To complete the discussion of clustering, we will review the clustering equations that are used in the cluster analysis.   There're really only three computations.   In the first one, we compute the spatial variance of an individual trajectory between itself and the cluster mean.   So in this case the individual trajectory, P, these are the endpoint positions, the vector endpoint positions, is compared, or the difference between the endpoint position and the cluster mean position along each point, along the trajectory, is computed and summed.   So this represents the spatial variance of an individual trajectory from its cluster mean. Then the cluster spatial variance is just the sum of the spatial variance of each of the trajectories within the cluster.   And the total spatial variance is just the sum of the cluster spatial variances over all clusters.

As was mentioned previously, the calculation starts with each trajectory being assigned to its own cluster.   And for instance, the positions along the first trajectory are compared to every other trajectory and the spatial variance is computed for all those pairs.   The pair with the minimum spatial variance is then merged to become the new cluster.   And so now we have a cluster with two trajectories and all the others only have one.   So which with each iteration, the minimum spatial, the pair with the minimum spatial variance is found, and then that cluster or that trajectory would be merged with the appropriate cluster. And this process continues until the only one cluster remains.

Now the graphic to illustrate why the spatial variance goes

up at the end, right, you can picture when the calculation starts since each trajectory is composed of a cluster with only one trajectory in it and the spatial variance would essentially be zero.   But toward the end let's assume we have remaining two clusters, cluster #1 and cluster #2, and that would be the blue line here, where CM is the cluster mean trajectory and the spatial variance for cluster number one is relatively small, it's essentially the mean squared difference between each of the trajectories within cluster #1 and in the same way, cluster #2, the spatial variance for the cluster #2 was relatively small.   But if we were to combine, let's say the last remaining two clusters into another cluster, now merging #1 and #2, you can see this red line would represent the cluster mean trajectory, the final cluster mean trajectory.   And you can see why the total spatial variance would go up, because now the variance is computed for each one of these grey trajectories, the difference between those and the cluster mean, which is going to be much larger now than it was previously, when there were just two clusters.   Because at this point we merged disparate clusters. So the clustering process is a way to put together trajectories that have similar characteristics and you do not have to predetermine downwind sectors or lay out some kind of arbitrary grid.

And that concludes the discussion of clusters.