# On the vertical structure and propagation of marine heatwaves in the Eastern Pacific

**Eike E. Köhn[1], Meike Vogt[1], Matthias Münnich[1], and Nicolas Gruber[1,2]**

[1]Environmental Physics, Institute of Biogeochemistry and Pollutant Dynamics, ETH Zurich, 8092 Zürich, Switzerland
[2]Center for Climate Systems Modeling, ETH Zurich, 8092 Zürich, Switzerland

**Key Points:**

- About a third of marine heatwaves (MHW) in the Eastern Pacific extend below the mixed layer. 10 % of MHWs extend on average deeper than 150m.
- 59 % of deep-reaching MHWs (dMHWs) are partly invisible at the surface, as they either deepen, shoal or exhibit multi-surfacing behaviour.
- As initial tests indicate that the surface-based detection of dMHWs seems feasible, dMHWs might be detectable using remote sensing data.

Corresponding author: E. E. Köhn, `eike.koehn@usys.ethz.ch`

**Abstract**

Marine heatwaves (MHWs) have been recognized as a serious threat to marine life, yet, most studies so far have focused on the surface only. Here, we investigate the vertical dimension and propagation of surface MHWs in the Eastern Pacific using results from a high-resolution hindcast simulation (1979 to 2019), performed with the Regional Ocean Modeling System. We detect MHWs using a seasonally varying percentile threshold on a fixed baseline and track their vertical propagation across the upper 500 m. We find that nearly a third ($\sim 29\,\%$) of the MHWs extend beyond the surface mixed layer depth (MLD). On average, these deep-reaching MHWs (dMHWs) extend to 110 m below the MLD and last five times longer than MHWs that are confined to the mixed layer (184 vs. 36 days). The dMHWs can cause stronger temperature anomalies at depth than at the surface (maximum intensity of 5.0 °C vs. 1.9 °C). This general subsurface MHW intensification even holds when scaling the temperatures with the respective local variability. A clustering of dMHWs reveals that 41 % of them are block-like, i.e., continually remain in contact with the sea surface, 24 % propagate downward, 20 % propagate upward, while 15 % appear at the surface multiple times. Although the water column MHW duration, intensity and severity are only moderately correlated with their corresponding surface-based MHW characteristics, dMHWs have the potential to be detected from the surface. Our study can help to augment the remote sensing-based monitoring of upper ocean exposure to MHWs.

## Plain Language Summary

Periods of extremely warm water temperatures, referred to as marine heatwaves (MHWs), pose a serious threat to marine life. While many studies have analyzed MHW properties at the sea surface, little is known about their subsurface nature. We here use a new approach to define MHWs throughout the water column and study the vertical structure and propagation behaviour of MHWs in the Eastern Pacific based on a high-resolution numerical hindcast simulation (1979 to 2019). We find that nearly a third of MHWs that are discernible at the sea surface reach below the layer of well-mixed waters near the sea surface, the so-called mixed layer. These deep-reaching MHWs last on average longer than mixed layer confined MHWs (184 vs. 36 days) and are marked by larger temperature threshold exceedances (maximum intensity of 5.0 °C vs. 1.9 °C). We furthermore identify four distinct vertical propagation patterns associated with the deep-reaching MHWs: 41 % continually affect the sea surface, 24 % propagate downwards, 20 % propagate upward, while 15 % appear at the surface multiple times. Overall, these analyses show that MHWs can affect the water column to a larger degree than diagnosed based on the sea surface only.

## 1 Introduction

The observed warming of the global ocean has increased already markedly the frequency, duration, and intensity of marine heatwaves (MHWs) (Oliver et al., 2018), commonly defined as prolonged periods of unusually high temperatures (IPCC SROCC, 2022). MHWs affect marine organisms and ecosystems profoundly (Wernberg et al., 2013; Cavole et al., 2016; Smale et al., 2019; Sen Gupta et al., 2020), as well as the economies depending on services provided by these systems (Cheung et al., 2021). Further ocean warming will increase the key impact metrics of MHW manifold (Frölicher et al., 2018), thus increasing the threats to the future health of the oceans. It is thus not surprising that the study of MHW has experienced a rapid surge in attention by the scientific community, policymakers, and the public (Collins et al., 2019).

So far, most MHW studies have focused on the surface only. This is largely a consequence of the easy accessibility of high resolution satellite sea surface temperature (SST) data, while no corresponding observation-based product is available to study subsurface

extremes (Oliver et al., 2021). But even model-based studies have rarely analyzed MHWs beyond the surface (e.g. Amaya et al., 2023), such that the vertical structure of surface MHWs has remained largely elusive (Gruber et al., 2021). Gaining more insights into the vertical extent and structure of MHWs is crucial to better assess the MHW drivers and the mechanisms sustaining them (Oliver et al., 2021). A look beyond the surface is also critically important to assess and understand the impacts of MHWs on upper ocean ecosystems. The vertical MHW structure can be particularly relevant for animals at higher trophic levels whose vertical habitats extend beyond the mixed layer (ML) or euphotic zone and which often perform diurnal vertical migration (Steinberg & Landry, 2017; Bianchi & Mislan, 2016). Deep extending MHWs may limit organisms in their ability to use vertical migration as a strategy to avoid the impact of extremely high surface ocean temperatures.

While a basin scale assessment of the depth structure of MHWs has not been conducted to date, a few studies have analyzed the vertical extent of individual events and linked them to different drivers. For example, Elzahaby and Schaeffer (2019) analyzed vertical temperature anomaly profiles from Argo floats during surface MHWs off Eastern Australia. The authors identify and link upper ocean MHWs (shallower than 150 m depth) to physical driving processes at the ocean surface, while associating deep-reaching MHWs (deeper than 800 m) to warm core eddies. Similarly, Scannell et al. (2020) and Johnson et al. (2022) used Argo float data to investigate the vertical extent and propagation of the Northeast Pacific *Blob* heatwave, i.e., the strong surface MHW in the Northeastern Pacific that lasted for multiple years (2013-2016) Bond et al. (2015); Di Lorenzo and Mantua (2016) and was likely even a compound extreme (Gruber et al., 2021). Scannell et al. (2020) demonstrated that the heat anomalies from this event were lingering for years at depth affected ecosystems for much longer periods than diagnosed from the surface only (see also (Jackson et al., 2018; Freeland & Ross, 2019; Holser et al., 2022). Due to the wide spacing ($\mathcal{O}$ of 100 km) of the Argo float profiles and their 10 day repeat cycle, these authors had to average the data considerably in time and space, thereby losing much detail. The low temporal resolution was overcome in the studies by Schaeffer and Roughan (2017) and Hu et al. (2021) by using high-frequency mooring time series throughout the upper ocean off Eastern Australia and in the western tropical Pacific, respectively to demonstrate that MHWs can reach well below the ML in conjunction with downwelling winds and wind-driven mixing. But the single point nature of the mooring data and the very limited number of mooring sites strongly limits this approach.

Results from ocean models can overcome these observational limitations, but the number of studies having investigated the vertical structure and propagation of MHWs is rather small. Notable exceptions are the studies of Ryan et al. (2021) and Großelindemann et al. (2022) who used regional models to elucidate the depth structure of MHWs off Western Australia and on the Northwest Atlantic Shelf, respectively. They highlighted the important role of thermocline depth variability and (eddy driven) current anomalies in driving deep-reaching MHWs.

Extending the study of MHWs beyond the surface poses some challenges. For example, the classical "Eulerian" perspective on MHWs is not well suited to track any vertical propagation behaviour, as it identifies and groups adjacent days of extremely warm temperatures solely in the temporal dimension (Gruber et al., 2021). A few studies have therefore extended the definition of marine extremes by grouping extreme states in time *and* space (Di Biagio et al., 2020; Desmet et al., 2022). We adopt here a so far unexplored one-dimensional water column perspective to study the vertical propagation of MHWs that affect the sea surface (Scannell et al., 2020). In this perspective, extremes are defined by grouping adjacent extreme states in the temporal and vertical dimension, which allows extremes to have a vertical extent and to propagate vertically while staying fixed at one horizontal location. This perspective permits to elucidate how deep MHWs reach and for how long they affect the upper ocean water column at a particular horizontal

location. Furthermore, the ability to track the vertical propagation of MHWs allows us to assess whether the vertical propagation of the extreme signal, as found by Scannell et al. (2020) during the *Blob*, is a common behaviour of MHWs. This view can be instructive to better understand the biological impacts of MHWs, particularly regarding resident species that are exposed to extreme temperatures across many depth levels and species that perform diurnal vertical migration (DVM, e.g., Bianchi and Mislan (2016)).

In this study we aim to increase our understanding of the vertical structure of MHWs and thus learn more about their driving mechanisms (Holbrook et al., 2019; Sen Gupta et al., 2020; Vogt et al., 2022). The Eastern Pacific (EP) is used as a pilot region since it harboured some of the strongest MHWs in the last decades, including the *Blob* MHW from 2013 to 2016 (Di Lorenzo & Mantua, 2016), the "Blob2.0" MHW in 2019 (Amaya et al., 2020) and many MHWs associated with El Niño conditions (Holbrook et al., 2019). We hence analyze the vertical structure of MHWs in a hindcast simulation of the EP. First, we want to understand how deep MHWs extend into the ocean interior below the ML. Our second goal is to understand how the subsurface structure of MHWs influences their general characteristics, that are commonly studied from the satellite perspective at the sea surface only. In this context, we aim to understand how representative surface-derived MHW characteristics such as duration or intensity are for MHWs that occupy the water column beyond the surface ML depth (MLD). Finally, our here introduced methodology to define MHWs in the water column further allows us to study the vertical propagation behaviour of MHWs across the EP basin.

## 2 Data and Methods

### 2.1 Model hindcast

We used the UCLA-ETH version of the Regional Ocean Modeling System (ROMS, Marchesiello et al. (2003); Shchepetkin and McWilliams (2005)) to perform the model hindcast. We employed a telescopic model grid focused on the tropical Southeast Pacific, but covering the entire Pacific basin (Fig. 1a). In this grid, the resolution is finest along the Peruvian coast (∼4 km) and decreases towards the western Pacific, with the coarsest resolution of ∼40 km occurring south of Australia (Fig. S1). This model setup, the initialization, and much of the simulation protocols follows the work of Köhn et al. (2022). The most important change is that we have improved the numerical representation of advection in the model by using an isoneutral advection scheme, and by employing a $3^{\text{rd}}$ order WENO advection scheme in the horizontal direction and a $5^{\text{th}}$ order scheme for the vertical dimension (Liu et al., 1994; Shu, 1998)). We also updated the atmospheric forcing from ERA interim (Dee et al., 2011) to ERA5 (Hersbach et al., 2020). For further details the reader is referred to Köhn et al. (2022). Here we provide a summary of the simulation protocol.

After a 20-year model spin-up from World Ocean Atlas 2018 (Boyer et al., 2018), we performed a hindcast simulation from 1979 to 2019 forced with ERA5 reanalysis data (Hersbach et al., 2020), to which we applied the Drakkar Forcing Set (DFS5.2) correction (Dussin et al., 2016) for the incoming shortwave and outgoing longwave radiation (following Desmet et al. (2022)). The model spin-up was undertaken with a normal-year forcing (Large & Yeager, 2004), as described in Köhn et al. (2022). Along the open boundaries in the Southern Ocean, we use SODA3.4.2 reanalysis (Carton et al., 2018) as time-varying boundary conditions. We integrated the model from 1979 onward with a timestep of 10 min and wrote out daily averages of the state variables across all depths. This yielded over the 1979 to 2019 period a total of 14 975 days of model output. We then regridded the model data from the bathymetry following vertical coordinate (i.e., s-levels, Song and Haidvogel (1994)) to 37 fixed depth levels (i.e., z-levels) in the upper 500 m, while maintaining the increasing resolution towards the surface (Fig. S2 in the Supplementary Information).

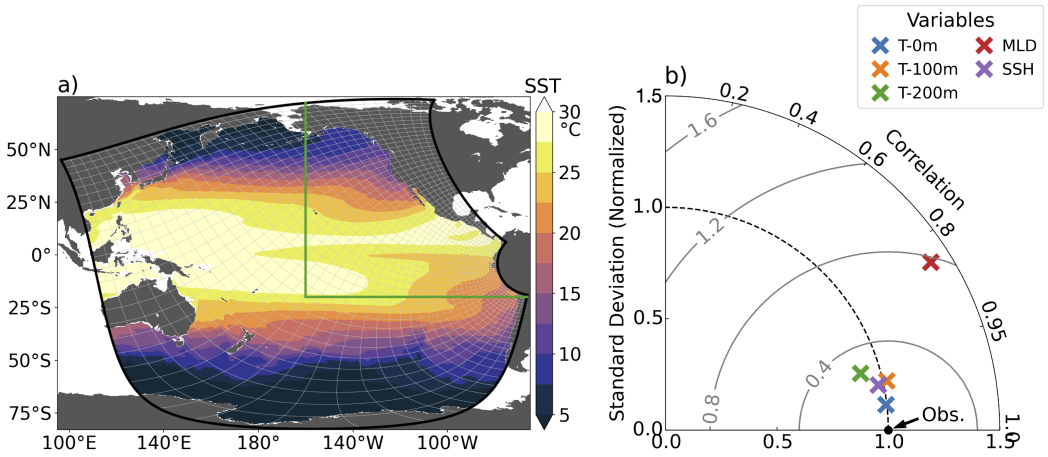



**Figure 1.** a) Model grid with SST averaged over full model hindcast. Grey lines show every $20^{\text{th}}$ grid point of the native model grid. b) Taylor Diagram of model SST, MLD and SSH based on monthly climatologies computed over the entire hindcast and compared to monthly observational climatologies interpolated onto the ROMS grid. The data feeding into the Taylor Diagrams is limited to the focus region of the study, i.e., the Eastern Pacific east of 160°W and north of 20°S, delimited by the green lines in panel a.

To reduce the data size, we subsequently downsampled the model output by averaging the regridded data to 3x3 horizontal grid boxes. This way we reduced the output's nominal horizontal resolution, i.e., to ∼12 km off Peru, but maintained the telescopic grid structure (Fig. 1a). We limited all following analyses to the downsampled data in the upper 500 m of the water column. From the regridded and downsampled model output, we calculated the MLD using the density threshold criterion of Holte et al. (2017), that is a density change relative to the surface density value corresponding to a temperature change of 0.2 °C.

We limit the study domain to the EP region east of 160°W and north of 20°S (Fig. 1a) where simulated local SST biases are mostly smaller than 0.5 °C (Fig. S3a). In this area monthly climatologies of the modeled temperature, MLD and sea surface height (SSH) fields (calculated from 1979 to 2019) distribution (Fig. 1b) show a high spatio-temporal correlation and comparable standard deviations with observed monthly climatologies (Fig. 1b). A more in-depth evaluation of model biases, variability and trends can be found in Section 3 and in the Supplementary Information. These analyses reveal overall a high fidelity of the model in capturing the mean state of the Pacific and its main pattern of variability, giving us high confidence in the robustness of the results generated by this model (Fig. 1b).

### 2.2 Detecting MHWs across the upper ocean

In order to study the vertical dimension of MHWs, we take a vertical one-dimensional perspective to identify MHW events across the upper 500 m. In this water column perspective, extreme temperatures are grouped in time and in depth to MHWs that can hence propagate vertically throughout the upper water column. We limit our analyses to those MHWs that were in contact with the surface at least once in their lifetime, and among those we mostly focus on the deep-reaching MHWs that extend below the MLD. In the following, we explain the individual steps, starting from the detection of extreme tem-

peratures (Sec. 2.2.1) over the joining of individual extreme cells to form MHW events (Sec. 2.2.2), to the definition of the key metrics (Sec. 2.3).

#### *2.2.1 Detecting extreme temperatures*

From the daily hindcast temperature data, we calculated at each horizontal and vertical grid point the seasonally varying MHW threshold temperature $T_{\text{thresh}}$ based on the local temperature time series following the methodology of Hobday et al. (2016) without de-trending the temperature data. To this end, we calculated the $90^{\text{th}}$ percentile threshold based on the 30-year reference period 1982 to 2011. We then detected extreme temperatures ($T$) by requiring $T > T_{\text{thresh}}$ for each grid point and day. We obtain a four-dimensional Boolean array $\mathbf{B}$ spanning across all 14 975 hindcast days (daily output from 1979 to 2019), 37 depth levels and $335 \times 232$ downsampled horizontal grid points.

To discard short extreme "heat spikes", we applied morphological operations to the Boolean array $\mathbf{B}$, consisting of a binary closing followed by a binary opening in the temporal dimension (see also Scannell et al. (2021)). Both operations were executed with a 5-day box-kernel. This temporal filtering of $\mathbf{B}$ ensures a minimum MHW duration of 5 days (congruent with previously defined minimum MHW durations, Hobday et al. (2016)) and guarantees that MHWs are at least 5 days apart from each other. Temporally isolated extreme spikes that last shorter than 5 consecutive days were thus removed (set to *False*), while extremes that are separated by less than 5 non-extreme days were merged.

#### *2.2.2 Defining MHWs in the water column*

We then grouped all extreme grid cells that are adjacent to each other in both time *and* depth to form a coherent MHW in the water column (Fig. 2b). In order to be adjacent, extreme grid cells had to share either a corner or an edge. In this one-dimensional water column detection, MHWs have a vertical extent and can grow, shrink and move vertically over time. This differs fundamentally from the usual (Eulerian) perspective, where extreme grid cells are only grouped in time at a given depth. We additionally identify MHWs using this classical approach at the sea surface (surface-only detection) and refer to them hereafter as "surface-only MHWs" (Fig. 2a). It is important to note that in our study here, all MHWs (and surface-only MHWs) are fixed to one horizontal location and cannot propagate laterally, as we do not connect adjacent extreme grid cells horizontally.

In the final step, we excluded all MHWs that never touch the sea surface, i.e., that do not contain any extreme surface grid cells. This excluded about 78 % of the detected one-dimensional MHWs in the upper 500 m water column (total MHW count 6 339 969) and 67 % of all extreme temperature grid cells. By retaining only the MHWs that can be identified at the sea surface, we can relate our results to the many previous studies that focused on MHWs at the sea surface (Oliver et al., 2018). We further separated the retained MHWs into a subset that remains within the ML (referred to as surface MHWs, sMHWs) and those MHWs that reach deeper into the ocean interior, referred to as deep-reaching MHWs (dMHWs, Fig. 3). In order to be classified as a dMHW, while excluding MHWs that move only shortly and slightly beneath the MLD (see Sec. 2.1), we required the MHWs to extend on average below a threshold defined by the instantaneous MLD plus its climatological standard deviation, which is seasonally varying and calculated over the 1982 to 2011 period (see dashed blue line in Fig. 2). These dMHWs are the main focus of our study.

### 2.3 MHW characteristics

To characterize the MHWs, a set of properties was calculated for each event (Tab. 1 and Fig. 2). The introduction of the vertical dimension allows for the definition of many

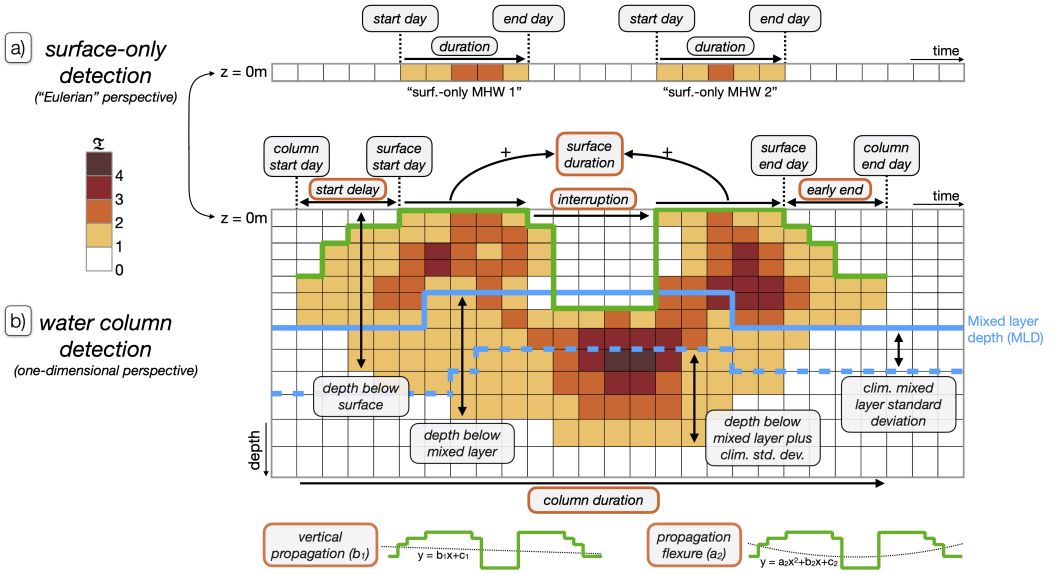



**Figure 2.** Detecting MHWs by aggregating extreme grid cells in two different ways. Panel a shows the classical surface-only detection (Eulerian perspective), while panel b) shows the one-dimensional water column detection of MHWs. Each box indicates a grid point in the time (panel a) and time-depth (panel b) dimension, i.e. the *event grid*. Colored boxes indicate extreme conditions (yellow to brown colors indicating increasing intensity), while white boxes indicate non-extreme conditions. Temperatures are expressed in a variability scaled form ($\mathfrak{T}$, see Sec. 2.3), with $\mathfrak{T} > 1$ indicating the presence of extreme temperatures. Panel a shows two individual surface MHWs, while panel b shows one single coherent and vertically propagating MHW. The solid and dashed blue line indicate the instantaneous MLD and the instantaneous MLD plus the climatological MLD standard deviation, respectively. The green line marks the upper boundary of the MHW. At the bottom of panel b the definition of the "vertical propagation" and the "propagation flexure" metrics are visualized, based on the upper MHW boundary, i.e., the green line. Grey text boxes contain labels and MHW characteristics as listed in Table 1. Text boxes outlined in orange highlight MHW metrics that are used in the clustering of deep-reaching MHWs.

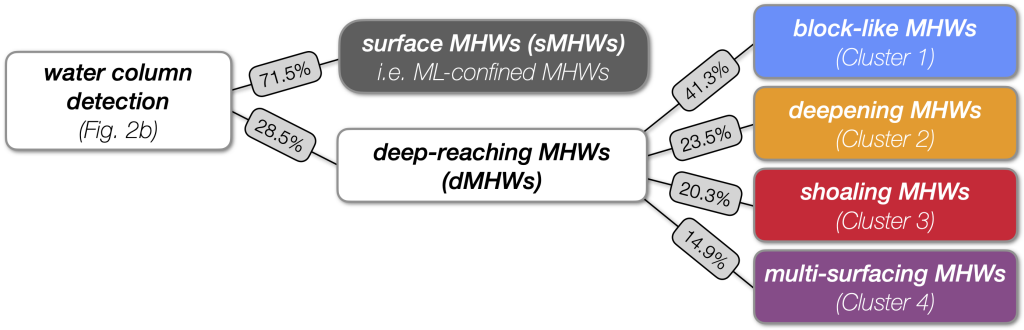



**Figure 3.** Overview of MHW definitions and distinctions based on the water column detection of MHWs. In a first step, MHWs are distinguished on the basis of their attained mean depth, especially with regard to the mixed layer depth. The deep reaching MHWs (dMHWs), i.e., those extending significantly below the mixed layer, are further split up into four different clusters based on their vertical propagation behaviour.

more characteristics than in the classical surface-only perspective (e.g. Hobday et al., 2016), which can be grouped into two classes: a) additional surface characteristics and b) column characteristics (Tab. 1).

| **MHW surface characteristics** | Units | **MHW column characteristics** | Units |
|---|---|---|---|
| surface start day | date | column start day | date |
| surface end day | date | column end day | date |
| surface duration (${}^{\mathrm{s}}D$) | days | column duration (${}^{\mathrm{c}}D$) | days |
| surface intensity (${}^{\mathrm{s}}I$ or ${}^{\mathrm{s}}\mathfrak{I}$)$^{\dagger}$ | °C or - | column intensity (${}^{\mathrm{c}}I$ or ${}^{\mathrm{c}}\mathfrak{I}$)$^{\dagger}$ | °C or - |
| surf. severity (${}^{\mathrm{s}}S = {}^{\mathrm{s}}I_{\mathrm{mean}} \times {}^{\mathrm{s}}D$) | °C×days | col. severity (${}^{\mathrm{c}}S = {}^{\mathrm{c}}\mathfrak{I}_{\mathrm{mean}} \times {}^{\mathrm{c}}D$) | days |
| number of interruptions | - | depth below surface/MLD$^{\dagger}$ | m |
| max. interruption duration | days | vertical propagation$^{\ddagger}$ | m day$^{-1}$ |
| max. consec. days at surf. | days | propagation flexure$^{\ddagger\ddagger}$ | m$^2$ day$^{-2}$ |
| | | mean fraction in ML | - |
| | | start delay at surface$^{*}$ | days |
| | | early ending at surface$^{**}$ | days |

**Table 1.** Overview over characteristics for the MHWs defined across the water column as shown in Figure 2b. For metrics marked by a $^{\dagger}$ the mean and maximum across each event is calculated. $^{\ddagger}$Vertical propagation is calculated as linear regression slope through upper MHW boundary. $^{\ddagger\ddagger}$Propagation flexure is calculated as coefficient of the squared term in quadratic fit through upper MHW boundary. $^{*}$Start delay at surface is calculated as surface start day minus column start day. $^{**}$Early ending at surface is calculated as column end day minus surface end day.

For the MHW characteristics at the sea surface, we mostly used the same extreme event characteristics as Hobday et al. (2016): start and end days, duration (${}^{\mathrm{s}}D$), mean and maximum intensity (${}^{\mathrm{s}}I_{\mathrm{mean}}$ & ${}^{\mathrm{s}}I_{\mathrm{max}}$), as well as severity (${}^{\mathrm{s}}S = {}^{\mathrm{s}}I_{\mathrm{mean}} \times {}^{\mathrm{s}}D$). For the calculation of these surface characteristics, we only considered the extreme surface grid boxes within a water column-spanning MHW. For instance, for the surface duration we summed the number of days the MHW affects the sea surface (Fig. 2b). In addition to expressing intensities in absolute temperature exceedances above the threshold ($I$ in °C), we calculated variability scaled MHW intensities ($\mathfrak{I}$, unitless). The variability scaled intensity $\mathfrak{I}$ is diagnosed on the basis of the unitless temperature index $\mathfrak{T}$, which corresponds to the continuous form of the MHW categories introduced by Hobday et al. (2018). We therefore computed $\mathfrak{T}$ following Sen Gupta et al. (2020) as:

$$\mathfrak{T} = \frac{T - T_{\mathrm{clim}}}{T_{\mathrm{thresh}} - T_{\mathrm{clim}}}, \tag{1}$$

where, $T_{\mathrm{clim}}$ is the climatological temperature and $T_{\mathrm{thresh}}$ is the seasonally varying $90^{\mathrm{th}}$ percentile threshold temperature. $\mathfrak{T} > 1$ thus indicates that the instantaneous temperature $T$ is larger than $T_{\mathrm{thresh}}$. Expressing temperatures during MHWs using the variability scaled $\mathfrak{T}$ allows for an unbiased comparison of MHW intensities across different locations, as threshold exceedances are generally proportional to the underlying variance of the temperature time series (Oliver et al., 2021). We calculated $I_{\mathrm{mean/max}}$ and $\mathfrak{I}_{\mathrm{mean/max}}$ as the mean or maximum of absolute threshold exceedances and $\mathfrak{T}$ values throughout each event, respectively. Additional surface characteristics are the number of interruptions at the surface (while being connected in the subsurface), the maximum duration of a surface interruption and the maximum number of consecutive extreme days at the surface.

Column characteristics describe how MHWs extend through the water column and how they move vertically with time (Fig. 2, Tab. 1). We defined the column duration (${}^{c}D$) as the total number of days that a MHW affects the water column, regardless of its vertical position. We calculated the temporal mean and maximum MHW depth relative to the surface ($\delta^{\text{surf.}}_{\text{mean}}$, $\delta^{\text{surf.}}_{\text{max}}$) and relative to the MLD ($\delta^{\text{MLD}}_{\text{mean}}$, $\delta^{\text{MLD}}_{\text{max}}$). We therefore defined the instantaneous MHW depth to be the depth of the lowest extreme grid cell below the surface and below the MLD, respectively. We characterized the vertical propagation of a MHW as the slope of a linear function fitted to the temporal evolution of the depth of the upper MHW boundary (Fig. 2). The focus on the upper MHW boundary thereby relates the vertical MHW propagation directly to the sea surface. We further described temporal changes in the vertical propagation behaviour of the MHW with the coefficient of the second order term in a quadratic function fitted to the MHW's upper boundary (i.e., the flexure, Fig. 2). We further estimated for each day of the MHW, what fraction of the vertical MHW extent is located in the surface ML and calculated the temporal mean fraction over the full column duration. Based on the surface and column characteristics of MHWs, we further derived secondary characteristics such as the MHW's start delay or its early ending at the surface. The mean column intensity (${}^{c}\mathfrak{I}_{\text{mean}}$) was computed by averaging all $\mathfrak{T}$ values from all grid cells of an individual MHW. The maximum intensity (${}^{c}\mathfrak{I}_{\text{max}}$) was taken to be the maximum $\mathfrak{T}$ value at any grid point throughout the MHW, regardless of its depth. The column severity (${}^{c}S$) was calculated as the product of the mean column intensity (${}^{c}\mathfrak{I}_{\text{mean}}$) and the column duration (${}^{c}S = {}^{c}\mathfrak{I}_{\text{mean}} \times {}^{c}D$).

To jointly characterize all MHWs occurring at one horizontal location, we computed "composite" statistics of the respective characteristic across all locally occurring MHWs. These composite statistics are either a composite mean (denoted by an overbar $\overline{X}$), a composite maximum (denoted by a tilde $\widetilde{X}$) or a composite minimum (denoted by a hat $\widehat{X}$).

### 2.4 Evaluation of modeled temperature extremes

Given the absence of observational constraints for evaluating MHWs in the subsurface at the basin scale, we are limited to the comparison of the model simulated surface MHWs with observations. In order to construct an observation-based set of surface MHWs that is directly comparable to the simulated set, we applied the same methodology as outlined in Section 2.2.1-2.2.2 to the daily satellite observations of SST from the AVHRR OISSTv2 dataset (Reynolds et al., 2007). This dataset is provided at 0.25° resolution and covers the period from 1982 to 2019.

In the tropical Pacific, we nonetheless compared simulated temperature time series to observed temperature time series during the extreme warming event of the 1997 to 1998 El Niño (see Supplementary Information). We therefore used temperature data from five moorings from the TAO/TRITON mooring array in the tropical Pacific between 155°W and 95°W (McPhaden et al., 2010). This permits us to extend the evaluation of the model to depth, with El Niño-related variability being a strong constraint given its strong impact on subsurface MHWs.

### 2.5 Clustering of deep-reaching MHWs

In order to group the dMHWs according to their vertical propagation behaviour, we used a k-means clustering algorithm (Pedregosa et al., 2011). As we are particularly interested in the vertical propagation behaviour of dMHWs, making them partially invisible at the ocean surface, we performed the clustering on MHW characteristics that reflect the temporal evolution of the upper MHW boundary (see Table 1 and the Supplementary Information for more detail on the clustering procedure).

These characteristics consist of a subset of the characteristics introduced in Table 1 and Figure 2, as well as their combinations, namely: a) the ratio of surface to column duration, b) the ratio of surface start delay to the column duration, c) the ratio of surface early ending to the column duration, d) the vertical propagation, e) the vertical propagation flexure, f) the ratio between the duration of the longest interruption at the surface and the column duration. Hence, the chosen characteristics focus on the upper boundary of each identified MHW and how it relates to the sea surface. Between all characteristics the rank correlations are generally below (above) 0.6 (−0.6) (Fig. S11). In order to eliminate collinearity between individual clustering features (Dormann et al., 2013), we performed a principal component analysis on the six characteristics, whereby each of them was scaled across all events. We identified only the first three principal components (PCs) to have eigenvalues above 1, which together explain 77.5 % of the variance in MHW dynamics (Fig. S12). The fourth PC still explains 14.4 % of the variance, but its eigenvalue is below 1. We hence continued to only use the standardized first three PCs as clustering features (Kaiser's rule, Kaiser (1960)). To find the optimal number of clusters, we performed the k-means clustering on the first three PCs with varying numbers of clusters, ranging from two to 19. Using the Calinski-Harabasz score (Caliński & Harabasz, 1974), the Davies-Bouldin score (Davies & Bouldin, 1979) and the Elbow method (i.e., the within-cluster sum of squared distances), we identified the optimal number of clusters to be four (Fig. S13). For the final analysis, we thus clustered the first three PCs using four clusters.

### 2.6 Sensitivity analyses

To test the sensitivity of the detected MHWs and their characteristics to our methodological choices, we conducted the following sensitivity analyses: We test for a) the effect of the chosen vertical resolution (37 vs. 19 z-levels in the upper 500 m), b) the effect of horizontal coarsening during downsampling (averaging 3x3 vs. 5x5 grid boxes), c) the effect of the horizontal downsampling methodology (meanpooling vs. maxpooling), d) the effect of the employed MHW temperature threshold percentile ($90^{\text{th}}$ vs. $95^{\text{th}}$), e) the effect of the employed threshold reference period (1982 to 2011 vs. 1990 to 2019), f) the effect of the temporal filtering of the Boolean array $\mathbf{B}$ (5 day kernel vs. no filtering) and g) the effect of the analysis period (1979 to 2019 vs. 1982 to 2019). In each sensitivity case, we solely altered the respective methodological choice, while keeping all other choices as in the reference case, which is used throughout the main manuscript. For a better overview, the different choices are summarized in Figure S18. We find that our results are generally robust with regards to the methodological choices and only differ in expected manners between sensitivity cases (see Supplementary Information).

We further tested the robustness of the MHW clustering with respect to the omission of 10 % to 99 % of detected MHWs and to the omission of individual MHW characteristics feeding into the principal component analysis. We therefore assess the degree of agreement in cluster labeling between the standard case and each sensitivity case, using Cohen's Kappa coefficient ($\kappa$, Cohen (1960), see Section 3.3 of Supplementary Information). $\kappa = 1$ indicates perfect agreement, while $\kappa = 0$ suggests agreement based on random labeling. We find that the clustering consistently produces the four clusters even under the omission of 99 % of the detected MHWs ($\kappa > 0.97$, Fig. S21a). The clustering is also robust to the elimination of individual MHW characteristics (Fig. S21b). In most cases, $\kappa > 0.7$. Only for the omission of the "start delay" and the "early end" at the surface (relative to the MHW column duration, see Tab. 1), $\kappa$ drops to around 0.61 to 0.67, indicating the important role of these characteristics in the clustering procedure. Applying the clustering methodology outlined in Section 2.5 to the detected MHWs in the other MHW detection sensitivity cases (Fig. S18), further produces structurally similar clusters (not shown). Together, these results indicate that the clustering into four different vertical propagation behaviours is robust with respect to our methodological choices and the underlying data.

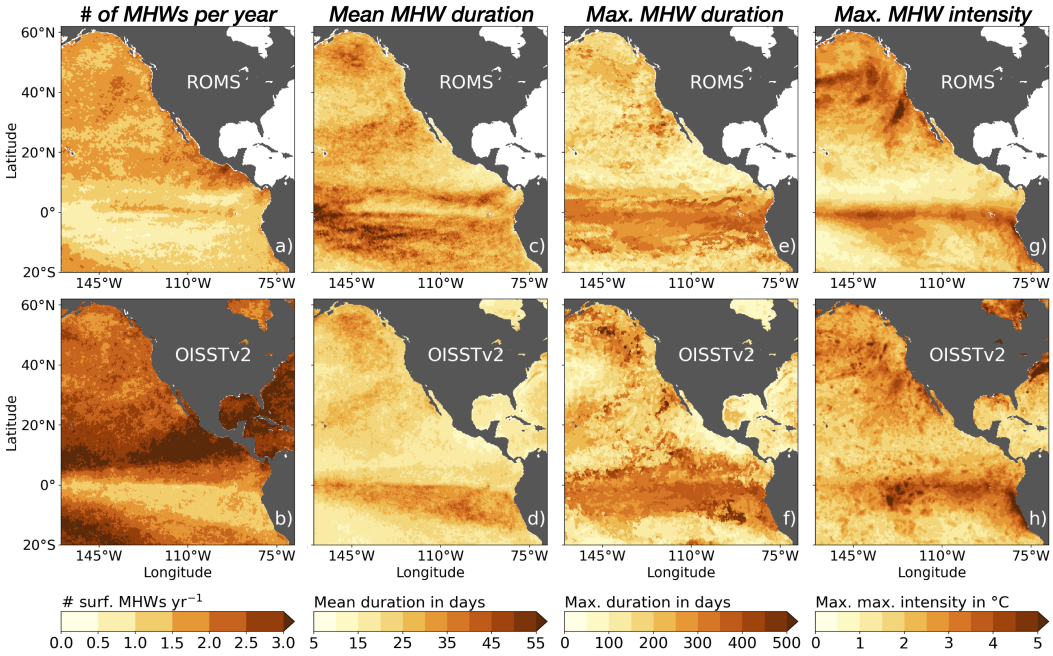



**Figure 4.** Comparison between simulated and observed MHW (surface-only detection, Fig. 2a) occurrences, durations and maximum intensities. The top row shows the results for simulated MHWs, the bottom row for MHWs detected in the OISSTv2 dataset (Reynolds et al., 2007). Panels a) and b) shows the number of MHWs per year. Panels c) and d) show the composite mean MHW duration. Panels e) and f) show the composite maximum MHW duration. Panels g) and h) shows the composite maximum of maximum MHW intensities (expressed in absolute terms, that is in °C).

## 3 Model evaluation

We find that the model captures the spatial structure of the number of surface-only MHW (Fig. 2a) occurrences per year relatively well (Pearson's correlation of $r_P = 0.51$), but underestimates the average number of surface-only MHWs per year (n) with an average value across the EP study area (spatial average denoted by $\langle \dots \rangle$) of $\langle n \rangle_{\text{Mod.}} = 1.35 \pm 0.38$ vs. $\langle n \rangle_{\text{Obs.}} = 2.25 \pm 0.60$ events per year. In contrast, the model has a tendency to overestimate the composite mean surface-only MHW duration $\overline{D}$ in the EP study area ($\langle \overline{D} \rangle_{\text{Mod.}} = 30.3 \pm 7.4$ vs. $\langle \overline{D} \rangle_{\text{Obs.}} = 23.4 \pm 5.5$ days). This overestimation is most prominent in the central tropical Pacific. However, the model reproduces the composite maximum duration well ($\langle \widetilde{D} \rangle_{\text{Mod.}} = 208.5 \pm 78.5$ vs. $\langle \widetilde{D} \rangle_{\text{Obs.}} = 233.8 \pm 89.3$ days) and also reproduces the spatial distribution with the longest events generally occurring in the tropical EP. The modeled composite maximum of the event maximum intensities agree well with observations with $\langle \widetilde{I_{\text{max}}} \rangle_{\text{Mod.}} = 2.16 \pm 0.87$°C vs. $\langle \widetilde{I_{\text{max}}} \rangle_{\text{Obs.}} = 2.48 \pm 0.75$°C. The composite intensity maxima spatially correlate with $r_P = 0.71$ between model and observations and are highest in the tropical EP (locally exceeding 4 °C).

These evaluation results give us confidence that the model reproduces the observed extreme temperature variability at the ocean surface relatively well. The underestimation of the number and overestimation of duration of surface-only MHWs is a common deficiency of ocean models and is commonly attributed to unresolved variability as well as potential interpolation artefacts in the high-resolution observational SST data sets (Frölicher et al., 2018; Pilo et al., 2019; Gruber et al., 2021). An evaluation of extreme

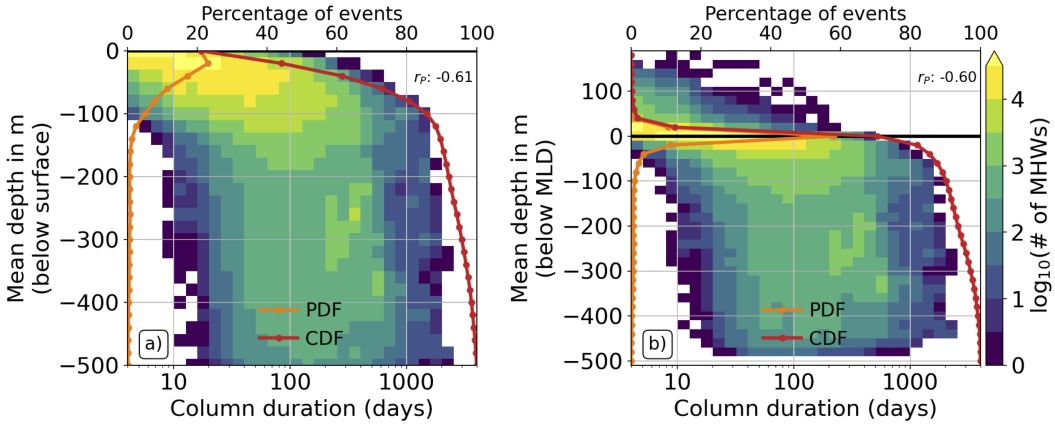



**Figure 5.** 2D histograms of MHW mean depth vs. MHW column duration. The colored squares show the binned number of occurrences of the respective MHWs. Panel a) shows the mean depth below the surface ($\delta^{\text{surf.}}_{\text{mean}}$) on the y-axis, while panel b) shows the mean depth below the MLD ($\delta^{\text{MLD}}_{\text{mean}}$). The solid lines show the distribution of MHWs only with respect to depth (using the same 20 m depth bins as in the 2D histogram). The orange line shows the occurrence distribution as a probability distribution function (PDF) in % of all events. The red line shows the summed occurrence distribution from the surface downwards, that is as a cumulative distribution function (CDF). The black horizontal lines indicate the reference depth in each case (i.e., the sea surface and MLD). The top right of each panel shows the Pearson correlation between the respective MHW characteristics.

warming events in the subsurface is challenging, due to the lack of observational high resolution data at the basin scale. Nevertheless, the simulated subsurface temperature evolution during the 2013 to 2016 *Blob* and the 1997 El Niño match with observations in the respective EP locations (see Sec. 4.3 and Fig. S10). We thus deem the model simulation suitable to extend the view on surface MHWs into the ocean interior. For further evaluation, e.g., of model temperature biases, trends and differences in simulated and observed temperature variability, the reader is referred to the Supplementary Information.

## 4 Results

Over the 41 years of simulation (1979 to 2019) we detect a total of 1 400 170 (one-dimensional) MHWs with at least one surface expression in the study area of the EP. A surface-only detection would find 27 % more events, i.e., 1 773 029, since it would count a resurfacing event as two separate events, while our one-dimensional perspective counts it as one (as depicted in the example sketches in Fig. 2).

### 4.1 Vertical extent of MHWs

The majority of the MHWs are quite shallow, with 50 % reaching a mean depth of 36 m or less ($\delta^{\text{surf.}}_{\text{mean}}$, Fig. 5a). 90 % of MHWs are on average confined to the upper 150 m and less than 1 % reach a depth of 450 m or more below the surface. However, given that extreme conditions at the surface usually tend to extend to the MLD, it is more insightful to analyze the MHW mean depth relative to the MLD ($\delta^{\text{MLD}}_{\text{mean}}$) (Fig. 5b). Indeed, we find that only 13 % of all MHWs are confined to waters shallower than the MLD, with this fraction likely associated with conditions where the active mixing is substantially

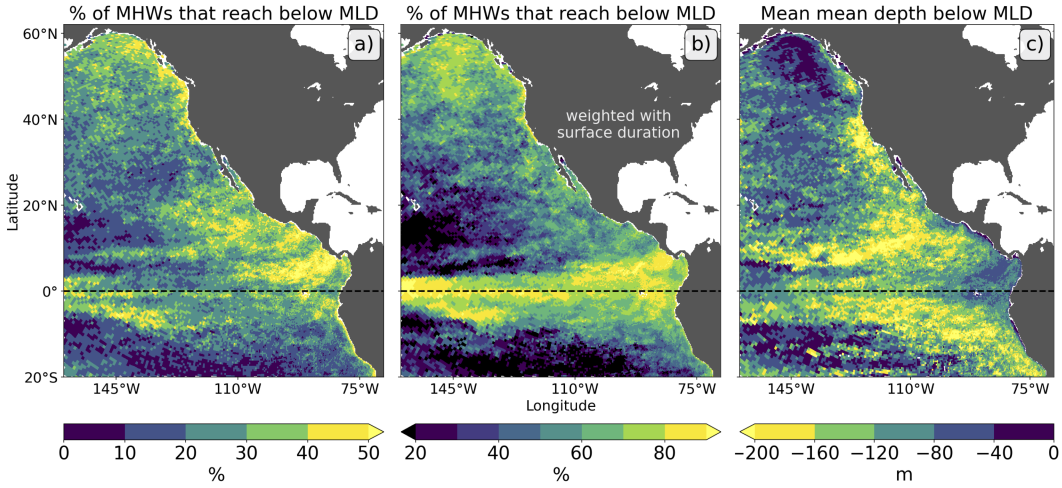



**Figure 6.** Mapped depth characteristics of dMHWs, that is MHWs that reach on average deeper than one (seasonally varying) climatological MLD standard deviation below the instantaneous MLD. a) Map of local fraction of dMHWs; b) same as a) but weighted with the respective MHW surface durations (weighted fraction can be interpreted as likelihood that a randomly chosen extreme surface temperature is part of a dMHW); c) composite mean average depth below the MLD of all dMHWs. In all panels, the dashed black line indicates the Equator.

shallower than the MLD. Including these events, a total of 71.5 % of all MHWs are on average confined to the ML. In this classification as ML-confined MHWs (abbreviated as sMHWs), we pardon short or low-amplitude excursions of the MHW below the MLD, by allowing them to reach on average less than one (seasonally varying) standard deviation of MLD (on average about 10 m) below the varying MLD itself (see Fig. 2b). Thus, as they do not substantially leave the ML, we find that 71.5 % of all MHWs are fundamentally accountable for by MHW analyses based on mixed layer or sea surface temperature data. The remaining 28.5 % of MHWs reach on average one MLD standard deviation or more beneath the MLD. In the following we will classify those MHWs as deep-reaching MHWs (abbreviated as dMHWs). About a third of those dMHWs (11 % of all MHWs) fall into the mean depth range from 10 m to 30 m below the MLD. The remaining two thirds (18 % of all MHWs) extend on average by more than 30 m on average below the MLD. The dMHWs are of particular interest to us, since they are partially decoupled from the surface.

We find a general link between the MHW mean depth and the event column duration ($r_P \approx -0.6$), with longer events tending to reach deeper below the surface and the MLD (Fig. 5). MHWs lasting shorter than ${}^{\mathrm{c}}D = 10$ days are mostly limited to the upper 100 m and ML, while longer lasting events often reach below the ML. For example, 35 % of events that last between ${}^{\mathrm{c}}D = 100$ and 125 days, have a mean depth of more than 110 m below the surface.

The deep-reaching MHWs occur primarily in the eastern tropical Pacific and along a coastal strip extending northward from Southern California up to Alaska, with regions where 40 % of all MHWs are dMHWs (Fig. 6a). In contrast, in the Subtropics, only 10 % to 30 % of the MHWs are deep-reaching. When we take into consideration that dMHWs tend to last longer (see Fig. 5), i.e., when we weight each MHW by its surface duration, the regional pattern of the dMHW becomes even more pronounced (Fig. 6b). With this weighting, in most of the tropical EP more than 80 % of all extremely warm temperatures at the sea surface are associated with dMHWs and in the subpolar North Pacific

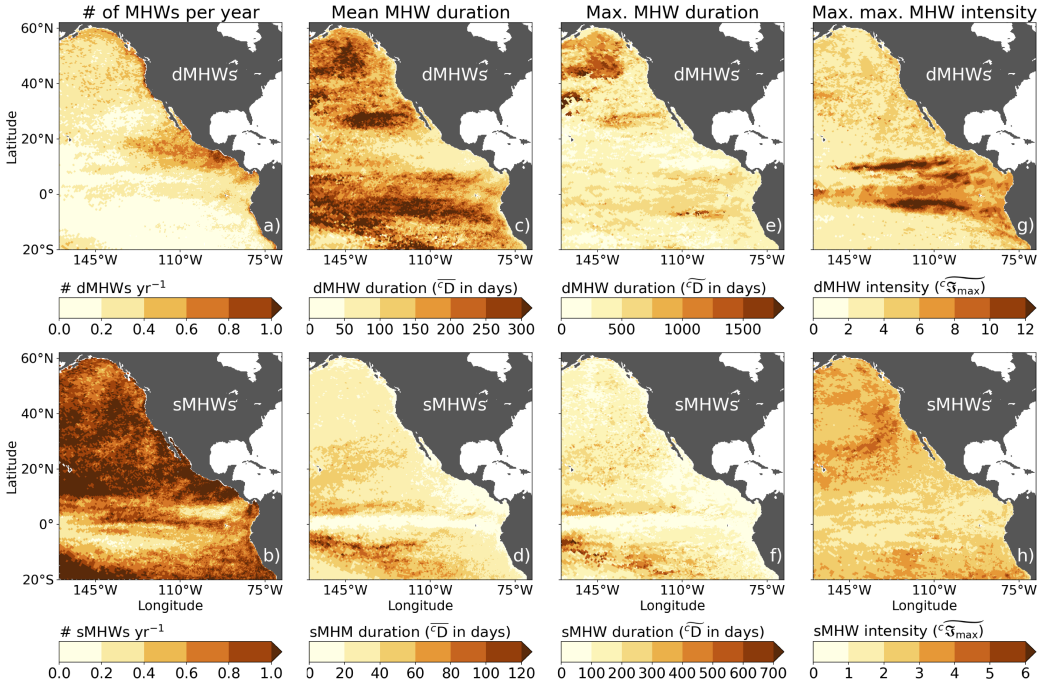



**Figure 7.** Comparison between deep-reaching (dMHW, upper row) and surface MHW (sMHW, lower row) occurrences, durations and intensities. Panels a) and b) shows the number of events per year. Panels c) and d) illustrate the composite mean event duration. Panels e) and f) illustrate the composite maximum event duration. Panels g) and h) show the composite maximum of the maximum intensity recorded for each event (in terms of $\mathfrak{I}$).

this fraction exceeds 70 % (Fig. 6b). This contrasts with the subtropical gyres, where the weighted probabilities are less than 30 %.

Averaged across the EP, we find that dMHWs reach on average 110 m below the MLD ($\langle\overline{\delta^{\mathrm{MLD}}_{\mathrm{mean}}}\rangle$, Fig 6c). However, there are coherent regional differences. The dMHWs tend to reach deepest along the West Coasts of the Americas, especially in the California Current System, and in the low-latitude EP, in particular along ∼10° latitude on both hemispheres. In these regions, the dMHWs extend, on average, substantially deeper than 160 m below the MLD, i.e., $\overline{\delta^{\mathrm{MLD}}_{\mathrm{mean}}} < -160\,\mathrm{m}$ (Fig 6c). However, along the Equator, where we see a high occurrence of dMHWs, these dMHWs do not reach particularly deep on average, but only to around $\overline{\delta^{\mathrm{MLD}}_{\mathrm{mean}}} \approx -120\,\mathrm{m}$. In the subtropical gyres, the few MHWs that qualify as dMHWs, reach generally around 40 m to 80 m below the MLD. The dMHWs also remain relatively shallow in the Alaska gyre, with depth of typically less than 50 m below the MLD.

### 4.2 Core characteristics of deep-reaching MHWs

Averaged across the EP study area, we find $\langle n\rangle = 0.26 \pm 0.15$ dMHWs per year compared to $0.8 \pm 0.27$ sMHWs per year, i.e., deep-reaching MHW occur about three times less often than sMHWs, consistent with the ratio of their total counts (Fig. 7a,b). The highest number of dMHWs per year are found in the eastern tropical North Pacific and along the North American coast (up to 0.8 dMHWs per year, cf. Fig. 6a). While occurring much less often than sMHWs, dMHWs last five times longer, with a mean dMHW column duration of $\langle\overline{^{c}D}\rangle = 184 \pm 82$ days compared to $\langle\overline{^{c}D}\rangle = 36 \pm 17$ days for the

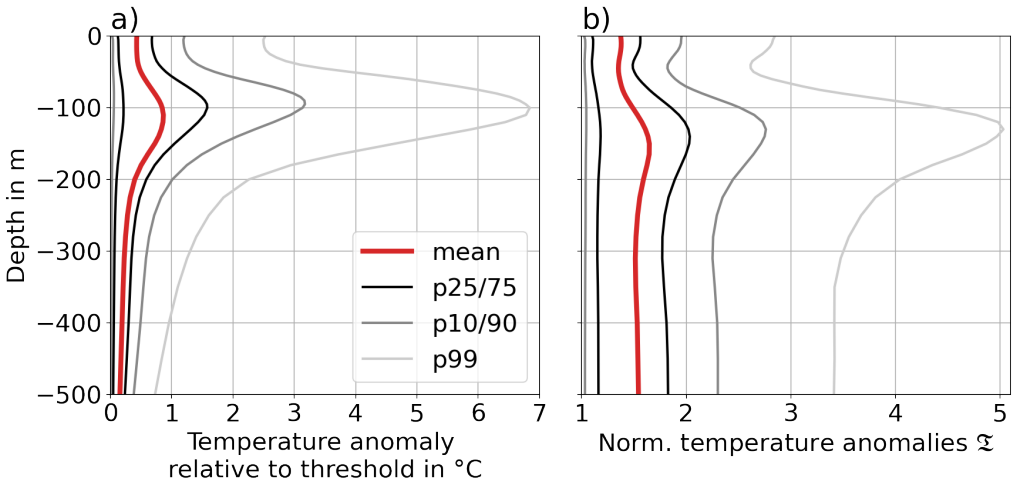



**Figure 8.** Mean profiles of MHW related temperature anomalies across depth. Left panel shows absolute temperature anomalies relative to the extreme threshold (i.e., absolute intensities), while right panel shows variability scaled temperature anomalies (i.e., in terms of the variability scaled temperature $\mathfrak{T}$). Mean profiles are calculated by averaging the (scaled) temperature anomalies across all extreme grid cells associated with MHWs throughout the hindcast. (Red) orange lines show the (grid cell size weighted) mean profile averaged over all extreme grid cells. The black (grey) lines show the $25^{\text{th}}/75^{\text{th}}$ ($10^{\text{th}}/90^{\text{th}}$, $99^{\text{th}}$) percentile levels for each depth.

sMHWs. The dMHWs tend to last especially long in the tropical Pacific and the subpolar North Pacific, where the composite mean column durations ($^{\text{c}}\overline{D}$) often exceed 250 days (Fig. 7c,d). The longest dMHWs are found in the subpolar North Pacific, with composite maximum column durations ($^{\text{c}}\widetilde{D}$) of more than 3 years, which exceed the maximum sMHW duration in the same region by over 2 years (Fig. 7e,f). These extremely long dMHWs are associated with the *Blob* heatwave event in the subpolar NP (see discussions below). The maximum column intensities of dMHWs are substantially higher than of sMHWs ($\langle ^{\text{c}}\overline{I_{\text{max}}} \rangle = 5.0\,^{\circ}\text{C}$ vs. $\langle ^{\text{c}}\overline{I_{\text{max}}} \rangle = 1.9\,^{\circ}\text{C}$). The maximum dMHW intensities scaled with the background variability locally exceed $^{\text{c}}\mathfrak{I}_{\text{max}} \geq 10$ in the tropical EP, with marked bands of extremely high intensities along the Equator and along ~10°N (Fig. 7g). These extremely high intensities occur below the ML, as the same regions do not stand out when analyzing the intensities of sMHWs (Fig. 7h, Fig. 4). In the subtropical and subpolar regions, the maximum dMHW column intensities are only slightly higher than the maximum sMHW intensities (generally less than 2 $\mathfrak{T}$ units difference).

The MHWs tend to intensify in the subsurface. Aggregated across all MHWs (dMHWs and sMHWs) we find the strongest temperature threshold exceedances at around ~100 m depth (Fig. 8a). Below 175 m, temperature anomalies (in absolute terms) drop on average below the surface anomaly amplitudes and decrease steadily down to 500 m. However, expressing the temperature anomalies in terms of the variability scaled $\mathfrak{T}$ (Fig. 8b) shows that MHWs tend to be subsurface intensified throughout most of upper 500 m of the water column, with $\mathfrak{T}$ values of around 1.5, which is slightly higher than the average surface intensities ($\mathfrak{T} \approx 1.4$).

In summary, incorporating the depth dimension in the MHW definition and allowing MHWs to propagate vertically reveals a set of intense, and long lasting MHWs that

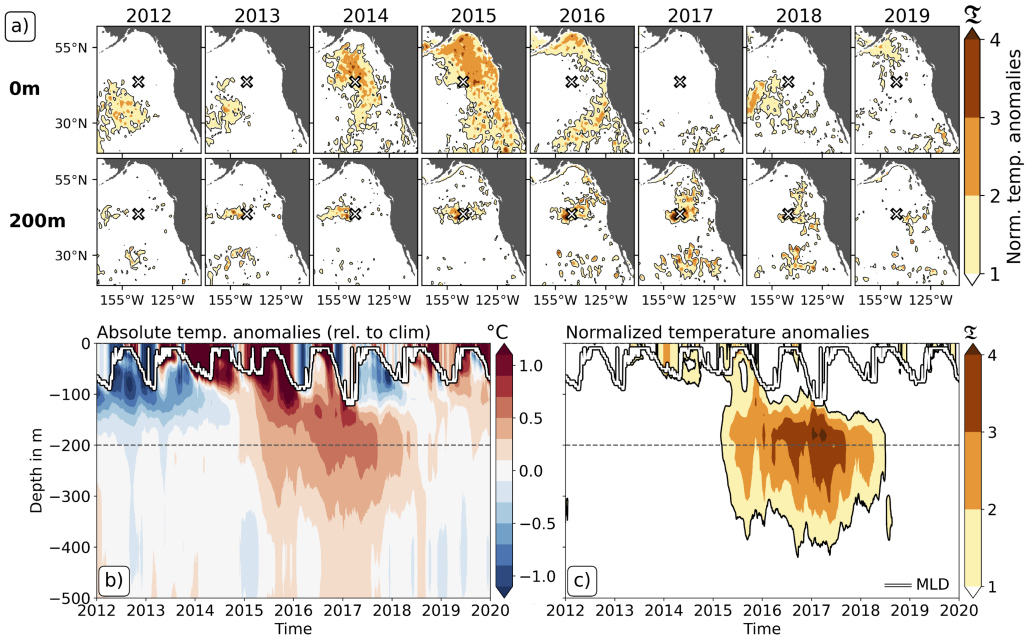



**Figure 9.** The *Blob* in the hindcast simulation. Maps in panel a) show snapshots of the variability scaled temperature $\mathfrak{T}$ taken in March of the years 2012 to 2019 at the surface (upper row) and at 200 m depth (lower row). Panels b) and c) show time-depth sections of the temperature anomaly relative to the climatology (panel b) and the corresponding $\mathfrak{T}$ (panel c) at 145°W, 44°N (indicated by the cross in the maps of panel a). White (dashed) lines in panels b) and c) indicate the MLD (200 m depth).

extend well below the ML. Especially in the subpolar North Pacific and the tropical eastern Pacifc, the difference of MHW durations and intensities between the sMHWs and dMHWs is striking. These results highlight that MHWs can be extremely intense and long-lasting in the subsurface, while being invisible or inconspicuous at the ocean surface.

### 4.3 The *Blob* as prime example of a long-lasting, deep-reaching MHW

Given the rich literature on the long-lasting *Blob* in the Northeast Pacific (Bond et al., 2015; Freeland & Ross, 2019; Gruber et al., 2021), our finding above that this event was also associated with extremely long-lasting subsurface MHWs exceeding 1000 days (Fig. 7e) warrants a more detailed investigation. This permits us also to illustrate our one-dimensional water column concept with a concrete and well studied example. Our approach is aided by the fact that our simulated *Blob* reproduces the observed vertical propagation behaviour rather well (Scannell et al., 2020) (cf. Fig. 9b and their Fig. 4b). This is illustrated by the time series from 145°W and 44°N, that is the grid point closest to the study location of Scannell et al. (2020) (Fig. 9b) and also the grid location where we found one of the longest dMHWs. This MHW starts at the sea surface in early 2015. The associated positive temperature anomaly, which is initially strongest at the sea surface, moves subsequently downward below the ML, where it lingers at around 200 m depth well into 2018, that is for over three years. At 200 m depth, the associated coherent MHW signal appears to slowly move eastward between 2013 and 2019 (lower panels in Fig. 9a). While the subsurface temperature anomalies are generally weaker than those at the sea

surface (∼0.5 °C vs. >1.1 °C), the intensities in terms of the variability scaled temperature $\mathfrak{T}$ show much higher values at around 200 m depth ($\mathfrak{T} > 4$ vs. $\mathfrak{T} \approx 2$), owing to the generally reduced temperature variability in the subsurface. This reduces the denominator in the definition of $\mathfrak{T}$, leading to higher values of $\mathfrak{T}$ (Eq. 1). Hence, the simulation shows a subsurface intensification of the MHW, relative to the local variability.

### 4.4 Linking the vertical MHW structure to surface characteristics

As most observations of MHWs are restricted to the sea surface, the question arises to which degree the water column MHW characteristics can be inferred from their surface characteristics. To study this link, we associate each detected surface-only MHW with its corresponding one-dimensional MHW in the water column. By definition, each water column detected MHW comprises the extreme grid cells of at least one surface-only MHW. In this context, multiple surface-only MHWs may be associated with the same water column MHW.

Individual characteristics of surface-only MHWs such as duration, maximum intensity or severity are only moderately correlated with the same characteristics for one-dimensional MHWs in the water column (Fig. S22). Thus, a direct estimate of individual MHW characteristics based on surface-only MHW properties is challenging. Yet, the comparison between dMHW and sMHW characteristics shows that dMHWs show different characteristics than the ML-confined sMHWs (Fig. 7). Thus, through the use of a logistic regression model, we might be able to identify those characteristics that allow us to state, with a certain probability whether a surface detected MHW is a deep-reaching one. To this end, we first link the surface-only MHW characteristics (duration, maximum intensity and severity) to the vertical characteristics of the associated one-dimensional MHW in a binary form (i.e., 1 if the associated MHW is a dMHW, 0 if the associated MHW is a sMHW) and then fit a logistic regression model to the respective data pairs across all detected surface-only MHWs (Fig. 10).

As expected from sMHWs generally being much shorter than dMHWs (see Section 4.2), we find that short surface-only MHWs are more likely to be associated with sMHWs than with dMHWs (Fig. 10a), while surface-only MHWs that last longer than 49 days are more likely to be associated with dMHWs. If we used this simple (uni-variate) logistic regression model across the entire EP, 68 % of predicted dMHWs are true positives, while 32 % of surface-only MHWs are mistakenly predicted to be dMHWs. If we fit such a model regionally, i.e., just for the California Current System or the equatorial EP, the uni-variate logistic regression model could already predict 72 % (76 %) of all events correctly (Fig. S23).

An even better prediction is possible if severity ($D \times \mathfrak{I}_{\text{mean}}$) was used as single variate predictor (Fig. 10c). In this case MHWs detected at the surface with a severity ($S = \mathfrak{I}_{\text{mean}} \times D$) of 71 days have a more than 50 % probability of being a dMHW. This regression then permits to detect dMHW correctly 70 % of the times. Again, a regionalization of the logistic model further enhances this fraction to 74 % (77 %) for the California Current System (Equatorial Pacific, Fig. S23). In contrast, maximum intensity ($I_{\text{max}}$) as predictor variable yields only true dMHW predictions in 59 % of all cases, with $I_{\text{mean}} = 1.17$ °C as critical distinction value (Fig. 10b). sMHWs, i.e., MHWs that are confined to the surface ML, are correctly predicted with a similar skill as dMHWs. Using any of the three surface-only MHW characteristics, sMHWs are correctly predicted in approximately two thirds of all cases.

These uni-variate logistic regression models suggest that surface-only MHW characteristics have a good potential to detect the presence/absence of a deep-reaching MHW, especially if prior knowledge about the spatial distribution of the deep-reaching MHW is taken into account. Future studies are required to explore this link in more detail and build more complex, multivariate statistical models that could also include other dynam-

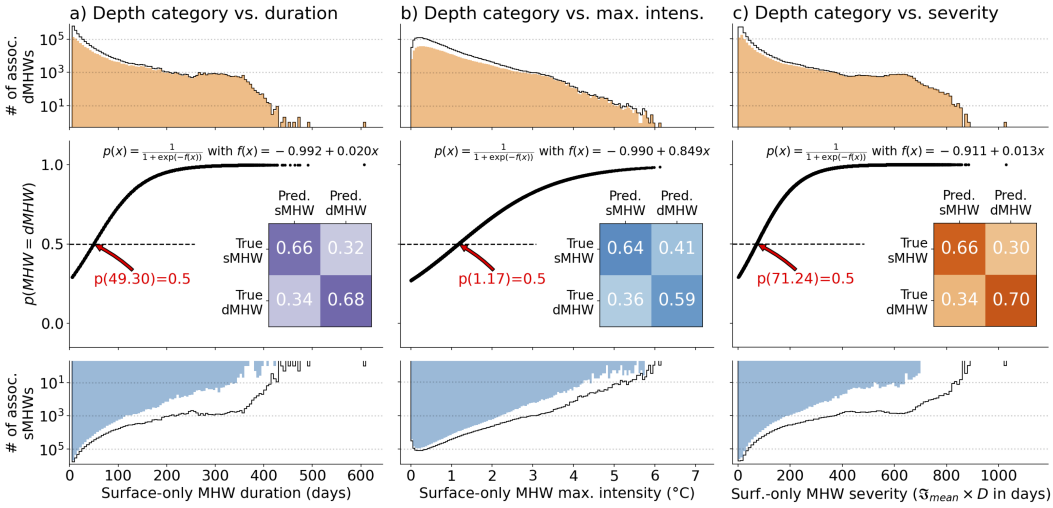
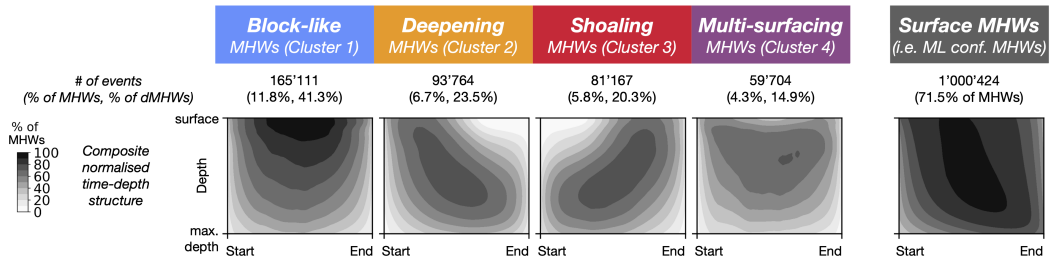



**Figure 10.** Linking surface-only detected MHW properties to the depth extent of the associated water column MHWs. Panels a, b and c show logistic regressions between the surface-only detected MHW characteristics (duration, maximum intensity and severity, respectively) and the associated binary distinction between sMHW and dMHW. Each panel contains a histogram of the shallow (blue) and deep-reaching (orange) MHWs. Thin histogram lines indicate in both cases the histogram of the totality of MHWs. Thick black line shows the logistic regression based on the linear logit function f(x). Red text indicates locations where probabilities of the logistic regression correspond to 0.5. In the middle of each panel, a confusion matrix shows for the statistical model fit which fraction of predicted sMHWs (dMHWs) are true sMHWs (dMHWs) (columns sum to 1).




| Characteristics | p10 | mean | p50 | p90 | p10 | mean | p50 | p90 | p10 | mean | p50 | p90 | p10 | mean | p50 | p90 | p10 | mean | p50 | p90 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Mean depth bel. surface ($\delta^{surf.}_{mean}$, m) | -310 | -125 | -80 | -21 | -377 | -156 | -108 | -32 | -412 | -177 | -117 | -32 | -339 | -169 | -130 | -36 | -76 | -32 | -25 | 0 |
| Mean depth below MLD ($\delta^{MLD}_{mean}$, m) | -267 | -91 | -37 | -8 | -321 | -111 | -54 | -12 | -366 | -138 | -68 | -12 | -297 | -127 | -79 | -13 | -7 | 2 | -0.2 | 10 |
| Column duration ($^{c}D$, days) | 17 | 142 | 76 | 376 | 26 | 116 | 73 | 246 | 23 | 113 | 69 | 244 | 54 | 191 | 152 | 354 | 6 | 28 | 16 | 63 |
| Surface to column duration ratio ($^{s}D/^{c}D$), no unit) | 0.62 | 0.80 | 0.81 | 0.97 | 0.10 | 0.32 | 0.31 | 0.54 | 0.10 | 0.32 | 0.30 | 0.56 | 0.21 | 0.44 | 0.45 | 0.65 | 0.45 | 0.81 | 0.90 | 1 |
| Vertical propagation (m day$^{-1}$) | -0.63 | -0.03 | 0.0 | 0.49 | -2.26 | -1.00 | -0.66 | -0.10 | 0.04 | 1.22 | 0.72 | 3.00 | -0.42 | -0.03 | -0.04 | 0.38 | -1.43 | -0.25 | 0.0 | 0.52 |
| Mean fraction in ML (no unit) | 0.12 | 0.36 | 0.37 | 0.58 | 0.04 | 0.23 | 0.21 | 0.43 | 0.03 | 0.18 | 0.16 | 0.38 | 0.05 | 0.26 | 0.25 | 0.47 | 0.49 | 0.76 | 0.77 | 1 |
| Max intensity ($\mathfrak{I}_{max}$, no unit) | 1.4 | 2.79 | 2.09 | 5.46 | 1.40 | 2.25 | 1.97 | 3.39 | 1.37 | 2.39 | 1.97 | 3.99 | 1.57 | 3.02 | 2.53 | 4.96 | 1.13 | 1.49 | 1.39 | 1.97 |
| Severity ($\mathfrak{I}_{mean} \times D$, days) | 21 | 222 | 102 | 659 | 30 | 161 | 92 | 350 | 26 | 157 | 85 | 359 | 64 | 274 | 204 | 529 | 7 | 34 | 18 | 79 |

**Figure 11.** Vertical propagation clusters for all dMHWs. The image panels show the structure of the four different cluster types using heatmaps calculated based on all MHWs of the respective cluster normalized in depth and time. Below, statistics calculated based on the MHW characteristics for each cluster is shown. These statistics include the mean, $10^{\text{th}}$, $90^{\text{th}}$ and $50^{\text{th}}$ percentile (median). For comparison, the rightmost panel and statistics column show the analoguous results obtained for the sMHWs, which are not part of the four other clusters.

ical variables, such as sea surface height or wind speed. Such models could furthermore explore the surface-based prediction of specific dMHW characteristics, such as their intensity or duration.

### 4.5 Vertical propagation behaviour of deep-reaching MHWs

The example of the North Pacific *Blob* (Sec. 4.3) shows that the extreme signal associated with dMHWs can undergo a vertical propagation throughout the MHW lifetime. To better understand this vertical propagation behaviour and to distinguish between different vertical propagation patterns, we cluster all dMHWs as described in Section 2.5. We identify four clusters of dMHWs (Fig. 11), that can be described by their overall time-depth structure as: a) *block-like* events, b) *deepening* events, c) *shoaling* events, d) *multi-surfacing* events. While the clusters are well separated, the diversity of event characteristics within each cluster is large (see statistics and violin plots in Fig 11, S14, S15, S17).

The block-like cluster, making up 41.3 % of all dMHWs, is characterized by relatively shallow events (cluster averages denoted by $[...]$) of $[\delta^{\text{surf.}}_{\text{mean}}] = -125\,\text{m}$, $[\delta^{\text{MLD}}_{\text{mean}}] = -91\,\text{m}$ but also by events that are relatively long-lasting $[^{c}D] = 142$ days. Throughout their lifetime, these events are quasi-permanently visible at the ocean surface, with the cluster average surface to column duration ratio being relatively high, i.e. $[^{s}D/^{c}D] = 0.80$. In contrast, the deepening events, which make up 23.5 % of dMHWs, affect the surface only during 32 % of their average column duration of $[^{c}D] = 116$ days ($[^{s}D/^{c}D] = 0.32$). These events exhibit on average a downward propagation of $-1.00$ m day$^{-1}$ and reach on average down to $[\delta^{\text{surf.}}_{\text{mean}}] = -156\,\text{m}$ ($[\delta^{\text{MLD}}_{\text{mean}}] = -111\,\text{m}$). Shoaling events (20.3 % of all dMHWs) show an average upward propagation of 1.22m day$^{-1}$ and also affect the

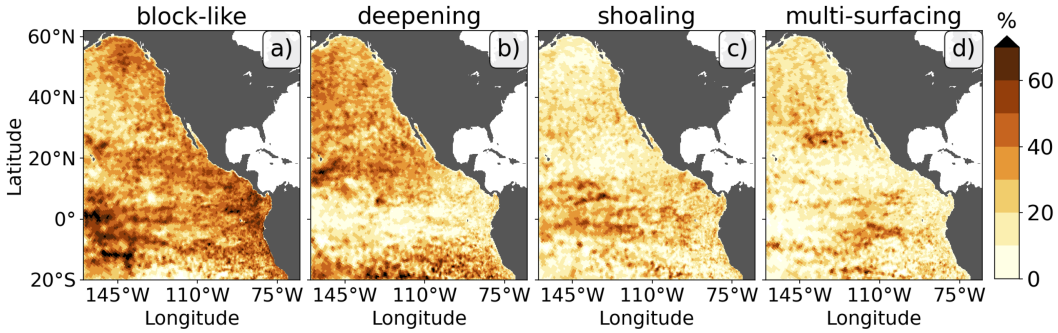



**Figure 12.** Spatial distribution of the contribution of the different MHW clusters to dMHWs, i.e. MHWs that reach on average more than one climatological MLD standard deviation below the instantaneous MLD. The fields are smoothed with a 3×3 Gaussian kernel. The unsmoothed four panels add up to 100 % of all dMHWs.

surface ocean only during 32 % of their $[^{c}D] = 113$ days lifespan. On average, shoaling events reach down to $[\delta^{\text{surf.}}_{\text{mean}}] = -177$ m and substantially below the MLD ($[\delta^{\text{MLD}}_{\text{mean}}] = -138$ m). The multi-surfacing cluster contains 14.9 % of all dMHWs. These events are the longest-lasting ($[^{c}D] = 191$ days) and reach on average 169 m below the surface (127 m below the MLD). They affect the sea surface on average during 44 % of their lifetime. In summary, we find that nearly two thirds of the dMHWs (∼ 59 %) show a behaviour that is not block-like, i.e., that are located for a substantial part of their lifetime purely below the sea surface and are thus not visible in any surface-only based detection algorithm.

The different clusters show distinct spatial distributions (Fig. 12). The block-like dMHWs occur throughout the EP and in particular in the regions of the dMHW hotspots (see Fig. 7a). Deepening MHWs hardly occur in the tropical EP, but make up around a third of dMHWs in the subtropical and subpolar EP (Fig. 12). In contrast, shoaling MHWs are mostly present at low latitudes between 10°N and 10°S. Multi-surfacing MHWs mainly occur in the tropical EP, but also account for more than 20 % of dMHWs in many regions of the subtropical and subpolar North Pacific. These identified regional differences in the occurrence of the different vertical propagation patterns of dMHWs indicate that different drivers are at play in their generation.

## 5 Discussion

In this study, we analyzed the vertical structure and propagation behaviour of MHWs that affect the sea surface, by extending the classical MHW definition (e.g. Hobday et al. (2016)) to incorporate the vertical dimension and by thus furnishing MHWs with a vertical extent. As a proving ground, we applied this methodology to output from a model hindcast simulation in the EP (1979 to 2019). But we consider the findings and insights gained here to be of relevance also in other regions of the world's oceans.

### 5.1 The vertical structure and propagation of MHWs

Our findings show that while the majority of MHWs detected at the ocean surface are limited to the ML (Fig. 5,7a,b), about one third (∼ 29 %) of these reach on average more than one climatological MLD standard deviation below the MLD. Thus, one out of three times the additional buoyancy and increased water column stratification associated with a surface MHW, which hinders the downward mixing of the temperature

anomaly signal below the ML (Oliver et al., 2021), can be overcome. When the MHW signal is below and partially decoupled from the ML, the resulting deep-reaching MHWs also tend to last longer. This matches the findings of Elzahaby and Schaeffer (2019), who also found an increase in MHW durations with increasing penetration depths in the Tasman Sea.

Our results show that MHWs reaching below the ML occur mostly in the eastern tropical (North) Pacific as well as along the American west coast (Fig. 6a,b). So far, no analysis of MHW depths exists across the full EP to compare with our findings shown in Fig. 6. Yet, regional observations and analyses of MHWs appear to support our results. For instance, tropical EP MHWs are often driven by El Niño events (Holbrook et al., 2019). Since El Niño events go along with a strong thermocline warming (e.g., Enfield, 2001), it is reasonable that the here detected MHWs show a propensity to reach below the ML. This is further corroborated by the results of Hu et al. (2021), who find in mooring array data that MHWs in the thermocline of the equatorial Western Pacific often occur in conjunction with surface MHWs. Similary, the increased proportion of dMHWs along the American coastline is perhaps associated with coastally trapped planetary waves that lead to local warming and downwelling conditions (Frischknecht et al., 2015; Wei et al., 2021). Lastly, the here detected deep extent of the North Pacific *Blob* down to $\sim$ 400 m (Fig. 9) agrees with the previously documented vertical extent of strong subsurface warm anomalies (Freeland & Ross, 2019; Scannell et al., 2020).

Next to revealing the MHW depth, our here employed methodology allows us to analyse the subsurface characteristics of MHWs. We find that deeper-extending MHWs tend to last longer and to show higher maximum intensities (Fig. 5, 7, 8, 11). Compared to the ML-confined sMHWs, the dMHWs last substantially longer ($\Delta\langle\overline{^{c}D}\rangle = 148$ days) and show stronger intensities ($\Delta\langle\overline{^{c}I_{\max}}\rangle = 3.1\,°C$, Fig. 7). As an example of a dMHW, we find that the extremely long subsurface *Blob* persistence described by Scannell et al. (2020) was unique in the EP. Nowhere else do we identify such long dMHW durations of over 3 years (Fig. 7e, 9), which are furthermore substantially longer than the maximum surface-only MHW durations of less than one year in the same region (Fig. 4e). In the tropical EP, the difference between the maximum surface-only MHW and dMHW duration is less pronounced, but still around 200 days (comparing Fig. 4f and Fig. 7e). There, maximum dMHW column durations of around 500 days are at the upper bound of time scales associated with El Niño events (Okumura & Deser, 2010), which are known to be the main driver of MHWs in this region (Holbrook et al., 2019). These examples demonstrate that surface-only descriptions of MHWs can dramatically underestimate for how long and with which intensity the upper ocean water column is subject to extreme conditions.

The surface-only perspective on MHWs cannot *per se* distinguish whether the associated extreme signal is confined to the ML or whether it extends also below the ML. The here chosen approach to define MHWs in the water column appears as a useful tool to study this link. Based on simple uni-variate logistic regression models, we find that the characteristics of surface-only MHWs carry some predictive capacity with regards to the vertical extent of the associated (water column) MHW. For instance, a prediction of a dMHW, using surface-only MHW severity as predictor, is correct in 70 % of all cases. These results encourage further studies to explore the link between surface-only MHW characteristics and the vertical structure of the associated MHW. Such links can be used to predict the vertical structure of MHWs that are detected with remote sensing data of sea surface temperature. Additional predictor variables that are available from remote sensing at high spatio-temporal resolution, such as sea surface height, salinity or heat and momentum fluxes might support the development of (multivariate) statistical models (Su et al., 2018) that predict the vertical extent of MHWs when detected at the sea surface. Such statistical models could furthermore be improved by a better dynamical understanding of the MHW drivers. The coherent spatial patterns in the dMHWs oc-

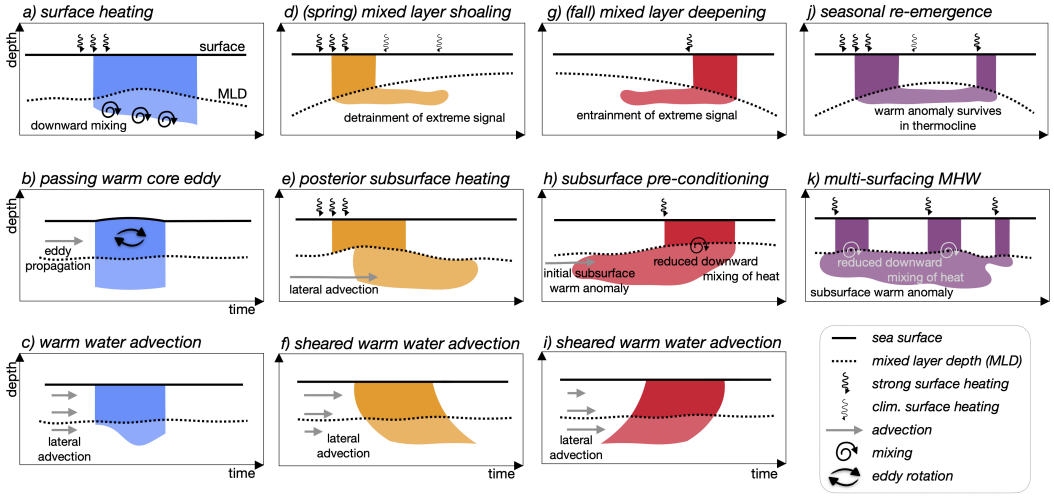



**Figure 13.** Summary of different vertical propagation patterns and their potential driving processes. Colors are associated with clusters as color-coded in Fig. 11, i.e. *block-like*, *deepening*, *shoaling* and *multi-surfacing* from left to right.

currence likelihood, as well as the mean depth of those dMHWs (Sec. 4.1, Fig. 6) suggest that similar driving mechanisms lead to such deep reaching MHWs. For instance, the deep reaching dMHWs along 10°N, which reach on average $\sim 200$ m below the ML (Fig. 6c), and which spatially coincide with the very high intensities of dMHWs (Fig. 7g), are in the alleyway of the westward propagating anticyclonic Tehuantepec eddies (Palacios & Bograd, 2005), which are associated with strong positive temperature anomalies mainly in the thermocline (Purkiani et al., 2022). Hence, the identification of (anticyclonic) mesoscale eddies might further support correct predictions of deep reaching MHWs (Elzahaby & Schaeffer, 2019). Eventually, the established statistical relationships need to be put to test using observational data, i.e., using high-resolution (satellite-derived) sea surface records of MHWs and associated temperature measurements across the water column.

Lastly, the fact that MHWs can affect the water column much longer than diagnosed at the sea surface (see examples given above), highlights the important role of vertical propagation of the MHW signal (Scannell et al., 2020). Using the water column definition of MHWs, we find four clusters with distinct vertical propagation behaviour for the dMHWs (Sec. 4.5). While block-like events remain quasi-permanently at the sea surface, the other three identified dMHW types ($\sim$59 % of dMHWs, $\sim$17 % of all MHWs) affect the sea surface on average only 32 % to 44 % of their lifetime (Fig. 11). Hence, for 17 % of all MHWs, up to two thirds of the MHW duration can be missed when relying solely on surface-only MHW characteristics.

### 5.2 Potential drivers shaping the different vertical MHW structures

While a full driver analysis of the different vertical MHW structures goes beyond the scope of this study, we briefly discuss how the different propagation patterns in the extreme signal could be generated. Figure 13 provides therefore a non-exhaustive overview of potential driving mechanisms.

MHWs driven by atmospheric heat fluxes can lead to dMHWs, if the heat anomaly is transferred below the MLD. This can either occur through downward mixing (Fig. 13a) or through the detrainment process associated with variability in the MLD (Fig. 13d,

Scannell et al. (2020)). In this process, anomalous warm water detrains from a deep (winter) ML and is incorporated into the thermocline during the (seasonal) shoaling of the MLD (Alexander & Deser, 1995). Thus, the MHW signal lingers in the thermocline, while temperatures can return to normal levels in the ML above, forming a deepening MHW. The finding that deepening MHWs occur primarily in spring/early summer (not shown) and only in the subtropical and subpolar EP, where the MLD undergoes a marked seasonal cycle (Fig. 12b), points towards a potential role of this detrainment process. In the opposite sense, a thermocline extreme warm anomaly could be entrained into the ML during the (fall time) MLD deepening, leading to a shoaling MHW (Fig. 13g), an inference also supported by these events occurring predominantly in late summer/fall (not shown). A combination of springtime MHW detrainment out of the ML and fall time MHW re-entrainment into the ML would describe a full re-emergence cycle (and qualify as a multi-surfacing MHW, Fig. 13j), which has long been noted for its role in affecting wintertime ML temperatures of the Northeast Pacific (Namias & Born, 1970; Alexander & Deser, 1995). Hence, in extra-tropical regions the seasonal variability in MLD and the associated detrainment (and (re-)entrainment) processes, are potential drivers behind vertically propagating MHWs.

MHWs driven by lateral advection are not necessarily restricted to the ML, but can span across the water column (Elzahaby & Schaeffer, 2019; Oliver et al., 2021; Großelindemann et al., 2022). Hence, depending on the vertical nature of the lateral cross-gradient temperature advection, the local MHW signal can start simultaneously across the water column (Fig. 13c) or initiate at the surface or at depth (Fig. 13f and i, respectively). Downward isotherm displacement associated with adiabatic heaving during the passage of a warm core eddy can similarly lead to deep reaching MHWs, that likely qualify as block-like MHWs (Fig. 13b). The dynamic (sub-surface) equatorial current system (Kessler, 2006) and the trajectories of warm-core eddies are thus potential hotspots for advection driven MHWs with different vertical propagation signatures.

The presence of subsurface MHW signals might furthermore precondition the formation of a MHW in the ML above, by reducing the downward mixing of any excess heat introduced into the ML. Depending on whether the MHW extends only once or multiple times in the ML, the overall MHW could take on a shoaling pattern (Fig. 13h) or a multi-surfacing pattern (Fig. 13k). A potential region, where this process plays a role is the equatorial Pacific, where we find elevated occurrences of shoaling events (Fig. 12c). In the tropical EP, where MHWs are tightly connected to El Niño events (Holbrook et al., 2019), we discern a shoaling behaviour of the warm anomaly signal which initiates in the subsurface during the 1997 to 1998 El Niño event (Fig. S10). Further supporting the role of initial subsurface warming in driving shoaling events, Vogt et al. (2022) find an important contribution of reduced vertical mixing of heat during the onset phase of MHWs in the tropical EP.

### 5.3 Biological implications

The here presented extended view on MHWs provides new insights into the hazard MHWs represent to marine life. While we find that the majority of MHWs represent an elevated heat stress purely in the ML, 29 % of MHWs can directly impact organisms in the thermocline below. The pronounced and deep reaching MHWs can adversely impact species' fitness due to limited nutrient/food availability or increased metabolic demands (Smith et al., 2023). This effect could be aggravated by lowered oxygen availability, causing even a compound extreme event (Gruber et al., 2021). At the same time, the anomalous deep warming of temperature stratified water columns can open thermal windows for vertically migrating marine species that are otherwise limited in their range by cold subsurface temperatures (Seibel & Birk, 2022), creating potential winners. But to fully evaluate the exposure of marine organisms to such extremes, it is important to also consider the (active and passive) displacement of the organisms through their up-

per ocean habitat (Hofmann Elizondo & Vogt, 2022). For instance, the (diurnal) vertical migration performed by planktonic and nektonic species strongly influences at which times and depths the organisms are exposed to extreme conditions.

Considering the vertical extent of MHWs also has implications for the MHW driven displacement of species. Based on sea surface temperature data only, Jacox et al. (2020) estimated that organisms exposed to MHWs need to horizontally move tens to thousands of km, in order to recover the original climatological conditions. This surface confined view on species displacement neglects the possibility of vertical displacements. In response to MHW conditions at the sea surface, organisms could transiently shift their vertical position to cooler subsurface waters, under the condition that the other habitat-forming biotic and abiotic factors such as oxygen concentrations, food availability, light, etc. are favourable (Reygondeau et al., 2013; Wishner et al., 2013; Seibel & Birk, 2022). In this context, the vertical MHW extent as well as other (biogeochemical) extremes in the water column (Pörtner & Farrell, 2008; Bednaršek et al., 2018; Gruber et al., 2021; Burger et al., 2022; Köhn et al., 2022) can influence the feasibility of this vertical displacement.

### 5.4 Caveats

In this study, we analyze vertical structures and propagation behaviours of MHWs, using output from a hindcast simulation performed with the regional ocean model ROMS. Similar to other models (e.g., Pilo et al. (2019)), our simulation shows biases in the duration and frequency of surface-only MHWs (Sec. 3, Fig. 4), despite the model's general skill to reproduce the observed temperature field (Fig. 1b, Supplementary Information). Beyond the sea surface, our evaluation of simulated temperature extremes is restricted to individual MHWs, such as the *Blob* (Fig. 9) and the 1997 to 1998 El Niño event (Fig. S10), due to scarce high resolution observational data. The realistic reproduction of the subsurface structure of these MHWs gives us confidence in the model's skill in reproducing the subsurface extreme temperature variability.

The purely one-dimensional approach to define MHWs does not capture the full three-dimensional evolution of MHWs over time. As such, we here analyze MHWs independently from one another, even if they occur in horizontally neighbouring locations. Furthermore, next to moving vertically through the water column, MHW signals can move (be advected) horizontally. Depending on the nature of the lateral movement, the lateral displacement of the MHW signals might rectify locally as vertically propagating MHWs (see Fig. 13f,i, Sec. 5.2). Similarly, a lateral movement of the extreme signal in the seasonally developing thermocline might lead to the break-up of a re-emergence cycle into a local deepening MHW and a shoaling MHW downstream. As such, Tak et al. (2021) found that the combination of the re-emergence phenomenon and lateral advection can affect sea surface MHW statistics downstream of the North Pacific subtropical mode water formation site. To additionally account for the lateral coherence and movements of MHWs, a full three-dimensional quasi-Lagrangian tracking of extremes would be required (Desmet et al., 2022).

Lastly, it is important to note, that the propagation of the statistical MHW signal is not necessarily bound to the movement of a particular water mass, but can move and connect with ease across isopycnals. Hence, extreme conditions that occur simultaneously in the ML and below due to different drivers, can appear as one coherent MHW (e.g. as in Fig. 13e). An analysis of heat fluxes during the MHWs, could shed more light on the mechanisms driving the vertical structure and propagation of MHWs.

## 6 Conclusion

As the commonly used surface-only perspective on MHWs has so far not addressed the vertical structure of MHWs, we here extended the classical MHW definition (e.g. Hobday

et al. (2016)) to incorporate the vertical dimension. This new perspective furnishes MHWs with a vertical extent and furthermore allows for the study of the vertical propagation behaviour of MHWs. We explored this new approach to study the vertical structure of MHWs by using daily output from a high-resolution numerical hindcast simulation (1979-2019) in the Eastern Pacific.

We find that one third of the MHWs extend below the ML and are partially (up to two thirds of their lifetime) undetectable at the sea surface. On average, the dMHWs (deep-reaching MHWs) last five times longer than their ML-confined counterparts (sMHWs) and are subsurface intensified, likely due to spatial displacements of sharp temperature gradients within the thermocline. Initial tests further show that the characteristics of MHWs diagnosed at the sea surface carry predictive skill regarding the presence/absence of a deep-reaching MHW. These results suggest that there is potential for the detection of deep-reaching MHWs from remote sensing. Nevertheless, model-derived relationships between MHW surface characteristics and the subsurface MHW structure need to be put to test with observations, by matching (satellite-derived) sea surface temperatures with subsurface hydrographic data, for instance from Argo floats (Su et al., 2018).

The here used approach to define MHWs allowed for an explicit one-dimensional tracking of the MHW signal. We find a variety in different vertical propagation behaviours. While the ML-confined sMHWs and block-like dMHWs generally dominate, we find a substantial fraction of dMHWs that shows a net upward or downward propagation or even a multi-surfacing behaviour. The existence of shoaling and multi-surfacing MHWs suggests, that studying subsurface warm anomalies can help to anticipate the development of surface MHWs, particularly in the tropical EP. However, further analyses will be needed in order to understand under what conditions MHW signals can move between the surface ML and the ocean interior.

**Acknowledgments**

This work has been funded by the Swiss National Science Foundation under grant agreement No. 175787 (Project X-EBUS: "Extreme Ocean Weather Events and their Role for Ocean Biogeochemistry and Ecosystems in Eastern Boundary Upwelling Systems"), by the Swiss Federal Institute of Technology Zurich (ETH Zurich), and by the European Union's Horizon 2020 research and innovation programme under grant agreement No. 820989 (project COMFORT). We thank Fabio Benedetti for helpful inputs regarding the clustering. We thank the Swiss Supercomputing Center (CSCS) and the ETH ITS Scientific IT Services for providing the infrastructure and support to run the model simulations on Euler, the HPC cluster of ETH Zurich. All analyses are performed in python and make use of the numpy, scipy, scikit-learn, pandas, astropy, xarray, sparse and matplotlib packages. The TAO/TRITON data was retrieved on Feb 23, 2022 from the Global Tropical Moored Buoy Array Program (`https://www.pmel.noaa.gov/tao/drupal/disdel/`). All data underlying the plots can be found under the following repository: `https://doi.org/10.3929/ethz-b-000612228` which can be accessed under the additional handle `http://hdl.handle.net/20.500.11850/612228`.

## References

Alexander, M. A., & Deser, C. (1995). A Mechanism for the Recurrence of Wintertime Midlatitude SST Anomalies. *Journal of Physical Oceanography*, *25*(1), 122–137. doi: 10.1175/1520-0485(1995)025⟨0122:AMFTRO⟩2.0.CO;2

Amaya, D. J., Jacox, M. G., Alexander, M. A., Scott, J. D., Deser, C., Capotondi, A., & Phillips, A. S. (2023). Bottom marine heatwaves along the continental shelves of North America. *Nature Communications*, *14*(1), 1038. doi: 10.1038/s41467-023-36567-0

Amaya, D. J., Miller, A. J., Xie, S.-P., & Kosaka, Y. (2020). Physical drivers of

the summer 2019 North Pacific marine heatwave. *Nature Communications*, *11*, 1903. doi: 10.1038/s41467-020-15820-w

Bednaršek, N., Feely, R. A., Beck, M. W., Glippa, O., Kanerva, M., & Engström-Öst, J. (2018). El Niño-Related Thermal Stress Coupled With Upwelling-Related Ocean Acidification Negatively Impacts Cellular to Population-Level Responses in Pteropods Along the California Current System With Implications for Increased Bioenergetic Costs. *Frontiers in Marine Science*, *5*, 486. doi: 10.3389/fmars.2018.00486

Bianchi, D., & Mislan, K. A. S. (2016). Global patterns of diel vertical migration times and velocities from acoustic data. *Limnology and Oceanography*, *61*(1), 353–364. doi: 10.1002/lno.10219

Bond, N. A., Cronin, M. F., Freeland, H., & Mantua, N. (2015). Causes and impacts of the 2014 warm anomaly in the NE Pacific. *Geophysical Research Letters*, *42*(9), 3414–3420. doi: 10.1002/2015GL063306

Boyer, T. P., Garcia, H. E., Locarnini, R. A., Zweng, M. M., Mishonov, A. V., Reagan, J. R., ... Smolyar, I. V. (2018). *World Ocean Atlas 2018.* NOAA National Centers for Environmental Information. Dataset. https://accession.nodc.noaa.gov/NCEI-WOA18.

Burger, F. A., Terhaar, J., & Frölicher, T. L. (2022). Compound marine heatwaves and ocean acidity extremes. *Nature Communications*, *13*(1), 4722. doi: 10.1038/s41467-022-32120-7

Caliński, T., & Harabasz, J. (1974). A dendrite method for cluster analysis. *Communications in Statistics*, *3*(1), 1–27. doi: 10.1080/03610927408827101

Carton, J. A., Chepurin, G. A., & Chen, L. (2018). SODA3: A New Ocean Climate Reanalysis. *Journal of Climate*, *31*(17), 6967–6983. doi: 10.1175/JCLI-D-18-0149.1

Cavole, L. M., Demko, A. M., Diner, R. E., Giddings, A., Koester, I., Pagniello, C. M. L. S., ... Franks, P. J. S. (2016). Biological Impacts of the 2013–2015 Warm-Water Anomaly in the Northeast Pacific: Winners, Losers, and the Future. *Oceanography*, *29*(2), 273–285. doi: 10.5670/oceanog.2016.32

Cheung, W. W. L., Frölicher, T. L., Lam, V. W. Y., Oyinlola, M. A., Reygondeau, G., Sumaila, U. R., ... Wabnitz, C. C. C. (2021). Marine high temperature extremes amplify the impacts of climate change on fish and fisheries. *Science Advances*, *7*(40), eabh0895. doi: 10.1126/sciadv.abh0895

Cohen, J. (1960). A Coefficient of Agreement for Nominal Scales. *Educational and Psychological Measurement*, *20*(1), 37–46. doi: 10.1177/001316446002000104

Collins, M., Sutherland, M., Bouwer, L., Cheong, S.-M., Frölicher, T., Jacot Des Combes, H., ... Tibig, L. (2019). Extremes, Abrupt Changes and Managing Risk. In H.-O. Pörtner et al. (Eds.), *IPCC Special Report on the Ocean and Cryosphere in a Changing Climate* (First ed., pp. 589–655). Cambridge, UK and New York, NY, USA: Cambridge University Press. doi: 10.1017/9781009157964

Davies, D. L., & Bouldin, D. W. (1979). A Cluster Separation Measure. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *PAMI-1*(2), 224–227. doi: 10.1109/TPAMI.1979.4766909

Dee, D. P., Uppala, S. M., Simmons, A. J., Berrisford, P., Poli, P., Kobayashi, S., ... Vitart, F. (2011). The ERA-Interim reanalysis: Configuration and performance of the data assimilation system. *Quarterly Journal of the Royal Meteorological Society*, *137*(656), 553–597. doi: 10.1002/qj.828

Desmet, F., Gruber, N., Köhn, E. E., Münnich, M., & Vogt, M. (2022). Tracking the Space-Time Evolution of Ocean Acidification Extremes in the California Current System and Northeast Pacific. *Journal of Geophysical Research: Oceans*, *127*(5). doi: 10.1029/2021JC018159

Di Biagio, V., Cossarini, G., Salon, S., & Solidoro, C. (2020). Extreme event waves in marine ecosystems: An application to Mediterranean Sea surface chloro-

phyll. *Biogeosciences*, *17*(23), 5967–5988. doi: 10.5194/bg-17-5967-2020

Di Lorenzo, E., & Mantua, N. (2016). Multi-year persistence of the 2014/15 North Pacific marine heatwave. *Nature Climate Change*, *6*, 1042–1047. doi: 10.1038/nclimate3082

Dormann, C. F., Elith, J., Bacher, S., Buchmann, C., Carl, G., Carré, G., ... Lautenbach, S. (2013). Collinearity: A review of methods to deal with it and a simulation study evaluating their performance. *Ecography*, *36*(1), 27–46. doi: 10.1111/j.1600-0587.2012.07348.x

Dussin, R., Barnier, B., Brodeau, L., & Molines, J. M. (2016). *The Making Of the Drakkar forcing set DFS5* (Tech. Rep.). Grenoble, France: LGGE. https://www.drakkar-ocean.eu/publications/reports.

Elzahaby, Y., & Schaeffer, A. (2019). Observational Insight Into the Subsurface Anomalies of Marine Heatwaves. *Frontiers in Marine Science*, *6*, 745. doi: 10.3389/fmars.2019.00745

Enfield, D. B. (2001). Evolution and historical perspective of the 1997–1998 El Niño–Southern Oscillation event. *Bulletin of Marine Science*, *69*(1), 7–25.

Freeland, H., & Ross, T. (2019). 'The Blob' - or, how unusual were ocean temperatures in the Northeast Pacific during 2014-2018? *Deep Sea Research Part I: Oceanographic Research Papers*, *150*, 103061. doi: 10.1016/j.dsr.2019.06.007

Frischknecht, M., Münnich, M., & Gruber, N. (2015). Remote versus local influence of ENSO on the California Current System. *Journal of Geophysical Research: Oceans*, *120*(2), 1353–1374. doi: 10.1002/2014JC010531

Frölicher, T. L., Fischer, E. M., & Gruber, N. (2018). Marine heatwaves under global warming. *Nature*, *560*, 360–364. doi: 10.1038/s41586-018-0383-9

Großelindemann, H., Ryan, S., Ummenhofer, C. C., Martin, T., & Biastoch, A. (2022). Marine Heatwaves and Their Depth Structures on the Northeast U.S. Continental Shelf. *Frontiers in Climate*, *4*, 857937. doi: 10.3389/fclim.2022.857937

Gruber, N., Boyd, P. W., Frölicher, T. L., & Vogt, M. (2021). Biogeochemical extremes and compound events in the ocean. *Nature*, *600*, 395–407. doi: 10.1038/s41586-021-03981-7

Hersbach, H., Bell, B., Berrisford, P., Hirahara, S., Horányi, A., Muñoz-Sabater, J., ... Thépaut, J.-N. (2020). The ERA5 global reanalysis. *Quarterly Journal of the Royal Meteorological Society*, *146*(730), 1999–2049. doi: 10.1002/qj.3803

Hobday, A. J., Alexander, L. V., Perkins, S. E., Smale, D. A., Straub, S. C., Oliver, E. C. J., ... Wernberg, T. (2016). A hierarchical approach to defining marine heatwaves. *Progress in Oceanography*, *141*, 227–238. doi: 10.1016/j.pocean.2015.12.014

Hobday, A. J., Oliver, E. C. J., Sen Gupta, A., Benthuysen, J. A., Burrows, M. T., Donat, M. G., ... Smale, D. A. (2018). Categorizing and Naming Marine Heatwaves. *Oceanography*, *31*(2), 162–173. doi: 10.5670/oceanog.2018.205

Hofmann Elizondo, U., & Vogt, M. (2022). Individual-based modeling of shelled pteropods. *Ecological Modelling*, *468*, 109944. doi: 10.1016/j.ecolmodel.2022.109944

Holbrook, N. J., Scannell, H. A., Sen Gupta, A., Benthuysen, J. A., Feng, M., Oliver, E. C. J., ... Wernberg, T. (2019). A global assessment of marine heatwaves and their drivers. *Nature Communications*, *10*, 2624. doi: 10.1038/s41467-019-10206-z

Holser, R. R., Keates, T. R., Costa, D. P., & Edwards, C. A. (2022). Extent and Magnitude of Subsurface Anomalies During the Northeast Pacific Blob as Measured by Animal-Borne Sensors. *Journal of Geophysical Research: Oceans*, *127*(7), e2021JC018356. doi: 10.1029/2021JC018356

Holte, J., Talley, L. D., Gilson, J., & Roemmich, D. (2017). An Argo mixed layer climatology and database. *Geophysical Research Letters*, *44*(11), 5618–5626. doi: 10.1002/2017GL073426

Hu, S., Li, S., Zhang, Y., Guan, C., Du, Y., Feng, M., ... Hu, D. (2021). Observed strong subsurface marine heatwaves in the tropical western Pacific Ocean. *Environmental Research Letters*, *16*(10), 104024. doi: 10.1088/1748-9326/ac26f2

IPCC SROCC. (2022). *The Ocean and Cryosphere in a Changing Climate: Special Report of the Intergovernmental Panel on Climate Change* (First ed.). Cambridge University Press. doi: 10.1017/9781009157964

Jackson, J. M., Johnson, G. C., Dosser, H. V., & Ross, T. (2018). Warming From Recent Marine Heatwave Lingers in Deep British Columbia Fjord. *Geophysical Research Letters*, *45*(18), 9757–9764. doi: 10.1029/2018GL078971

Jacox, M. G., Alexander, M. A., Bograd, S. J., & Scott, J. D. (2020). Thermal displacement by marine heatwaves. *Nature*, *584*, 82–86. doi: 10.1038/s41586-020-2534-z

Johnson, G. C., Hosoda, S., Jayne, S. R., Oke, P. R., Riser, S. C., Roemmich, D., ... Xu, J. (2022). Argo—Two Decades: Global Oceanography, Revolutionized. *Annual Review of Marine Science*, *14*(1), 379–403. doi: 10.1146/annurev-marine-022521-102008

Kaiser, H. F. (1960). The Application of Electronic Computers to Factor Analysis. *Educational and Psychological Measurement*, *20*(1), 141–151. doi: 10.1177/001316446002000116

Kessler, W. S. (2006). The circulation of the eastern tropical Pacific: A review. *Progress in Oceanography*, *69*(2-4), 181–217. doi: 10.1016/j.pocean.2006.03.009

Köhn, E. E., Münnich, M., Vogt, M., Desmet, F., & Gruber, N. (2022). Strong Habitat Compression by Extreme Shoaling Events of Hypoxic Waters in the Eastern Pacific. *Journal of Geophysical Research: Oceans*, *127*(6), e2022JC018429. doi: 10.1029/2022JC018429

Large, W. G., & Yeager, S. (2004). *Diurnal to decadal global forcing for ocean and sea-ice models: The data sets and flux climatologies* (Tech. Rep.). (No. NCAR/TN-460+STR) University Corporation for Atmospheric Research. doi: 10.5065/D6KK98Q6

Liu, X.-D., Osher, S., & Chan, T. (1994). Weighted Essentially Non-oscillatory Schemes. *Journal of Computational Physics*, *115*(1), 200–212. doi: 10.1006/jcph.1994.1187

Marchesiello, P., McWilliams, J. C., & Shchepetkin, A. (2003). Equilibrium Structure and Dynamics of the California Current System. *Journal of Physical Oceanography*, *33*(4), 753–783. doi: 10.1175/1520-0485(2003)33⟨753:ESADOT⟩2.0.CO;2

McPhaden, M. J., Busalacchi, A. J., & Anderson, D. L. T. (2010). A TOGA Retrospective. *Oceanography*, *23*(3), 86–103. doi: 10.5670/oceanog.2010.26

Namias, J., & Born, R. M. (1970). Temporal coherence in North Pacific sea-surface temperature patterns. *Journal of Geophysical Research*, *75*(30), 5952–5955. doi: 10.1029/JC075i030p05952

Okumura, Y. M., & Deser, C. (2010). Asymmetry in the Duration of El Niño and La Niña. *Journal of Climate*, *23*(21), 5826–5843. doi: 10.1175/2010JCLI3592.1

Oliver, E. C. J., Benthuysen, J. A., Darmaraki, S., Donat, M. G., Hobday, A. J., Holbrook, N. J., ... Sen Gupta, A. (2021). Marine Heatwaves. *Annual Review of Marine Science*, *13*(1), 313–342. doi: 10.1146/annurev-marine-032720-095144

Oliver, E. C. J., Donat, M. G., Burrows, M. T., Moore, P. J., Smale, D. A., Alexander, L. V., ... Wernberg, T. (2018). Longer and more frequent marine heatwaves over the past century. *Nature Communications*, *9*, 1324. doi: 10.1038/s41467-018-03732-9

Palacios, D. M., & Bograd, S. J. (2005). A census of Tehuantepec and Papagayo eddies in the northeastern tropical Pacific. *Geophysical Research Letters*, *32*,

L23606. doi: 10.1029/2005GL024324

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., ... Duchesnay, É. (2011). Scikit-learn: Machine learning in Python. *the Journal of machine Learning research*, *12*, 2825–2830. doi: 10.48550/ARXIV.1201.0490

Pilo, G. S., Holbrook, N. J., Kiss, A. E., & Hogg, A. M. (2019). Sensitivity of Marine Heatwave Metrics to Ocean Model Resolution. *Geophysical Research Letters*, *46*(24), 14604–14612. doi: 10.1029/2019GL084928

Pörtner, H. O., & Farrell, A. P. (2008). Physiology and Climate Change. *Science*, *322*(5902), 690–692. doi: 10.1126/science.1163156

Purkiani, K., Haeckel, M., Haalboom, S., Schmidt, K., Urban, P., Gazis, I.-Z., ... Vink, A. (2022). Impact of a long-lived anticyclonic mesoscale eddy on seawater anomalies in the northeastern tropical Pacific Ocean: A composite analysis from hydrographic measurements, sea level altimetry data, and reanalysis model products. *Ocean Science*, *18*(4), 1163–1181. doi: 10.5194/os-18-1163-2022

Reygondeau, G., Longhurst, A., Martinez, E., Beaugrand, G., Antoine, D., & Maury, O. (2013). Dynamic biogeochemical provinces in the global ocean. *Global Biogeochemical Cycles*, *27*(4), 1046–1058. doi: 10.1002/gbc.20089

Reynolds, R. W., Smith, T. M., Liu, C., Chelton, D. B., Casey, K. S., & Schlax, M. G. (2007). Daily High-Resolution-Blended Analyses for Sea Surface Temperature. *Journal of Climate*, *20*(22), 5473–5496. doi: 10.1175/2007JCLI1824.1

Ryan, S., Ummenhofer, C. C., Gawarkiewicz, G., Wagner, P., Scheinert, M., Biastoch, A., & Böning, C. W. (2021). Depth Structure of Ningaloo Niño/Niña Events and Associated Drivers. *Journal of Climate*, *34*(5), 1767–1788. doi: 10.1175/JCLI-D-19-1020.1

Scannell, H. A., Abernathey, R., Busecke, J., Gagne, D. J., Thompson, L., & Whitt, D. (2021). *Ocetrac.* Zenodo. doi: 10.5281/ZENODO.5102928

Scannell, H. A., Johnson, G. C., Thompson, L., Lyman, J. M., & Riser, S. C. (2020). Subsurface Evolution and Persistence of Marine Heatwaves in the Northeast Pacific. *Geophysical Research Letters*, *47*(23), e2020GL090548. doi: 10.1029/2020GL090548

Schaeffer, A., & Roughan, M. (2017). Subsurface intensification of marine heatwaves off southeastern Australia: The role of stratification and local winds. *Geophysical Research Letters*, *44*(10), 5025–5033. doi: 10.1002/2017GL073714

Seibel, B. A., & Birk, M. A. (2022). Unique thermal sensitivity imposes a cold-water energetic barrier for vertical migrators. *Nature Climate Change*, *12*, 1052–1058. doi: 10.1038/s41558-022-01491-6

Sen Gupta, A., Thomsen, M., Benthuysen, J. A., Hobday, A. J., Oliver, E., Alexander, L. V., ... Smale, D. A. (2020). Drivers and impacts of the most extreme marine heatwave events. *Scientific Reports*, *10*, 19359. doi: 10.1038/s41598-020-75445-3

Shchepetkin, A. F., & McWilliams, J. C. (2005). The regional oceanic modeling system (ROMS): A split-explicit, free-surface, topography-following-coordinate oceanic model. *Ocean Modelling*, *9*(4), 347–404. doi: 10.1016/j.ocemod.2004.08.002

Shu, C.-W. (1998). Essentially non-oscillatory and weighted essentially non-oscillatory schemes for hyperbolic conservation laws. In *Advanced Numerical Approximation of Nonlinear Hyperbolic Equations* (Vol. 1697, pp. 325–432). Berlin, Heidelberg: Springer Berlin Heidelberg. doi: 10.1007/BFb0096355

Smale, D. A., Wernberg, T., Oliver, E. C. J., Thomsen, M., Harvey, B. P., Straub, S. C., ... Moore, P. J. (2019). Marine heatwaves threaten global biodiversity and the provision of ecosystem services. *Nature Climate Change*, *9*, 306–312. doi: 10.1038/s41558-019-0412-1

Smith, K. E., Burrows, M. T., Hobday, A. J., King, N. G., Moore, P. J., Sen Gupta,

A., ... Smale, D. A. (2023). Biological Impacts of Marine Heatwaves. *Annual Review of Marine Science*, *15*(1), 119–145. doi: 10.1146/annurev-marine-032122-121437

Song, Y., & Haidvogel, D. (1994). A Semi-implicit Ocean Circulation Model Using a Generalized Topography-Following Coordinate System. *Journal of Computational Physics*, *115*(1), 228–244. doi: 10.1006/jcph.1994.1189

Steinberg, D. K., & Landry, M. R. (2017). Zooplankton and the Ocean Carbon Cycle. *Annual Review of Marine Science*, *9*(1), 413–444. doi: 10.1146/annurev-marine-010814-015924

Su, H., Huang, L., Li, W., Yang, X., & Yan, X.-H. (2018). Retrieving Ocean Subsurface Temperature Using a Satellite-Based Geographically Weighted Regression Model. *Journal of Geophysical Research: Oceans*, *123*(8), 5180–5193. doi: 10.1029/2018JC014246

Tak, Y.-J., Song, H., & Cho, Y.-K. (2021). Impact of the reemergence of North Pacific subtropical mode water on the multi-year modulation of marine heatwaves in the North Pacific Ocean during winter and early spring. *Environmental Research Letters*, *16*(7), 074036. doi: 10.1088/1748-9326/ac0cad

Vogt, L., Burger, F. A., Griffies, S. M., & Frölicher, T. L. (2022). Local Drivers of Marine Heatwaves: A Global Analysis With an Earth System Model. *Frontiers in Climate*, *4*, 847995. doi: 10.3389/fclim.2022.847995

Wei, X., Li, K.-Y., Kilpatrick, T., Wang, M., & Xie, S.-P. (2021). Large-Scale Conditions for the Record-Setting Southern California Marine Heatwave of August 2018. *Geophysical Research Letters*, *48*(7). doi: 10.1029/2020GL091803

Wernberg, T., Smale, D. A., Tuya, F., Thomsen, M. S., Langlois, T. J., de Bettignies, T., ... Rousseaux, C. S. (2013). An extreme climatic event alters marine ecosystem structure in a global biodiversity hotspot. *Nature Climate Change*, *3*, 78–82. doi: 10.1038/nclimate1627

Wishner, K. F., Outram, D. M., Seibel, B. A., Daly, K. L., & Williams, R. L. (2013). Zooplankton in the eastern tropical north Pacific: Boundary effects of oxygen minimum zone expansion. *Deep Sea Research Part I: Oceanographic Research Papers*, *79*, 122–140. doi: 10.1016/j.dsr.2013.05.012