Contents lists available at ScienceDirect

Data in Brief

Data Article

Tropical western and central Pacific marine heatwave data calculated from gridded sea surface temperature observations and CMIP6

Jules B. Kajtar [a,b,*], Vanessa Hernaman [c], Neil J. Holbrook [a,b], Paola Petrelli [a,b]

[a] *Institute for Marine and Antarctic Studies, University of Tasmania, Hobart, Tasmania, Australia*
[b] *Australian Research Council Centre of Excellence for Climate Extremes, University of Tasmania, Hobart, Tasmania, Australia*
[c] *Climate Science Centre, CSIRO Oceans and Atmosphere, Aspendale, Victoria, Australia*

ARTICLE INFO

ABSTRACT

Processed marine heatwave metrics are provided for the tropical western and central Pacific Ocean region (120°E-140°W, 40°S-15°N). The metrics are computed from daily sea surface temperature (SST) data, from both observations and models. The observed marine heatwave data are calculated from NOAA 0.25° daily Optimum Interpolation Sea Surface Temperature (OISST) over the period 1982–2019. The modelled marine heatwave data are from analysis of 18 model simulations as part of the Coupled Model Intercomparison Project, Phase 6 (CMIP6) over the period 1982–2100, where two future scenarios have been analysed. The marine heatwave data are provided on a grid point basis across the domain. Marine heatwave timeseries metrics are also provided for three case study regions: Fiji, Samoa, and Palau.

https://doi.org/10.1016/j.dib.2021.107694

Specifications Table

Subject	Environmental Science
Specific subject area	Observed, modelled, and projected marine heatwaves in the tropical western and central Pacific Ocean
Type of data	NetCDF Marine heatwave metrics
How data were acquired	Daily sea surface temperature (SST) data were sourced from a uniformly gridded observational dataset and model outputs from participating organisations in the Coupled Model Intercomparison Project, Phase 6 (CMIP6). Marine heatwave statistics were computed from these data using the marineHeatWaves python module: https://github.com/ecjoliver/marineHeatWaves.
Data format	Analysed
Parameters for data collection	A warm water event is typically defined as a marine heatwave if it exceeds a given temperature threshold for five or more days [1]. Daily SST data provide a relevant, robust, and useful measure for the calculation of extreme warm water events in the upper ocean [1,2]. Here, daily SST observations over multiple decades were desired. Model simulations and projections from the latest generation of global climate simulations were favoured.
Description of data collection	Observed sea surface temperature data analysed were the NOAA 0.25° daily Optimum Interpolation Sea Surface Temperature (OISST), version 2.1 [3]. Observations are available from 1 September 1981 to present day. Coupled Model Intercomparison Project, Phase 6 (CMIP6) daily sea surface temperature data were obtained from the National Computational Infrastructure (NCI) node of the Earth System Grid Federation (ESGF). Marine heatwaves metrics were computed using the marineHeatWaves python module: https://github.com/ecjoliver/marineHeatWaves. All source data are freely and publicly available.
Data source location	Observed primary data source: Metadata and data access link are available at 10.25921/RE9P-PT57. SST data are provided in a single file per day, but they were analysed from concatenated files (one year of data per file) made available at the National Computational Infrastructure (NCI). Observed data analysis period: 1 January 1982 to 31 December 2019. Model primary data source: Searched for and accessed through the NCI ESGF portal at https://esgf.nci.org.au/search/cmip6-nci. Primary sources and metadata locations provided in the main text. Search for analysed models can be refined with the following selections: • Experiment ID: historical, ssp126, ssp585 • Grid Label: gn • Frequency: day • Variable: tos Model data analysis period: 1 January 1982 to 31 December 2100. Region analysed: 120°E-140°W, 40°S-15°N (tropical western and central Pacific Ocean). Reference climatology period: 1 January 1995 to 31 December 2014
Data accessibility	Public repository Repository name: Zenodo Data identification number: 10.5281/zenodo.5069012 Direct URL to data: https://www.doi.org/10.5281/zenodo.5069012
Related research article	N.J. Holbrook, V. Hernaman, S. Koshiba, J. Lako, J.B. Kajtar, P. Amosa, A. Singh, Impacts of marine heatwaves on tropical western and central Pacific Island nations and their communities, Glob. Planet. Change, 208 (2022) 103,680. https://www.doi.org/10.1016/j.gloplacha.2021.103680 [4].

Value of the Data

- Marine heatwaves can have devasting impacts on marine ecosystems, with flow-on effects to human society. Such impacts are of particular importance to Pacific Island countries that rely heavily on marine resources. The data here provide marine heatwave metrics for the tropical western and central Pacific Ocean region, including counts of marine heatwave days, intensities, and durations over the observational period (1982–2019). Counts of marine heatwave days are also provided from CMIP6 models, from the historical period, through two possible future greenhouse gas emissions scenarios (1982–2100).
- Operators in the coastal and marine sectors in the tropical western and central Pacific Ocean region may have use of these data. Such sectors might include subsistence and commercial fishing, diving, aquaculture, fisheries management, tourism, conservation management and policy development, amongst a host of other sectors dependant on marine heatwave observations and projected changes.
- The data provided can be used to identify historical links between marine heatwaves and other observed impacts, including environmental, or societal, such as human health. Further evaluation of model performance can be conducted by determining the level of accuracy in simulating warm temperature extremes. Projections of marine heatwaves may also be used to evaluate levels of future climate change risk to relevant sectors in the region, facilitating adaptation planning.

1. Data Description

The primary data are sea surface temperature (SST), which were analysed from both observations and global climate models. In both cases, daily mean SSTs were obtained and processed from 1 January 1982, across a broad region of the Pacific Ocean (120°E-140°W, 40°S-15°N; Fig. 1). Marine heatwave metrics were computed at each grid point within the region. The region spans beyond the tropics, but the descriptor 'tropical western and central Pacific' reflects that the case
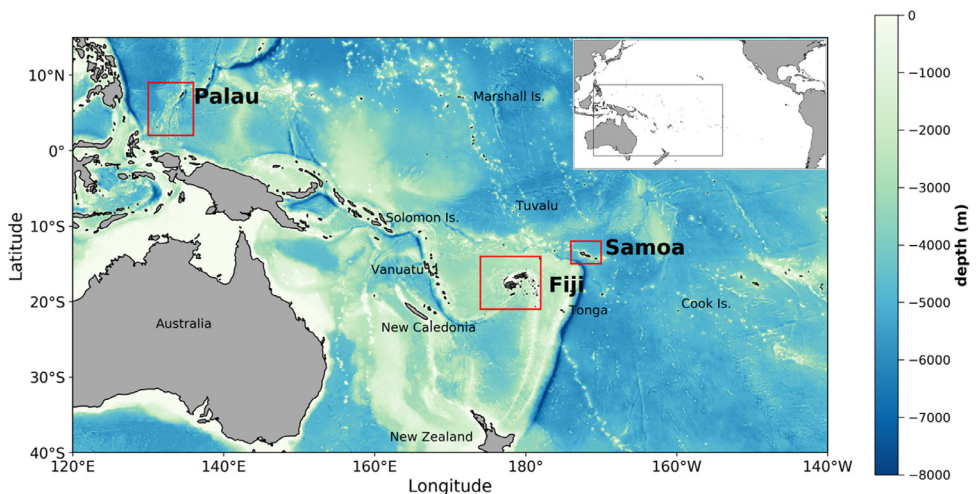

Fig. 1. Domain of the tropical western and central Pacific Ocean region. Case study regions for Palau, Fiji, and Samoa are denoted by red boxes. The inset shows the region in the context of the wider Pacific Ocean. Figure adapted from Fig. 1 of [4].

study island nations of Fiji, Samoa, and Palau all lie within the tropics. Analysis was also conducted on area-weighted mean SST for the three case study regions (Fig. 1):

- Fiji: 174°E-178°W, 21°S-14°S
- Samoa: 174°W-170°W, 15°S-12°S
- Palau: 130°E-136°W, 2°N-9°N

1.1. Observational data

Observational data are from the U.S. National Oceanic and Atmospheric Administration (NOAA) 0.25° daily Optimum Interpolation Sea Surface Temperature (OISST) dataset, version 2.1 [3]. The dataset was generated on a regular global grid from a range of sources, including Advanced Very High Resolution Radiometer (AVHRR) satellite data, ships, buoys, and Argo floats. Gaps in the data are infilled using optimum interpolation, providing a spatially complete dataset. Version 2.1 of the dataset has significant improvements over version 2.0 for January 2016 onward [5].

The data were provided by NOAA in single netCDF files for each calendar day. Australia's National Computational Infrastructure (NCI) provided an extra level of processing, by concatenating and storing the data in single netCDF files for each calendar year. Those annual data files were processed here. They are not publicly available, but equivalent annual data files are available from the NOAA Physical Sciences Laboratory (PSL) at https://psl.noaa.gov/data/gridded/data.noaa.oisst.v2.highres.html.

The observational data were analysed through to 31 December 2019, thus resulting in 38 complete calendar years. The following data and marine heatwave metrics are provided at every grid point within the region (file: *mhw_cats/mhw_cats.pac_is.NOAA_OISST.AVHRR.v2–1_modified.nc*):

- Annual mean SST (in °C; variable name: *sst_mean*).
- The total number of detected marine heatwaves (*mhw_total_events*).
- Annual counts of detected marine heatwaves (*mhw_count*).
- Annual counts of days per year spent in each marine heatwave category (*mhw_cats_dpy*). In this output, lower categories do not include the counts of days spent in more extreme categories. For example, if a marine heatwave is Moderate (Category I) for 5 days, then Strong (Category II) for 10 days, then Moderate for 2 days, it will be recorded as 7 days in Moderate and 10 days in Strong (see Section 2).
- Annual mean of the maximum intensity of each marine heatwave, given as a sea surface temperature anomaly with respect to the seasonal climatology (*mhw_intensity*).
- Annual mean duration of marine heatwaves (in days; *mhw_duration*).
- Linear temporal trends of the annual counts, intensities, and durations (per year; *mhw_count_tr, mhw_intensity_tr*, and *mhw_duration_tr*, respectively).

For the three case study regions, area-weighted means of the SST on grid points within the Fiji, Samoa, and Palau regions were computed. The three daily SST index timeseries are provided (file: *sst_indices/sst_indices.pac_is.NOAA_OISST.AVHRR.v2–1_modified.nc*). Marine heatwave metrics were then computed from these timeseries, and the following output is provided for each case study region (file: *mhw_ts/mhw_ts.pac_is.NOAA_OISST.AVHRR.v2–1_modified.nc*):

- Annual mean SST (in °C; variable name: *sst_mean*).
- The total number of detected marine heatwaves (*mhw_total_events*).
- Annual counts of days per year spent in each marine heatwave category (*mhw_cats_dpy*). In this output, lower categories do not include the counts of days spent in more extreme categories (see Section 2).

The 20-year mean of total marine heatwave days per year over the period 1982–2001 is shown in Fig. 2a. Total annual marine heatwave days are computed by summing the days in each category (Moderate, Strong, Severe, and Extreme).
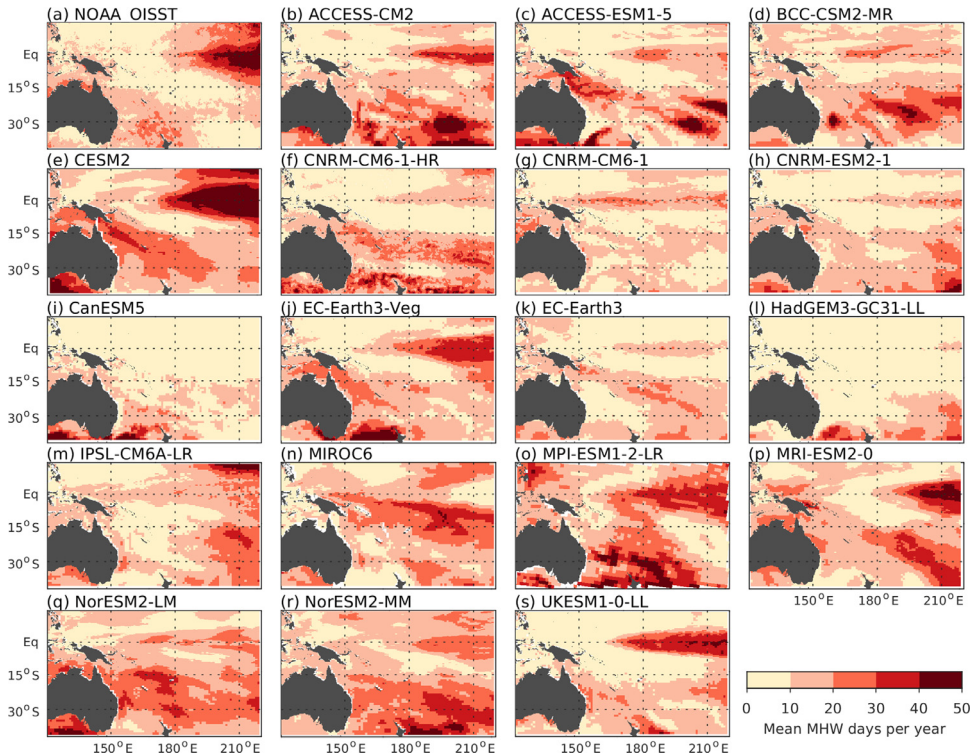
Fig. 2. Mean marine heatwave days over the period 1982–2001, computed from variable *mhw_cats_dpy*. (a) Observational data (file: *mhw_cats/mhw_cats.pac_is.NOAA_OISST.AVHRR.v2–1_modified.nc*). (b–s) Individual CMIP6 models (files: *mhw_cats/mhw_cats.pac_is.<model>.<scenario>.<variant>.nc*). Each dataset is plotted on its native grid. Total annual marine heatwave days are obtained by summing the number of days in each marine heatwave category (Moderate, Strong, Severe, and Extreme) at each grid point, and then the 20-year mean is computed. Panel (a) is adapted from Fig. 4a of [4], and the remaining panels show results in each of the models that contributed to the CMIP6 model mean shown in Fig. 8a of [4].

1.2. Model data

Model outputs from those participating in the Coupled Model Intercomparison Project, phase 6 (CMIP6) were analysed. CMIP6 is the current generation of global climate simulations that have informed the 6th Assessment Report of the Intergovernmental Panel on Climate Change (IPCC AR6; https://www.ipcc.ch/assessment-report/ar6/). Daily SST fields (variable name: *tos*) from native model grids (grid label: *gn*) were analysed from the *historical* set of experiments, together with two future shared socioeconomic pathways (SSPs) from the Scenario Model Intercomparison Project (ScenarioMIP) [6]. The future scenarios represent a low greenhouse gas emissions scenario: *SSP1–2.6*, and a high emissions scenario: *SSP5–8.5*. One ensemble member from each model and for each experiment was analysed. The *historical* experiments conclude at the end of year 2014, at which point the future scenarios were appended, thus resulting in two datasets for each model, each spanning the period 1982 to 2100. The requirement of having daily SST from each of the three experiments resulted in a set of 18 models with available data at the time of initial analysis (May 2020; Table 1).

The model data were analysed over the period 1 January 1982 through to 31 December 2100, thus resulting in 119 complete calendar years. The data are provided in two files for each model: *historical* with *SSP1–2.6* extension (*historical+ssp126*), and *historical* with *SSP5–8.5* exten-

sion (*historical+ssp585*) The following data are provided at every model native grid point within the region (files: *mhw_cats/mhw_cats.pac_is.<model>.<scenario>.<variant>.nc*; see Table 1 for model, scenario, and variant labels):

- Annual mean SST (in °C) (variable name: *sst_mean*).
- The total number of detected marine heatwaves (*mhw_total_events*).
- Annual counts of days per year spent in each marine heatwave category (*mhw_cats_dpy*). In this output, lower categories do not include the counts of days spent in more extreme categories (see Section 2).

The 20-year mean of total marine heatwave days per year over the period 1982–2001 is shown in Fig. 2b–s, for each CMIP6 model on its native grid.

As with the observational data, area-weighted means of the SST on native grid points within the Fiji, Samoa, and Palau regions were computed. The three daily SST index timeseries are provided for each model and scenario (files: *sst_indices/sst_indices.pac_is.<model>.<scenario>.<variant>.nc*). The data were area-weighted using the cell area data provided by each modelling group (variable name: *areacello*). The cell area data, trimmed to the tropical western and central Pacific Ocean region are also made available here (files: *grid/areacello.pac_is.<model>.<scenario>.<variant>.nc*). The *scenario* and *variant* labels of the cell area data do not necessarily match those of the SST data, but *areacello* is the same across all realisations for a single model.

Marine heatwave metrics were computed from area-weighted mean SST timeseries, and the following outputs are provided for each case study region (files: *mhw_ts/mhw_ts.pac_is.<model>.<scenario>.<variant>.nc*):

- Annual mean SST (in °C) (variable name: *sst_mean*).
- The total number of detected marine heatwaves (*mhw_total_events*).
- Annual counts of days per year spent in each marine heatwave category (*mhw_cats_dpy*). In this output, lower categories do not include the counts of days spent in more extreme categories (see Section 2).

Model and observational data were processed further to simplify the figure generation for [4]. The data were interpolated to the observational grid (i.e., at regular 0.25° resolution), and averaged over four 20-year periods: 1982–2001, 2000–2019, 2020–2039, 2040–2059. One file was generated for each scenario (files: *processed/mhw_cats.pac_is.cmip6.<scenario>.nc*), and the observational data are present in each. The following outputs are provided:

- Four blocks of 20-year mean SST (in °C) (variable name: *sst_mean*).
- Four blocks of 20-year means of days per year spent in each marine heatwave category (*mhw_cats*). In this output, the counts of days spent in more extreme categories are included in the counts of days spent in lower categories. For example, the counts of Moderate days include the counts of days spent in the Strong, Severe, or Extreme category (see Section 2).

The case study region data, for models and observations, and both scenarios, were also packaged into a single file (file: *processed/mhw_cats.pac_is.cmip6.timeseries.nc*). The following output is provided:

- Annual mean SST (in °C) (variable name: *sst_idx*).
- Annual counts of days per year spent in each marine heatwave category (*mhw_cats*). As above, the counts of days spent in more extreme categories are included in the counts of days spent in lower categories (see Section 2).

2. Experimental Design, Materials and Methods

Marine heatwave metrics were computed from the primary source daily sea surface temperature (SST) data using the marineHeatWaves python module (https://github.com/ecjoliver/

marineHeatWaves). The code uses the Hobday et al. definition [1] of a marine heatwave, which is a discrete, prolonged anomalously warm water event that lasts for five or more days, with temperatures warmer than the 90th percentile relative to climatological values. Two excursions above the threshold separated by less than three days are considered a single event. The seasonal climatology and the seasonally varying 90th percentile threshold were uniquely computed for each calendar day and grid point as follows. To increase the number of sample points, the climatology and threshold were computed over an 11-day window centred on a given day, sampling each year within the selected climatological period. After computing the mean and threshold for each calendar day, 31-day moving means were applied to each year-long timeseries in order to give smoother seasonally varying climatologies. This follows the procedure in [2].

Marine heatwaves can be characterised by their duration (number of consecutive days above the 90th percentile threshold), maximum intensity (maximum temperature above the climatological mean during the event), and their categorisation of intensity [7]. Categorisations of marine heatwaves are based on multiples of the marine heatwave intensity threshold, i.e., the temperature difference between the climatological mean and the climatological 90th percentile (see Fig. 2 in [7]). In this way, the category ranges are equally spaced, and temperature anomalies between 1 and 2 × of the marine heatwave intensity threshold are denoted as Moderate, between 2 and 3 × as Strong, between 3 and 4 × as Severe, and >4 × as Extreme. This dataset is focused on providing counts of days per year recorded in each marine heatwave category.

A 20-year baseline period (1995–2014) was used for the seasonal climatological mean and seasonally varying 90th percentile threshold. This baseline period is the IPCC AR6 recommendation for assessing projected changes relative to the historical period.

The python code for computing marine heatwave metrics, along with the MATLAB code for post-processing are available at https://github.com/jbkajtar/mhw_pacific/. A CSV file containing all CMIP6 model details and references is also available at that location.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships which have or could be perceived to have influenced the work reported in this article.

CRediT Author Statement

Jules B. Kajtar: Conceptualization, Methodology, Investigation, Software, Data curation, Writing – original draft, Writing – review & editing; **Vanessa Hernaman:** Conceptualization, Methodology, Investigation, Writing – review & editing; **Neil J. Holbrook:** Conceptualization, Methodology, Investigation, Writing – review & editing; **Paola Petrelli:** Data curation, Validation, Writing – review & editing.

Acknowledgments

References

[1] A.J. Hobday, L.V. Alexander, S.E. Perkins-Kirkpatrick, D.A. Smale, S.C. Straub, E.C.J. Oliver, J.A. Benthuysen, M.T. Burrows, M.G. Donat, M. Feng, N.J. Holbrook, P.J. Moore, H.A. Scannell, A. Sen Gupta, T. Wernberg, A hierarchical approach to defining marine heatwaves, Prog. Oceanogr. 141 (2016) 227–238, doi:10.1016/j.pocean.2015.12.014.

[2] E.C.J. Oliver, M.G. Donat, M.T. Burrows, P.J. Moore, D.A. Smale, L.V. Alexander, J.A. Benthuysen, M. Feng, A. Sen Gupta, A.J. Hobday, N.J. Holbrook, S.E. Perkins-Kirkpatrick, H.A. Scannell, S.C. Straub, T. Wernberg, Longer and more frequent marine heatwaves over the past century, Nat. Commun. 9 (2018) 1324, doi:10.1038/s41467-018-03732-9.

[3] B. Huang, C. Liu, V.F. Banzon, E. Freeman, G. Graham, B. Hankins, T.M. Smith, H.M. Zhang, NOAA 0.25-degree daily optimum interpolation sea surface temperature (OISST), Version 2.1, 2020. 10.25921/RE9P-PT57. Dataset accessed on 23 June 2021.

[4] N.J. Holbrook, V. Hernaman, S. Koshiba, J. Lako, J.B. Kajtar, P. Amosa, A. Singh, Impacts of marine heatwaves on tropical western and central Pacific Island nations and their communities, Glob. Planet. Change 208 (2022) 103680, doi:10.1016/j.gloplacha.2021.103680.

[5] B. Huang, C. Liu, V. Banzon, E. Freeman, G. Graham, B. Hankins, T. Smith, H.M. Zhang, Improvements of the daily optimum sea surface temperature (DOISST) - version 2.1, J. Clim. 34 (2021) 2923–2939.

[6] B.C. O'Neill, C. Tebaldi, D.P. Van Vuuren, V. Eyring, P. Friedlingstein, G. Hurtt, R. Knutti, E. Kriegler, J.F. Lamarque, J. Lowe, G.A. Meehl, R. Moss, K. Riahi, B.M. Sanderson, The scenario model intercomparison project (ScenarioMIP) for CMIP6, Geosci. Model Dev. 9 (2016) 3461–3482, doi:10.5194/gmd-9-3461-2016.

[7] A.J. Hobday, E.C.J. Oliver, A. Sen Gupta, J.A. Benthuysen, M.T. Burrows, M.G. Donat, N.J. Holbrook, P.J. Moore, M.S. Thomsen, T. Wernberg, D.A. Smale, Categorizing and naming marine heatwaves, Oceanography 31 (2018) 162–173, doi:10.5670/oceanog.2018.205.