



The Weather Extremes Archive (WEXA)

Patrick Martineau¹, Hidehiro Fujio¹, Ingo Richter¹, Ayako Yamamoto², and Masami Nonaka¹

¹Japan Agency for Marine-Earth Science and Technology, Yokohama, Japan

²Rikkyo University, Tokyo, Japan

Correspondence to: Patrick Martineau (pmartineau@jamstec.go.jp)

Abstract. The Weather Extremes Archive (WEXA) is a global, event-based dataset designed to systematically detect, track, and document persistent weather extremes. In its initial release, WEXA focuses on heat waves and cold spells identified from 2-m air temperature in ERA5 reanalysis data from 1979 to present. Cold and warm extremes are defined using seasonally varying 10th and 90th percentile thresholds derived from a fixed 1991–2020 climatological baseline. Periods of at least three consecutive extreme days are grouped into coherent spatiotemporal heat waves or cold spell events using three-dimensional connected component labelling. WEXA delivers a global catalogue of all detected individual heat wave and cold spell events, providing standardized diagnostics of their timing, duration, spatial extent, and intensity. The archive is openly accessible at <https://doi.org/10.5281/zenodo.20550615> and <https://wexa.jamstec.go.jp>, permitting the exploration of event statistics and visualizations through an interactive analysis tool.

1 Introduction

Temperature extremes exert severe socio-economic impacts including strain on energy supply, damage to infrastructure, disruption to agriculture, adverse effects on natural ecosystems as well as posing substantial risk to human health (Chen et al., 2013; Lei et al., 2022; Pantavou et al., 2008; Perkins, 2015). The severity of these impacts typically scales with their duration, with prolonged events disproportionately leading to larger consequences (Anderson and Bell, 2011; Li et al., 2024; Miller et al., 2021). Depending on the polarity of the persistent anomalies, prolonged events are typically classified as cold waves (or cold spells) and heat waves (or hot spells).

Given the substantial socioeconomic impacts of heat waves and cold spells, a consistent and robust framework for defining and quantifying these temperature extremes is essential. In the literature, heat waves and cold spells are typically, though not universally, defined as periods during which temperatures exceed or fall below specified thresholds (e.g., the 90th percentile for heat waves and the 10th percentile for cold spells) for several consecutive days (typically at least three) (Bunting et al., 2024; Perkins, 2015). Definitions, however, differ widely in the choice of temperature variable (e.g., daily maximum, minimum, mean, or apparent temperature) and in the methods used to quantify event magnitude. While some studies focus on a single aspect such as intensity, duration, or frequency to characterize event magnitude, others employ composite indices that



integrate multiple aspects of heat-wave behaviour (Russo et al., 2015). Since the review paper of Perkins (2015), more variants of heat-wave definitions have been proposed, often based on broadly similar detection principles (Lavaysse et al., 2018; Perkins-Kirkpatrick and Lewis, 2020; Russo et al., 2015). Similar to heat waves, cold spells are frequently defined based on duration requirements typically ranging from two to six days, and thresholds set at specific percentiles (e.g., 5th or 10th) or standard deviations (e.g., -1.25 to -2 SD; Table 1).

Building on the earlier standardization efforts by Frich et al. (2002), the Expert Team on Climate Change Detection, Monitoring and Indices (ETCCDI), which has since concluded its formal mandate, established a suite of 27 climate extreme indices to ensure consistency in climate extreme analysis (Alexander et al., 2006). While these indices, such as the Warm Spell Duration Index (WSDI), which quantifies the occurrence of consecutive days of extreme heat, are instrumental for assessing long-term trends, they are typically aggregated at monthly resolution (Dunn et al., 2020). Consequently, they often mask the precise timing, spatial evolution, and life cycle of individual events. This limitation hinders a deeper physical understanding of extremes and the full assessment of their socio-economic impacts.

Several recent initiatives have sought to provide more detailed information on thermal extremes. The Thermal Trace application, developed by the Copernicus Climate Change Service (C3S) and ECMWF, enables interactive analysis of global hourly thermal stress data based on “feels-like” temperature, a metric closely related to human thermal comfort (<https://climate.copernicus.eu/thermal-trace-decades-heat-and-cold-stress-data-your-fingertips>). However, this application does not comprehensively catalogue historical extreme events. Another initiative, the Copernicus European Drought Observatory (EDO) Heat and Cold Wave Index (HCWI; <https://data.jrc.ec.europa.eu/dataset/c9e06815-0c52-4ff5-b596-89ec169996eb>), aims to analyse the spatial and temporal evolution of temperature-related hazards over Europe (Lavaysse et al., 2018). In this framework, heat waves are defined as periods of three or more consecutive days during which both daily minimum and maximum temperatures exceed the calendar-day 90th percentile, with an analogous definition applied to cold spells. This effort, however, is limited to the European sector.

To the best of our knowledge, no existing initiative provides an up-to-date global catalogue of heat waves and cold spells that explicitly tracks their spatiotemporal evolution - that is, their detailed development over time, including changes in affected areas as events propagate and decay. While several methodological frameworks for tracking spatiotemporal extremes have been proposed (Lo et al., 2021; Osmolska et al., 2025), they have not yet been applied to create a comprehensive, operational global inventory. Addressing this limitation, the objective of the dataset presented here is to introduce a global, event-based dataset designed to systematically detect, track, and document the spatiotemporal evolution of these persistent heat waves and cold spells. It is based on established, percentile-based definitions of heat waves and cold spells (90th and 10th percentiles, respectively, with 3-day persistence requirement), applied to ERA5 reanalysis data (Hersbach et al., 2020). The dataset is produced in near-real time, with a latency of only a few days, and is made accessible through a web platform



(<https://wexa.jamstec.go.jp>), which provides pre-computed diagnostics from 2020 onward as well as an interactive analysis system archiving events from 1979 onward (<https://wexa.jamstec.go.jp/home/interactive>). This interactive system enables filtering and analysis of events based on various search criteria. By providing a comprehensive, spatiotemporally resolved catalogue of temperature extremes, this dataset aims to accelerate scientific discovery, support decision-makers with advanced statistics on extreme-event occurrence, and facilitate accurate media reporting of recent extreme weather events.

Table 1: Threshold and minimum duration criteria of some cold-spell definitions reported in the scientific literature. SD denotes a definition based on a departure for the mean value measured in standard deviations and % denotes a definition based on percentiles.

Reference	Threshold	Minimum duration (days)
Alam et al. (2023)	10%	2
De Vries et al. (2012)	10%	5
Osmolska et al. (2025)	5%	3
Smith and Sheridan (2019)	-1.25 SD	5
Tringa et al. (2022)	5%	3
Vavrus et al. (2006)	-2 SD	2

2 Detecting extremes

2.1 Data

For the detection of temperature extremes, temperature at 2 meters above the surface (hereafter T_{2m} ; units of K) is obtained from the ERA5 reanalysis (Hersbach et al., 2020) at hourly intervals on a $0.5^\circ \times 0.5^\circ$ horizontal grid.

2.2 Definition of extremes

With a focus on persistent weather extremes, the first step is to compute daily averages from the hourly data. This reduces high-frequency noise associated with short-term weather variability and lowers computational costs for subsequent analyses. Next, we determine the thresholds used to identify temperature extremes. These thresholds correspond to the 10th and 90th percentiles of the temperature distribution calculated over a fixed 30-year baseline period (1991–2020) as per WMO recommendations (WMO, 2017). These thresholds have often been used to define cold spells and heat waves (Bunting et al., 2024; Lavaysse et al., 2018; Perkins, 2015). The percentiles are estimated separately for each grid point and calendar day, using a 7-day moving window to smooth sampling variability arising from the limited sample size. The thresholds thus obtained are referred to as T_{2m}^{10} and T_{2m}^{90} , denoting respectively the 10th and 90th percentiles of T_{2m} , and are functions of calendar day (d), longitude (λ) and latitude (φ). Figure 1 illustrates T_{2m}^{90} averaged in December-January-February (DJF) and June-July-August (JJA) and its departure from the seasonal median temperature, while Figure 2 shows the same for T_{2m}^{10} . It can be seen



that the thresholds fluctuate in response to seasonal temperature changes, with warmer threshold temperatures in the summer hemisphere compared to the winter hemisphere, as well as to regional differences such as warmer temperatures in the tropics compared to the extratropics. We also note that the threshold mirrors the typical range of temperature at each location, with larger differences from the seasonal median over land masses where temperature variations tend to be larger in comparison to ocean bodies (Fig. 1c,d and Fig 2c,d).

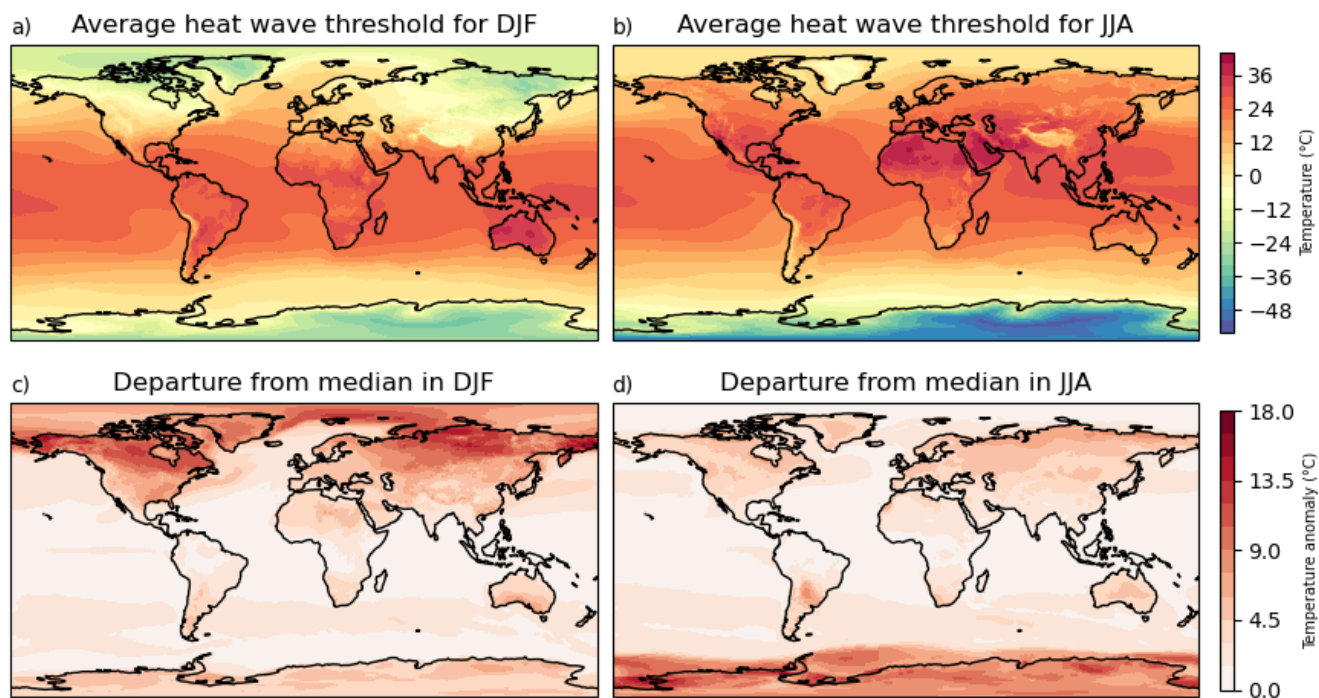


Figure 1: (top) The threshold to define warm extremes, T_{2m}^{90} , is shown for (left) DJF and (right) JJA. (bottom) The difference of these thresholds from the seasonal median is shown for the same seasons. Thresholds are defined from the 1991-2020 baseline period.

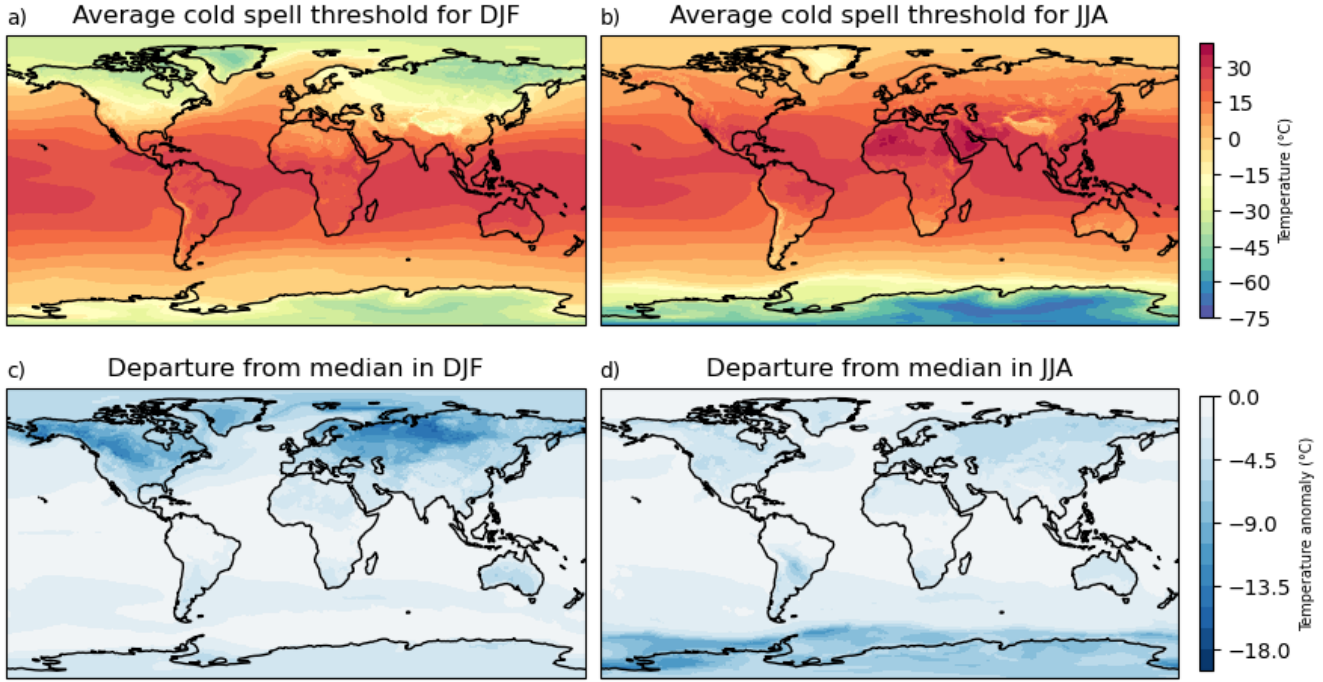


Figure 2: Same as Fig. 1 but for the threshold of cold extremes, T_{2m}^{10} .

The occurrence of temperature extremes is evaluated at each time step (t), longitude (λ), and latitude (φ) by comparing T_{2m} to the threshold value for the corresponding calendar day ($d(t)$). Warm extremes (WE) and cold extremes (CE) are defined as follows:

$$WE(t, \lambda, \varphi) = \begin{cases} 1, & T_{2m}(t, \lambda, \varphi) > T_{2m}^{90}(d(t), \lambda, \varphi) \\ 0, & \text{otherwise} \end{cases} \quad (1)$$

$$CE(t, \lambda, \varphi) = \begin{cases} 1, & T_{2m}(t, \lambda, \varphi) < T_{2m}^{10}(d(t), \lambda, \varphi) \\ 0, & \text{otherwise} \end{cases} \quad (2)$$

WE and CE are binary indicators, where a value of 1 indicates the presence of an extreme event and 0 indicates its absence. The temporal average of WE and CE yields the frequency of warm and cold extremes, which, by construction, is 10% over the baseline period.

Next, we identify persistent extreme events at each grid point, i.e. heat waves ($HW(t, \lambda, \varphi)$) and cold spells ($CS(t, \lambda, \varphi)$). Here, a heat wave (HW) is defined as any period in which three or more consecutive days are classified as warm extremes (i.e., $WE = 1$). This minimum duration is often used to identify heat waves and cold spells (Bunting et al., 2024; Lavaysse et al., 2018; Perkins, 2015). All days within such a sequence are labelled as part of a heat wave ($HW = 1$). Cold spells (CS) are identified with CE in the same manner. All other days are labelled with a value of 0. In contrast to daily extremes (WE and



CE), the frequency of *HW* and *CS* does not average to 10% over the baseline period. Rather, 10% represents an upper bound, while the actual occurrence depends on the likelihood of observing consecutive extremes. This likelihood varies across regions and seasons depending on the persistence of the weather systems driving these extremes and processes leading to interactions between the lower atmosphere, land, and ocean surfaces (Fig. 3). For instance, persistent extremes tend to be less frequent over tropical rainfall zones and storm-track sectors, particularly in the winter hemisphere. In contrast, they occur more frequently over the eastern tropical Pacific, possibly modulated by the El Niño-Southern Oscillation (ENSO), and over subtropical high-pressure zones like the Sahara and southern Africa.

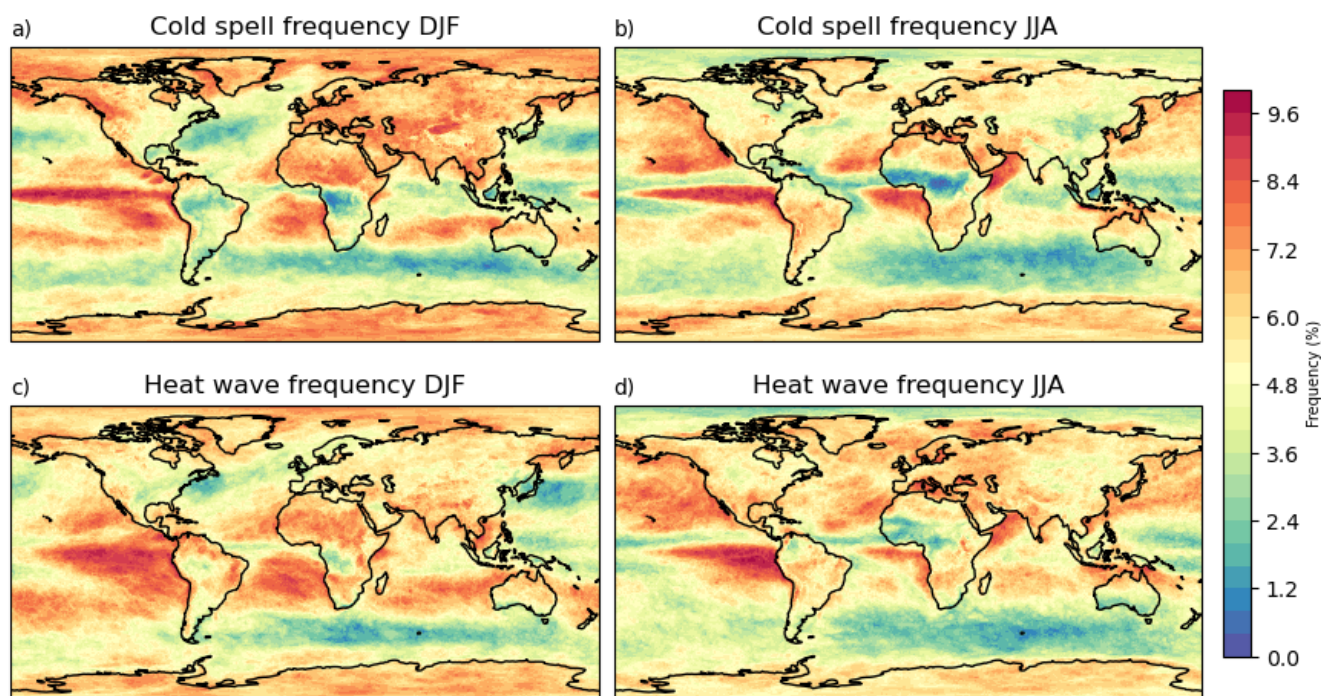


Figure 3: Frequency of (top) cold spells (CS) and (bottom) heat wave (HW) events in (left) DJF and (right) JJA for the 1991-2020 baseline period.

2.2 Grouping extremes into spatiotemporal events

The method described above for detecting heat waves and cold spells provides detailed spatiotemporal information on the occurrence of these weather extremes. Driven by the evolution and movement of atmospheric weather systems, these extremes tend to form spatially and temporally coherent patterns (Lo et al., 2021; Osmolska et al., 2025). Thus, as a next step, we group these coherent patterns into *HW* and *CS* spatiotemporal events. To systematically identify and group such contiguous regions, we apply connected component labelling (CCL), which enables the classification of neighbouring grid points into distinct



spatiotemporal extreme events based on their spatial and temporal continuity. CCL is well-established method in image processing (Rosenfeld and Pfaltz, 1966) and has been used for climate applications (Knippertz et al., 2015), including the grouping of marine heat waves into spatiotemporally coherent events (Bonino et al., 2023). CCL is applied here to the three-dimensional (t, λ, φ) binary arrays of heat waves (HW) and cold spells (CS) to identify individual spatiotemporal coherent events. This method was implemented using the *label* function from the *skimage.measure* module of the *scikit-image* Python library (Van Der Walt et al., 2014). While more advanced spatiotemporal tracking methodologies have recently been proposed (Lo et al., 2021; Osmolska et al., 2025), CCL's relative simplicity facilitates reproducibility and improves the transparency and interpretability of the detected events database. We use a version of CCL that connects grid points or time steps that are directly adjacent or diagonally adjacent.

Figure 4 illustrates the CCL procedure for identifying heat wave (HW) and cold spell (CS) spatiotemporal events. The situation shown is not necessarily realistic, but it serves to demonstrate how the CCL rules operate. On day 0, three spatial clusters are identified: two HW clusters in the west and one mixed cluster in the east containing both CS and HW grid points. According to the CCL rules, all HW grid points in the western clusters are assigned the same identification number (#1). They are grouped together because (1) they are adjacent either zonally, meridionally, or diagonally; and (2) they are temporally connected, since the northwestern and southwestern clusters merge on day 1. In the eastern region, the CS and HW grid points are assigned different identification numbers (#2 and #3, respectively) because they represent distinct event types. On day 1, a new HW event appears to the southwest (#4), and the HW event #3 shifts north-eastward, illustrating that the diagonal-adjacency criterion also applies in the three-dimensional space (t, λ, φ) . The cold-spell event detected on day 0 (#2) splits into two separate clusters on day 1. Despite becoming disconnected, both clusters retain the same identification number because they were previously part of a single event.

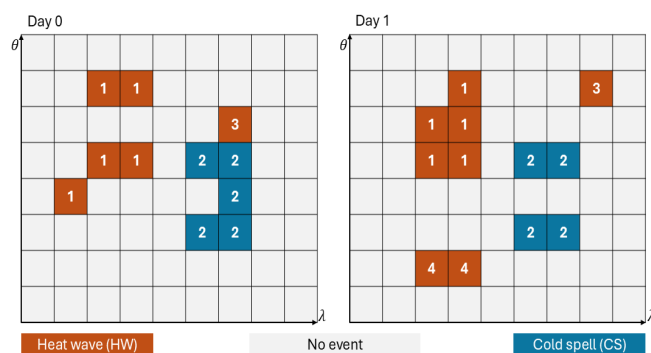


Figure 4: Schematic diagram of the labelling of heat wave and cold spell spatiotemporal events with the CCL algorithm. Event labelling is illustrated with latitudes (θ) longitude (λ) maps for two consecutive days (Day 0 and Day 1). Orange and blue shading represents grid cells with $HW=1$ and $CS=1$, respectively. The numbers indicated at each grid cell correspond to unique event identification numbers. The situation illustrated assumes that none of these events merge afterwards.



A realistic illustrative example of spatiotemporal event detection using CCL is shown in Fig. 5. As expected from the CCL algorithm, spatiotemporally-connected extremes are grouped into single events. At the onset of the event, two distinct regional clusters, one to the south and another to the north, are observed. Despite being initially separated, these clusters are grouped into a single event since they merge later in their lifetime (at $t=3$). Similarly, although not shown, some events may later split into multiple sectors. It is nonetheless meaningful to treat these separate regions as part of the same event, as they are typically generated by a common large-scale atmospheric circulation feature and therefore likely have similar formation and dissipation processes.

The unique integer labels assigned to each spatiotemporal coherent event thus identified by the CCL are archived as functions of time, longitude, and latitude in NetCDF files, which are distributed via a dedicated server at <https://www.jamstec.go.jp/RID/thredds/catalog/testAll/WEXA/catalog.html>.

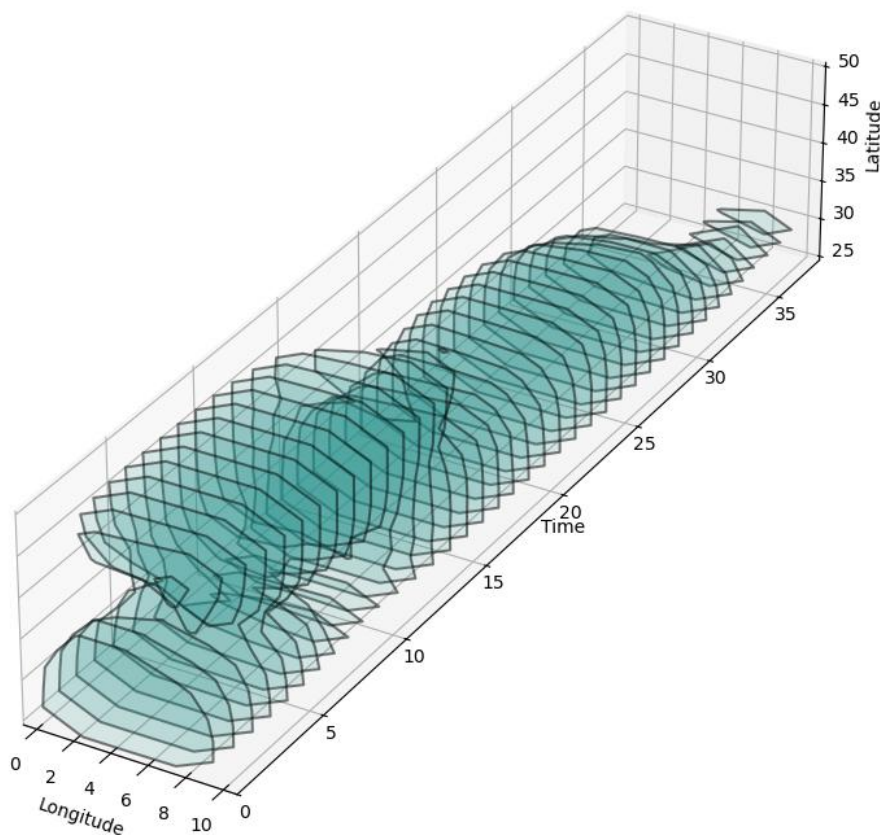


Figure 5: Illustrative example of spatiotemporal event detection based on connected component labelling. Applying connected component labelling on identified heat waves (HW) with dimensions of longitude (λ), latitude (φ) and time (t) yields a volume of spatiotemporally connected grid points that share the same unique identification number. For each time step, the region where the event is labelled is shaded in green in the longitude-latitude plane and delimited with a black contour.



To synthesize the detection process, Fig. 6 provides a flowchart of the complete WEXA algorithm. This diagram summarizes the progression from raw data and threshold application to the final 3D grouping of events, providing a visual roadmap for the methodology described above.

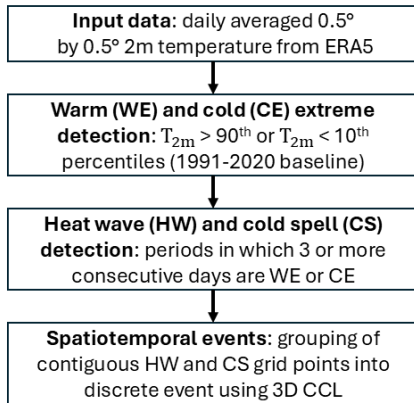


Figure 6: Flowchart of the WEXA event detection algorithm.

3 Statistics and visualizations

A standardized set of properties is reported for each spatiotemporal event and provided in JavaScript Object Notation (JSON) files, with individual diagnostics defined in Table 2. All the files are compiled into a single comprehensive archive hosted at <https://doi.org/10.5281/zenodo.20550615> (Martineau et al., 2026). Several of these diagnostics quantify event intensity using standardized T_{2m} anomalies. To obtain standardized anomalies, T_{2m} anomalies are first computed by subtracting the climatological mean for the corresponding calendar day (d) from the daily T_{2m} values

$$T'_{2m}(t) = T_{2m}(t) - \bar{T}_{2m}(d(t)) \quad (3)$$

where overbar denotes a calendar-day mean. Standardized temperature anomalies are subsequently obtained by dividing daily temperature anomalies by the standard deviation for each calendar day

$$\tilde{T}_{2m}(t) = T'_{2m}(t) / \sigma_{clim}(d(t)) \quad (4)$$

where σ_{clim} denotes the standard deviation for a particular calendar day. To reduce sampling noise in both the mean and the standard deviation, $\bar{T}_{2m}$ and σ_{clim} are smoothed using a 7-day running mean centered on the target day.

Table 2: Event properties provided in JSON files for each spatiotemporal event. The JSON field is an identifier used to access each variable in the dataset.

Properties	JSON field	Description
Event ID	event_id	A unique identifier based on: beginning_date, end_date, longitude, and latitude
Event label	event_label	Integer-valued grid identifier that uniquely labels each detected extreme event in space and time, consistent with the corresponding labels stored in the NetCDF dataset
Start date	start_date	First date on which grid points are under the influence of an event



End date	end_date	Last date on which grid points are under the influence of an event
Duration	duration	Time elapsed (days) from the start date to the end date
Status	status	An ongoing event denotes an event that lasts up to the latest data available. These events are subject to change as newer data are added.
Center latitude	center_lat	Latitude of the maximum time-integrated standardized T_{2m} anomalies. Only T_{2m} of the same polarity as the events are considered (positive T_{2m} for a HW, for instance). The centre is thus representative of both the severity and persistence of anomalies.
Center longitude	center_lon	Same as above but for longitude
Center latitude formatted	center_latf	Formatted version of center latitude
Center longitude formatted	center_lonf	Formatted version of center longitude
Westernmost longitude	western_lon	Westernmost longitude affected by event
Easternmost longitude	eastern_lon	Easternmost longitude affected by event
Southernmost latitude	southern_lat	Southernmost latitude affected by event
Northernmost latitude	northern_lat	Northernmost latitude affected by event
Event type	event_type	Type of the event detected
Detection variable	detection_var	Variable used for the detection scheme
Area	area	The area (in km^2) encompassing all grid points under the influence of an event during its lifetime.
Countries affected	countries_affected	The World Countries database (https://hub.arcgis.com/datasets/esri::world-countries/) is used to obtain a list of all the countries affected by an event during its lifetime.



Continents affected	continents_affected	The World Countries database (https://hub.arcgis.com/datasets/esri::world-countries/) is used to obtain a list of all the continents affected by an event during its lifetime.
IPCC regions	IPCC_regions	List of all IPCC regions affected. More details about these regions are provided at https://www.ipcc.ch/report/ar6/wg1/downloads/factsheets/IPCC_AR6_WGI_Regional_Fact_Sheet_Introduction.pdf and the shapefiles describing these regions were obtained from https://github.com/IPCC-WG1/Atlas/tree/devel/reference-regions
Oceans basins affected	oceans_affected	World_Seas_IHO_v3.shp dataset is used to obtain a list of all the ocean basins affected by an event during its lifetime.
Polygons	polygons	Polygon vertices describing the boundary of the event at each time step
Mean anomaly amplitude	mean_amp	Mean T_{2m} anomaly over entire duration
Max anomaly amplitude	max_amp	Maximum T_{2m} anomaly recorded over entire duration
Evolution of maximum T2m	evo_maxamp	The daily evolution of maximum T_{2m}
Evolution of mean T2m	evo_areamean	The daily evolution of mean T_{2m}
Evolution of maximum T2m anomalies	anom_evo_maxamp	The daily evolution of maximum T_{2m} anomalies
Evolution of mean T2m anomalies	anom_evo_areamean	The daily evolution of mean T_{2m} anomalies
Evolution of maximum standardized	anom_scaled_evo_maxamp	The daily evolution of maximum standardized T_{2m} anomalies



ed T2m anomalies		
Evolution of mean standardiz ed T2m anomalies	anom_scaled_evo_ar eamean	The daily evolution of mean standardized T_{2m} anomalies

195
196
197
198
199
200
201

202

203
204
205
206
207
208
209
210
211
212
213
214
215
216
217

All spatiotemporal events detected in the database with an area exceeding 50,000 km², approximately the size of a 2.5° × 2.5° grid cell at 50°N, and that occurred after January 1st 2020 are subjected to detailed examination through the advanced production of illustrative figures. These events are allotted a special page on the archive’s webpage (<https://wexa.jamstec.go.jp/>). Smaller events and those occurring prior to 2020 are not given their own page, but are nonetheless accessible on the interactive analysis system described later.

3.1 Precomputed diagnostic figures

By default, three figures are generated for each spatiotemporal event. As a representative example, we present these figures for the extreme heat wave known as the 2021 Western North America heat wave, which occurred in June 2021 (ID: *2021-May-29_to_2021-Jul-20_241.5_49.0*). This event shattered temperature records across multiple cities in the western United States and Canada by substantial margins (Philip et al., 2022; Thompson et al., 2022). It caused significant excess mortality with 619 heat-related deaths in British Columbia (Tetzlaff et al., 2024) and contributed to the devastating wildfires which destroyed the town of Lytton, British Columbia (Philip et al., 2022).

The first figure (Fig. 7) summarizes the severity of the event by displaying the most intense 3-day rolling mean standardized temperature anomalies ($\tilde{T}_{2m}$) observed over its lifetime (maximum for heat waves, as shown here, and minimum when this figure is produced for cold spells). The 3-day rolling mean is used to capture persistent temperature anomalies and is consistent with the persistence criterion applied in the detection of extreme events. The event primarily affected western Canada, the western United States, and parts of Mexico, with additional impacts in the northeastern United States (green hatching) with $\tilde{T}_{2m}$ exceeding 4σ over the West (color shading). According to our definition, the event center (magenta star) was located near the Canada-US border, reflecting the intensity and persistence of anomalies at that location, as the event centre is defined by the maximum time-integrated $\tilde{T}_{2m}$ (Table 2).

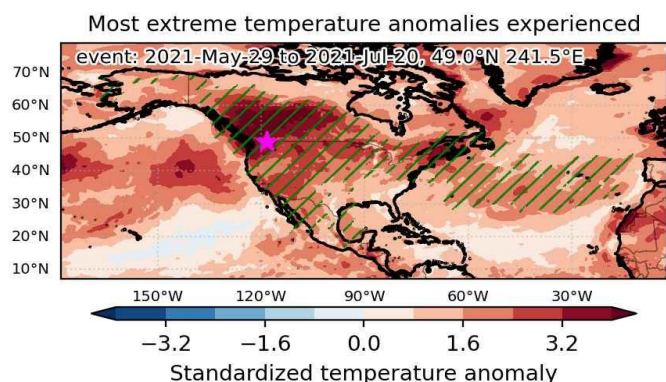


Figure 7: Map of the warmest 3-day mean $\tilde{T}_{2m}$ (shading) observed during the 2021 Western North America heat wave. The area defined as being under the influence of the event is hatched in green. The event center is shown with a magenta star and the event ID (2021-May-29 to 2021-Jul-20, 49.0°N 241.5°E) is shown at the top of the map.

The second figure (Fig. 8.) illustrates the temporal evolution of 2-m temperature at the event centre. Three quantities are shown: the raw temperature (T_{2m}), the temperature anomaly (T'_{2m}) and the standardized anomaly ($\tilde{T}_{2m}$). They are overlaid on the evolution of their respective probability distributions for the 1991-2020 baseline period. Values may fall outside the 0-100 percentile range for events occurring outside the baseline period, which is increasingly common for recent record-breaking heat waves under ongoing global warming. During the illustrated heat wave, the western part of Canada-United States border experienced an initial short-lived episode of exceptionally warm temperatures in early June, followed by a gradual reintensification culminating in a record-breaking peak in late June with $\tilde{T}_{2m}$ exceeding 4σ .

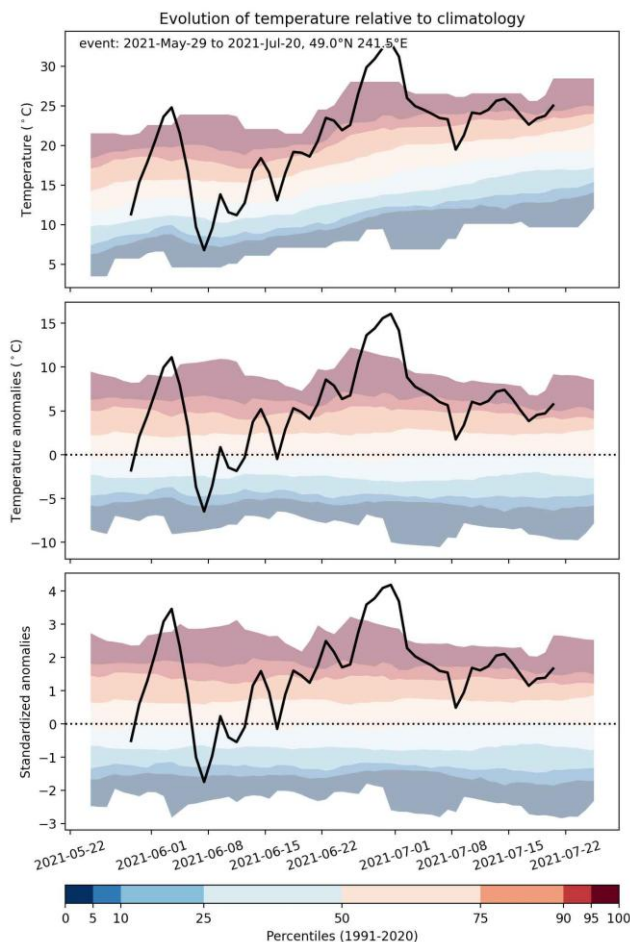


Figure 8: Temporal evolution of temperature at the event centre. From top to bottom, T_{2m} , T'_{2m} and $\tilde{T}_{2m}$ are shown (black lines). The corresponding probability distributions for the 1991-2020 baseline period are illustrated with colour shadings delimited by the 0, 5, 10, 25, 50, 75, 90, 95, and 100th percentiles.

The third figure (Fig. 9) depicts the temporal evolution of the event, showing both $\tilde{T}_{2m}$ and the spatial extent of the affected region from onset to termination. The maps indicate that the event initially developed as two centers of activity: one over western Canada and the western United-States, and another over the western North Atlantic in late May. These centers expanded and intensified, merging in early June. Subsequently, the eastern lobe split and propagated into the central North Atlantic, while a third lobe developed over Mexico around June 6 and gradually migrated northward. This southern feature eventually amplified and contributed to the peak intensity over Western Canada and the United States in late June, after which the event gradually dissipated.

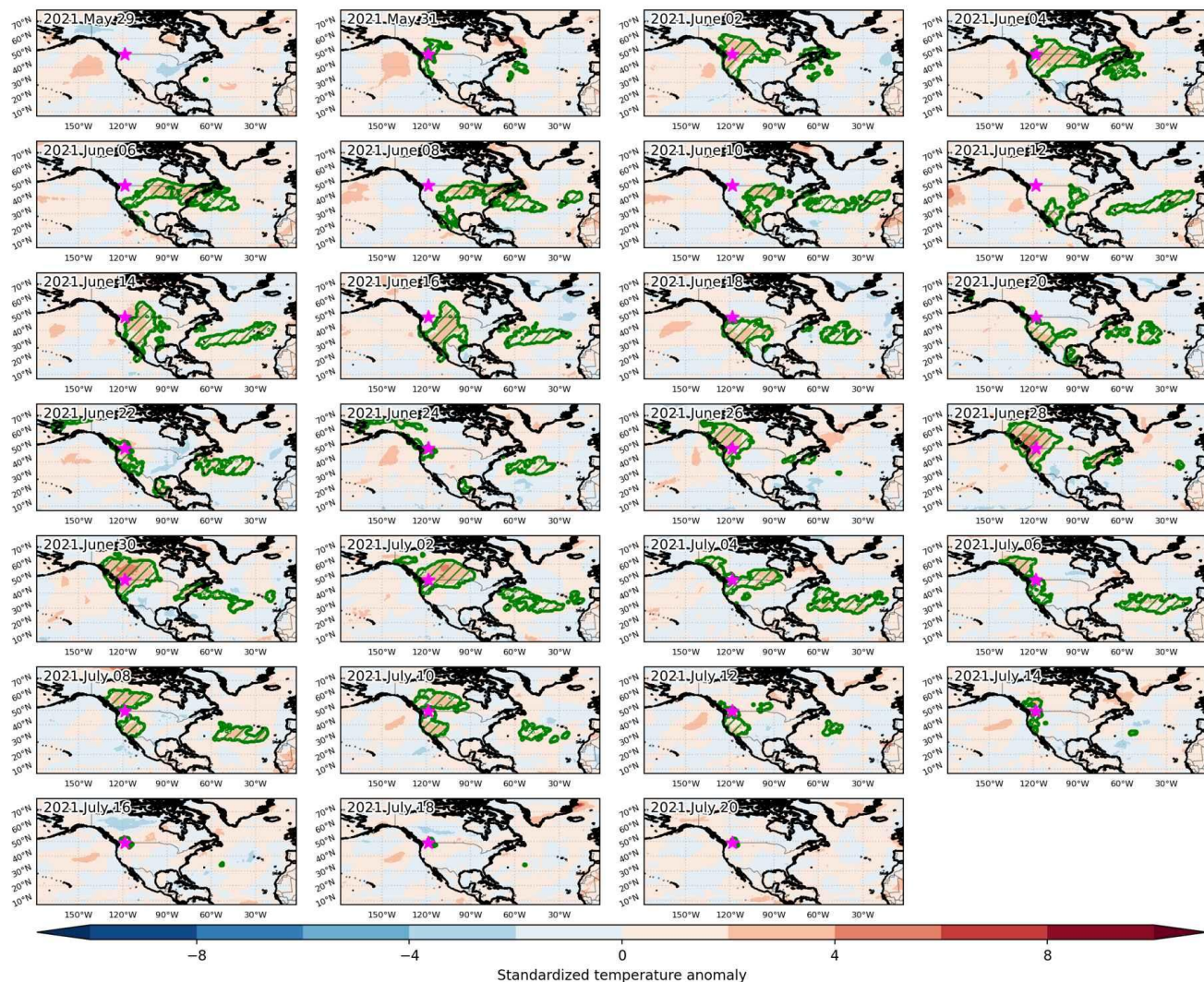


Figure 9: Maps of $\tilde{T}_{2m}$ (color shadings) for selected days throughout the event. Areas under the event's influence are indicated by green contours and hatching. Dates are shown in each panel and the event centre is marked with a magenta star.

3.2 Interactive visualization tool

In addition to the pre-computed figures and statistics, an interactive visualization and analysis tool is available at <https://wexa.jamstec.go.jp/home/interactive/>. The tool enables users to search and display spatiotemporal events in the database according to various criteria, including affected countries, year and month of occurrence, event type, and severity (Fig. 10). It includes a search panel, a map displaying areas affected by events, visualizations of event frequencies and weather conditions, a table summarizing event properties, and several diagnostic figures illustrating event characteristics and temporal trends. The interface is highly interactive, allowing users to rescale charts and maps, as well as select events from the figures for further analysis.

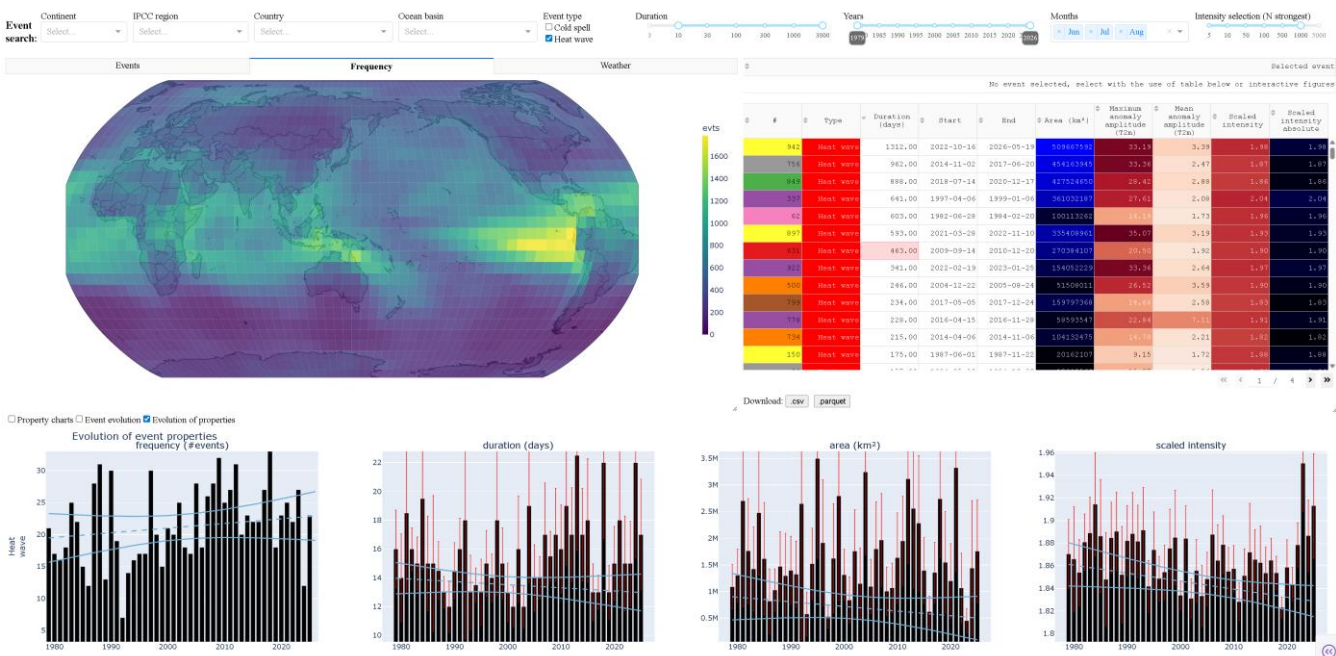


Figure 10: Screenshot of WEXA’s interactive visualization and analysis system. In this example, the system displays the 1000 strongest worldwide heat wave events occurring in June, July, and August, that lasted 10 days or longer between 1979 and 2026. The event selection panel is shown at the top. Below, a map on the left illustrates the frequency of the selected events, while a table on the right lists their properties. Below the table are two buttons that allow the user to download properties of the listed events in *csv* and *parquet* formats. The lower section presents charts depicting the frequency and evolution of event characteristics over time.

As an example of the platform’s more advanced analysis capabilities, we present a set of diagnostics that become available upon selecting an event (Fig. 11), either by selecting one from the event table or from a map showing the areas affected by all events that match the specified search criteria (Fig. 11a). For illustrative purposes, here we show the weather conditions that occurred on June 28 2021 during the 2021 Western North America Heat Wave that was discussed in the previous section. To find this event, users can select, for instance, the 50 strongest heat wave events that have affected Canada in June 2021. Upon applying these search criteria, the “Events” map displays the sectors affected by all events found (Fig. 11a). For this specific example, the targeted heat wave is shaded in green.

Once this event is selected, it becomes possible to visualize the associated meteorological conditions (Fig. 11b) by clicking the “Weather” tab of the map panel. From this view it is possible to choose both the date (here June 28, 2021) and the variable to display (here T'_{2m} ; choices include T_{2m} , T'_{2m} , and $\tilde{T}_{2m}$). Selecting an individual grid point from this map (here illustrated for 51.25°N, 126.25°W, a location in Southern British Columbia) reveals additional diagnostics, such as the temporal evolution of T_{2m} and T'_{2m} (Fig. 11c), as well as a histogram of historical T_{2m} for the corresponding calendar day, showcasing the



probability of the specific event (Fig. 11d). We can infer from these visualisations that the peak 2-m temperature anomaly of this specific event, T'_{2m} , was about 12.5 °C above normal, which exceeded all previously recorded T_{2m} for that time of the year.

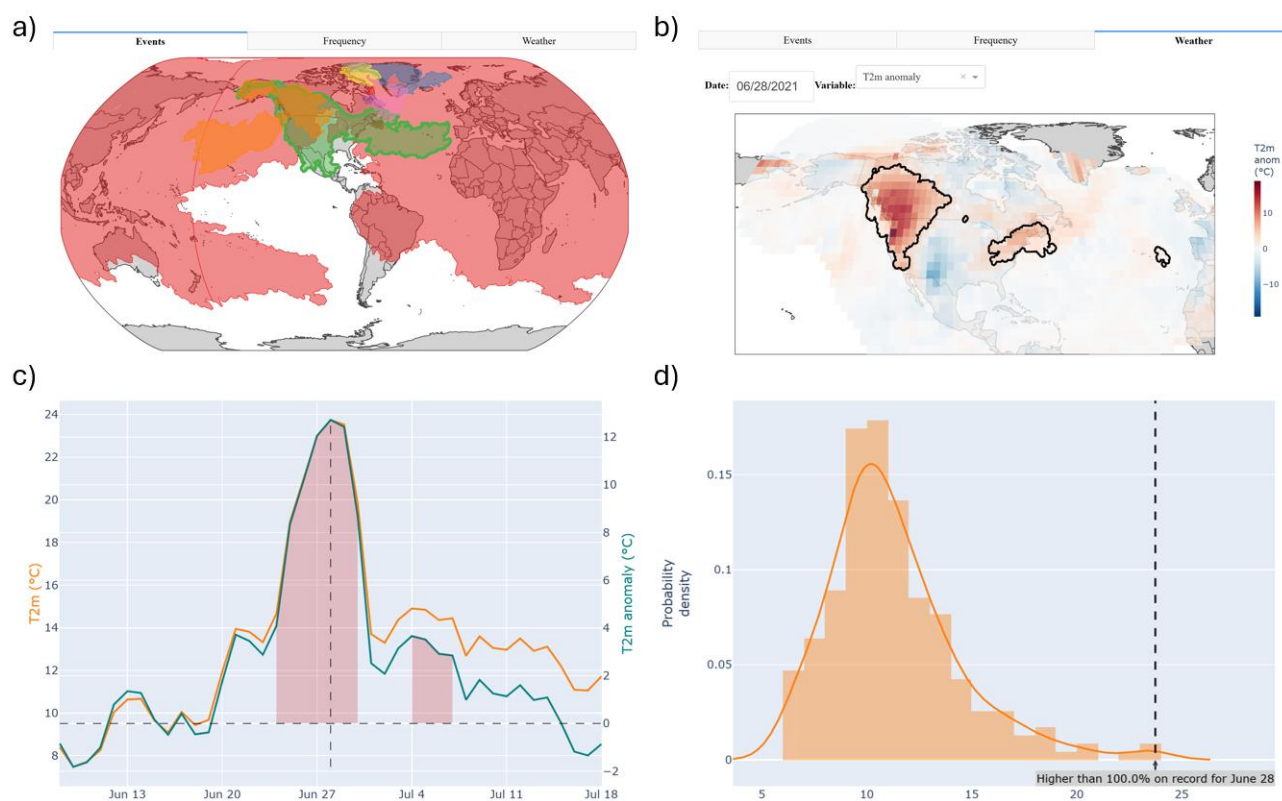


Figure 11: Screenshots of WEXA's advanced interactive analysis capabilities. a) The spatial extent of events corresponding to specific search criteria is shown with coloured shadings. b) 2-m temperature anomalies (T'_{2m} ; shading) for a selected date during an event, with the event's area contoured in black. c) 2-m temperature (T_{2m} ; orange) and its anomalies (T'_{2m} ; green) centred on the selected date for a selected geographical coordinate. The periods defined as heat waves are shaded in red. Cold spells would be shown in blue. d) Histogram of T_{2m} from 1979 to present (orange bars) with the corresponding Gaussian kernel density estimate (orange line) overlaid. The T_{2m} experienced on the selected date and location is indicated with a vertical black dashed line.

4 A note on exceptionally persistent and large-scale spatiotemporal mega-events and the extreme heat waves of 2025

Some detected spatiotemporal events exhibit exceptionally long durations. The most extreme case identified in our dataset persisted from October 15, 2022, to March 27, 2026, spanning 1,259 days. Notably, 27 March 2026 corresponds to the final day available at the time of the current analysis; the event may therefore extend beyond this date. As discussed below, this event affected a substantial fraction of the globe and thus can be classified as a *mega-event*.



Figure 12 illustrates the evolution of this event. It originated over central Africa (Fig. 12a) in the mid-October 2022, and moved toward North Africa approximately two months later (Fig. 12b). By January 2023, the event encompassed multiple regions across the Northern Hemisphere (Fig. 12c), reflecting either fragmentation of the original structure or the merging of previously distinct centres of action. By June 2023, the event exhibited centers of activity in both hemispheres (Fig. 12d), with a particularly strong presence in the equatorial eastern Pacific. One year later (Fig. 12e), the event continued to display multiple centres of action, especially across the tropics and subtropics.

A particularly notable phase occurred in August 2025 (Fig. 12f), when large parts of Europe were affected. During this period, the 2025 European heat wave set numerous temperature records across European cities and led to substantial socioeconomic impacts, including widespread wildfires, disruptions to education systems, and excess mortality in the thousands (Clarke et al., 2025). At the same time, warm anomalies developed over the eastern North Pacific, propagated eastward, and contributed to a major heat wave over western Canada (Fig. 12g) (Rapid Extreme Weather Attribution system: Top heat events of summer 2025, 2026). By early 2026 (Fig. 12h), the mega-event maintained a strong and geographically extensive global footprint. A detailed day-by-day evolution is provided in Supplementary Video 1, and readers are encouraged to explore the detailed evolution of this mega-event using the interactive analyser described above.

The emergence of such mega-events is consistent with expectations under a warming climate. As the atmosphere warms, a larger fraction of the Earth's surface is expected to exceed fixed heat-wave thresholds, leading to increases in both the spatial extent and temporal persistence of heat waves. This, in turn, raises the likelihood of large, spatiotemporally connected structures such as the one illustrated in Fig. 12. The effect may be particularly pronounced for marine heat waves, which are projected to become quasi-permanent by the late 21st century due to substantial increases in sea surface temperatures (Oliver et al., 2019).

One possible strategy to moderate the detection of the apparent increase in event duration and extent would be to partition mega-events into sub-events based on stages of their life cycle (e.g., splitting, temporary dissipation, and reintensification). However, event subdivision criteria would inevitably involve a degree of subjectivity. For this reason, we do not subdivide events in the present dataset. Instead, we recommend that users analyse the structure of these complex events by examining their day-to-day evolution and adjust/crop them to their specific usage.

A second option would be to adopt a stricter definition of extremes (e.g., the 95th percentile instead of the 90th percentile). This would reduce the spatial extent of detected extremes and decrease the likelihood of large-scale clustering into persistent mega-events. However, this approach has two limitations: (1) it would lead to under-detection of events in the earlier portion of the record, and (2) it would not eliminate mega-events under continued warming, but merely delay their emergence.



A third approach is to allow the baseline climatology to evolve over time (evolving baseline), thereby reducing exceedances driven by long-term warming. Such shifting-baseline frameworks have been proposed for marine heat waves (Smith et al., 2025). However, this approach remains debated, as it may misleadingly suggest that the frequency of extremes is unchanged under a globally warming climate (Sen Gupta, 2023). In this study, we therefore report events identified using a fixed baseline (1991-2020). Nevertheless, evolving-baseline approaches may be appropriate for specific applications, and its inclusion in future dataset versions remains under consideration.

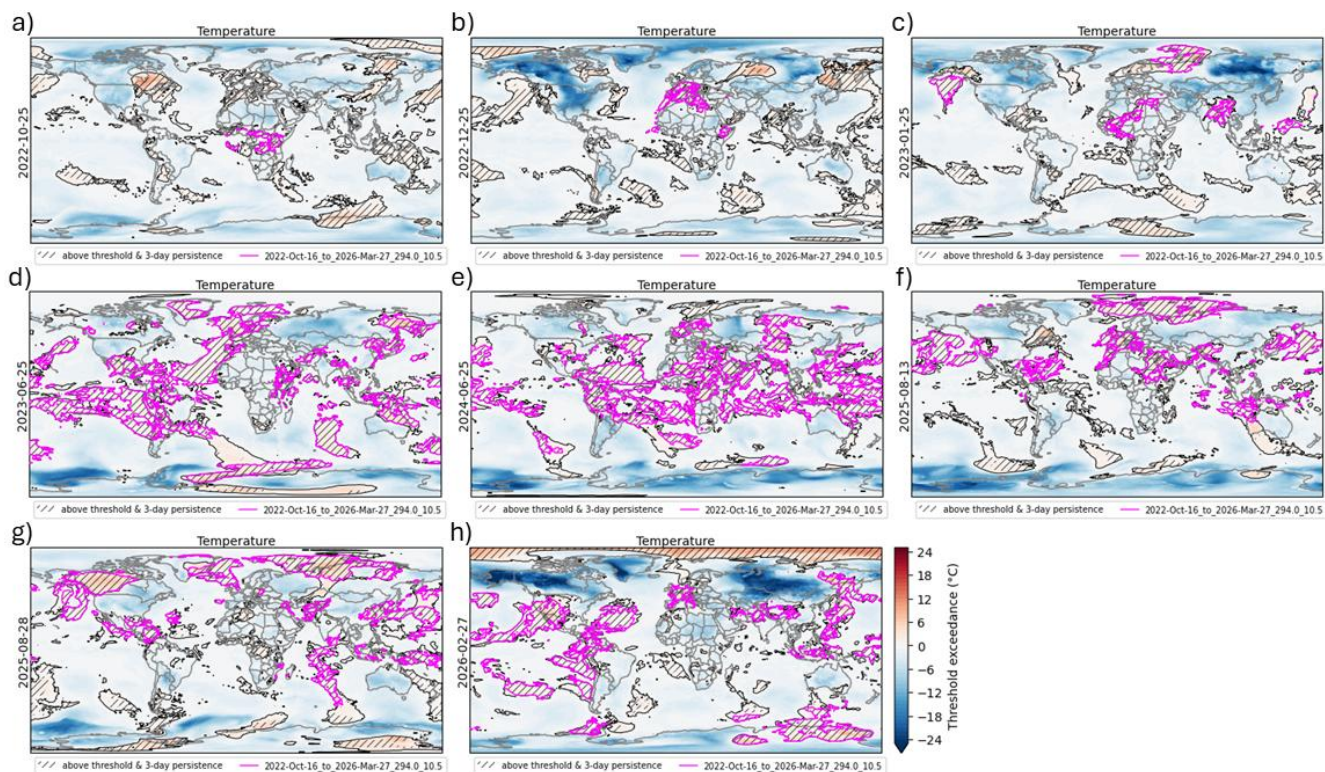


Figure 12: Selected snapshots illustrating the evolution of the 2022-Oct-16 to 2026-Mar-27 294.0_10.5 event. Shading represents the threshold exceedance ($T_{2m} - T_{2m}^{90}$), with black contours marking the 0-exceedance boundary. Regions of positive exceedance are outlined in black while those satisfying the persistence criterion are hatched in black. The sectors included in the 2022-Oct-16 to 2026-Mar-27 294.0_10.5 event are highlighted with magenta contours.

5 Summary

The Weather Extremes Archive (WEXA) is a global, event-based dataset designed to systematically identify, characterize, and visualize persistent temperature extremes, with an initial focus on heat waves and cold spells. Using 2-m temperature from ERA5 reanalysis data, extremes are defined relative to seasonally- and spatially-varying percentile thresholds derived from a



fixed 1991–2020 climatological baseline period. Persistent events are detected by identifying consecutive daily extremes and subsequently grouped into spatiotemporally coherent heat wave and cold spell events using connected component labelling in three-dimensional space (time, latitude, longitude).

For each spatiotemporal event, WEXA provides a standardized and comprehensive set of diagnostics describing its timing, duration, spatial extent, intensity, and evolution. Event intensity is quantified using both raw and standardized temperature anomalies, enabling meaningful comparisons across regions and seasons. Events exceeding a minimum spatial extent and occurring since 2020 are further documented with dedicated web pages and precomputed figures illustrating their severity, temporal development, and spatial progression.

In addition to the static archive, WEXA offers an interactive visualization and analysis platform that allows users to search, filter, and explore events based on multiple criteria, including location, season, event type, and intensity. Advanced diagnostics enable users to examine meteorological conditions at selected times and locations, view historical temperature distributions, and place individual events within a broader climatological context.

By providing a consistent, transparent, and user-friendly framework for analysing persistent temperature extremes, WEXA serves as a valuable resource for climate monitoring, event attribution, and impact-oriented studies. In addition, it may also serve as an information resource for the public. Access to the various components of WEXA is summarized in Table 3. Future releases will extend the archive to include additional categories of hydroclimatic extremes, such as marine heat waves and cold spells, further broadening its applicability.

Table 3: Summary of WEXA’s components and access method

Component	Description	URLs
Website	Main website of WEXA from which all other components are accessible and where pre-processed diagnostics of all events starting from 2020 are archived. Event properties stored in JSON format are accessible from these pages.	https://wexa.jamstec.go.jp (main website) https://wexa.jamstec.go.jp/home/recent-extreme-weather-events/ (most recent events)
Interactive analyser	Interactive analysis system for extreme events. It is possible to	https://wexa.jamstec.go.jp/home/interactive/



generate .csv and .parquet files
 containing event properties after
 applying search criteria.

NetCDF Data server for the gridded <https://www.jamstec.go.jp/RID/thredds/catalog/testAll/WEXA/catalog.html>
 server NetCDF files containing event
 labels.

Digital The data is permanently <https://doi.org/10.5281/zenodo.20550615>
 repository archived on Zenodo

364

365 **6 Data availability**

366 In addition to an official permanent repository for WEXA’s data hosted at <https://doi.org/10.5281/zenodo.20550615>
 367 (Martineau et al., 2026), all data can be accessed from WEXA’s official website (<https://wexa.jamstec.go.jp>). See Table 3 for
 368 more information on the different components of WEXA.

369 **Author contribution**

370 PM, IR, AY, and MN contributed to the study conceptualization. PM developed the event detection code and interactive
 371 analyzer. HF configured and managed the servers supporting the website and interactive analyzer. All authors contributed to
 372 the writing and revision of the manuscript.

373 **Competing interests**

374 The authors declare that they have no conflict of interest

375

376 **Acknowledgements**

377

378 We are grateful to Ryo Furue and Hisashi Nakamura for their valuable comments and suggestions regarding the database.
 379 Large language models were utilized solely to improve the English grammar, syntax, and phrasing of this manuscript. All
 380 intellectual content, data analysis, and scientific conclusions were generated entirely by the authors.



References

- Alam, Md. M., Mahtab, A. S. M., Ahmed, M. R., and Hassan, Q. K.: A Percentile Method to Determine Cold Days and Spells in Bangladesh, *Applied Sciences*, 13, 7030, <https://doi.org/10.3390/app13127030>, 2023.
- Alexander, L. V., Zhang, X., Peterson, T. C., Caesar, J., Gleason, B., Klein Tank, A. M. G., Haylock, M., Collins, D., Trewin, B., Rahimzadeh, F., Tagipour, A., Rupa Kumar, K., Revadekar, J., Griffiths, G., Vincent, L., Stephenson, D. B., Burn, J., Aguilar, E., Brunet, M., Taylor, M., New, M., Zhai, P., Rusticucci, M., and Vazquez-Aguirre, J. L.: Global observed changes in daily climate extremes of temperature and precipitation, *J. Geophys. Res.*, 111, 2005JD006290, <https://doi.org/10.1029/2005JD006290>, 2006.
- Anderson, G. B. and Bell, M. L.: Heat Waves in the United States: Mortality Risk during Heat Waves and Effect Modification by Heat Wave Characteristics in 43 U.S. Communities, *Environ Health Perspect.*, 119, 210–218, <https://doi.org/10.1289/ehp.1002313>, 2011.
- Bonino, G., Masina, S., Galimberti, G., and Moretti, M.: Southern Europe and western Asian marine heatwaves (SEWA-MHWs): a dataset based on macroevents, *Earth Syst. Sci. Data*, 15, 1269–1285, <https://doi.org/10.5194/essd-15-1269-2023>, 2023.
- Bunting, E. L., Tolmanov, V., and Keellings, D.: What is a heat wave: A survey and literature synthesis of heat wave definitions across the United States, *PLOS Clim.*, 3, e0000468, <https://doi.org/10.1371/journal.pclm.0000468>, 2024.
- Chen, R., Wang, C., Meng, X., Chen, H., Thach, T. Q., Wong, C.-M., and Kan, H.: Both low and high temperature may increase the risk of stroke mortality, *Neurology*, 81, 1064–1070, <https://doi.org/10.1212/WNL.0b013e3182a4a43c>, 2013.
- Clarke, B., Konstantinoudis, G., Pinto, I., Barnes, C., Keeping, T., Otto, F., Gasparrini, A., Masselot, P., Mistry, M., Vicedo-Cabrera, A., and Theokritoff, E.: Climate change tripled heat-related deaths in early summer European heatwave, Grantham Institute report, 2025.
- De Vries, H., Haarsma, R. J., and Hazeleger, W.: Western European cold spells in current and future climate, *Geophysical Research Letters*, 39, 2011GL050665, <https://doi.org/10.1029/2011GL050665>, 2012.
- Dunn, R. J. H., Alexander, L. V., Donat, M. G., Zhang, X., Bador, M., Herold, N., Lippmann, T., Allan, R., Aguilar, E., Barry, A. A., Brunet, M., Caesar, J., Chagnaud, G., Cheng, V., Cinco, T., Durre, I., De Guzman, R., Htay, T. M., Wan Ibadullah, W. M., Bin Ibrahim, M. K. I., Khoshkam, M., Kruger, A., Kubota, H., Leng, T. W., Lim, G., Li-Sha, L., Marengo, J., Mbatha, S., McGree, S., Menne, M., De Los Milagros Skansi, M., Ngwenya, S., Nkrumah, F., Oonariya, C., Pabon-Caicedo, J. D., Panthou, G., Pham, C., Rahimzadeh, F., Ramos, A., Salgado, E., Salinger, J., Sané, Y., Sopaheluwakan, A., Srivastava, A., Sun, Y., Timbal, B., Trachow, N., Trewin, B., Van Der Schrier, G., Vazquez-Aguirre, J., Vasquez, R., Villarroel, C., Vincent, L., Vischel, T., Vose, R., and Bin Hj Yussof, M. N.: Development of an Updated Global Land In Situ-Based Data Set of Temperature and Precipitation Extremes: HadEX3, *JGR Atmospheres*, 125, e2019JD032263, <https://doi.org/10.1029/2019JD032263>, 2020.
- Frich, P., Alexander, L., Della-Marta, P., Gleason, B., Haylock, M., Klein Tank, A., and Peterson, T.: Observed coherent changes in climatic extremes during the second half of the twentieth century, *Clim. Res.*, 19, 193–212, <https://doi.org/10.3354/cr019193>, 2002.
- Hersbach, H., Bell, B., Berrisford, P., Hirahara, S., Horányi, A., Muñoz-Sabater, J., Nicolas, J., Peubey, C., Radu, R., Schepers, D., Simmons, A., Soci, C., Abdalla, S., Abellan, X., Balsamo, G., Bechtold, P., Biavati, G., Bidlot, J., Bonavita, M., De Chiara, G., Dahlgren, P., Dee, D., Diamantakis, M., Dragani, R., Flemming, J., Forbes, R., Fuentes, M., Geer, A., Haimberger, L., Healy, S., Hogan, R. J., Hólm, E., Janisková, M., Keeley, S., Laloyaux, P., Lopez, P., Lupu, C., Radnoti, G., de Rosnay, P.,



- 420 Rozum, I., Vamborg, F., Villaume, S., and Thépaut, J.: The ERA5 global reanalysis, *Quarterly Journal of the Royal*
421 *Meteorological Society*, 146, 1999–2049, <https://doi.org/10.1002/qj.3803>, 2020.
- 422 Knippertz, P., Evans, M. J., Field, P. R., Fink, A. H., Liousse, C., and Marsham, J. H.: The possible role of local air pollution
423 in climate change in West Africa, *Nature Clim Change*, 5, 815–822, <https://doi.org/10.1038/nclimate2727>, 2015.
- 424 Lavaysse, C., Cammalleri, C., Dosio, A., Van Der Schrier, G., Toreti, A., and Vogt, J.: Towards a monitoring system of
425 temperature extremes in Europe, *Nat. Hazards Earth Syst. Sci.*, 18, 91–104, <https://doi.org/10.5194/nhess-18-91-2018>, 2018.
- 426 Lei, J., Chen, R., Yin, P., Meng, X., Zhang, L., Liu, C., Qiu, Y., Ji, J. S., Kan, H., and Zhou, M.: Association between Cold
427 Spells and Mortality Risk and Burden: A Nationwide Study in China, *Environ Health Perspect*, 130, 027006,
428 <https://doi.org/10.1289/EHP9284>, 2022.
- 429 Li, Z., Fan, Y., Xu, Z., Ho, H. C., Tong, S., Huang, C., Bai, Z., Gai, Y., Cheng, W., Hu, J., Feng, Y., Zheng, H., Wang, N., Ni,
430 J., Pan, G., Hossain, M. Z., Su, H., and Cheng, J.: Exceptional heatwaves and mortality in Europe: Greater impacts since the
431 coronavirus disease 2019 outbreak, *Environmental Pollution*, 363, 125058, <https://doi.org/10.1016/j.envpol.2024.125058>,
432 2024.
- 433 Lo, S.-H., Chen, C.-T., Russo, S., Huang, W.-R., and Shih, M.-F.: Tracking heatwave extremes from an event perspective,
434 *Weather and Climate Extremes*, 34, 100371, <https://doi.org/10.1016/j.wace.2021.100371>, 2021.
- 435 Martineau, P., Richter, I., Yamamoto, A., Nonaka, M., and Fujio, H.: Weather Extremes Archive,
436 <https://doi.org/10.5281/ZENODO.20550616>, 2026.
- 437 Miller, S., Chua, K., Coggins, J., and Mohtadi, H.: Heat Waves, Climate Change, and Economic Output, *Journal of the*
438 *European Economic Association*, 19, 2658–2694, <https://doi.org/10.1093/jeea/jvab009>, 2021.
- 439 Oliver, E. C. J., Burrows, M. T., Donat, M. G., Sen Gupta, A., Alexander, L. V., Perkins-Kirkpatrick, S. E., Benthuisen, J. A.,
440 Hobday, A. J., Holbrook, N. J., Moore, P. J., Thomsen, M. S., Wernberg, T., and Smale, D. A.: Projected Marine Heatwaves
441 in the 21st Century and the Potential for Ecological Impact, *Front. Mar. Sci.*, 6, 734, <https://doi.org/10.3389/fmars.2019.00734>,
442 2019.
- 443 Osmolska, W., Chemel, C., Maycock, A., and Field, P.: A methodology for tracking cold spells in space and time: development,
444 evaluation and applications, *Weather Clim. Dynam.*, 6, 1221–1240, <https://doi.org/10.5194/wcd-6-1221-2025>, 2025.
- 445 Pantavou, K., Theoharatos, G., Nikolopoulos, G., Katavoutas, G., and Asimakopoulos, D.: Evaluation of thermal discomfort
446 in Athens territory and its effect on the daily number of recorded patients at hospitals' emergency rooms, *Int J Biometeorol*,
447 52, 773–778, <https://doi.org/10.1007/s00484-008-0170-7>, 2008.
- 448 Perkins, S. E.: A review on the scientific understanding of heatwaves—Their measurement, driving mechanisms, and changes
449 at the global scale, *Atmospheric Research*, 164–165, 242–267, <https://doi.org/10.1016/j.atmosres.2015.05.014>, 2015.
- 450 Perkins-Kirkpatrick, S. E. and Lewis, S. C.: Increasing trends in regional heatwaves, *Nat Commun*, 11, 3357,
451 <https://doi.org/10.1038/s41467-020-16970-7>, 2020.
- 452 Philip, S. Y., Kew, S. F., Van Oldenborgh, G. J., Anslow, F. S., Seneviratne, S. I., Vautard, R., Coumou, D., Ebi, K. L., Arrighi,
453 J., Singh, R., Van Aalst, M., Pereira Marghidan, C., Wehner, M., Yang, W., Li, S., Schumacher, D. L., Hauser, M., Bonnet,
454 R., Luu, L. N., Lehner, F., Gillett, N., Tradowsky, J. S., Vecchi, G. A., Rodell, C., Stull, R. B., Howard, R., and Otto, F. E. L.:
455 Rapid attribution analysis of the extraordinary heat wave on the Pacific coast of the US and Canada in June 2021, *Earth Syst.*
456 *Dynam.*, 13, 1689–1713, <https://doi.org/10.5194/esd-13-1689-2022>, 2022.



- 457 Rosenfeld, A. and Pfaltz, J. L.: Sequential Operations in Digital Picture Processing, J. ACM, 13, 471–494,
 458 <https://doi.org/10.1145/321356.321357>, 1966.
- 459 Russo, S., Sillmann, J., and Fischer, E. M.: Top ten European heatwaves since 1950 and their occurrence in the coming decades,
 460 Environ. Res. Lett., 10, 124003, <https://doi.org/10.1088/1748-9326/10/12/124003>, 2015.
- 461 Sen Gupta, A.: Marine heatwaves: definition duel heats up, Nature, 617, 465–465, [https://doi.org/10.1038/d41586-023-01619-](https://doi.org/10.1038/d41586-023-01619-4)
 462 4, 2023.
- 463 Smith, E. T. and Sheridan, S. C.: The influence of atmospheric circulation patterns on cold air outbreaks in the eastern United
 464 States, Intl Journal of Climatology, 39, 2080–2095, <https://doi.org/10.1002/joc.5935>, 2019.
- 465 Smith, K. E., Sen Gupta, A., Amaya, D., Benthuyssen, J. A., Burrows, M. T., Capotondi, A., Filbee-Dexter, K., Frölicher, T.
 466 L., Hobday, A. J., Holbrook, N. J., Malan, N., Moore, P. J., Oliver, E. C. J., Richaud, B., Salcedo-Castro, J., Smale, D. A.,
 467 Thomsen, M., and Wernberg, T.: Baseline matters: Challenges and implications of different marine heatwave baselines,
 468 Progress in Oceanography, 231, 103404, <https://doi.org/10.1016/j.pocean.2024.103404>, 2025.
- 469 Rapid Extreme Weather Attribution system: Top heat events of summer 2025: [https://climatedata.ca/news/rapid-extreme-](https://climatedata.ca/news/rapid-extreme-weather-attribution-system-top-heat-events-of-summer-2025/)
 470 [weather-attribution-system-top-heat-events-of-summer-2025/](https://climatedata.ca/news/rapid-extreme-weather-attribution-system-top-heat-events-of-summer-2025/), last access: 5 June 2026.
- 471 Tetzlaff, E. J., Goulet, N., Gorman, M., Richardson, G. R., Enright, P. M., Henderson, S. B., and Kenny, G. P.: Media-Based
 472 Post-Event Impact Analysis of the 2021 Heat Dome in Canada, Environ Health Insights, 18, 11786302241276669,
 473 <https://doi.org/10.1177/11786302241276669>, 2024.
- 474 Thompson, V., Kennedy-Asser, A. T., Vosper, E., Lo, Y. T. E., Huntingford, C., Andrews, O., Collins, M., Hegerl, G. C., and
 475 Mitchell, D.: The 2021 western North America heat wave among the most extreme events ever recorded globally, Sci. Adv.,
 476 8, eabm6860, <https://doi.org/10.1126/sciadv.abm6860>, 2022.
- 477 Tringa, E., Tolika, K., Anagnostopoulou, C., and Kostopoulou, E.: A Climatological and Synoptic Analysis of Winter Cold
 478 Spells over the Balkan Peninsula, Atmosphere, 13, 1851, <https://doi.org/10.3390/atmos13111851>, 2022.
- 479 Van Der Walt, S., Schönberger, J. L., Nunez-Iglesias, J., Boulogne, F., Warner, J. D., Yager, N., Gouillart, E., and Yu, T.:
 480 scikit-image: image processing in Python, PeerJ, 2, e453, <https://doi.org/10.7717/peerj.453>, 2014.
- 481 Vavrus, S., Walsh, J. E., Chapman, W. L., and Portis, D.: The behavior of extreme cold air outbreaks under greenhouse
 482 warming, Int. J. Climatol., 26, 1133–1147, <https://doi.org/10.1002/joc.1301>, 2006.
- 483 WMO: WMO Guidelines on the Calculation of Climate Normals, World Meteorological Organization Geneva, Switzerland,
 484 <https://doi.org/https://library.wmo.int/idurl/4/55797>, 2017.

485