Using Clustering to Establish Climate Regimes from PCM Output

Forrest Hoffman, Robert Oglesby**, William W. Hargrove*, David Erickson**

**Oak Ridge National Laboratory*   ***NASA Marshall Space Flight Center*

Introduction

A new statistical clustering technique was used to analyze output from the Parallel Climate Model (PCM, Warren Washington). Five 100−year simulations of "business−as−usual" (BAU) scenarios were compared. This analysis considers three PCM output variables: surface temperature, precipitation, root−zone soil water. Only land was considered in the clustering analysis.

The copious output (about 1200 monthly maps per run) is difficult to decipher. The long−term climatic trend of interest is masked by the magnitude of the seasonal cycles and large inter−annual variability.

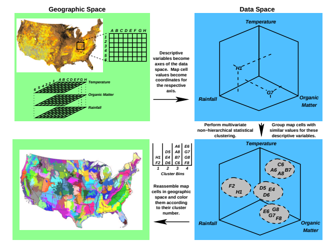
Multivariate Clustering

Multivariate clustering is the division or classification of objects into groups or categories based on the similarity of their properties.

Non−hierarchical clustering produces a single level of division of objects into some (often pre−determined) number of groups.

Multivariate Geographic Clustering (MGC) employs non−hierarchical statistical clustering to the classification of geographic areas.

Multivariate Spatio−Temporal Clustering (MSTC) is an application of Multivariate Geographic Clustering across space and through time.



B05.12 B05.15 B05.18 B06.06 B06.09

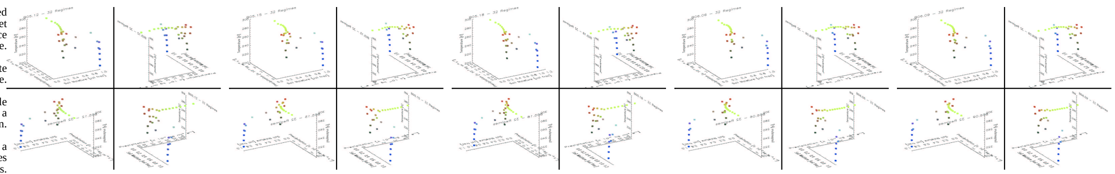
Clustered Climate Regimes

The clustering process establishes an exhaustive set of occupied climate regimes (i.e., the cluster centroids) which define the subset of phase space occupied by the simulated atmosphere/land surface at all points in space and time.

Any geographic location will exist in only one of those climate regimes at any single point in time.

The PCM model simulates a relatively small subset of all possible combinations in the full climate space, even globally through a 100−year run.

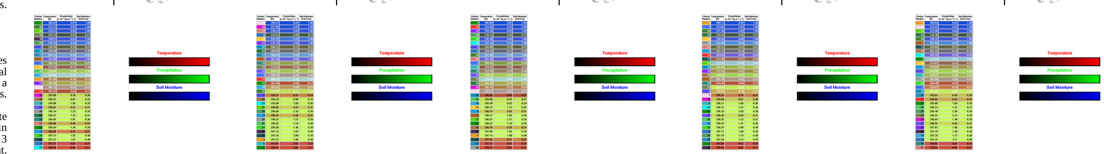
The clustering analysis consistentlyidentifies and establishes a very similar (but not identical) set of recurring climate regimes within each of the 5 BAU scenarios.



Climate Regime Tables and Maps

The centroid coordinate of each of the clustered climate regimes gives the synoptic conditions of that regime, in the original measurement units. Each clustered climate regime is shown in a table for each of the 5 BAU runs.

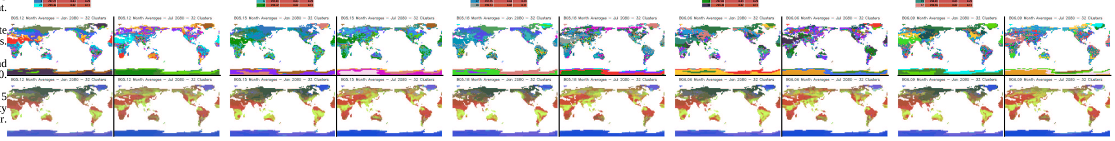
The first column uses random colors for each clustered climate regime, also shown in the top row of maps. Remaining columns in the table are shown in similarity colors, where each of the 3 variables contributes a red, green, and blue component.



The bottom row of maps depict the same clustered climate regimes but are colored using similarity colors.

The first column of maps is January 2080, and the second column is July 2080.

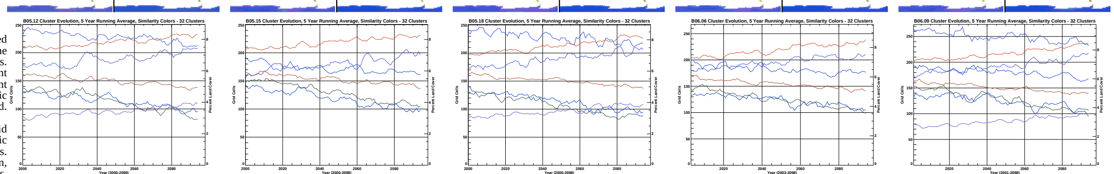
Randomly colored cluster maps differ widely across the 5 independently−analyzed BAU runs, but the similarity colored maps are very similar.



Area Changes

Because the same clustered sets of conditions are identified through time, we can plot changes in the geographic area of the globe in a particular clustered climatic regime as it evolves. Many of the 32 clustered climate regimes remain relatively constant in area throughout the run. Here we have eliminated these constant regimes to show only climate regimes experiencing large climatic shifts during the simulation period.

All 5 BAU runs indicate growth in the area of hot, dry, arid climatic regimes, and decreases in area of the coldest arctic regimes, which are replaced by cold summer conditions. These area changes are consistent with increased desertification, and a warming of polar regions.

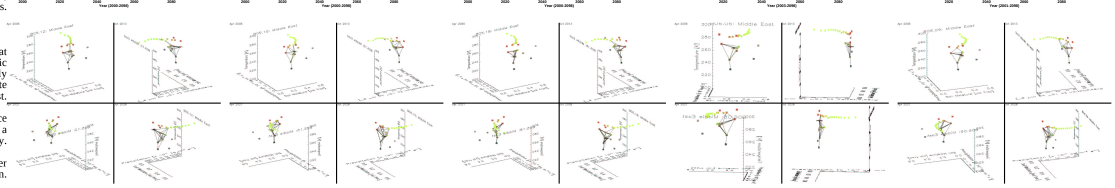


Climate Trajectories

A geographic location exists in only one of the climate regimes at any point in time. By incrementing time, any single geographic point will trace out a trajectory or orbit among successively occupied climatic regimes in climate phase space. The climate trajectories shown here are for a location in the Middle East.

A "spider" representing the simulated atmosphere/land surface sequentially moves among the climatic regimes, leaving a thickening "web" outlining the trajectory.

When a geographic location adopts a climate regime that it never previously occupied, a climatic shift has occurred at that location.

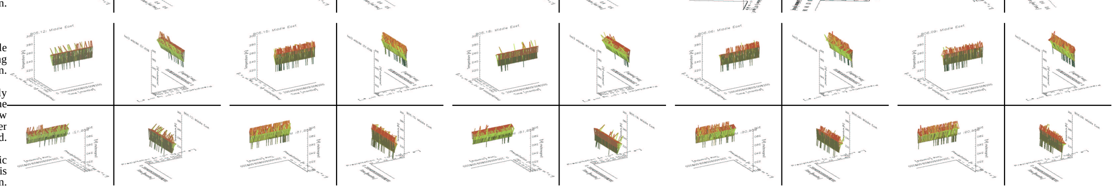


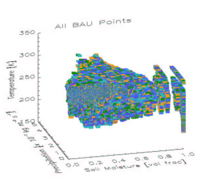
Climate Manifolds

Tracing out the entire seasonal and annual trajectory of a single spot on the globe yields a "manifold" in state space representing the shape of the predicted climate occupancy for that location.

Early in the simulation period, this Middle East location regularly adopts the cold, darkly colored regime at the bottom. Later in the simulation, this geographic location travels upwards to a new climate regime which is hotter and drier than it has ever previously visited.

The clustering analysis has established the long−term climatic shift as well as showing the seasonal and annual cycles for this location.





All Points Plotted in Three−Space

When every monthly data point from all 5 BAU runs is plotted in the 3−dimensional climate state space, it is clear that a fairly small region of this space on the left side is densely packed with points. At the right, discrete values of high soil moisture result in planes of points. Points are colored by BAU run.

On the left panel, each of the 5 BAU runs was analyzed individually and separately. On this panel, however, all 5 BAU runs were taken together and divided into a single, common set of climate regimes in the same clustering analysis.

All 5 BAU runs are thus placed into a common set of climate regimes which permit the direct comparison of the individual runs within the same climate space.

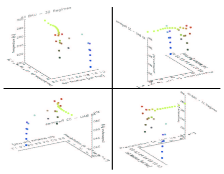
Thus, a parallel set of results can be shown on this panel as on the opposite one, but these results are now shown IN TERMS OF ALL BAU RUNS COMBINED FOR DIRECT COMPARISON.

Conclusions

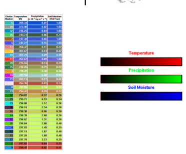
Cluster analysis is a powerful data mining tool which provides a common basis for comparison across space and time for multiple climate simulation runs. Because it runs efficiently on a parallel supercomputer, it can be used to reveal long−term patterns in large multivariate data sets. The clustering technique statistically establishes a common and exhaustive set of occupied climate regimes within climate space. Cluster regimes are defined in terms of the original measurement units for each variable used.

The regimes defined by the clustering process were very similar, both in the analysis of individual runs and in a composite analysis of multiple runs taken together. Cluster analysis can be used to compare an entire ensemble of multiple runs of the same scenario (as done here), two ensembles of different scenarios, and even results from two different climate models.

B05.12 B05.15 B05.18 B06.06 B06.09

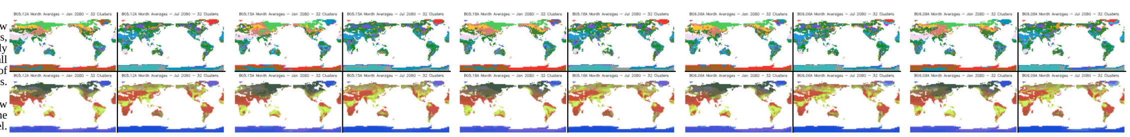


All 5 BAU Runs Clustered into a Single Common Set of Climate Regimes for Direct Comparison



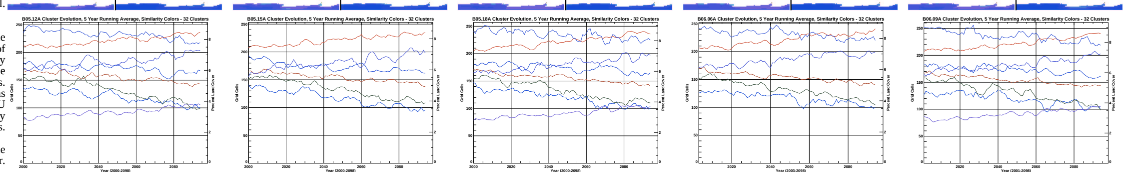
Now the randomly−colored maps in the top row are very similar across the different BAUs, and maps from different BAUs can be directly compared with each other. This is because all BAUs are now classified into a common set of clustered climate regimes.

The similarity colored maps in the bottom row still look identical, as they did in the separate analyses on the left panel.

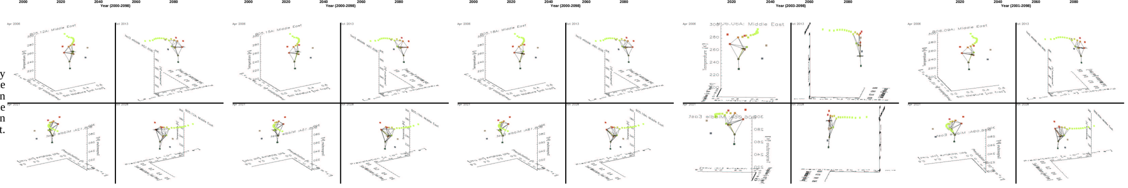


Although there are slight differences in the changing area plots for the common set of regimes across BAUs, it is easy to identify the same clustered regime lines (drawn in the same similarity color) on each of the graphs. There are differences in area inflections within the same cluster regime across BAU runs, however, highlighting the variability among predictions.

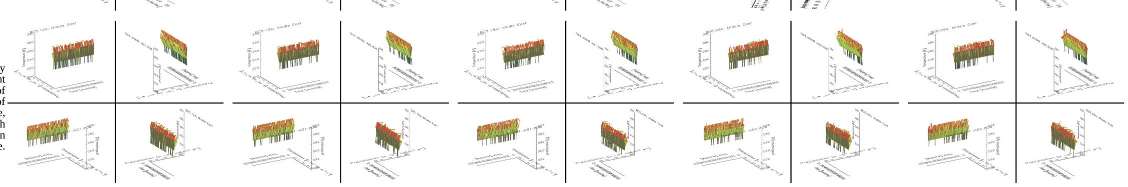
The B05 and B06 runs are not distinguishable from each other.



These plots show each single BAU trajectory relative to the common set of climate regions. Because the two B06 runs start in different years, their trajectories for the same date are shown at a different rotation point.



On the trajectory manifolds, it becomes easy to see the variability across the different runs, particularly in the prediction of extreme regimes. Changes in frequency of visitation fo the coldest (bottommost) regime, for example, varies among BAUs, even though all BAUs show a decreasing rate of visitation to this coldest regime through time.



Five Climate Trajectories in a Common Climate Space

Now the 5 BAUs can each be represented by 5 different spiders as they move within a single common climate space. Each spider is colored differently; when any two spiders occupy the same climate regime, the overlapping spiders are colored black.

Trajectories are drawn with the similarity color of the climate regime to which the spider has just moved, but the links subsequently change to the color of the spider that traversed them most frequently. Web segments between climate regimes also become thicker with repeated use. Trajectories will become thick and black if they represent seasonal orbits commonly used by many spiders.

The multiple spiders are often co−incident on the same climate regime in January and July, at the climatic extremes of the year, but spread out across multiple climate regimes in spring and fall months. Spiders frequently appear on opposite sides of the kite−shaped seasonal orbit, but rejoin at the top and bottom of the kite. Thus, the BAUs are similar with regard to temperature, but have a greater spread with respect to soil moisture. The multiple spiders spread out more among regimes as the simulation period progresses.