1. Motivation

In previous work, climate models considered by the IPCC-AR4 were given a performance score based on their simulation of present-day mean climate. An overall score was calculated from mean errors in 37 different variables. The question addressed here is what if there are a large number of variables that are similar to each other? In this case, the scores will be biased in favor of those variables.

2. Goals

We analyze the similarity amongst errors in different variables. Our goal is to calculate an objective score of overall model performance by adjusting weights given to errors in different variables before computing the final result.

3. Cluster Analysis of Variable Errors

To gain insight into variable similarity, we apply hierarchical clustering to the variable errors shown in Fig 1. The outcome of the clustering is shown in Fig 2.

4. Error Visualization

We project the variable errors (in the 25 models) onto a 2-dimensional space using Principal Component Analysis and visualize the clusters in color.

5. Key-Variables

In order to understand which variables are most representative for a given cluster, we determine the similarity of variable errors within their respective clusters.

6. Improved Variable Weights

We explored different methods to assign appropriate weights to the variable errors. Here, we show weights determined by a “center-of-mass” method; weights are derived by the change of center-of-mass of the entire data set when removing a particular variable.

7. Model Performance Scores

We now apply our variable weights (shown in Fig 9) to compute a more objective model ranking when compared to equal weighting.