The Community Land Model, version 5: Parameter Perturbation Experiment

D. Kennedy [1], Katherine Dagon[1], David M. Lawrence[1]

[1]Climate and Global Dynamics Laboratory, NCAR, Boulder, CO, USA.

Key Points:

- enter point 1 here
- enter point 2 here
- enter point 3 here

Corresponding author: Daniel Kennedy, djk2120@ucar.edu

Abstract
[enter your Abstract here]

1 Introduction

Water availability, land temperature extremes, fire risk, and crop productivity will all see impacts from climate change, and are among the many variables represented within the land component of Earth System Models. Understanding how such states respond to carbon dioxide concentration, and how we expect carbon dioxide concentration to evolve over time, is at the center of climate change research. Our certainty in climate model projections varies by domain and generally decreases with extended time horizons. Centennial-scale estimates of the cumulative terrestrial carbon sink have been especially challenging, with high uncertainty persisting across model generations (?, ?, ?). A portion of this uncertainty is irreducible, inherent to the challenge of predicting vegetation dynamics in a novel climate (?, ?). But with the ever-increasing observational basis from remote sensing, meteorological stations, flux towers, and field campaigns, we would expect predictions to become more skillful over time.

While model inter-comparison projects have had tremendous utility, it can be difficult to interpret the differences between models, or even between subsequent versions of the same model, due to the multiplicity of structural and parametric variations. Within the day-to-day model development process, there is a greater emphasis on hypothesis testing, substituting one parameterization at a time to understand the various consequences. However, the complexity of land models has increased to the point where understanding the full span of model function has become challenging (?, ?), and manual model calibration is becoming intractable (?, ?). Maintaining, improving, and interpreting complex land models benefits from thoughtful investment in software to automate and routinize important components of the development process, e.g. ? (?).

One such fundamental model development step is parameter sensitivity testing and/or parameter optimization. Because parameter values are uncertain, evaluating the effect of parameter values on model output is critical to the model development process (?, ?). When applied systematically, the result is a parameter perturbation ensemble (PPE), alternatively termed a perturbed physics ensemble. PPEs have played an important role in assessing projection uncertainty in climate models. They also form a basis for automated model calibration. Our goal in this project is to systematize parameter perturbation activity within our modeling framework, and develop the necessary tools and datasets to efficiently test parameter effects across the full suite of land model processes. In doing so, we have generated a large PPE, comprising thousands of parameter sensitivity tests. In this paper we present that dataset, how it was produced, and its potential applications.

We are working with the Community Land Model (CLM), the land component of the Community Earth System Model, which is developed and maintained at the National Center for Atmospheric Research (NCAR). This work follows up on ? (?), which began the process of systematically varying CLM parameters in service of automated model calibration. We extend that work by sampling a more exhaustive set of parameters across a broader range of land model processes. A set of 206 parameters were varied independently (one-at-a-time) across a range of climate scenarios. This dataset has already demonstrated great utility for diagnosing parameter effects, while the software and modeling infrastructure has greatly our expanded our capability to generate insights about CLM dynamics and uncertainties.

2 Experiment Description

2.1 Model description

The Community Terrestrial System Model (CTSM) is developed by the CESM Land Model Working Group and maintained at the National Center for Atmospheric Research (NCAR). This experiment utilizes the Community Land Model configuration of CTSM, version 5.1 (CLM5.1). The model source code and documentation are available online (`https://github.com/ESCOMP/CTSM`), as is a full model description (?, ?).

Relative to CLM5.0, version 5.1 includes minor bug fixes, parameter adjustments, and the implementation of biomass heat storage (?, ?). The PPE experiment required additional code modifications to programmatically vary the full suite of model parameters. The exact model code for this experiment is contained in a development tag (`https://github.com/ESCOMP/CTSM/tree/branch_tags/PPE.n11_ctsm5.1.dev030`). We utilized the biogeochemistry version of CLM in land-only mode, with the crop model turned off. The component set longname is:

`2000_DATM%GSWP3v1_CLM51%BGC_SICE_SOCN_SROF_SGLC_SWAV_SIAC_SESP`

2.2 Model spin-up

Model spin-up for the equilibration of carbon and nitrogen pools within biogeochemistry-enabled land models can consume up to 98% of computational time (?, ?). Depending on the evaluation criteria and model configuration, CLM5 requires between 800 and 2000 years (or more) to reach steady-state conditions (?, ?). Absent equilibrium, the drift towards steady state can obscure important model dynamics or features. For this reason, each member of the PPE requires independent spin-up.

To manage computational cost we leveraged the Matrix-CN spin-up mode recently implemented within CLM (?, ?). This new module utilizes a linearized simplification of CLM's biogeochemistry to significantly reduce spin-up time. Our spin-up protocol featured 20 years in accelerated decomposition mode, followed by 80 years of Matrix-CN, followed by 40 years of 'normal' mode, cycling over a ten-year forcing dataset. This was designed to achieve sufficiently equilibrated model states, while minimizing computational time. This spin-up methodology did not always reach full equilibration of deep soil carbon (beyond 1 meter). Certain inferences about deep soil carbon would therefore be subject to uncertainty due to spin-up concerns.

2.3 Sparsegrid

Another control on model cost is resolution. Most CLM simulations utilize nominal 1° resolution, which requires over 20,000 land grid cells. In order to manage computational cost, parameter perturbation experiments often use lower resolution, such as 4°x5° (?, ?). We used a clustering algorithm to achieve an alternative low resolution configuration.

Multivariate spatio-temporal clustering (MVSC) has been utilized to extract patterns of climatological significance from climate model output (?, ?) and applied to design a representativeness-based sampling network (?, ?). Instead of lowering resolution by coarsening a rectilinear grid, we used MVSC to strategically remove redundant grid cells, leaving only 400 grid cells that efficiently sample important model dynamics.

We used k-means clustering to identify groups of grid cells with similar climates based on a 2° transient simulation (1850-2014) using the CLM-PPE codebase. We selected one representative grid cell from each cluster to stand in for the entire cluster. The representative grid cell is whichever is located nearest the cluster centroid in climate space. The set of representative grid cells comprise a 'sparsegrid', which are used in lieu of a 'coarse' grid. A paint-by-number approach is used to recompose mapped output and compute global means, where the output from the representative grid cell is substituted for all members of the cluster cohort.

Clustering was based on a subset of 18 meaningful CLM variables (Table 1). The clustering algorithm analyzed 12 observations of each variable per grid cell, namely the mean and interannual variability computed for six 30-year climatology windows (1865-1894, 1895-1924, ... , 1985-2014). Clusters were delineated to equalize the multi-dimensional variance across the user-specified number of groups, k. We tested 15 values of k, ranging from 10 to 800. Based on ILAMB2.5 benchmarking (?, ?) against the full grid output, we opted for a 400-cluster sparsegrid, to balance computational cost against model fidelity (Supp Figure A1). Because our emphasis is on vegetated regions, we masked out Antarctica within the clustering algorithm, whereby we do not provide any output below 60°S.

Table 1. Clustering inputs

Climate forcing variables	Ecosystem state variables	Ecosystem flux variables
2m air temperature (TSA)	Leaf area index (TLAI)	Gross primary production (GPP)
Atmospheric rain (RAIN)	Ecosystem carbon (TOTECOSYSC)	Heterotrophic respiration (HR)
Atmospheric snow (SNOW)	Ecosystem nitrogen (TOTECOSYSN)	Autotrophic respiration (AR)
2m specific humidity (Q2M)	Soil ice (TOTSOILICE)	Net biome production (NBP)
Solar radiation (FSDS)	Soil liquid water (TOTSOILLIQ)	Total liquid runoff (QRUNOFF)
	Snow cover fraction (FSNO)	Sensible heat (FSH)
		Latent heat (EFLX_LH_TOT)

2.4 Experimental Design

In the preparation for this experiment, 206 CLM-BGC parameters were identified. We decided to vary each parameter independently, to a low and high value. To define parameter ranges we created an online spreadsheet and solicited domain-area experts to provide a minimum and maximum value for each parameter. In some cases literature values were directly utilized, but in the majority of cases, expert judgment was used. The spreadsheet, with literature references and parameter descriptions is available online and in appendix zqz.

Each simulation comprised 150 years, with the first 140 for spin-up, followed by a 10-year period for analysis. We opted for six different forcing scenarios to understand the intersection of parameter effects and climate change (Table 2). The GSWP3v1 reanalysis product (`http://hydro.iis.u-tokyo.ac.jp/GSWP3/`) served as our atmospheric forcing, and is the default forcing data for CLM5 (?, ?). We applied climate and CO_2 anomalies independently, in order to disentangle their effects on parameter rankings. Future and pre-industrial climate forcing datasets were prepared by adding GSWP3v1 anomalies from 2005-2014 to a derived climate change signal. We inferred the mean climate change signal using the CESM2 large ensemble experiment (?, ?), computed as the average of the difference between the period of interest and present day for the six atmospheric forcing variables.

2.5 Analyses

Biome analysis

Table 2. Forcing Scenarios

Name	Meteorology	CO_2 (ppmv)	N addition	Description
CTL2010	2005-2014	367	-	control experiment
C285	2005-2014	285	-	low CO_2
C867	2005-2014	867	-	high CO_2
AF1855	1851-1860	367	-	pre-industrial climate
AF2095	2091-2100	367	-	late century climate (SSP3-7.0)
NDEP	2005-2014	367	$5g/m^2$	enhanced nitrogen deposition

3 Results

A major goal of this project was to develop a fast configuration of CLM5-BGC that would allow for a large number of simulations given the available computational resources. Combining the CN-Matrix spinup approach with our sparsegrid formulation (see Section 2.3 for details) yielded a configuration approximately 500 times faster than the 1-degree configuration most often used for CLM simulations (Figure 1). There are 22648, 5666, and 1764 land gridcells in standard 1°, 2°, and 4°x5° CLM simulations, respectively, as compared to just 400 grid cells in the sparsegrid. Likewise, whereas previous spin-up methodologies required 1500 years or longer to satisfy equilibration criteria, the CN-matrix approach yielded satisfactory spin-up for our experiment within 140 years.

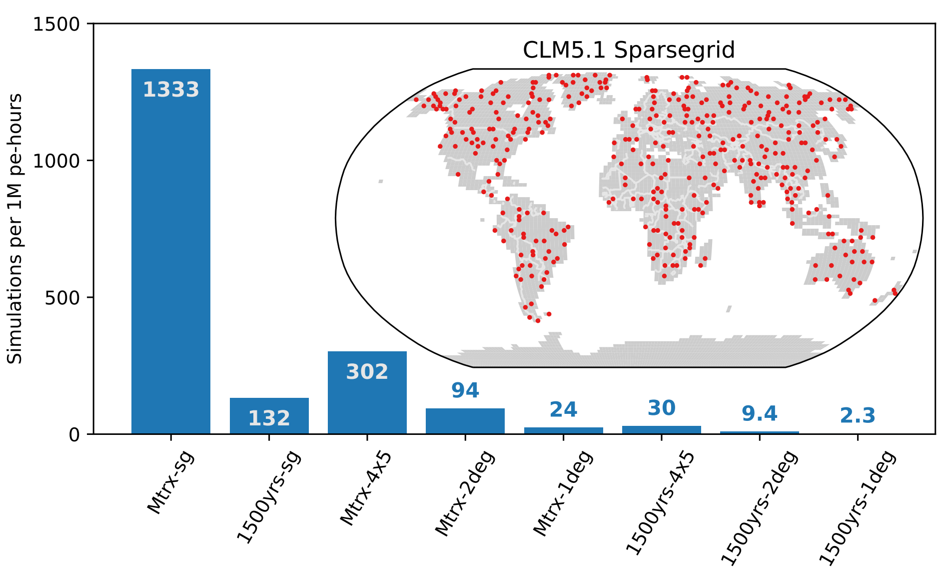

Figure 1. The approximate number of simulations afforded by 1 million core-hours for a range of CLM configurations. Configurations are labeled according to spin-up procedure and resolution, with 'sg' signifying sparsegrid. The inset map shows the locations of the 400 sparse grid cells. See Section 2 for spin-up and sparsegrid details. (My speed calculations here are a bit janky, assume linear scaling of cost for increasing time or space zqz)

Choosing the number of clusters to generate the sparsegrid involved balancing the computational savings against representational fidelity. Generally the more clusters, the stronger the relationship between output from the full grid and the sparsegrid. Accuracy of global photosynthesis could be achieved with a relatively small number of clusters, with R_2 >0.95 achieved with only 200 clusters, and RMSE<1PgC requiring approximately 700 clusters (Figure 2). We chose 400 clusters, based on ILAMB scores of a wide variety of output variables (Supp Figure A1)

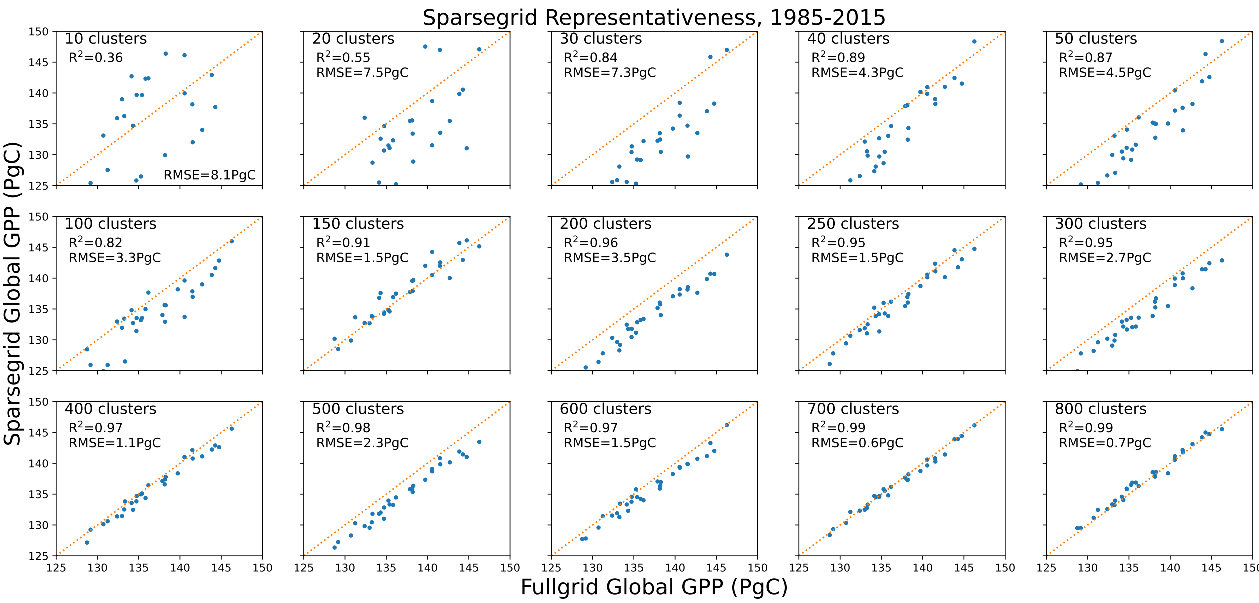

Figure 2. Sparsegrid vs fullgrid (2° resolution) global annual GPP across the last thirty years of a transient CLM5.1 simulation. We opted for 400 clusters to balance computational cost against representativeness.

Parameters were varied across a wide variety of CLM subcomponents (Figure 3). A total of 206 parameters were targeted for sensitivity testing. Many parameters were hard-coded, but have now been extracted to the CLM parameter file for easier manipulation. In some cases new parameters were introduced to allow perturbation of important processes. For example a sand perturbation factor ($sand\_pf$) was introduced in order to perturb the percent sand in the soil column without creating a new surface dataset. A spreadsheet detailing all of the parameters, and their ranges, is provided in the supplementary material and available online (zqz link).

4 Discussion

Major discussion points:

- need for speed
- parameter effects can be quite large
- parameter priors are difficult to establish
- mean and climate response variables may vary
- community resource
- invest in infrastructure

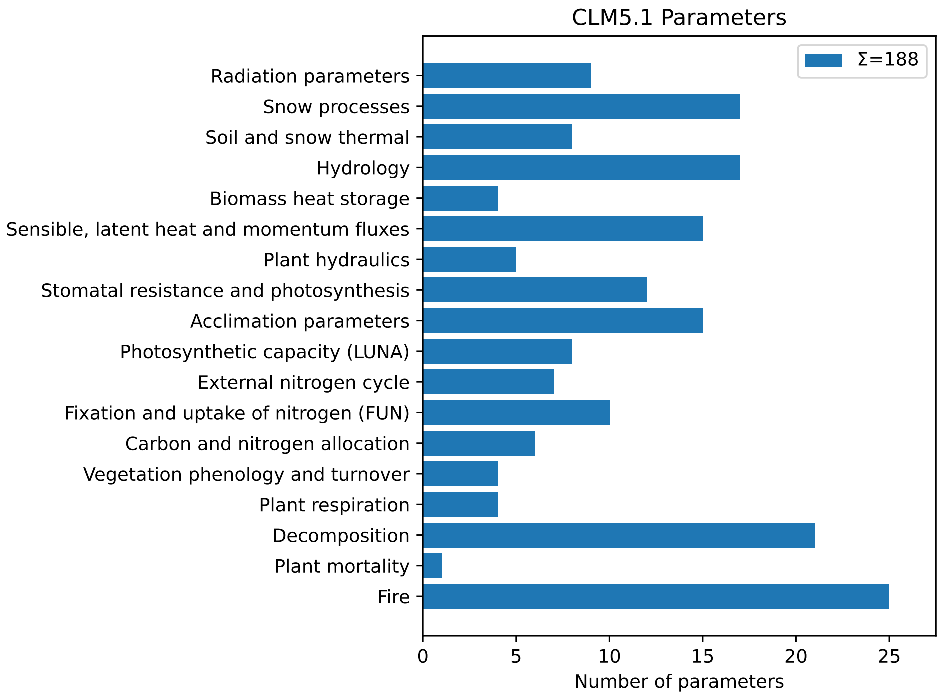

Figure 3. 206 parameters were identified and perturbed across the various domains of the land model.

5 Conclusion

Open Research Section

This section MUST contain a statement that describes where the data supporting the conclusions can be obtained. Data cannot be listed as "Available from authors" or stored solely in supporting information. Citations to archived data should be included in your reference list. Wiley will publish it as a separate section on the papers page. Examples and complete information are here: https://www.agu.org/Publish with AGU/Publish/Author Resources/Data for Authors

Acknowledgments

Enter acknowledgments here. This section is to acknowledge funding, thank colleagues, enter any secondary affiliations, and so on.

References

Collier, N., Hoffman, F. M., Lawrence, D. M., Keppel-Aleks, G., Koven, C. D., Riley, W. J., ... Randerson, J. T. (2018). The International Land Model Benchmarking (ILAMB) system: Design, theory, and implementation. *Journal of Advances in Modeling Earth Systems*, *10*(11), 2731-2754. Retrieved from `https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2018MS001354` doi: https://doi.org/10.1029/2018MS001354

Dagon, K., Sanderson, B. M., Fisher, R. A., & Lawrence, D. M. (2020). A machine learning approach to emulation and biophysical parameter estimation with the Community Land Model, version 5. *Advances in Statistical Climatology, Meteorology and Oceanography*, *6*(2), 223–244. Retrieved

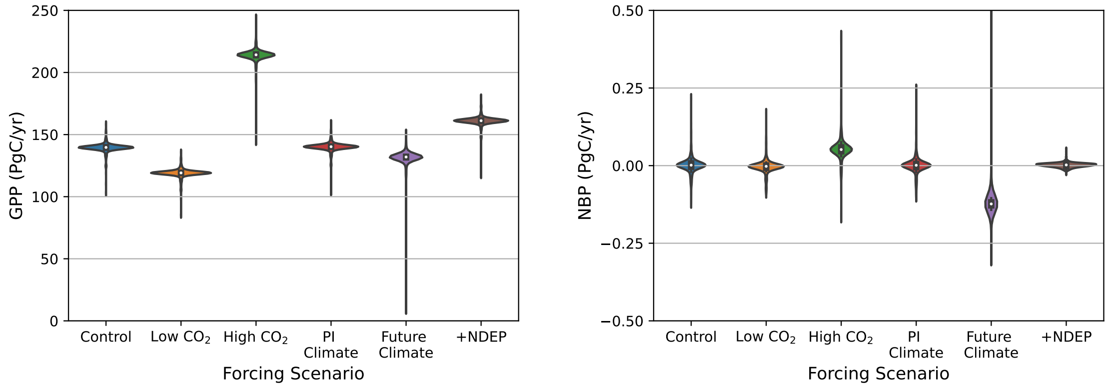

Figure 4. Distributions of global annual gross primary production (GPP) and net biome production (NBP) across the six forcing scenarios. See Section 2.4 for scenario descriptions. The distributions are concentrated around the default parameterization with long tails, indicating that a small proportion of parameters have large effects. Especially in the case of NBP, parameter effects can be larger than scenario effects.

from https://ascmo.copernicus.org/articles/6/223/2020/ doi: 10.5194/ascmo-6-223-2020

Hoffman, F. M., Hargrove, W. W., Erickson, D. J., & Oglesby, R. J. (2005). Using clustered climate regimes to analyze and compare predictions from fully coupled general circulation models. *Earth Interactions*, *9*(10), 1 - 27. Retrieved from https://journals.ametsoc.org/view/journals/eint/9/10/ei110.1.xml doi: https://doi.org/10.1175/EI110.1

Hoffman, F. M., Kumar, J., Mills, R. T., & Hargrove, W. W. (2013). Representativeness-based sampling network design for the state of alaska. *Landscape Ecology*, *28*(8), 1567–1586. Retrieved from https://doi.org/10.1007/s10980-013-9902-0 doi: 10.1007/s10980-013-9902-0

Lawrence, D. M., et al., & et al. (2019). The Community Land Model version 5: Description of new features, benchmarking, and impact of forcing uncertainty. *Journal of Advances in Modeling Earth Systems*, *in press*. doi: 10.1029/2018MS001583

Lu, X., Du, Z., Huang, Y., Lawrence, D., Kluzek, E., Collier, N., . . . Luo, Y. (2020). Full implementation of matrix approach to biogeochemistry module of clm5. *Journal of Advances in Modeling Earth Systems*, *12*(11), e2020MS002105. doi: https://doi.org/10.1029/2020MS002105

Rodgers, K. B., Lee, S.-S., Rosenbloom, N., Timmermann, A., Danabasoglu, G., Deser, C., . . . Yeager, S. G. (2021). Ubiquity of human-induced changes in climate variability. *Earth System Dynamics*, *12*(4), 1393–1411. Retrieved from https://esd.copernicus.org/articles/12/1393/2021/ doi: 10.5194/esd-12-1393-2021

Sun, Y., Goll, D. S., Huang, Y., Ciais, P., Wang, Y.-P., Bastrikov, V., & Wang, Y. (2023). Machine learning for accelerating process-based computation of land biogeochemical cycles. *Global Change Biology*, *n/a*(n/a). doi: https://doi.org/10.1111/gcb.16623

Swenson, S. C., Burns, S. P., & Lawrence, D. M. (2019). The impact of biomass heat storage on the canopy energy balance and atmospheric stability in the community land model. *Journal of Advances in Modeling Earth Systems*, *11*(1), 83-98. doi: https://doi.org/10.1029/2018MS001476

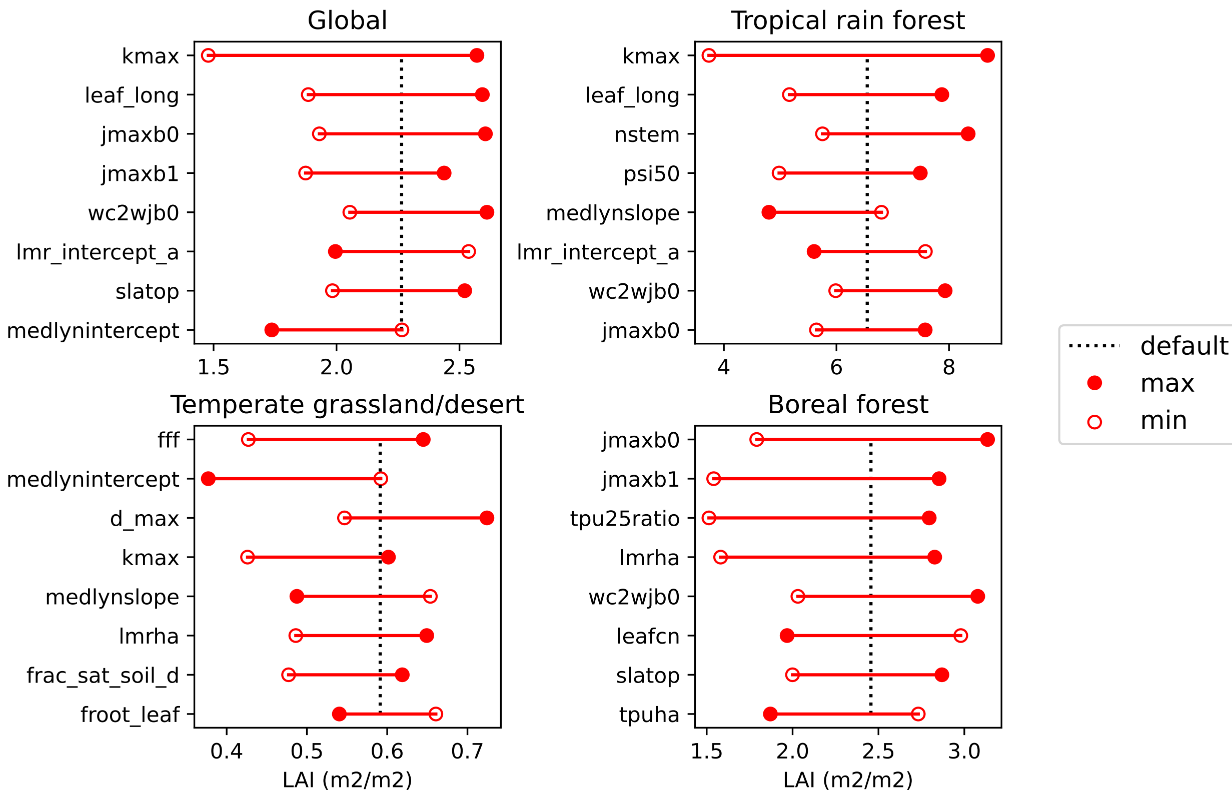

Figure 5. The eight most influential parameters on leaf area index within the CTL2010 ensemble, globally and within three biomes.

224 # Appendix A Supplementary Figures

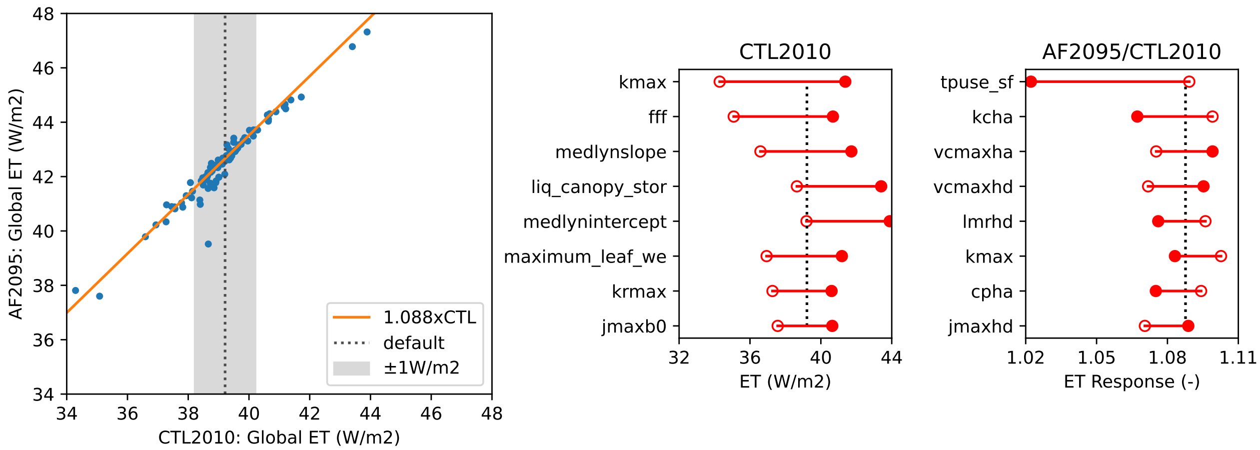

Figure 6. (a) Global evapotranspiration in the warm forcing scenario vs. our control experiment. The default parameterization ET enhancement is +8.8%, as shown with the orange line. The shading spans the CTL2010 default ET, plus or minus 1 W/m2. (b) The top 8 parameters governing global ET in the CTL2010 experiment. (c) The top 8 parameters governing ET response to the future climate anomaly forcing (AF2095). Parameters governing the response to temperature anomalies tend to differ from the parameters controlling present-day ET.

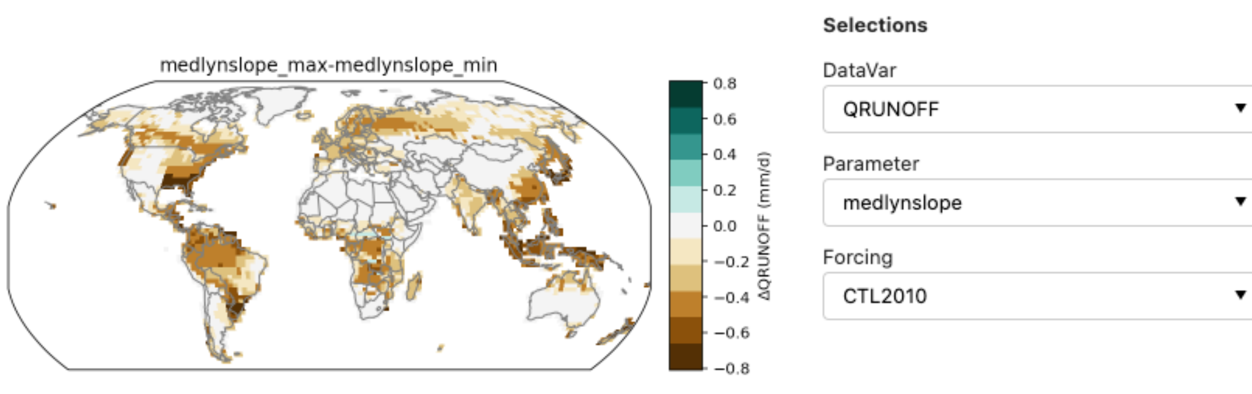

Figure 7. Screenshot of interactive diagnostic for exploring maps of parameter effects. In this case, the effect of medlynslope on runoff within the CTL2010 ensemble. Increasing medlynslope tends to increase transpiration and reduce runoff.

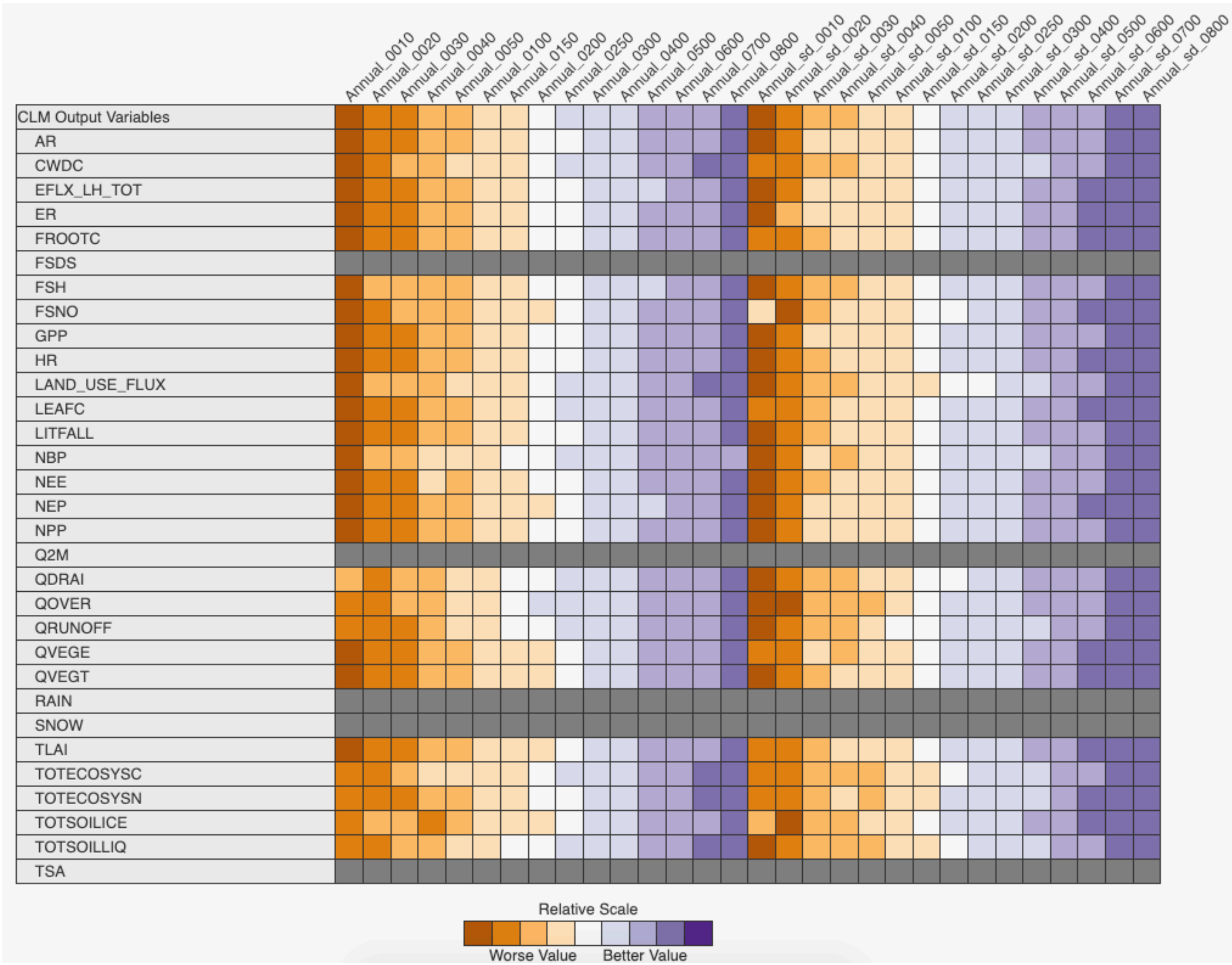

Figure A1. ILAMB 2.5 overall scores comparing remapped sparsegrid output to the full grid 2° model output. Column headings indicate the number of clusters, and whether annual means or annual means and standard deviations were used as input to the clustering algorithm. We should try to remove the gray bits (zqz). Full, interactive results are available at `www.ilamb.org/PPE/CLM/2021-02`. See main text Section 2.3 for clustering details.