A Dataset on subgrid surface microclimate and energy balance under the $4\times CO_2$ scenario

Tao Tang, Keer Zhang, Xuhui Lee,

School of the Environment, Yale University, New Haven, CT, USA

Corresponding author: Tao Tang, tao.tang@yale.edu

Abstract

This dataset was produced by a simulation of instant quadrupling of preindustrial CO_2 concentration with the CESM Version 2 (CESM2) global climate model. In this simulation, each plant functional type was assigned with an independent soil column, differing from the shared soil column configuration in the default CESM2 configuration. The dataset contain key variables produced by the land surface model of CESM2, at both grid and subgrid scales with monthly resolution. The data are archived in the netCDF4 format and are freely downloadable from Harvard Dataverse (https://doi.org/10.7910/DVN/LQXAT3). The model setup, simulation and output are described in this document. An application of this dataset can be found at Tang et al. (2022).

1. Model description

The CESM Version 2 (CESM2) is a fully coupled global climate model developed by the National Center for Atmospheric Research (NCAR) (Danabasoglu et al., 2020). It consists of models of the atmosphere, the ocean, sea-ice, the land, rivers and ocean waves. In the Community Land Model version 5 (CLM5), the land component of CESM2 model (Lawrence et al., 2019), the land surface is represented as a nested hierarchy of subgrid levels. The first subgrid level is the land unit, including vegetation, lake, urban, glaciers and crop, with a fraction assigned to each land unit. All subgrid land unit in a gridcell receive the same amount of atmospheric forcing, such as precipitation and incoming radiation. The second subgrid level is the column. The third subgrid level is plant functional type (PFT), which captures the biophysical and biogeochemical differences between broad categories of plants. In the default CLM5 configuration, the vegetated land unit is assigned a single column whereby all the PFTs in the unit share the same soil column in terms of water, nutrient use and soil heat exchange. Recent studies have shown that this shared soil column configuration can lead to unrealistic ground heat fluxes because a common soil temperature is artificially maintained for all PFTs (Schultz et al., 2016; Meier et al., 2018). In this simulation, each PFT in the vegetation unit was assigned its own soil column, following the method described by Schultz et al. (2016) (Figure 1). For the detailed description of CLM5, please refer to Lawrence et al. (2018).

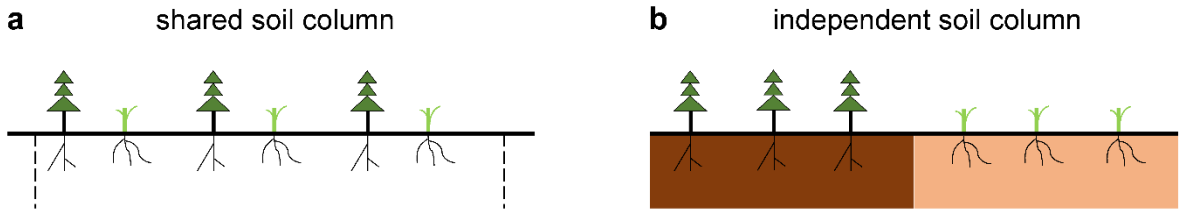

Figure 1. Schematic diagram of two plant functional types in shared soil column (a, default CLM5 setting) and independent soil column configuration (b, this simulation).

2. Model simulation

A horizontal resolution of 0.9° × 1.25° (lat × lon) was used in the model simulation and an independent soil column was assigned to each PFT in the vegetated land unit. The surface condition is held constant at the year of 1850. The spin-up of this experiment was initiated with

the active CLM5 model forced by GSWP3 atmospheric forcing data and was run for 160 years. Then the model was run for another 20 years in fully coupled mode as the control simulation. The perturbation was an instant quadrupling of the preindustrial CO_2 concentration ($4\times CO_2$, ~1140 ppm) with everything else being the same as control simulation, including the land use and land cover. The simulation was run in coupled mode for 119 years, with irrigation activity turned off.

3. Data description and availability

Both grid mean and subgrid output from CLM5 are archived. The subgrid data were aggregated to the five land units of CLM5 (vegetated, lake, ice, crop and urban). The vegetated land unit is similar to the primary and secondary land (psl) in the LUMIP protocol (Lawrence et al., 2016). The fractions of the five land units relative to gridcell is shown in Figure 2. The urban tile is the combination of three urban districts, tall building district, high density and medium density (Figure 3).

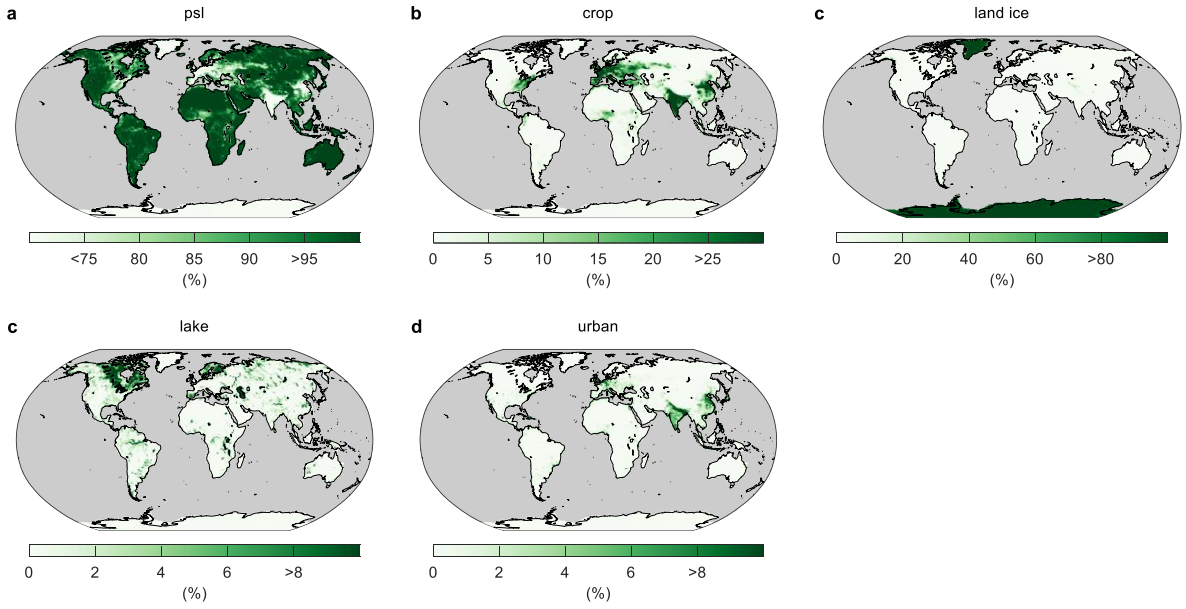

Figure 2. Fractions of each land unit relative to gridcell.

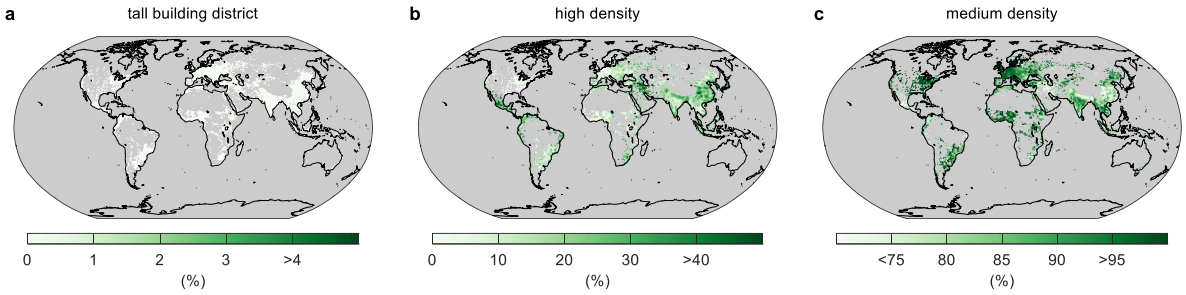

Figure 3. Fractions of each urban district relative to urban tile.

Table 1 lists the variables archived in this dataset, which are all commonly used variables in atmospheric and climate sciences, including temperature, humidity, circulation, components of the energy budget and the hydrological cycle. The dataset covers both the 20-year control simulation and the 119-year quadrupling experiment at monthly intervals. On top of these variables, data on some surface properties are also provided (Table 2). It is noted that there are 25 soil layers in total, but only the upper 20 layers are hydrologically and biogeochemically active. The deepest 5 layers are only included in the thermodynamical calculations (Lawrence et al., 2018). As a result, the output for *volumetric soil water (vsw)* has 20 layers only. The lake depth is shown in Figure 4. The land fraction is the area of land relative to the area of the whole gridcell.

Table 1. Output variable list, × indicates the output is available.

Variable	Short name	Unit	Grid	Subgrid
2 m air temperature	tas	K	×	×
2 m vapor pressure	huv	Pa	×	×
2 m wet-bulb temperature	wba	℃	×	×
Maximum 2 m air temperature	tasmax	K	×	×
Minimum 2 m air temperature	tasmin	K	×	×
Skin temperature	ts	K	×	×
Atmospheric air temperature at blending height	tab	K	×	
Latent heat flux	hfls	W m^{-2}	×	×
Sensible heat flux	hfss	W m^{-2}	×	×
Ground heat flux	hfdsl	W m^{-2}	×	×
Surface upwelling shortwave radiation	rsus	W m^{-2}	×	×
Surface upwelling longwave radiation	rlus	W m^{-2}	×	×
Surface downwelling shortwave radiation	rsds	W m^{-2}	×	
Surface downwelling longwave radiation	rlds	W m^{-2}	×	
Surface downwelling shortwave radiation clear-sky	rsdscs*	W m^{-2}	×	
Surface downwelling longwave radiation clear-sky	rldscs*	W m^{-2}	×	
Precipitation	pr	mm day^{-1}	×	
Atmospheric pressure at surface	ps	Pa	×	
Specific humidity at blending height	husb	kg kg^{-1}	×	
10 m wind speed	sfcWind	m s^{-1}	×	
Total runoff	mrro	mm day^{-1}	×	
surface runoff	mrros	mm day^{-1}	×	
Soil moisture in the upper 10 cm of soil (liquid + ice)	mrsos	kg m^{-2}	×	×
Volumetric soil water	vsw	mm^3 mm^{-3}	×	×
Total water storage	tws	mm	×	

*Note: rldscs and rsdscs are not available for the control run.

Table 2. Output for surface properties.

Property	Unit
Grid cell areas	km^2
Land unit fraction on grid cell	unitless
Soil layer thickness	m
Lake depth	m
Land/sea mask	1 for land; NaN for sea
Land fraction	unitless

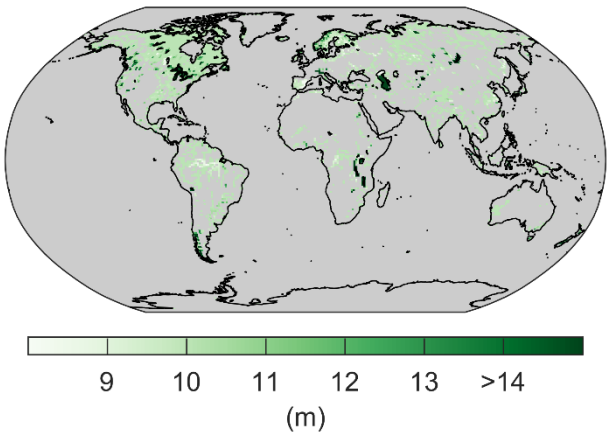

Figure 4. Lake depth.

The data is provided in the NetCDF-4 format. The surface properties are stored in one single file named as '*sfc_property.nc*'. The other variables are provided with one variable per file. Most of subgrid data are stored in one file per variable. However, some variables are split into several files based on simulation years due to the limitation of file size. The naming convention for the files is '*CO2x4/Control_shortname_gridmean/subgrid_yr.nc*'. As an example, file '*CO2x4_tas_subgrid_yr001_yr060.nc*' is the file storing subgrid 2-m air temperature of the five land units in the first 60 years of perturbation. Exceptions apply to *volumetric soil water (vsw)*: the subgrid data for the 119-year experiment are further split into psl and crop due to the file size of the data. The whole dataset is open-access and can be freely downloaded from the

Harvard Dataverse (https://doi.org/10.7910/DVN/LQXAT3). The total file size is roughly 99 GB, with 31 GB of grid mean data and 68 GB of subgrid data. The model code of the soil configuration is also available upon request.

Acknowledgments

We acknowledge the support from the US National Science Foundation (Grant AGS1933630 to XL and NS). High-performance computing support from Cheyenne (doi:10.5065/D6RX99HX) was provided by the NCAR's Computational and Information System Laboratory, sponsored by the National Science Foundation and other agencies.

REFERENCES

Danabasoglu, G., Lamarque, J. F., Bacmeister, J., Bailey, D. A., DuVivier, A. K., Edwards, J., Emmons, L. K., Fasullo, J., Garcia, R., Gettelman, A., Hannay, C., Holland, M. M., Large, W. G., Lauritzen, P. H., Lawrence, D. M., Lenaerts, J. T. M., Lindsay, K., Lipscomb, W. H., Mills, M. J., Neale, R., Oleson, K. W., Otto-Bliesner, B., Phillips, A. S., Sacks, W., Tilmes, S., van Kampenhout, L., Vertenstein, M., Bertini, A., Dennis, J., Deser, C., Fischer, C., Fox-Kemper, B., Kay, J. E., Kinnison, D., Kushner, P. J., Larson, V. E., Long, M. C., Mickelson, S., Moore, J. K., Nienhouse, E., Polvani, L., Rasch, P. J., & Strand, W. G. (2020). The community earth system model version 2 (cesm2). *Journal of Advances in Modeling Earth Systems, 12*(2), e2019MS001916. doi:10.1029/2019MS001916

Lawrence, D., Fisher, R., Koven, C., Oleson, K., Swenson, S., & Vertenstein, M. (2018). *Technical description of version 5.0 of the community land model (clm)*. Retrieved from https://www.cesm.ucar.edu/models/cesm2/land/CLM50_Tech_Note.pdf:

Lawrence, D. M., Fisher, R. A., Koven, C. D., Oleson, K. W., Swenson, S. C., Bonan, G., Collier, N., Ghimire, B., van Kampenhout, L., Kennedy, D., Kluzek, E., Lawrence, P. J., Li, F., Li, H., Lombardozzi, D., Riley, W. J., Sacks, W. J., Shi, M., Vertenstein, M., Wieder, W. R., Xu, C., Ali, A. A., Badger, A. M., Bisht, G., van den Broeke, M., Brunke, M. A., Burns, S. P., Buzan, J., Clark, M., Craig, A., Dahlin, K., Drewniak, B., Fisher, J. B., Flanner, M., Fox, A. M., Gentine, P., Hoffman, F., Keppel-Aleks, G., Knox, R., Kumar, S., Lenaerts, J., Leung, L. R., Lipscomb, W. H., Lu, Y., Pandey, A., Pelletier, J. D., Perket, J., Randerson, J. T., Ricciuto, D. M., Sanderson, B. M., Slater, A., Subin, Z. M., Tang, J., Thomas, R. Q., Val Martin, M., & Zeng, X. (2019). The community land model version 5: Description of new features, benchmarking, and impact of forcing uncertainty. *Journal of Advances in Modeling Earth Systems, 11*(12), 4245-4287. doi:10.1029/2018MS001583

Lawrence, D. M., Hurtt, G. C., Arneth, A., Brovkin, V., Calvin, K. V., Jones, A. D., Jones, C. D., Lawrence, P. J., de Noblet-Ducoudré, N., Pongratz, J., Seneviratne, S. I., & Shevliakova, E. (2016). The land use model intercomparison project (lumip) contribution to cmip6: Rationale and experimental design. *Geosci. Model Dev., 9*(9), 2973-2998. doi:10.5194/gmd-9-2973-2016

Meier, R., Davin, E. L., Lejeune, Q., Hauser, M., Li, Y., Martens, B., Schultz, N. M., Sterling, S., & Thiery, W. (2018). Evaluating and improving the community land model's sensitivity to land cover. *Biogeosciences, 15*(15), 4731-4757. doi:10.5194/bg-15-4731-2018

Schultz, N. M., Lee, X., Lawrence, P. J., Lawrence, D. M., & Zhao, L. (2016). Assessing the use of subgrid land model output to study impacts of land cover change. *Journal of Geophysical Research: Atmospheres, 121*(11), 6133-6147. doi:10.1002/2016JD025094

Tang, T., Lee, X., Zhang, K., Cai, L., Lawrence, D. M., & Shevliakova, E. (2023). Biophysical Impact of Land-Use and Land-Cover Change on Subgrid Temperature in CMIP6 Models, *Journal of Hydrometeorology*, 24(3), 373-388, doi:10.1175/JHM-D-22-0073.1