A Trajectory Analysis of Atmospheric Transport of Black Carbon Aerosols to Canadian High Arctic in Winter and Spring (1990-2005)

L. Huang[1], S.L. Gong[2,*], S. Sharma[3], D. Lavoué[2], C. Q. Jia[1]

[1] Department of Chemical Engineering and Applied Chemistry, University of Toronto, 200 College Street, Toronto, ON M5S 3E5 Canada

[2] Air Quality Research Division, Environment Canada, Toronto, ON M3H 5T4 Canada

[3] Climate Research Division, Environment Canada, Toronto, ON M3H 5T4 Canada

Correspondence to: S.L. Gong (sunling.gong@ec.gc.ca)

Supplement: Modified Dorling' algorithm

In Dorling's algorithm, much of the computational effort is devoted to the repeated updating of cluster mean trajectories. In order to assign one trajectory to its closest cluster, for example, the Euclidean distances between a specific trajectory and all cluster mean trajectories have to be computed and compared. The shortest Euclidean distance defines the most representative cluster. Then the cluster mean trajectory of any newly updated group needs to be calculated and updated again. This process involves the transformation of spherical coordinates into Cartesian coordinates and tends to be inefficient. More importantly, the mean trajectory calculated from a group of complex and slow-moving trajectories appears less representative, and can be misleading. Thus using such slow-moving centroids in determining which cluster a trajectory belongs to (as Dorling's algorithm described) can affect the accuracy of clustering results and/or increase the computational coast. Here, a new criterion of assigning trajectory to its closest cluster is proposed. In the new clustering algorithm, the distance between a trajectory and a cluster is computed as the average Euclidean distance between the trajectory to be assigned and all the trajectories in that cluster. Thus, the improper use of a slow-moving centroid in representing the center of a group of trajectories is avoided. Furthermore, the computational overhead due to iterative computing of the cluster-mean trajectories and evaluating trajectory-to-cluster distances is overcome. The new clustering algorithm based on Dorling’s algorithm is detailed as follows:

1. Calculate the Euclidean distance between every possible pair of trajectories.

2. Start from single-member clusters.

3. Assign each of the real trajectories to the cluster that is closest in terms of the average distance as previously defined. Update the total root mean squire deviation (RMSD). If RMSD decreases, the assignment is accepted; if not, then the assignment is rejected.

4. Repeat step 3 until all real trajectories are correctly assigned and no more assignment is required.

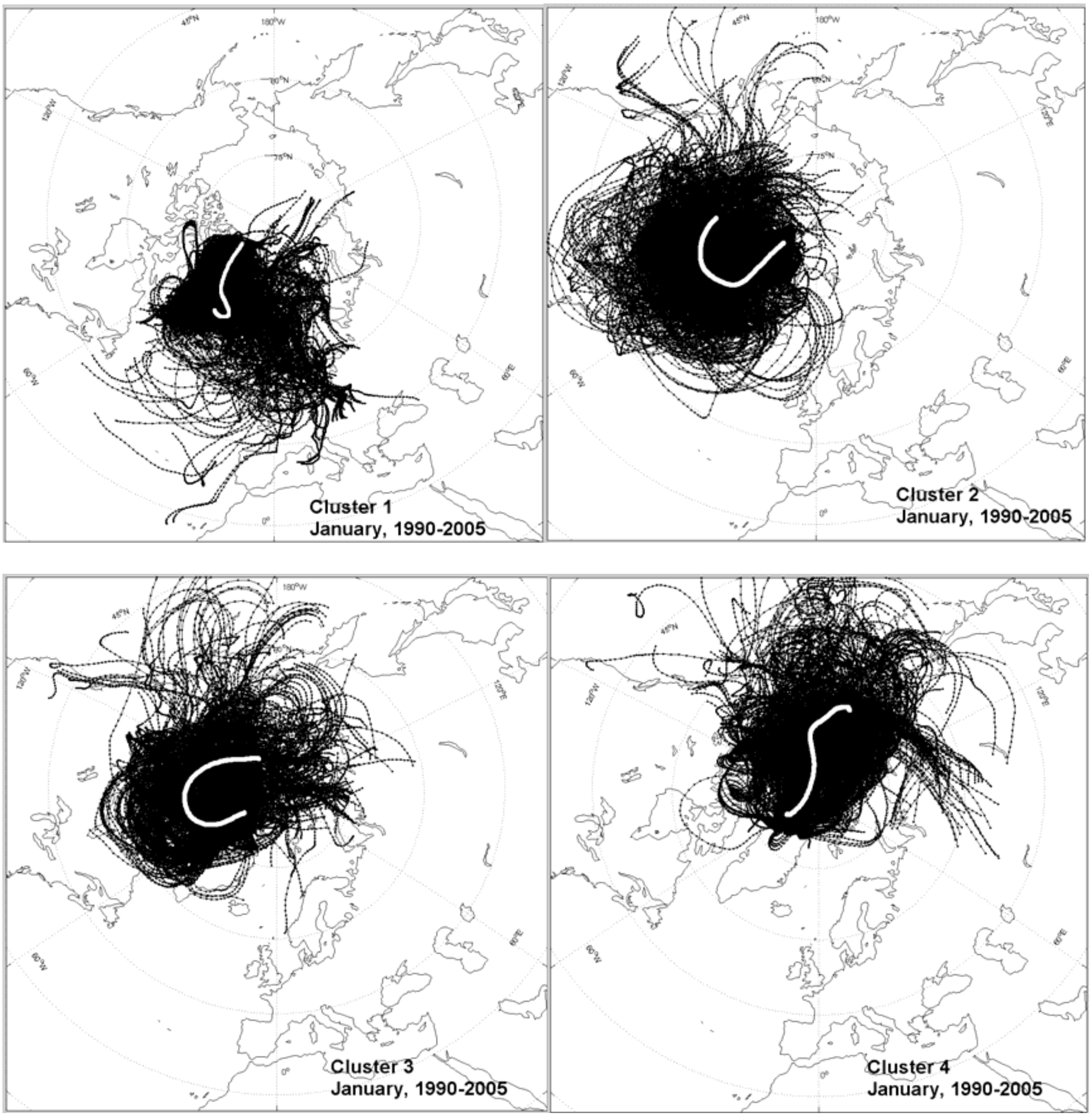
5. Calculate the final RMSD value with respect to the current cluster numbers.

6. Find and merge those two closest clusters. Recalculate the RMSD value.

7. Repeat steps 3 to 6 until a single cluster containing all real trajectories forms.

The number of transport patterns best represent the types of distinct pathways were defined by the clustering algorithm itself. As previously described, the total RMSD was calculated by step 4 in the new algorithm. At the end of clustering, the percentage change in this value can be plotted with respect to cluster number. Substantial change in such a plot indicates the merging of clusters of trajectories that are significantly different in terms of the wind directions and speeds. In this study, the change of 1.5% was assumed to be significant. So the cluster number before the unacceptable merging of trajectories is determined as the optimal number of clusters.

Supplement: Cluster-member plot



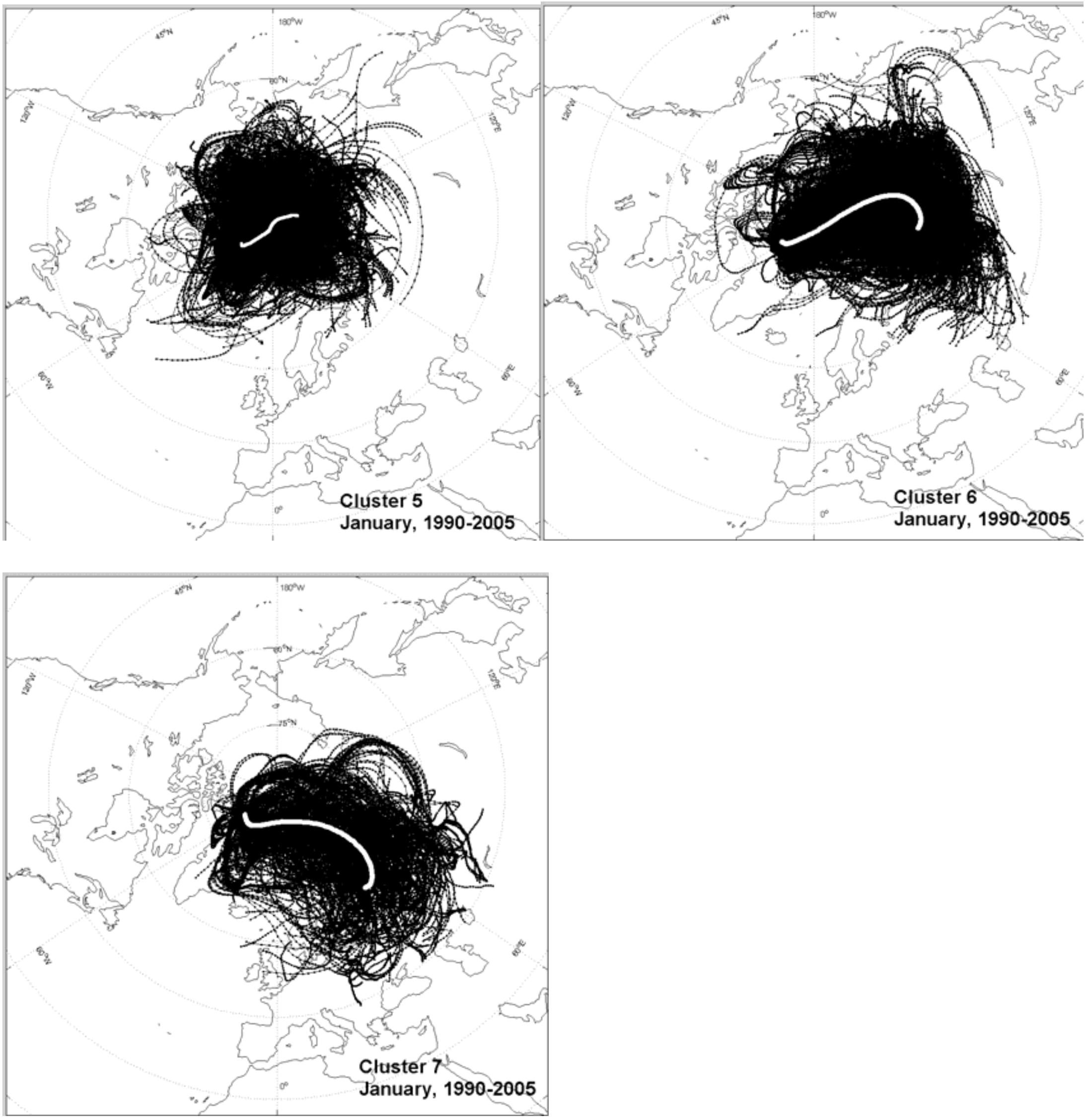
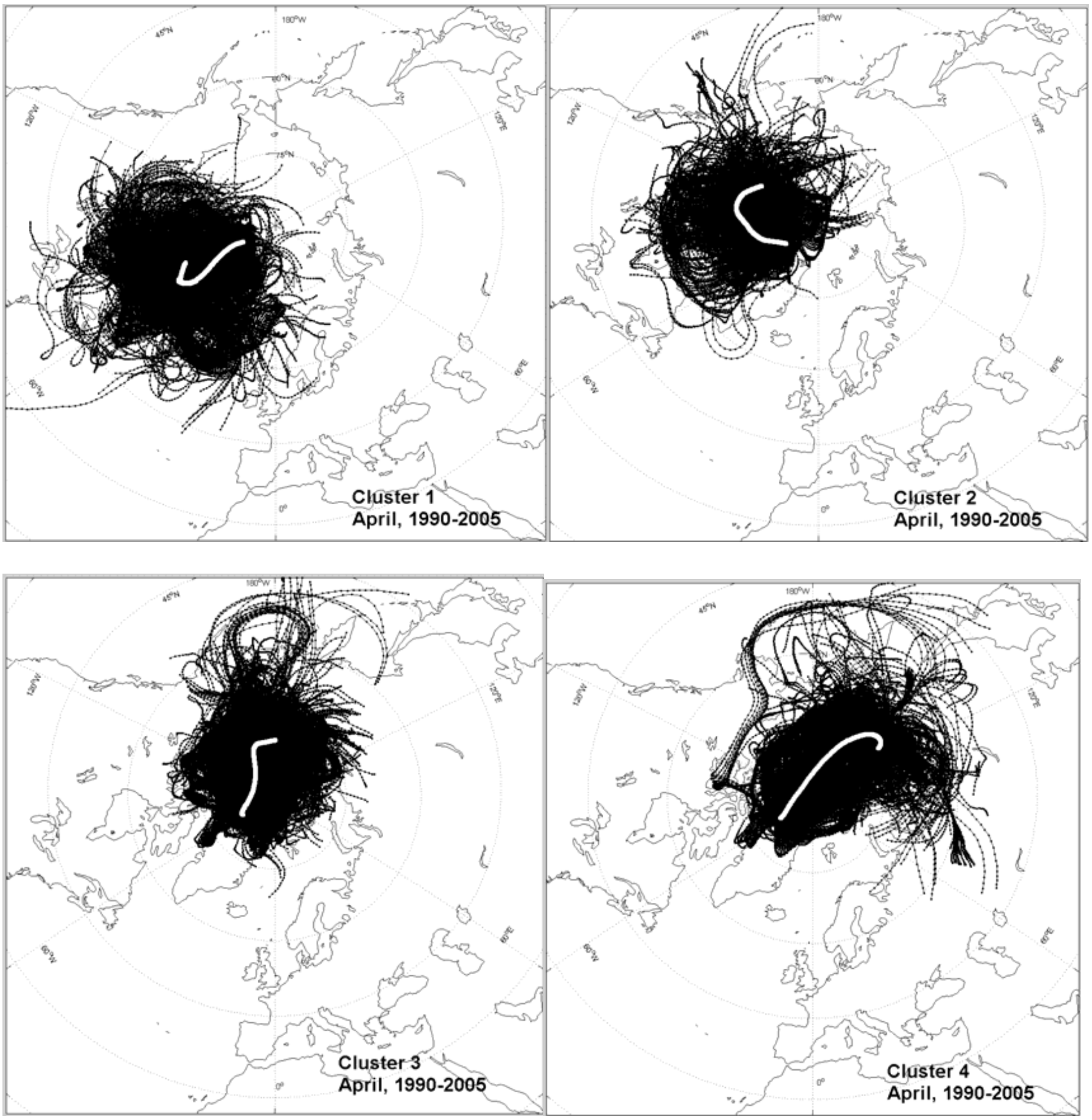


Figure S1. Cluster-member plot of 7 distinct clusters affecting Alert, Nunavut identified by cluster analysis for January, 1990-2005. Ten-day back trajectories arriving at Alert are shown in black lines, and the cluster-mean trajectories are shown in white lines.


Cluster 1
April, 1990-2005
Cluster 2
April, 1990-2005
Cluster 3
April, 1990-2005
Cluster 4
April, 1990-2005

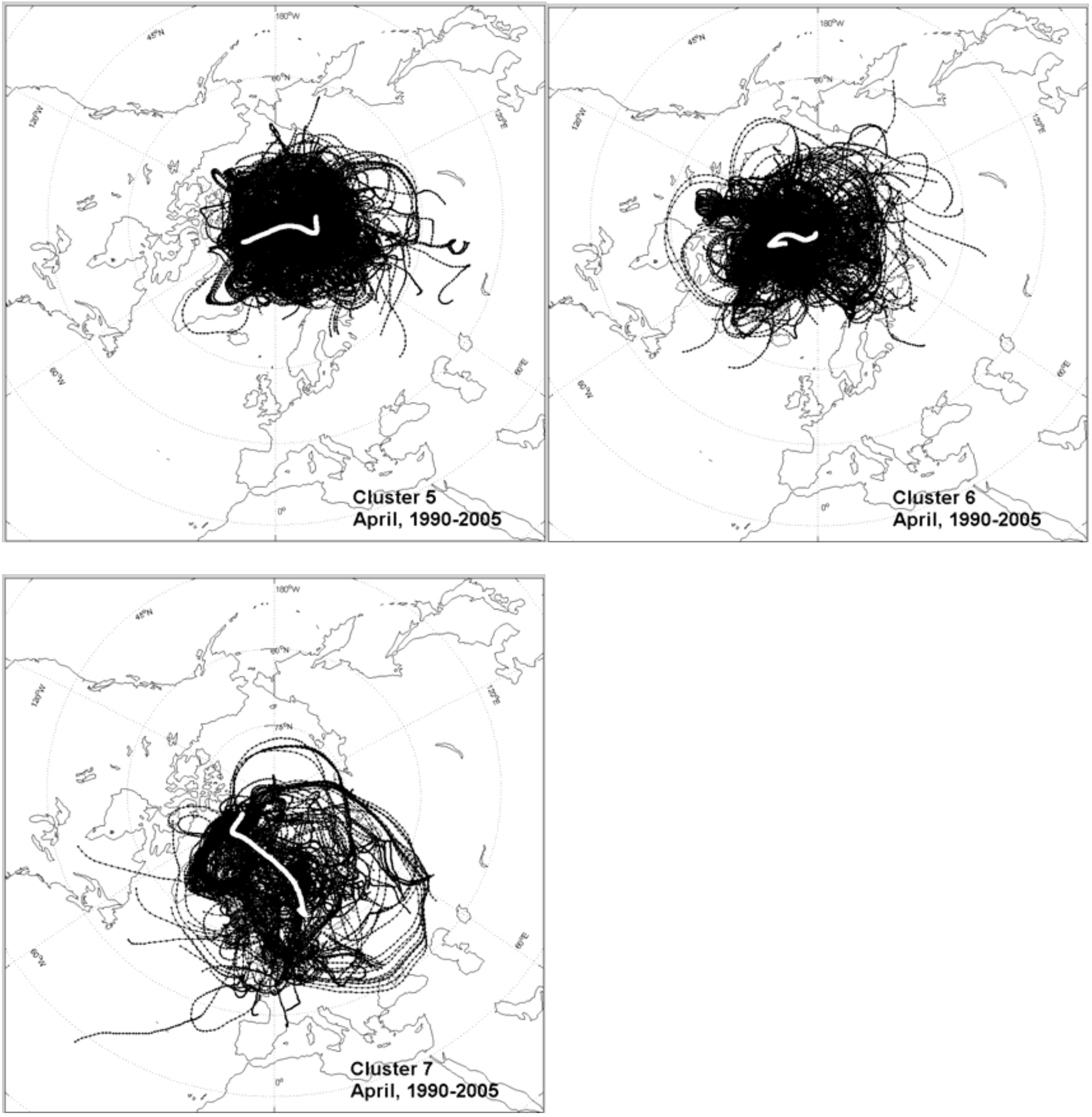


Figure S2. Same as Figure S1, but for April, 1990-2005.