The Journal of Open Source Software

helios: An R package to process heating and cooling degrees for GCAM

Mengqi Zhao [1], **Zarrar Khan** [2], **Kalyn Dorheim** [2], **and Chris Vernon** [1]

1 Pacific Northwest National Laboratory, Richland, WA, USA **2** Joint Global Change Research Institute, Pacific Northwest National Laboratory, College Park, MD, USA

DOI: 10.xxxxxx/draft

Software

- Review
- Repository
- Archive

Editor: Open Journals

Reviewers:

- @openjournals

Submitted: 01 January 1970
Published: unpublished

Summary

`helios` is an open-source R package that estimates population-weighted heating and cooling degree-hours (HDH and CDH) and degree-days (HDD and CDD) at various temporal (e.g., energy dispatch segments, monthly, yearly) and spatial scales (e.g., U.S. states, global political regions, countries). The degree hour and degree day outputs from `helios` are used to inform electricity demand load in the Global Change Analysis Model (GCAM) (Calvin et al., 2019) as well as in GCAM-USA (which is the version of GCAM with U.S. state-level details) (Binsted et al., 2022). `helios` uses a workflow with four steps: processing raw data; calculating heating and cooling degrees; visualizing performance diagnostics; and outputing results in various formats. There are two sources of widely-used climate data compatible with `helios`: (1) hourly climate data with 12-km resolution that are dynamically downscaled with the Weather Research and Forecasting (WRF) model and projected using a thermal global warming (TGW) approach (Jones et al., 2022); and (2) daily climate data with 0.5-degree resolution from the Coupled Model Intercomparison Project (CMIP) that is bias-adjusted and statistical downscaled by the Inter-Sectoral Impact Model Intercomparison Project (ISIMIP). In summary, `helios` is a model that standardizes methodology of heating and cooling degrees-hours and degree-days using publicly available data and advances the understanding of the impact of spatial and temporal temperature variability on building energy services.

Statement of Need

`helios` was developed to meet the increasing research interests to explore the spatial and temporal heterogeneity of climate impacts on sub-annual electricity demand from buildings. Ciscar & Dowling (2014) pointed out most integrated modeling systems designed to link human-Earth systems are unable to take advantage of the publicly available high resolution data to account for the impact of seasonal temperature change on energy system. To better fill in this gap, researchers have developed GCAM versions (e.g., GCAM-USA) to include power sector details at sub-annual and sub-national level (Wise et al., 2019). For example, Khan et al. (2021) used GCAM-USA to show that the temperature-induced heating and cooling demands can significantly affect sub-annual electricity demand profiles and peak electricity loads. Understanding the seasonal dynamics of electricity demand and capacity within multi-sector dynamics models is of importance to support future infrastructure planning (Binsted, 2022). We develop `helios` to bridge the gap between high resolution data and global scale models by facilitating the workflow in estimating population-weighted heating and cooling degrees. `helios` serves as a pre-processing tool of GCAM for researchers to capture the impact of sub-annual variation of different climate and socioeconomic scenarios on building energy demand.

Statement of Field

Heating and cooling degree days are commonly used as meteorological indices in the energy system to measure the temperature deviations from the reference temperature over time. These indices are widely calculated at point scale rather than spatial scale. With the increasing availability of spatially distributed climate data, few tools are developed to access and post-process the raw climate data format into tabular format, such as `Climate4R` (Iturbide et al., 2019). However, there are rarely well-documented and open-source tools that streamline the calculation of population-weighted HDD and CDD at user-defined spatiotemporal resolutions or electricity dispatch sectors defined in GCAM, directly using gridded climate and population data. `helios` is developed to integrate these workflows and standardize the output for easy usage within and beyond GCAM applications.

Design and Functionality

`helios` is designed to provide heating and cooling degrees to GCAM (or GCAM-USA) at two spatiotemporal scales: (1) HDH and CDH for the U.S. States for dispatch segments by building thermal service (for GCAM-USA); (2) annual HDD and CDD for GCAM's 32 geopolitical regions at 5-year time step by building thermal service (for GCAM-Regions). Beyond providing information for GCAM, `helios` can serve as a general tool to calculate heating and cooling degrees at a monthly time step and various spatial scales (e.g., country, basin). The use of `helios` requires users to provide information about the input files, such as the climate model source for climate data and the variable name for temperature. For example, the ISIMIP-CMIP data uses "tas" for the variable name for temperature while the WRF data uses "T2". Figure 1 shows the workflow for both GCAM-USA and GCAM-Regions. More details can be accessed in the `helios` documentation page on Github.

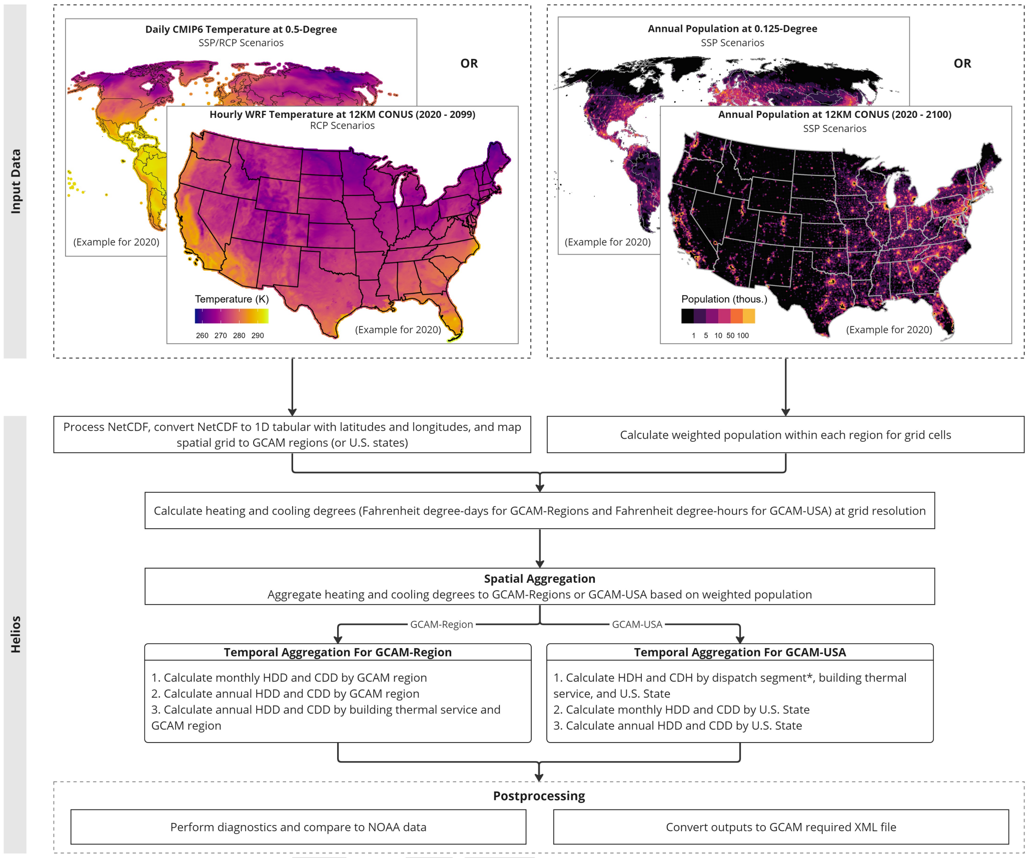

Figure 1: An example of the helios workflow using two types of input datasets (e.g., global data from CMIP6 and CONUS data from WRF). This demonstration showcases helios' capability to generate heating and cooling degrees by GCAM region or U.S. States, among other spatiotemporal scales.

Working with climate data can pose challenges, given large data sizes and diverse formats, spatiotemporal resolutions, data structures, and dimensions involved. The `helios` package provides functionalities that make it more convenient for users to manipulate climate data. `helios` provides easier access to various climate data types in a simplified format, facilitates the calculation of heating and cooling degrees using a standardized methodology, and ensures quality control through detailed diagnostics. There are five main functions provided by `helios`:

(1) `helios::read_ncdf` processes complex climate data (e.g., NetCDF) and converts to tabular data with latitude and longitude.
(2) `helios::read_population` processes population data and converts to the same resolution as the climate data if needed.
(3) `helios::hdcd` calculates heating and cooling degree-hours and degree-days at various spatial and temporal scales.
(4) `helios::diagnostics` visualizes the outputs and compares with observation data if available.
(5) `helios::save_xml` converts outputs to XML files, which is a format required by GCAM to calculate building energy demand.

Acknowledgements

This research was supported by the U.S. Department of Energy, Office of Science, as part of research in MultiSector Dynamics, Earth and Environmental System Modeling Program.

References

Binsted, M. (2022). An electrified road to climate goals. *Nature Energy*, *7*(1), 9–10. https://doi.org/10.1038/s41560-021-00974-8

Binsted, M., Iyer, G., Patel, P., Graham, N. T., Ou, Y., Khan, Z., Kholod, N., Narayan, K., Hejazi, M., Kim, S., Calvin, K., & Wise, M. (2022). GCAM-USA v5.3_water_dispatch: Integrated modeling of subnational US energy, water, and land systems within a global framework. *Geoscientific Model Development*, *15*(6), 2533–2559. https://doi.org/10.5194/gmd-15-2533-2022

Calvin, K., Patel, P., Clarke, L., Asrar, G., Bond-Lamberty, B., Cui, R. Y., Di Vittorio, A., Dorheim, K., Edmonds, J., Hartin, C., Hejazi, M., Horowitz, R., Iyer, G., Kyle, P., Kim, S., Link, R., McJeon, H., Smith, S. J., Snyder, A., … Wise, M. (2019). GCAM v5.1: Representing the linkages between energy, water, land, climate, and economic systems. *Geoscientific Model Development*, *12*(2), 677–698. https://doi.org/10.5194/gmd-12-677-2019

Ciscar, J.-C., & Dowling, P. (2014). Integrated assessment of climate impacts and adaptation in the energy sector. *Energy Economics*, *46*, 531–538. https://doi.org/10.1016/j.eneco.2014.07.003

Iturbide, M., Bedia, J., Herrera, S., Baño-Medina, J., Fernández, J., Frías, M. D., Manzanas, R., San-Martín, D., Cimadevilla, E., Cofiño, A. S., & Gutiérrez, J. M. (2019). The r-based climate4R open framework for reproducible climate data access and post-processing. *Environmental Modelling &Amp; Software*, *111*, 42–54. https://doi.org/10.1016/j.envsoft.2018.09.009

Jones, A. D., Rastogi, D., Vahmani, P., Stansfield, A., Reed, K., Thurber, T., Ullrich, P., & Rice, J. S. (2022). *IM3/HyperFACETS thermodynamic global warming (TGW) simulation datasets*. MultiSector Dynamics-Living, Intuitive, Value-adding, Environment. https://doi.org/10.57931/1885756

Khan, Z., Iyer, G., Patel, P., Kim, S., Hejazi, M., Burleyson, C., & Wise, M. (2021). Impacts of long-term temperature change and variability on electricity investments. *Nature Communications*, *12*(1). https://doi.org/10.1038/s41467-021-21785-1

Wise, M., Patel, P., Khan, Z., Kim, S. H., Hejazi, M., & Iyer, G. (2019). Representing power sector detail and flexibility in a multi-sector model. *Energy Strategy Reviews*, *26*, 100411. https://doi.org/10.1016/j.esr.2019.100411