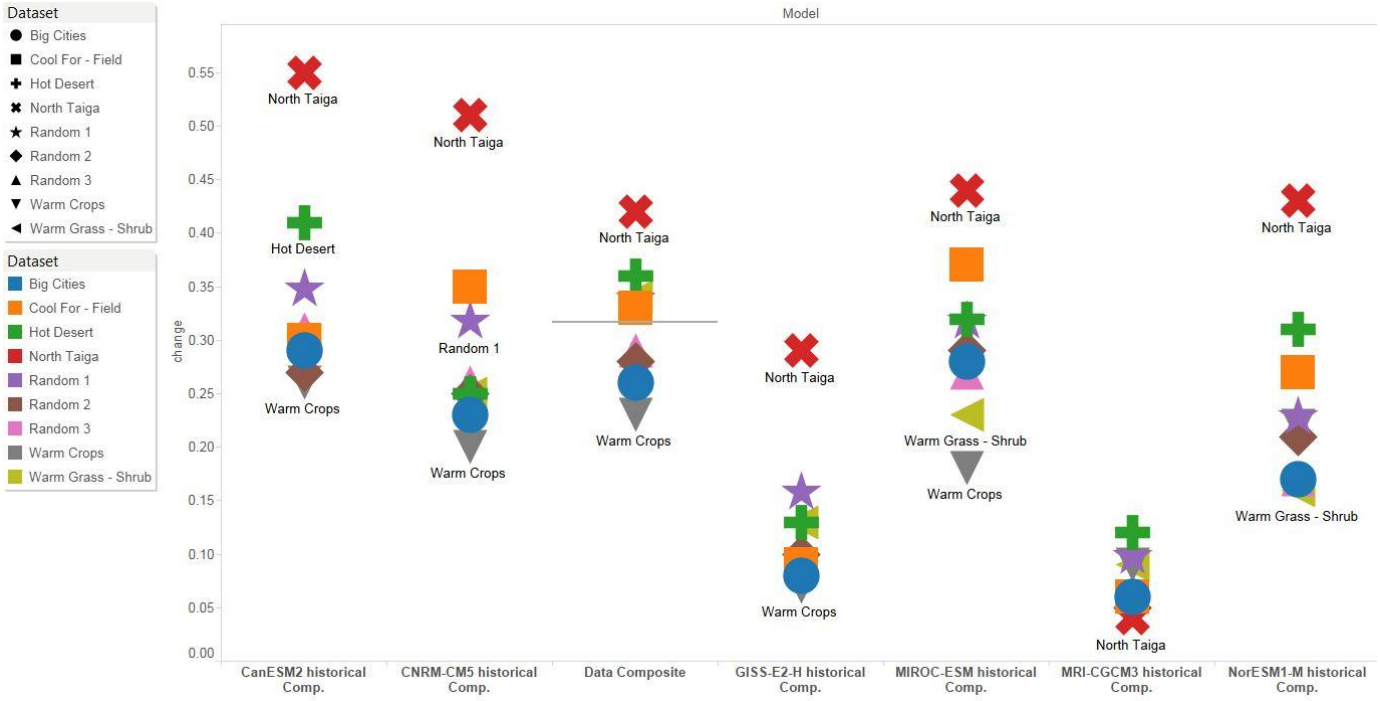
Examination of the Accuracy of Different Climate Models

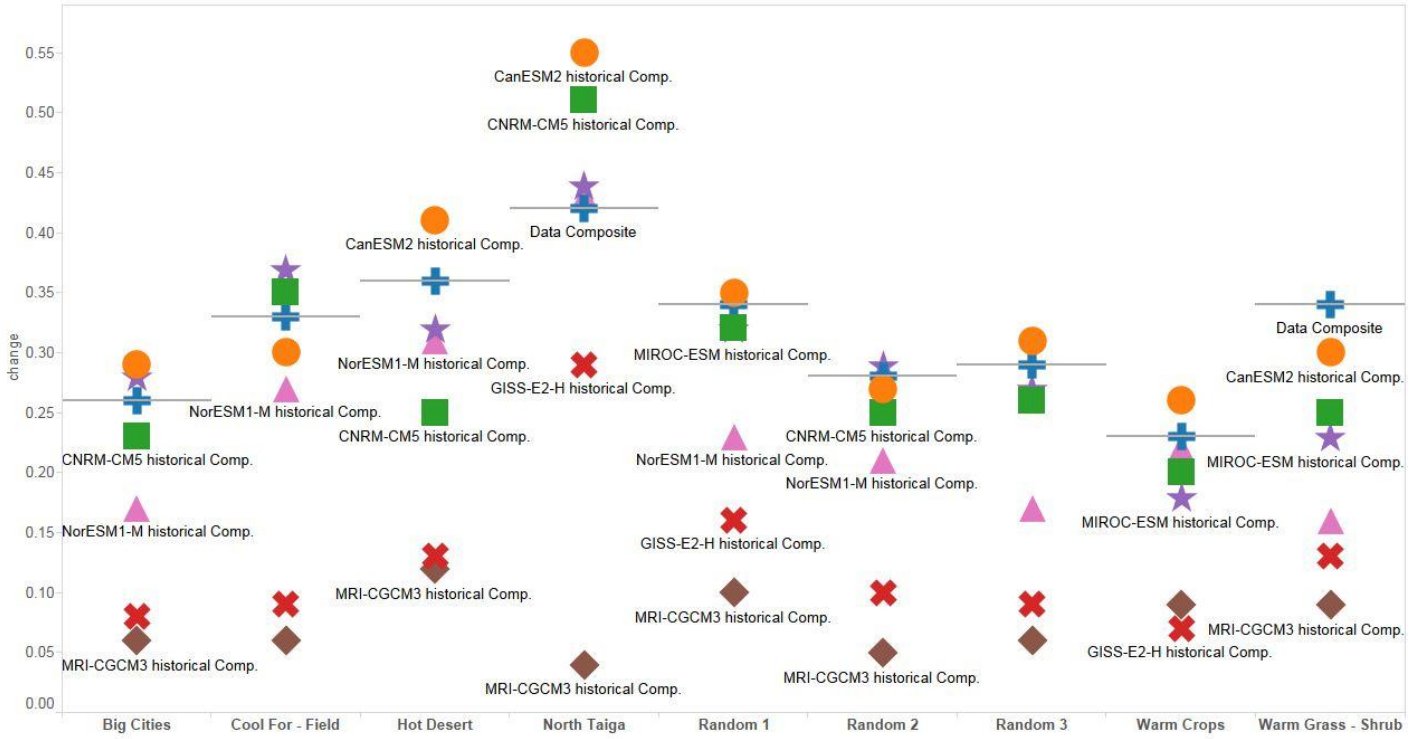
This paper will examine the accuracy of six climate models over a range of vegetation types and against a random set of weather stations. The vegetation types selected were designed to cover a variety of climates and the set of stations chosen within each type were chosen somewhat randomly, but also to provide geographic diversity. Similarly, for the random sets of stations, the first two thirds were selected via a random scroll through the station list and then the remaining stations were selected less randomly to cover all geographic regions.

This criteria enabled me to examine the accuracy of the different climate models in two different ways. A series of random sets of stations was used to assess the overall accuracy of the models. The sets of vegetation-type stations allowed for an examination of whether or not some models were more accurate for different types of climates. First, the data was normalized using the "Norm 1900-1950" button. Then a Composite was taken for that normalized data. Next, the Time Range was adjusted to 1970-2013 and the set of stations was Averaged together. Finally, the set of models was selected using the Historical model scenario. The resulting data was analyzed as shown below.

The first figure shows the averages for each model and the actual historical data. Both GISS-E2-H and MRI-CGCM3 consistently underestimated the actual deg C / decade temperature change for all of the data sets. The North Taiga model hindcast by MRI-CGCM3 is particularly low. CanESM2-M and MIROC-ESM appear to provide the best models for the historical data, but all four of the others seem to have done a good job.



The second figure displays the model data from the data set point of view. The historical data is labeled with a blue plus symbol and has a line through it for easy reference. Quick visual inspection shows that the GISS-E2 and MRI-CGCM3 models underperformed relative to the other four and their values were always much less than actual.



The other four required additional analysis. The chart below shows the difference between the actual average decade temperature change between the historical data and that of the models. CanESM2 and MIROC-ESM did the best job and CNRM-CM5 was slightly more inaccurate. NorESM1-M was even a bit more off, but was highly accurate for most data sets. CanESM2 typically predicted change rates slightly higher than actual, while the NorESM1-M numbers were always less than actual.

Data set	Hist Data	Can data	Can Diff	CNRM dat	CNRM diff	GISS	GISS diff	MIROCC	MIROCC diff	MRI	MRI Diff	Nor	Nor Diff
Warm Crops	0.23	0.26	0.03	0.2	0.03	0.07	0.16	0.18	0.05	0.09	0.14	0.22	0.01
Hot Desert	0.36	0.41	0.05	0.25	0.11	0.13	0.23	0.32	0.04	0.12	0.24	0.31	0.05
Cool For - Field	0.33	0.3	0.03	0.35	0.02	0.09	0.24	0.37	0.04	0.06	0.27	0.27	0.06
North Taiga	0.42	0.55	0.13	0.51	0.09	0.29	0.13	0.44	0.02	0.04	0.38	0.43	0.01
Warm Grass - Shrub	0.34	0.3	0.04	0.25	0.09	0.13	0.21	0.23	0.11	0.09	0.25	0.16	0.18
Random 1	0.34	0.35	0.01	0.32	0.02	0.16	0.18	0.32	0.02	0.1	0.24	0.23	0.11
Big Cities	0.26	0.29	0.03	0.23	0.03	0.08	0.18	0.28	0.02	0.06	0.2	0.17	0.09
Random 2	0.28	0.27	0.01	0.25	0.03	0.1	0.18	0.29	0.01	0.05	0.23	0.21	0.07
Random 3	0.29	0.31	0.02	0.26	0.03	0.09	0.2	0.27	0.02	0.06	0.23	0.17	0.12
			0.35		0.45		1.71		0.33		2.18		0.7

Conclusion: Based on the results of the data sets analyzed here, the GISS-E2 and MRI-CGCM3 models should be examined further to determine the reasons why their hindcasts performed so poorly. Both consistently and significantly under-predicted the rate of change. Perhaps random chance in the stations selected led to a larger than normal error level or maybe the models need to be adjusted in some ways. No vegetation-dependent differences between the models were obvious from these sets of data. Maybe additional cuts of the data from different vegetation types, particularly in other climate bands, would show some differences.