

¹ Heatwave Diagnostics Package: Efficiently Compute Heatwave Metrics Across Parameter Spaces

³ Cameron Cummins¹ and Geeta Persad¹

⁴ 1 Department of Earth and Planetary Sciences, Jackson School of Geoscience, The University of Texas
⁵ at Austin, Austin, TX, USA

DOI: [10.xxxxxx/draft](https://doi.org/10.xxxxxx/draft)

Software

- [Review ↗](#)
- [Repository ↗](#)
- [Archive ↗](#)

Editor: [Open Journals](#) ↗

Reviewers:

- [@openjournals](#)

Submitted: 01 January 1970

Published: unpublished

License

Authors of papers retain copyright and release the work under a Creative Commons Attribution 4.0 International License ([CC BY 4.0](#))

⁶ Summary

The heatwave diagnostics package (HDP) is a Python package that provides the climate research community with tools to compute heatwave metrics for the large volumes of data produced by earth system model large ensembles, across multiple measures of heat, extreme heat thresholds, and heatwave definitions. The HDP leverages performance-oriented design using xarray, Dask, and Numba to maximize the use of available hardware resources while maintaining accessibility through an intuitive interface and well-documented user guide. This approach empowers the user to generate metrics for a wide and diverse range of heatwave types across the parameter space.

¹⁵ Statement of Need

Accurate quantification of the evolution of heatwave trends in climate model output is critical for evaluating future changes in hazard. The framework for indexing heatwaves by comparing a time-evolving measure of heat against some seasonally-varying percentile threshold is well-established in the literature (Baldwin et al. (2019); Schoetter et al. (2015); Acero et al. (2024); Argüeso et al. (2016)). Metrics such as heatwave frequency and duration are commonly used in hazard assessments, but there are few centralized tools and no universal heatwave criteria for computing them. This has resulted in parameter heterogeneity across the literature and has prompted some studies to adopt multiple definitions to build robustness (Sarah E. Perkins (2015)). However, many studies rely on only a handful of metrics and definitions due to the excessive data management and computational burden of sampling a greater number of parameters (S. E. Perkins & Alexander (2013)). The introduction of large ensembles has further complicated the development of software tools, which have remained mostly specific to individual studies. Some generalized tools have been developed to address this problem, but they are often not explicitly optimized for evaluating the potential sensitivities of heatwave hazard to the choices of heat measure, extreme heat threshold, and heatwave definition.

Development of the HDP was started in 2023 primarily to address the computational obstacles around handling terabyte-scale large ensembles, but quickly evolved to investigate new scientific questions around how the selection of characteristic heatwave parameters may impact hazard analysis. The HDP can provide insight into how the spatial-temporal response of heatwaves to climate perturbations depends on the choice of heatwave parameters. While other software packages primarily focus on calculating heatwave metrics for individual parameter selections (e.g. heatwave3([Schlegel et al., 2024](#)), xclim([Bourgault et al., 2023](#)), ehfheatwaves([Loughran, 2021](#))), the HDP builds on these tools by optimizing the computation to evaluate metrics across large ranges of the parameter space.

40 Key Features

41 Extension of Xarray with Implementations of Dask and Numba

42 xarray is a popular Python package used for geospatial analysis and for working with the netCDF
 43 files produced by climate models. The HDP workflow is based around xarray and seamlessly
 44 integrates with the xarray.DataArray data structure. Parallelization of HDP functions is
 45 achieved through the integration of dask with automated chunking and task graph construction
 46 features built into the xarray library.

47 Heatwave Metrics for Multiple Measures, Thresholds, and Definitions

48 The “heatwave parameter space” refers to the span of measures, thresholds, and definitions
 49 that define individual heatwave “types” as described in Table 1.

Table 1: Parameters that define the “heatwave parameter space” and can be sampled using the HDP.

Parameter	Description	Example
Measure	The daily variable used to quantify heat.	Average temperature, minimum temperature, maximum temperature, heat index, etc.
Threshold	The minimum value of heat measure that indicates a “hot day.” The threshold can be constant or change relative to the day of year and/or location.	90th percentile temperature for each day of the year derived from observed temperatures from 1961 to 1990.
Definition	“X-Y-Z” where X indicates the minimum number of consecutive hot days, Y indicates the maximum number of non-hot days that can break up a heatwave, and Z indicates the maximum number of breaks.	“3-0-0” (three-day heatwaves), “3-1-1” (three-day heatwaves with possible one-day breaks)

50 The HDP allows the user to test a range of parameter values: for example, heatwaves that
 51 exceed 90th, 91st, ... 99th percentile thresholds for 3-day, 4-day, ... 7-day heatwaves. Four
 52 heatwave metrics that evaluate the temporal patterns in each grid cell are calculated for each
 53 measure and aggregated into a xarray.Dataset. Detailed descriptions of these metrics are
 54 shown in Table 2.

Table 2: Description of the heatwave metrics produced by the HDP.

Metric	Long Name	Units	Description
HWF	heatwave frequency	days	The number of heatwave days per heatwave season.
HWN	heatwave number	events	The number of heatwaves per heatwave season.
HWA	heatwave average	days	The average length of heatwaves per heatwave season.

Metric	Long Name	Units	Description
HWD	heatwave duration	days	The length of the longest heatwave per heatwave season.

55 Diagnostic Notebooks and Figures

56 The automatic workflow compiles a “figure deck” containing diagnostic plots for multiple
 57 heatwave parameters and input variables. To simplify this process, figure decks are serialized
 58 and stored in a single Jupyter Notebook separated into descriptive sections. Basic descrip-
 59 tions are included in markdown cells at the top of each figure. The HDPNotebook class in
 60 hdp.graphics.notebook is utilized to facilitate the generation of these Notebooks internally,
 61 but can be called through the API as well to build custom notebooks. An example figure of
 62 HWF from the sample figure deck is shown in Figure 1.

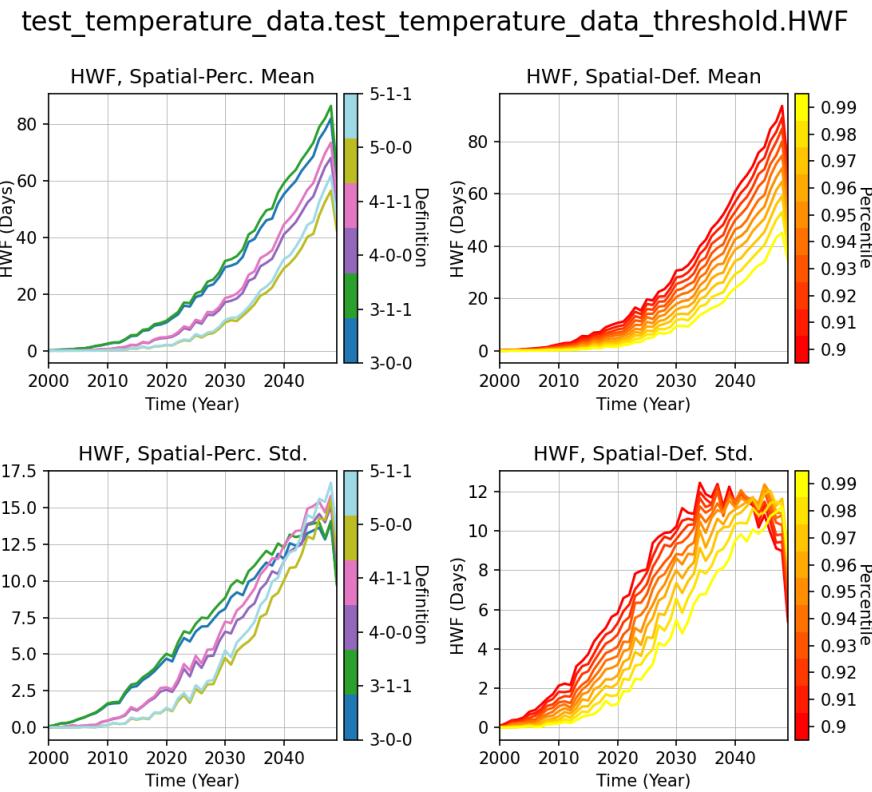


Figure 1: Example of an HDP standard figure deck

63 Ongoing Work

64 This package was used to produce the results featured in a research manuscript currently
 65 undergoing the peer-review process in a scientific journal. Updates to the HDP are ongoing.

66 Acknowledgements

67 We thank Dr. Tammas Loughran, Dr. Jane Baldwin, and Dr. Sarah Perkins-Kirkpatrick for their
68 work on developing the initial Python software and heatwave analysis framework that inspired
69 this project. Dr. Loughran's Python package is available on [GitHub](#). This work is partially
70 supported by the Modeling, Analysis, Predictions, and Projections Award Program under the
71 National Oceanic and Atmospheric Administration (Award Number NA23OAE4310601).

72 References

- 73 Acero, F. J., Fernández-Fernández, M. I., Carrasco, V. M. S., Parey, S., Hoang, T. T. H.,
74 Dacunha-Castelle, D., & García, J. A. (2024). *Changes in heat waves characteristics over*
75 *Extremadura (SW Spain)*. arXiv. <https://doi.org/10.48550/arXiv.2402.00514>
- 76 Argüeso, D., Di Luca, A., Perkins-Kirkpatrick, S. E., & Evans, J. P. (2016). Seasonal mean
77 temperature changes control future heat waves. *Geophysical Research Letters*, 43(14),
78 7653–7660. <https://doi.org/10.1002/2016GL069408>
- 79 Baldwin, J. W., Dessy, J. B., Vecchi, G. A., & Oppenheimer, M. (2019). Temporally Compound
80 Heat Wave Events and Global Warming: An Emerging Hazard. *Earth's Future*, 7(4),
81 411–427. <https://doi.org/10.1029/2018EF000989>
- 82 Bourgault, P., Huard, D., Smith, T. J., Logan, T., Aoun, A., Lavoie, J., Dupuis, É., Rondeau-
83 Genesse, G., Alegre, R., Barnes, C., Laperrière, A. B., Biner, S., Caron, D., Ehbrecht,
84 C., Fyke, J., Keel, T., Labonté, M.-P., Lierhammer, L., Low, J.-F., ... Whelan, C. (2023).
85 Xclim: Xarray-based climate data analytics. *Journal of Open Source Software*, 8(85), 5415.
86 <https://doi.org/10.21105/joss.05415>
- 87 Loughran, T. (2021). *tammasloughran/ehfheatwaves: First zenodo release* (Version v1.2).
88 Zenodo. <https://doi.org/10.5281/zenodo.5637520>
- 89 Perkins, Sarah E. (2015). A review on the scientific understanding of heatwaves—Their
90 measurement, driving mechanisms, and changes at the global scale. *Atmospheric Research*,
91 164–165, 242–267. <https://doi.org/10.1016/j.atmosres.2015.05.014>
- 92 Perkins, S. E., & Alexander, L. V. (2013). *On the Measurement of Heat Waves*. <https://doi.org/10.1175/JCLI-D-12-00383.1>
- 94 Schlegel, R. W., Pinto, G., Genevier, L., Cohen, J., Dasgupta, P., Meeker, E., Beaudin, E.,
95 & Smit, A. J. (2024). *heatwave3: Detect heatwaves and cold-spells within cubed data*.
96 <https://robwschlegel.github.io/heatwave3/index.html>
- 97 Schoetter, R., Cattiaux, J., & Douville, H. (2015). Changes of western European heat wave
98 characteristics projected by the CMIP5 ensemble. *Climate Dynamics*, 45(5), 1601–1616.
99 <https://doi.org/10.1007/s00382-014-2434-8>