Finally, a cluster analysis method has been applied based on the involvement of each model in the input feature. We wanted to see if we can cluster conceptual models based on the geographical location. For each pair of model, we counted the number of locations where both of the models have involvement in the input features. Based on this analysis on 24 conceptual models, a 24x24 matrix has been formed.

As we can see in Figure 1, there is a value for each pair of conceptual model where the value represent the involvement of that particular pair of conceptual model for all geographical locations. Based on this matrix of pairs, we generated the tuples of triples, quadruple, quintuple and so on for different thresholds. This process stops naturally after a certain tuples of sequences.

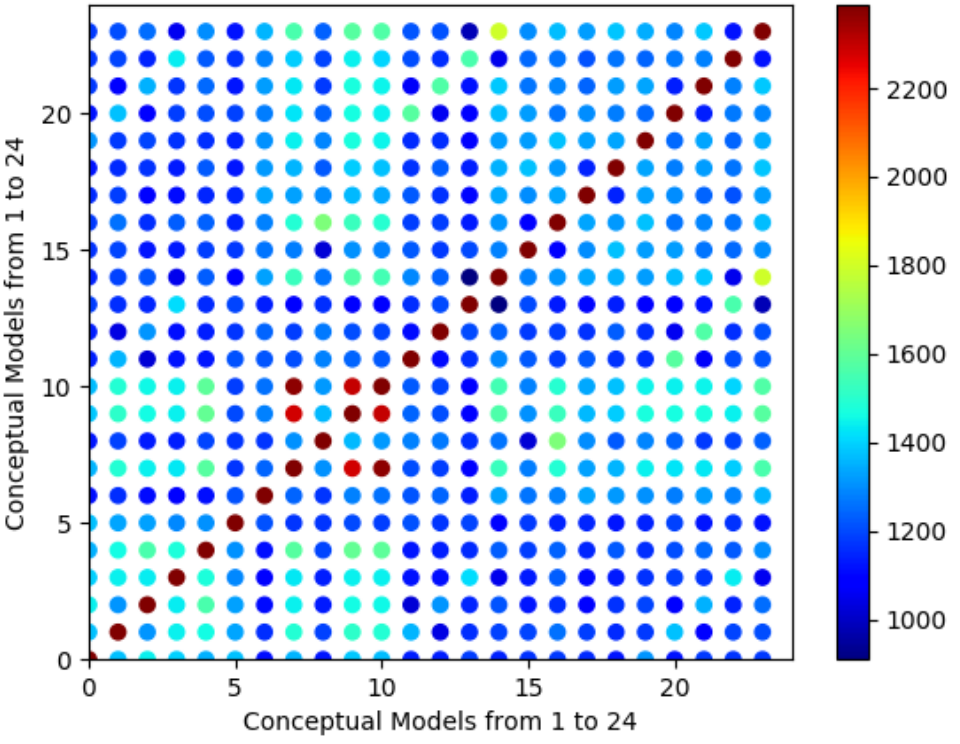

Figure 1: 24x24 Matrix that shows the involvement of each pair of conceptual model in all geographical location

Figure 2 illustrates all the clusters of conceptual models based on the geographical locations for a threshold 1300. That means, these clusters formed when a bunch of conceptual model involved in more that 1300 geographical locations. Here, values 1 to 4 represent the cluster numbers while the other values repre-

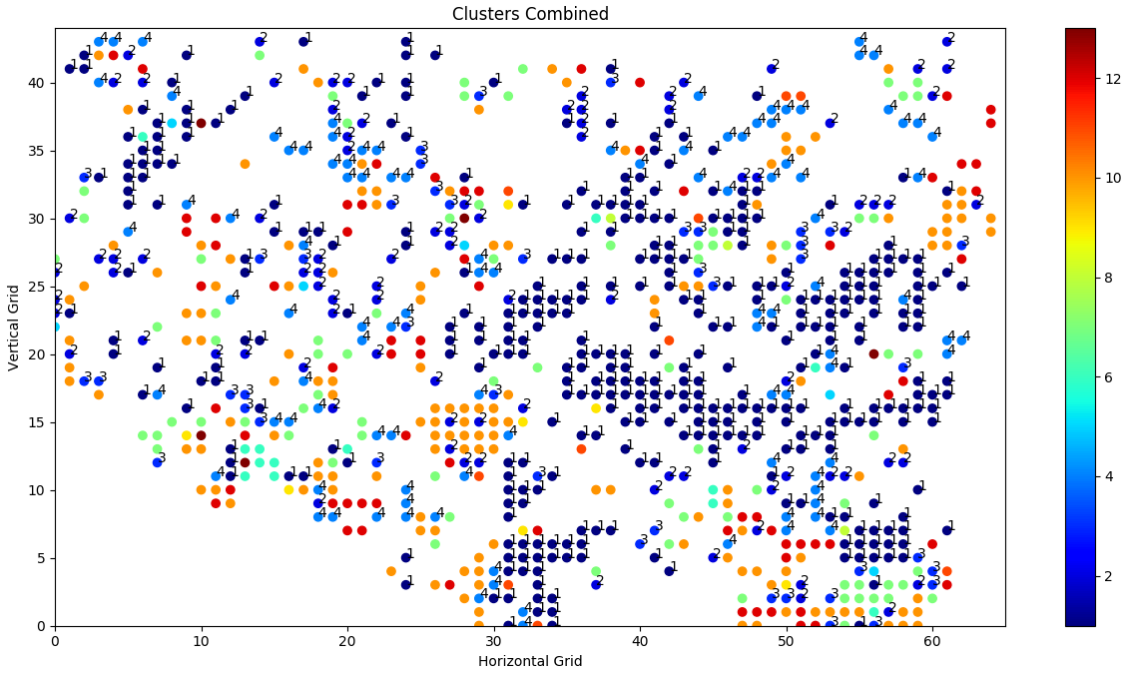

Figure 2: Clustering based on the involvement of each conceptual model in each geographical location

sents the overlaps of different clusters. In this figure the values that have been assigned for each overlap between clusters are listed in the table 1.

Table 1: The values that represent the overlapping of cluster in the figure 2

Values	Overlapping of Clusters
1	Only Cluster 1. No overlaps.
2	Only Cluster 2. No overlaps.
3	Only Cluster 3. No overlaps.
4	Only Cluster 4. No overlaps.
5	Cluster 1 and 2
6	Cluster 1, 2 and 3
7	Cluster 2 and 3
8	Cluster 1 and 3
9	Cluster 1, 2, 3 and 4
10	Cluster 2, 3 and 4
11	Cluster 1 and 4
12	Cluster 3 and 4

Table 2 shows the clusters of conceptual models that illustrate on the figure

2. The tuple of each cluster contains 8 conceptual models.

Table 2: Clusters of Conceptual Models based on the geographical location

Cluster Number	Cluster of Conceptual Models
1	(0, 1, 2, 3, 4, 7, 9, 10)
2	(7, 9, 10, 14, 18, 19, 21, 23)
3	(7, 9, 10, 14, 16, 18, 19, 23)
4	(7, 9, 10, 14, 16, 17, 19, 23)

For a better view and analysis, we plotted all the 4 clusters separately in figure 3. A careful analysis shows that cluster 1 has a strong involvement on the east side of the USA while rest of the three clusters have strong involvement on the west side of the USA. Cluster 2, 3, and 4 overlap in many geographical location. Again, these three clusters also have some non-overlapped area as well. In the first cluster of figure 3, a big patch is visible in the bottom middle of the map while there is no such patch in other clusters. That indicates a strong involvement of the 8 conceptual models of cluster 1 in this geographical location. All the conceptual models of a cluster have some common input factors that influenced the aggregation and helped clustering them based on the geographical location. This clustering analysis will lead us to understand the conceptual models in a better way and will provide a broad knowledge about the predictive power of conceptual models. In other word, clustering of conceptual models would make a better sense while aggregating different conceptual models and also while developing new conceptual models. Based on this clustering analysis, many decision can be taken about the input factors that influence the performance of different conceptual models which advance the atmospheric science.

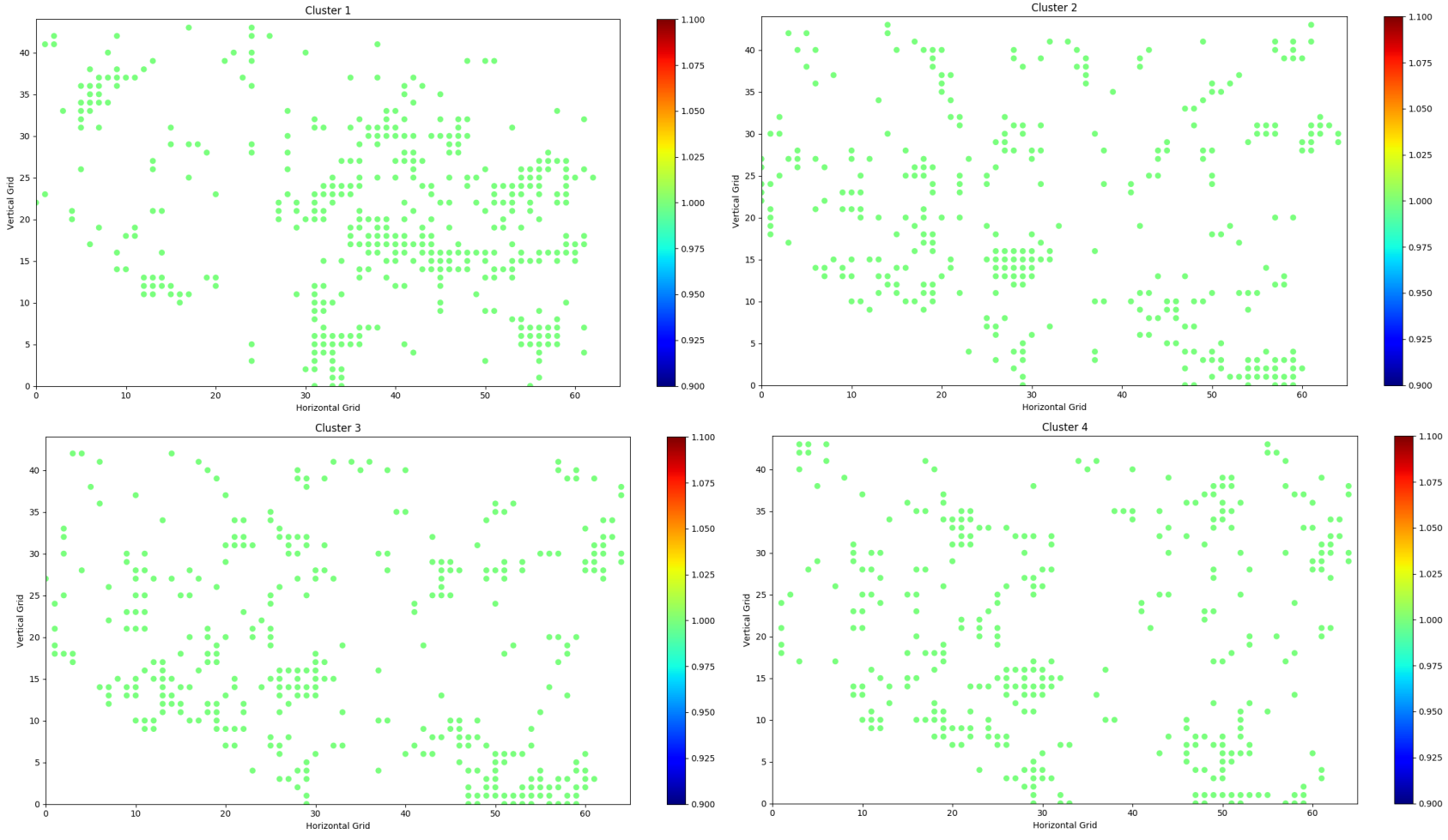

Figure 3: Geographical patches of all clusters separately