Supplementary File

Table 2. Mean percent bias & Correlation coefficient (R)					
a) GEOS-Chem Low-level	Cluster 1	Cluster 2	Cluster 3	Cluster 4	Cluster 5
Bias	- 0.051	0.020	0.068	- 0.026	- 0.039
R	0.53	0.55	0.51	0.61	0.55
b) GEOS-Chem Mid-level					
Bias	- 0.268	- 0.241	- 0.150	- 0.167	- 0.096
R	- 0.002	- 0.033	- 0.26	0.11	0.23
c) GEOS-CF Low-level					
Bias	0.139	0.189	0.340	0.241	0.197
R	0.74	0.60	0.56	0.61	0.54
d) GEOS-CF Mid-level					
Bias	- 0.143	- 0.044	0.021	0.008	0.112
R	0.43	0.14	- 0.19	0.21	0.74

Table S1. Calculated mean percent bias and correlation coefficient (R) by cluster. a) Low-level GEOS-Chem, b) Mid-level GEOS-Chem; c) Low-level GEOS-CF and d) Mid-level GEOS-CF results.

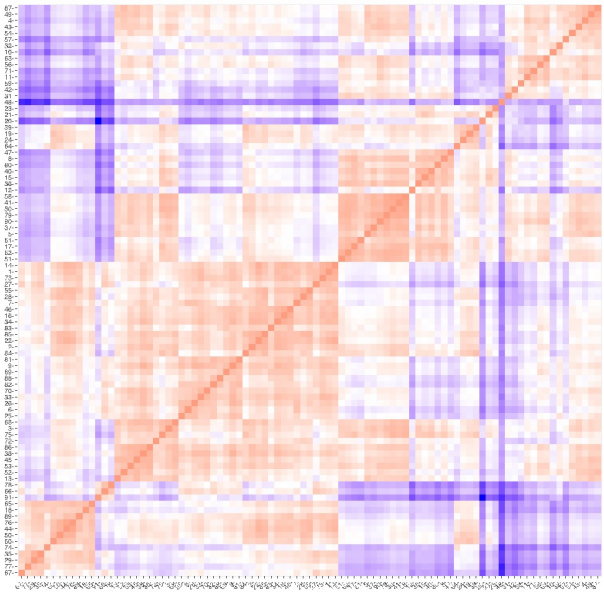

Figure S1. Visual assessment of cluster tendency (VAT) approach. Dataset high similarity (red) and low similarity (blue).

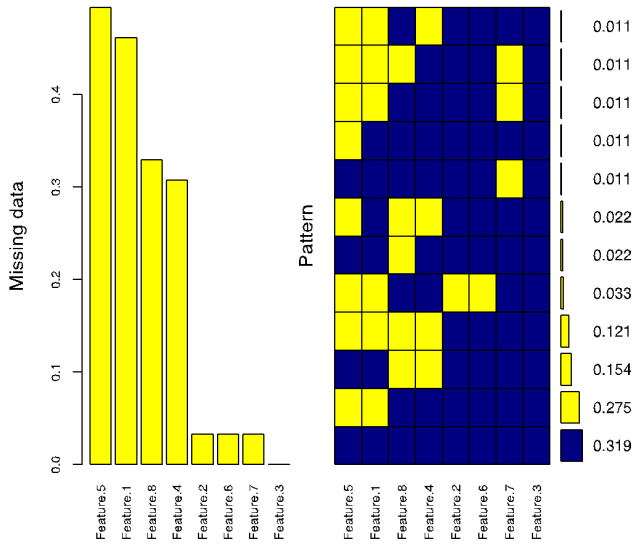

Figure S2. Percentage and pattern of missing data points by each feature used for clustering.

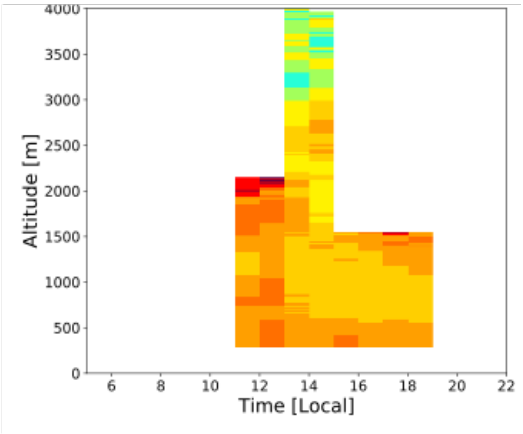

Figure S3. Results from the clustering: Cluster 6 which was assigned only one date (2018-06-17). Considered an outlier and was removed from the analysis.

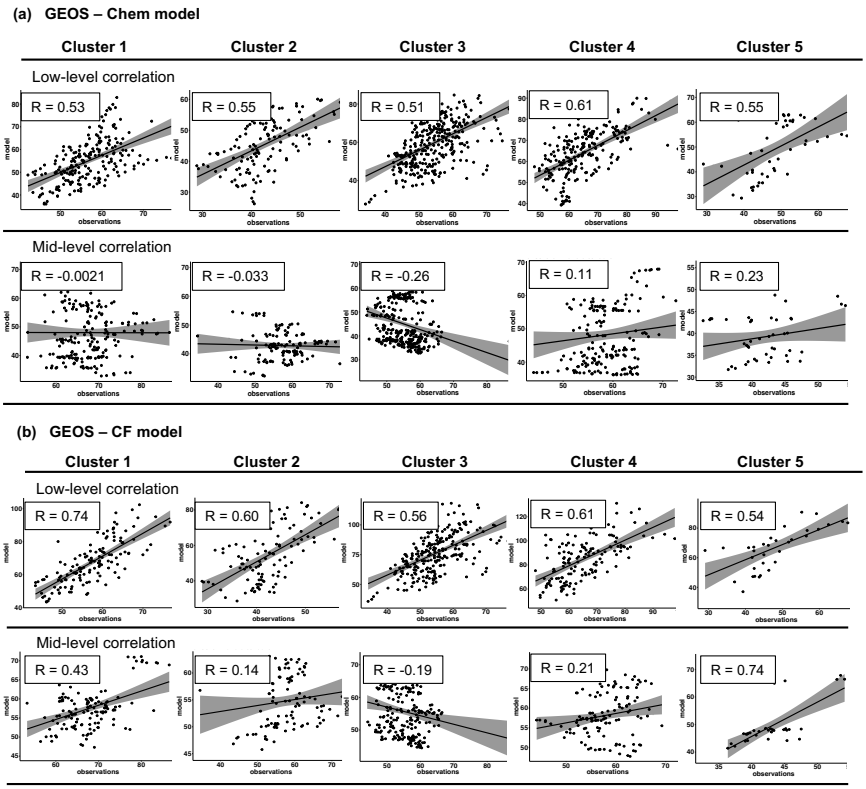

Figure S4. O_3 correlation between lidar observations and a) GEOS-Chem model simulation results and b) GEOS-CF model results by each cluster split by low-level (top panel) and mid-level (bottom panel).