4.67
PM	13.06	14.49
# Trajs	661	1263
# Trajs w. Poll	152	307

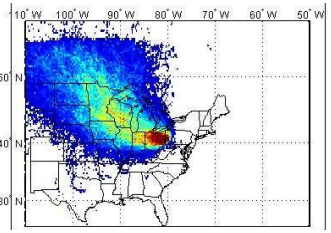
Cluster 5

Central Trajectory

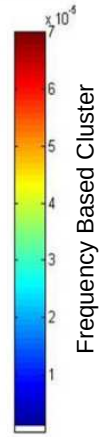
Frequency Based Cluster

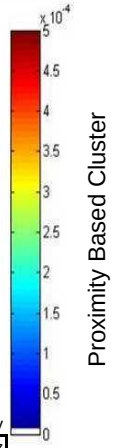


Proximity Based Cluster



	Frequency	Proximity
Sulfate	6.80	3.47
PM	19.56	12.82
# Trajs	560	1538
# Trajs w. Poll	107	324





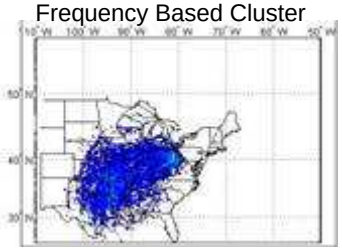
Pittsburgh All Trajectories 00-04, Top 10 Clusters

Modes defined at: R=12, 120hr BT, 500m Start ht, 7852 Valid Trajectories, 6394 Invalid
Reassigned Trajectories Based on 72hr BT, 500m Start Ht, 11330 Valid Trajectories

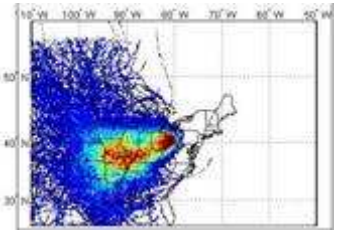
Cluster 6

Central Trajectory

Frequency Based Cluster



Proximity Based Cluster

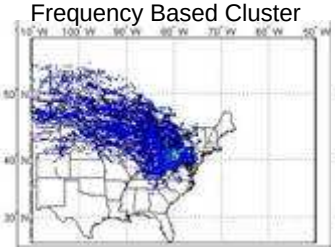


	Frequency	Proximity
Sulfate	2.89	5.95
PM	11.30	17.74
# Trajs	386	792
# Trajs w. Poll	79	212

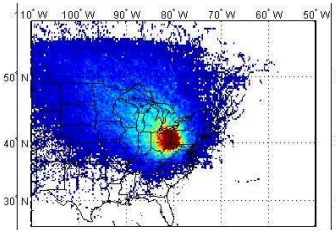
Cluster 7

Central Trajectory

Frequency Based Cluster



Proximity Based Cluster

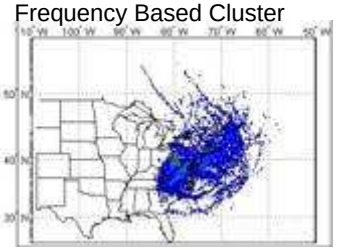


	Frequency	Proximity
Sulfate	10.53	7.23
PM	26.70	20.13
# Trajs	301	2168
# Trajs w. Poll	91	482

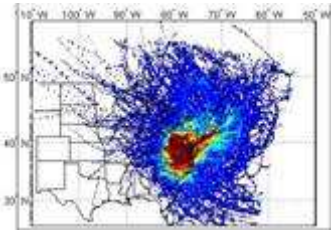
Cluster 8

Central Trajectory

Frequency Based Cluster



Proximity Based Cluster

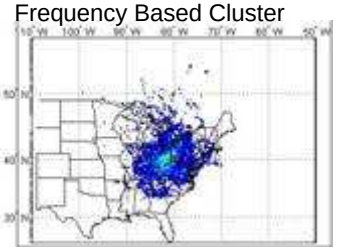


	Frequency	Proximity
Sulfate	4.24	5.94
PM	14.82	16.56
# Trajs	199	631
# Trajs w. Poll	51	153

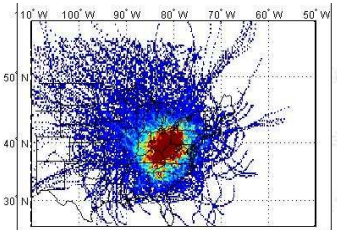
Cluster 9

Central Trajectory

Frequency Based Cluster



Proximity Based Cluster

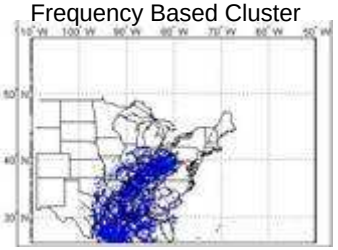


	Frequency	Proximity
Sulfate	2.78	3.28
PM	10.12	12.83
# Trajs	186	569
# Trajs w. Poll	39	160

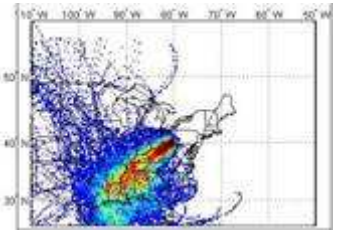
Cluster 10

Central Trajectory

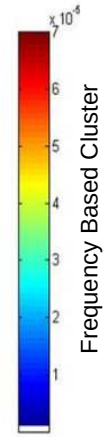
Frequency Based Cluster

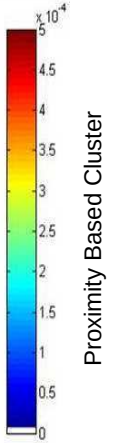


Proximity Based Cluster



	Frequency	Proximity
Sulfate	5.08	6.13
PM	13.91	19.03
# Trajs	179	456
# Trajs w. Poll	42	127

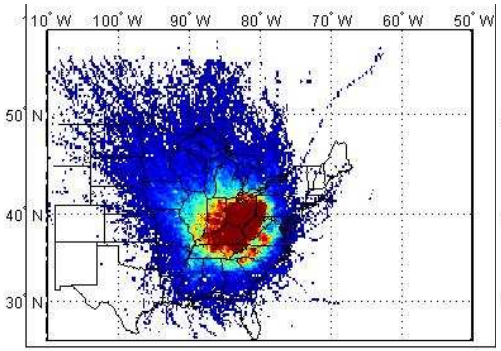


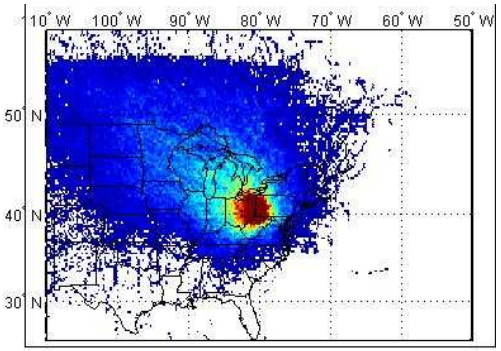


Pittsburgh All Trajectories 00-04, Top 10 Clusters

Modes defined at: R=12, 120hr BT, 500m Start ht, 7852 Valid Trajectories, 6394 Invalid
Reassigned Trajectories Based on 72hr BT, 500m Start Ht, 11330 Valid Trajectories

Highest Sulfate (Proximity)

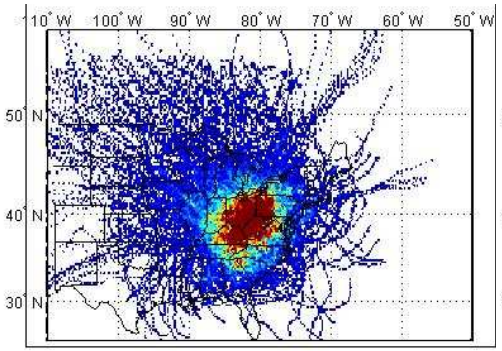


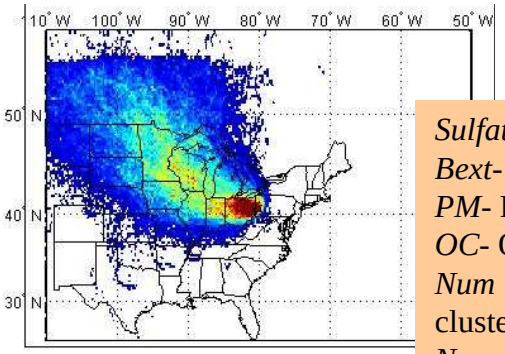


	Frequency	Proximity
Sulfate	7.64	8.12
PM	20.81	22.32
# Trajs	5383	1887
# Trajs w. Poll	1187	418

	Frequency	Proximity
Sulfate	10.53	7.23
PM	26.70	20.13
# Trajs	301	2168
# Trajs w. Poll	91	482

Lowest Sulfate (Proximity)





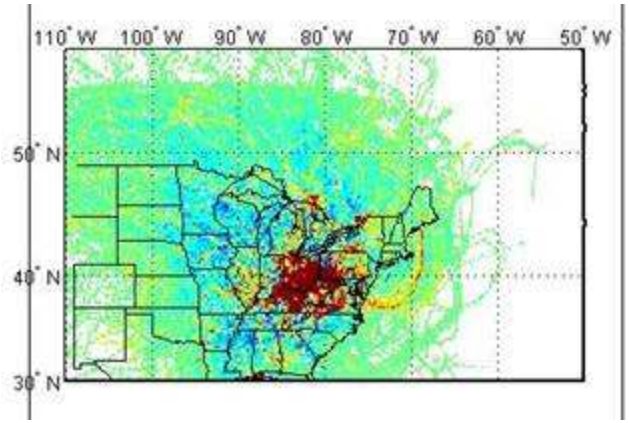
	Frequency	Proximity
Sulfate	2.78	3.28
PM	10.12	12.83
# Trajs	186	569
# Trajs w. Poll	39	160

	Frequency	Proximity
Sulfate	6.80	3.47
PM	19.56	12.82
# Trajs	560	1538
# Trajs w. Poll	107	324

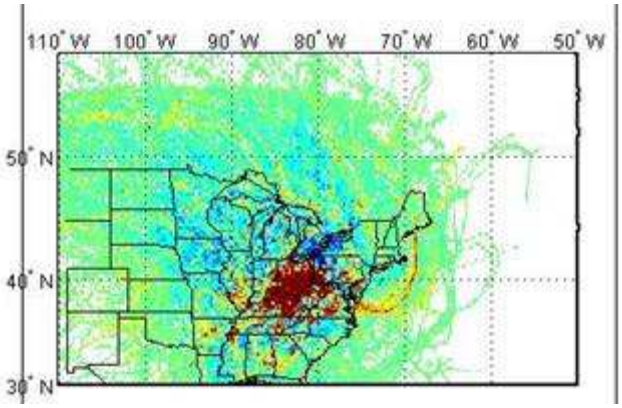
Sulfate- Sulfate ion Conc. (ug/m3)
Bext- Extinction (Mm-1)
PM- Particulate Matter Conc. (ug/m3)
OC- Organic Carbon Conc. (ug/m3)
Num Trajs- Number of trajectories in cluster
Num Trajs w. Poll- Number of trajectories in cluster with associated pollution measurement (Based on number of AQS ASPD samples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

Pittsburgh All Trajectories 00-04, Incremental Probability

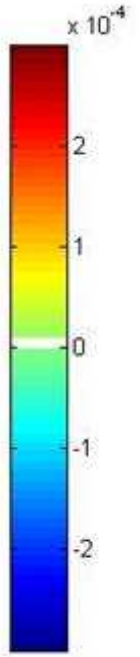
IP Based on Top10%, 500m



Sulfate

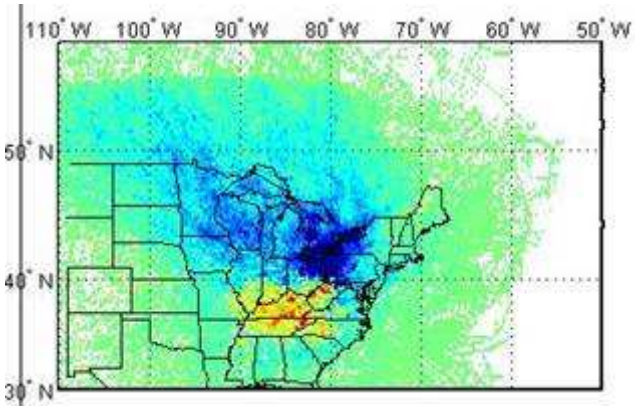


PM

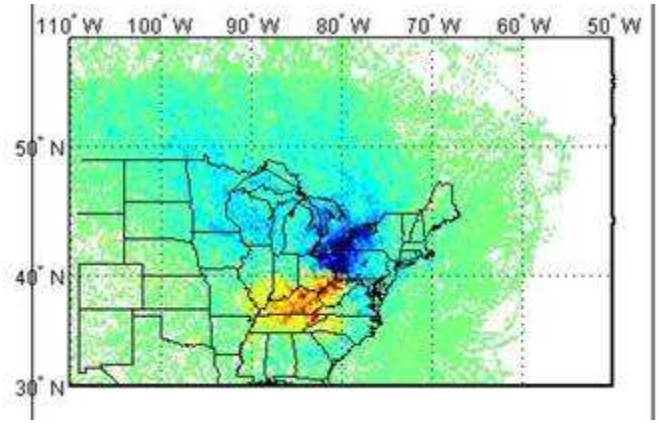


Pittsburgh All Trajectories 00-04, Cluster Weighted Probability

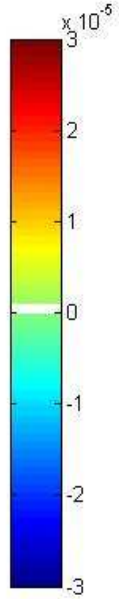
Calculated using Proximity Based Clusters, 500m



Sulfate

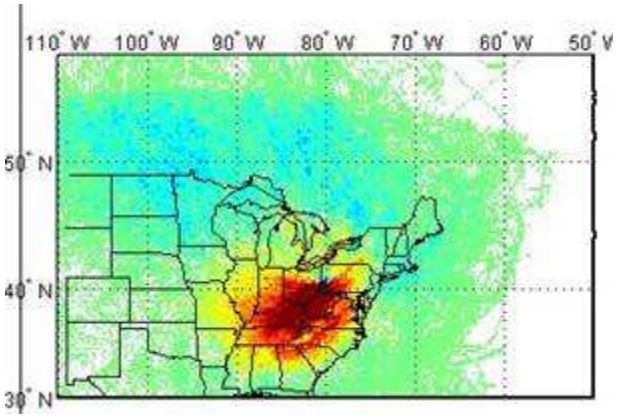


PM

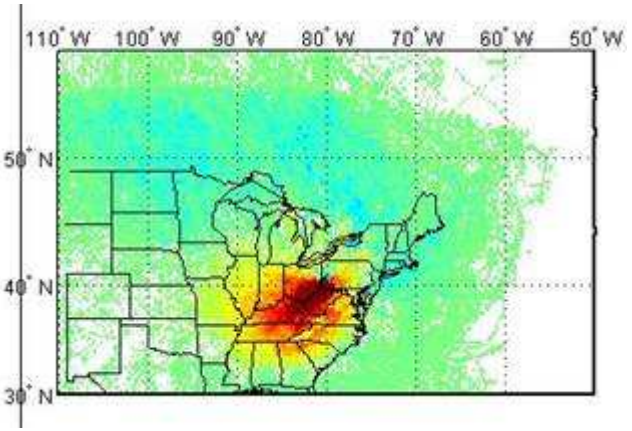


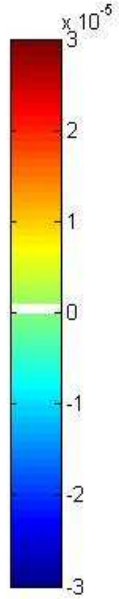
Pittsburgh All Trajectories 00-04, Cluster Weighted Probability

Calculated using Frequency Based Clusters, 500m



Sulfate





PM