Environmental Data Science (2022), 1: 1–7
doi:10.1017/xxxx

SUPPLEMENTARY MATERIALS

Supplementary Materials for: A Dependent Multi-model Approach to Climate Prediction with Gaussian Processes

Marten Thompson[1]*, Amy Braverman[2] and Snigdhansu Chatterjee[1]

[1]School of Statistics, University of Minnesota, Minneapolis, 55455, MN, United States
[2]Jet Propulsion Laboratory, California Institute of Technology, Pasadena, 91109, CA, United States
*Corresponding author. Email: thom7058@umn.edu

Received: 10 January 2022; **Revised:** 01 May 2022

1. CMIP6 Model Outputs

The spaghetti plot below depicts all 17 CMIP6 simulations considered in this analysis.

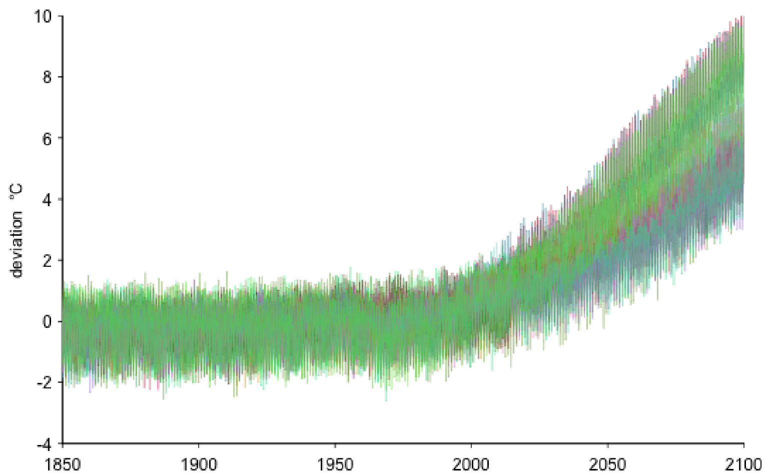

Figure 1. *De-seasoned output from all 17 CMIP6 simulation outputs. In the LOO analysis, these are each individually treated as the observed time series in order to assess our modelling approach.*

2. All LOO Cases

Each of the 17 CMIP6 simulations were treated as the "observed" dataset in order to assess our methodology's accuracy. The main paper depicted results for access-cm2, and identical analysis exists for the remaining 16 simulations, detailed below.

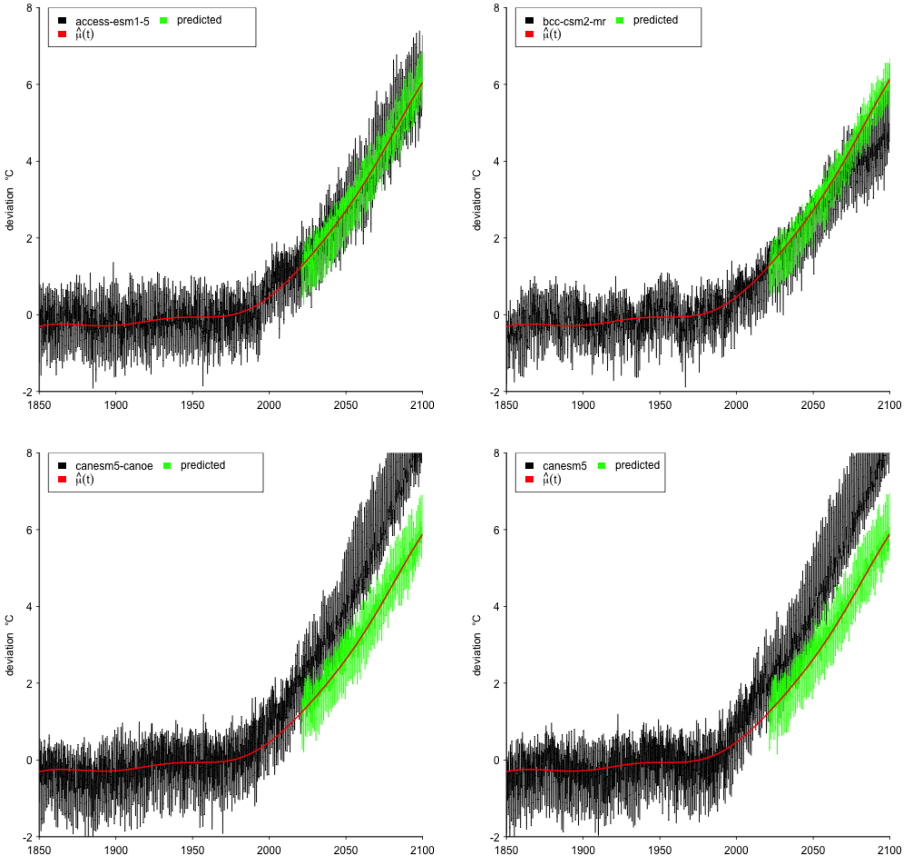

Figure 2. *Four separate analyses, where each simulation (black) had 2021-2100 held out and their values predicted by our methodology (green).*

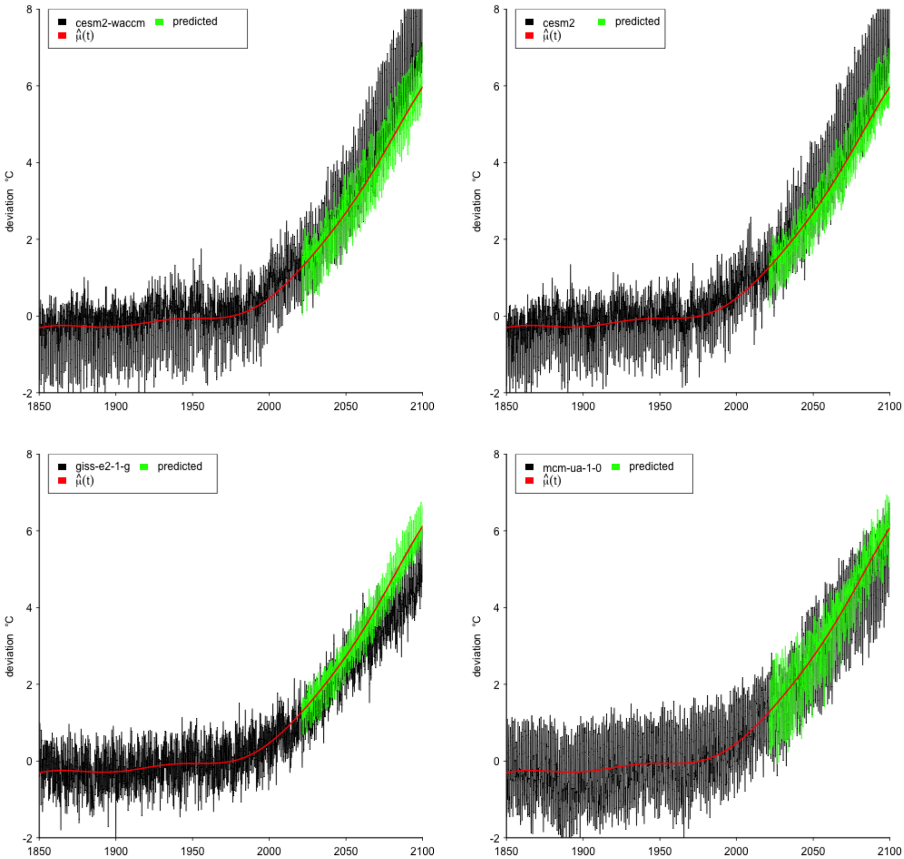

Figure 3. *Model simulations (black) and their corresponding predicted values (green) when 2021-2100 was held out.*

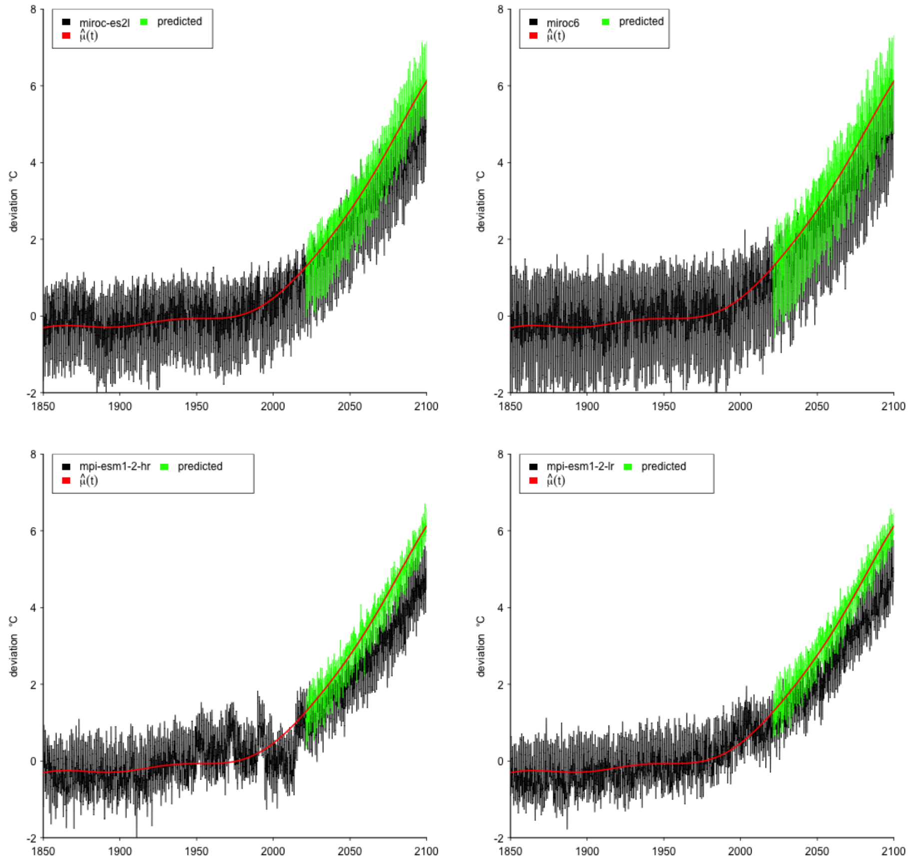

Figure 4. *Model simulations (black) and their corresponding predicted values (green) when 2021-2100 was held out.*

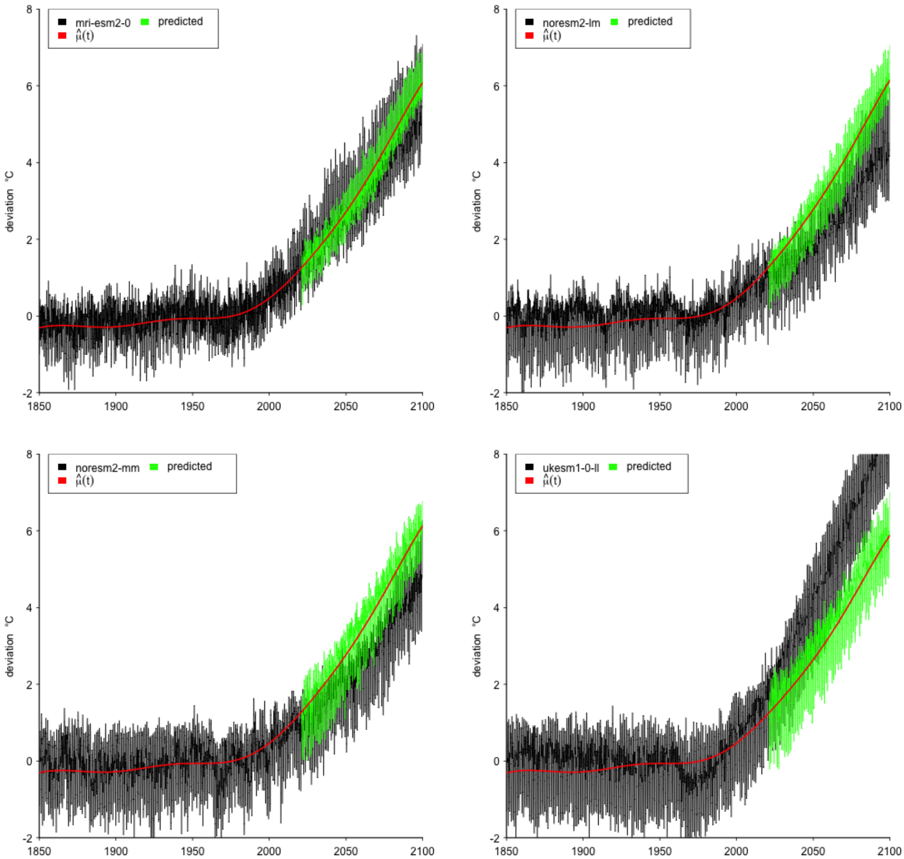

Figure 5. *Model simulations (black) and their corresponding predicted values (green) when 2021-2100 was held out.*

3. Kernel Choice

The choice kernel is worth attention as it defines the class of functions over which we place our Gaussian process prior. Since we have deseasoned our data, we may forgo explicitly encoding seasonality in the kernel (as in, for example, [1] Section 5.4). We also wanted to avoid placing any strict shape restrictions e.g. linear or polynomial. Finally, the Matérn $5/2$ kernel is a common alternative to the squared exponential that we employed, but our experimentation did not find a meaningful difference between the two (see Figure 6).

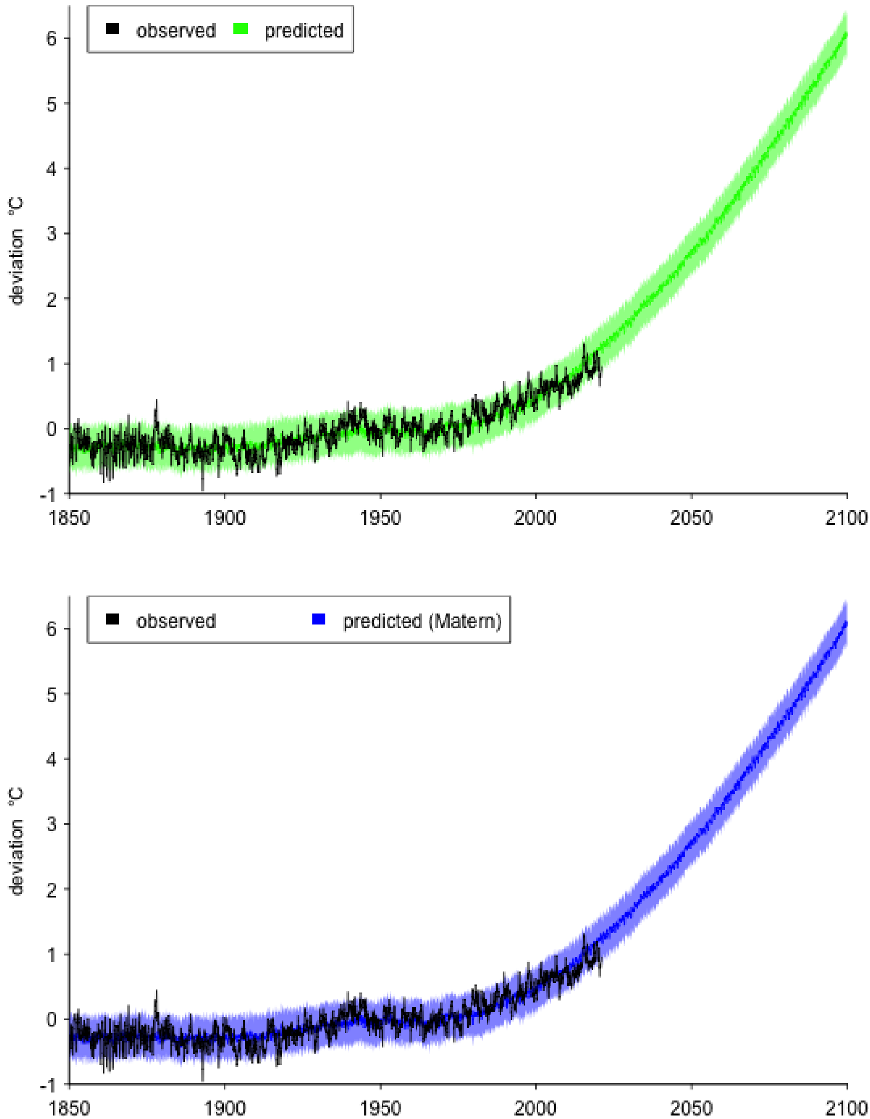

Figure 6. *The top panel recreates Figure 2 in the manuscript, where we predicted the observed global mean surface air temperature through 2100. We did so using our methodology, which employed a squared exponential kernel in its Gaussian process. The bottom panel depicts similar analysis with a Matérn* $5/2$ *kernel employed instead. These results, and similar for LOO analysis, did not discover a compelling difference between the two kernels, and we opted for the simpler squared exponential.*

References

[1] Christopher K Williams and Carl Edward Rasmussen. *Gaussian processes for machine learning*. Vol. 2. 3. MIT press Cambridge, MA, 2006.