



Comparing Surface and Subsurface Marine Heatwaves in Ocean Reanalyses

Vincenzo de Toma¹, Andrea Storto¹, Ronan McAdam², Antonietta Capotondi^{3,4}, Chunxue Yang¹

¹National Research Council, Institute of Marine Sciences, CNR-ISMAR, Via Fosso del Cavaliere 100, 00133 Rome, Italy

5 ²CMCC Foundation - Euro-Mediterranean Centre on Climate Change, Bologna, Italy

³Cooperative Institute for Research in Environmental Sciences, University of Colorado, Boulder, CO, USA

⁴Physical Sciences Laboratory, NOAA, Boulder, CO, USA

Correspondence to: Vincenzo de Toma (vincenzo.detoma@cnr.it)

Abstract. Marine heatwaves (MHWs) are increasingly recognised as major drivers of ecological and socio-economic
10 disruption, yet their subsurface expression remains poorly constrained due to sparse in situ observations. Here, we
characterise properties such as the intensity, duration, and heterogeneity of subsurface MHWs using the Copernicus Global
Ocean Reanalysis Ensemble Product (GREP). MHWs are detected across multiple depths (0.5 - 200 m) using a consistent
percentile-based thresholding, applied to three reanalysis members (CGLORSv7, GLORYS2v4, ORAS5) and evaluated
against well-documented historical events. All reanalyses successfully capture the timing and general evolution of major,
15 well-known surface expressions of MHW events when compared against the ESA Climate Change Initiative SST dataset in
given areas of interest, such as, for example, events of 2011 in Western Australia, of 2015 in the Northeast Pacific Blob, and
of 2006 in the Mediterranean Sea. Here, it is also shown that anomalies frequently penetrate below the mixed layer, down to
100–150 m. On a global scale, there is a positive bias of about 15 days in the mean duration for all the events with respect to
satellite SST data, consistent with the limited eddy-resolving capability of the eddy-permitting GREP reanalyses, the
20 additional smoothing introduced by ensemble averaging and the difference between the different representative depths of
SST in satellite observations and models. Global diagnostics reveal that MHWs extending into the subsurface are more
persistent, spatially heterogeneous, and characterised by sharper onset and decline rates than shallow events, particularly in
the tropics, western boundary currents, and the Southern Ocean. Mixed-layer variability strongly modulates the vertical
propagation of anomalies, underscoring the role of vertical mixing in shaping subsurface extremes. Our results highlight both
25 the value and limitations of current reanalyses for subsurface MHW characterisation, emphasising the need for
ensemble-based approaches to quantify uncertainty.

1 Introduction

Global warming manifests itself on a variety of spatial and temporal scales, ranging from large-scale motions and long-term
trends to small and rapid scales. Extreme events, commonly identified using threshold-based algorithms such as those
30 employed for marine heat waves (MHWs), occur regardless of the underlying system dynamics. However, climate change is



expected to change their statistical properties, including their intensities, durations and frequency. These changes are ubiquitous across all the climatic subdomains, including the atmosphere, land, rivers, lakes and oceans, with many cascading effects on the biosphere (Easterling et al., 2000; Capotondi et al., 2024).

Specifically regarding the ocean's domain, marine heatwaves (MHWs) are one type of these extreme events (Hobday et al., 2016), which can have disruptive effects on ocean ecosystems, like corals, kelp forests, and nesting areas of various species of animals, both globally and regionally speaking (Smith et al., 2023, Artana et al., 2026). The cascading socio-economic implications of MHWs on human-related activities such as fishing or aquaculture makes the detection and classification of MHWs and the understanding of their underlying mechanisms (Bond et al., 2015; Chen et al., 2014; Di Lorenzo and Mantua, 2016; Olita et al., 2007; Oliver et al., 2017; Tan et al., 2018; Wang et al., 2022; Smale et al., 2019; Garrabou et al., 2009; Thomson et al., 2015; Hughes et al., 2017; Mills et al., 2013; Cavole et al., 2016; Wernberg et al., 2013) particularly urgent.

Since the definition of MHWs definition was introduced (Hobday et al. 2016), they have been extensively detected and studied, primarily using satellite Sea Surface Temperature (SST) observations which, due to their spatial and temporal coverage/resolution, allowed for more robust statistics than those obtained with in situ measurements.

Investigations of subsurface events, on the other hand, are still in their infancy. Indeed, subsurface ocean temperature measurements from various platforms (e.g., buoys, drifters, stations and cruises) are more sparse, and have often coarser resolution (or even gaps) in space and time (Mishonov et al., 2024).

While a large body of research has focused on surface extreme events (Hobday et al., 2016; Oliver et al., 2018; Oliver et al., 2019; Sen Gupta et al., 2020; Holbrook et al., 2019; Frölicher et al., 2018), there is increasing interest in characterising and classifying MHW based on their vertical extent and structure (Sun et al., 2023, Yuan et al., 2024; Zhang et al., 2024; He et al., 2024; Chen et al., 2024; Köhn et al., 2024; Schaeffer et al., 2017; Scannell et al., 2020; Schaeffer et al., 2023). This interest and ongoing efforts are motivated by the recognition that key quantities, like air-sea exchanges, primary productivity, upper-ocean biogeochemical and turbulence processes, predominantly occur within the mixed layer, which in some regions can extend beyond 200 m below the ocean's surface.

Given the paucity and sparsity of in situ measurements from Argo floats or drifters, ocean reanalyses - which combine numerical models with data assimilation to provide physically consistent, gap-free 3D fields - represent one of the best alternatives for characterising subsurface variability.

Despite their known sensitivities to model physics, forcing, and data assimilation choices, they have been used in several cases as a reference dataset in intercomparison studies. For example, Sun et al., 2023 used reanalysis data in one of the first attempts at a systematic subsurface characterization; Amaya et al., 2023 used the GLORYs ocean reanalysis to characterize bottom MHWs along the North American shelf; Dayan et al., 2023 used reanalyses for characterizing MHWs in the Mediterranean Sea, and McAdam et al., 2023, used a suite of reanalyses as a reference to validate subsurface MHWs



65 seasonal forecasting systems. Given the proven usefulness of reanalysis products for MHW studies, especially in the subsurface, examining the consistency of different products is critical for an informed use of these data. Thus, in this study, we exploit the Copernicus Global Ocean Reanalysis Ensemble Products to conduct an intercomparison of subsurface MHW characteristics across different products. In Section 2, we describe all the data used and the analysis methods; in Section 3, we present the main results for the regional and global analysis; Section 4 follows with discussion and conclusions.

70 **2 Data and Methods**

2.1 ESACCI SST data

We use the European Space Agency Climate Change Initiative SST (ESACCI SST) data set v2.0, which provides global daily satellite-based SST data covering the period from September 1981 to present, distributed in the Climate Data Record (CDR). The data set is designed to provide a long-term, stable, low-bias CDR derived from different infrared sensors, that is, 75 the Advanced Very High Resolution Radiometer (AVHRR), Along Track Scanning Radiometer (ATSR), and Sea and Land Surface Temperature Radiometer (SLSTR) series of sensors (Merchant et al., 2014, 2019). We utilized the Level 4 product v2.0, which consists of spatially complete (optimally interpolated) maps of global daily average (built considering also daytime measurements) SST at 20 centimetres nominal depth, natively available on a $0.05^\circ \times 0.05^\circ$ regular grid (Donlon et al., 2012).

80 **2.2 GREP Ensemble**

Reanalysis data at $1/4^\circ$ degree horizontal resolution are available from the Copernicus Marine Environmental Monitoring Service (CMEMS) as the Global Ocean Reanalysis Ensemble Products (GREP): single products in the ensemble, respectively named CGLORSv7 from the Centro Mediterraneo per i Cambiamenti Climatici (CMCC), GLORYS2v4 from Mercator Ocean International (MOI), ORAS5 from the European Center for Medium Weather Forecast (ECMWF), have 85 been used here to compute also the ensemble statistics (mean and standard deviation). GREP data were used to investigate the subsurface properties of MHWs, assuming a good level of accuracy in their ability to represent what happens below the ocean surface (Storto et al., 2019). Further details about these products, including model versions, parameterisations, assimilation system, and assimilated data, reported in Table 1 are taken from the Product User Manual (Gonou et al., 2020) and the Quality Information Document (Bourdallé-Badie et al., 2024) available on the CMEMS portal. Details of particular 90 interest to this study are the surface forcing, the assimilation window, and the assimilated data. For example, ORAS5 has a shorter assimilation window and assimilates different satellite SST data with respect to the others (in-situ profiles are the same as CGLORSv7). This study is the outcome of the Copernicus Marine Service project on the evaluation of GREP data; therefore, we include only GREP data rather than all possible reanalysis products.



Reanalysis	Common features	Main differences between individual members		
		Model Version	Surface Forcing	SST Assimilation
CGLORSv7 (CMCC)	NEMO, ORCA1/4° 75 vertical levels	NEMO 3.4 LIM2	SST, SSS, SIC surface nudging	OceanVar (3Dvar) Large scale bias correction 7-day assimilation window Reynolds SST (Reynolds et al., 2007), EN4
GLORYS2v4 (MOI)	TKE Altimetry ERA: 1993-2015 ERAinterim and bulk formulae	NEMO 3.1 LIM2	No surface nudging precipitation, flux correction Climatological runoff + ice shelf and iceberg melting	SAM2 (SEEK) Large scale bias correction 7-day assimilation window Reynolds SST (Reynolds et al. 2007), CORA
ORAS5 (ECMWF)	Assimilated Observations: SST,SLA, T/S profiles, SIC Multivariate assimilation, monovariate for the SIC	NEMO 3.4.1 LIM2	Surface waves forcing, SST, SSS surface nudging	NEMOVAR (3Dvar) 5-day assimilation window HadISSTv2 SST prior to 2008, OSTIA afterwards, EN4

Table 1. Synthetic overview of the main inter-member common features and differences

2.3 ARMOR3D

ARMOR3D is a multi-observation oceanographic product that provides three-dimensional fields of temperature, salinity, dynamic height, and geostrophic currents for the global ocean. It is generated by combining satellite observations — altimetry (sea level) and sea surface temperature — with in situ vertical profiles, primarily from Argo floats, supplemented by XBT and CTD data. The methodology follows a "synthetic" approach: surface information is projected down through the water column using statistical regressions calibrated on the historical relationship between surface parameters and in situ profiles, generating synthetic temperature and salinity profiles; these are then merged with real observations through reduced-order optimal interpolation techniques (Guinehut et al., 2012). The product is distributed operationally within the Copernicus Marine Environment Monitoring Service (CMEMS), and we used it to validate spatially averaged anomalies within the Western Australia area of interest (details on its definition below).



2.4 Methodologies used

ESACCI SST and GREP Members were first remapped on a common $1^\circ \times 1^\circ$ regular grid with bilinear interpolation using the `cdo` command `remapbil` (Schulzweida et al., 2019)

The four-dimensional (i.e., time, latitude, longitude, depth) temperature data from GREP members were then used to examine the vertical characteristics of Marine Heat Waves (MHW). The widely used hierarchical approach of Hobday et al. (2016) was applied at selected model levels 1, 9, 15, 19, 25, 28, and 31, corresponding to depths of about 0.5, 10, 30, 50, 100, 150, and 200 meters from the ocean's surface. MHWs were detected independently on each of the selected levels and on ESACCI SST data.

<i>Area</i>	<i>Latitude</i>	<i>Longitude</i>	<i>Sub lat</i>	<i>Sub lon</i>	<i>Time Span</i>
<i>Western Australia</i>	<i>(40°S - 20°S)</i>	<i>(90°E - 120°E)</i>	<i>(32.5°S - 27.5°S)</i>	<i>(110°E - 115°E)</i>	<i>(2010-11-01, 2011-07-01)</i>
<i>Northeast Pacific Blob</i>	<i>(15°N - 70°N)</i>	<i>(180°W - 100°W)</i>	<i>(28°N - 55°N)</i>	<i>(140°W - 120 °W)</i>	<i>(2014-05-01, 2016-08-01)</i>
<i>Mediterranean Sea</i>	<i>(30°N - 50°N)</i>	<i>(0°E - 30°E)</i>	<i>(40°N - 42.5°N)</i>	<i>(6°E - 8.5°E)</i>	<i>(2006-04-01, 2006-11-01)</i>
<i>Tasman Sea</i>	<i>(50°S - 35°S)</i>	<i>(140°E - 160°E)</i>	<i>(46°S - 39°S)</i>	<i>(147.5°E - 157.5°E)</i>	<i>(2015-07-01, 2016-11-01)</i>
<i>Great Barrier Reef</i>	<i>(25°S - 5°S)</i>	<i>(135°E - 155°E)</i>	<i>(15°S - 5°S)</i>	<i>(140°E - 145°E)</i>	<i>(2015-11-15, 2016-07-15)</i>
<i>Santa Barbara</i>	<i>(30°N-37°N)</i>	<i>(125°W-115°W)</i>	<i>(31°N-34°N)</i>	<i>(121°W-118°W)</i>	<i>(2015-06-15, 2016-01-15)</i>
<i>Northwest Atlantic</i>	<i>(35°N-50°N)</i>	<i>(80°W-50°W)</i>	<i>(40°N-45°N)</i>	<i>(75°W-70°W)</i>	<i>(2011-11-01, 2012-08-01)</i>

Table 2. Summary of the selected Areas of Interest (AOI), including the domain and time span of the event. The latter was determined according to Hobday et al. (2018).

To understand the performance of GREP data at the global scale, we have selected several extreme events in different regions (Table 2 - we know that for example the Northeast Pacific Blob has a different origin - see Chen et al., 2023, but we align the event definition here as those defined in Hobday et al. 2018 for the AOI and event definition/naming). For results concerning each Area of Interest (AOI) defined in Table 2, we have adopted the Hobday et al. (2016) algorithm to produce the matrices of the MHW intensity and index of the category with dimensions (time, latitude, longitude) for each of the selected depths. Even if this required a verification on known events on the surface (against ESACCI SST data), having this level of detail allowed us to explore the temporal evolution of anomalies at depth, showing that, depending on the region considered, MHWs can extend beyond the MLD. For each GREP member and each of the selected depths, MHW detection was carried out on both the original (i.e., not-detrended) and linearly detrended data. The baseline period chosen for the climatology and associated 90th percentile threshold spans the whole data coverage (i.e., from 1993 to 2019, corresponding to the reanalysis time span). Histograms for threshold selection were built after applying a binomial smoothing with a [-



15,+15] days running window, and the climatology itself was built considering a window of [-5, +5] days centred around the climatological day. We classified the intensity of events following the standard Hobday et al. (2016) classification.

A further vertical categorisation between events shallower than 50 meters / deeper than 50 meters has been carried out, with this depth threshold motivated by the mean variability of the MLD in most of the AOI, and also considering the MLD seasonality over the analysis period (see Figure S1). We chose not to adopt a more complex categorisation (e.g., those proposed by Sun et al. (2023) or Malan et al. (2025), which consider the vertical structure of temperature throughout the water column) for two reasons. First, there is currently no consensus on a standard definition of subsurface MHW categories. We therefore opted for the simpler distinction described above, which is transparent, reproducible, and still provides useful insight into differences in MHW representation across the GREP reanalyses. Second, the upper 50 m of the ocean is both ecologically and observationally important, as it encompasses much of the sunlit surface layer where many marine organisms are concentrated and where the reanalyses are generally better constrained by the higher density of available observations than at greater depths.

This distinction also facilitates the interpretation of the results by separating differences that primarily affect the well-observed, ecologically relevant near-surface layer from those emerging in the deeper ocean, where observational constraints are weaker and reanalysis uncertainty (quantified via the ensemble spread) is generally larger.

All the above-mentioned codes, together with the ones used to produce results and figures within this manuscript, are available in a public GitHub repository (<https://github.com/vincenzodetoma/ROHAN>).

3 Results

As the first verification step, MHW detections on the first model level were used to assess the ability of each product to represent events that are known from the literature (a different baseline climatology period has been used here for these events compared to the one used in Hobday et al. 2018 because of the availability of the data). Details on the selected events are outlined in Table 2.

3.1 Surface Properties – Global

First, global MHW properties of the ensemble mean are validated at the surface by comparison with the ESA CCI SST. Here, the mean duration across all events detected in ESA CCI SST can be as long as 60 days in the tropical Pacific, and exceeds two weeks over 90% of the global domain (Figure 1a). The differences between the ensemble mean and ESA CCI SST, which we shall call for brevity here bias, show that the ensemble mean of GREP significantly overestimates (non-significant bias points, marked as grey shaded regions, where the bias exceeds two times the ensemble standard deviation appear as a small-scale, spike-like distribution, especially in the belt around 20S on the map in Figures 1b, 1e) mean duration ocean-wide, reaching peaks of 15 days in the Southern Ocean. The ensemble standard deviation reaches up to about 30 days in the Tropical Pacific area, but it is much lower in the rest of the ocean (less than 10 days - see Figure 1c). The maximum



intensity across all the events detected in ESA CCI SST shows peaks exceeding three degrees (with values as large as 6.5 °C in most of the areas where the colorbar saturates; Figure 1d) in most of the ocean, especially along western boundary currents and in eddy-rich, turbulent areas such as the Agulhas Leakage and south of the Brazilian Current. In these areas, the GREP ensemble mean significantly overestimates the peak values by up to 1.5°C (see Figure 1e, where the same logic as the ensemble spread of Figure 1b has been applied). Concurrently, these are also areas where the ensemble standard deviation of the maximum intensity across all the events exceeds 1°C, being lower than 0.30° in most of the remaining ocean regions worldwide. Overall, there is a good level of consistency between GREP members and the satellite data, likely because all GREP reanalyses assimilate satellite-derived SST data. At the extremes of the ensemble, ORAS5 overestimated the mean duration (by about +20 days in the Southern Hemisphere, at the crossing of the Southern Indian and Atlantic Ocean and the Antarctic Ocean) and the maximum intensity (by about +1.5°C in the Agulhas Leakage region) with respect to ESA CCI data (not shown). This systematic overestimation of MHW duration by the GREP ensemble mean is consistent with known limitations of eddy-permitting (0.25°) reanalyses in representing the mesoscale-eddy-driven heat-flux convergence that terminates surface warm anomalies (Bian et al., 2023). Comparisons of MHW metrics across ocean-model resolutions have shown that coarser, eddy-permitting configurations systematically detect fewer, weaker and longer-lasting events than eddy-resolving simulations or high-resolution observations, with the bias shrinking as resolution increases (Pilo et al., 2019). The resolution issue is also paired with the fact that surface MHW are detected on the first model level, which is about 50cm thick, and then compared against MHW detected in satellite observations, here representative of a 20cm depth SST. The same pattern has been reported for GREP specifically, where the ensemble mean underestimates MHW frequency and intensity while overestimating duration relative to ESA CCI SST (McAdam et al., 2023), and for individual eddy-permitting GREP members benchmarked against an eddy-resolving product and satellite altimetry (Paul et al., 2025). This resolution-driven damping of sub-monthly SST variability is likely compounded here by ensemble averaging across the four GREP systems, whose individual eddy and frontal realizations are not synchronized in time or space — an effect shown to bias MHW metrics in other multi-product ensemble means (Chevallard et al., 2026). This is consistent with the duration bias peaking in regions of high mesoscale activity (the Southern Ocean, the Agulhas Leakage, the Brazil Current), which coincide with the regions of highest ensemble spread (Fig. 1c).

3.2 Surface Properties – Regional

Figures 2 and 3 (together with S2-S13, which we report in the supplementary material) show SST anomalies and categories reached during peak phases of the MHW events listed in Table 2, compared with the ESA CCI SST. The Western Australia 2011 event is successfully classified as an extreme-category event in all three reanalysis datasets, with a consistent spatial pattern for the peak anomaly and an ensemble standard deviation that is quite low relative to the variability of the spatially averaged signal. Indeed, looking at the time series of Figure 3, the maximum category level remains the same even when adding/subtracting the ensemble standard deviation from its mean.



Region\model (detrended)	cglo	glor	oras	ensemble mean
Western Australia	0.94 (0.94)	0.92 (0.90)	0.94 (0.93)	<i>0.94 (0.94)</i>
Northeast Pacific Blob	0.95 (0.94)	0.93 (0.92)	0.96 (0.95)	<i>0.96 (0.96)</i>
Mediterranean Sea	0.97 (0.96)	0.92 (0.92)	0.98 (0.97)	<i>0.98 (0.98)</i>
Tasman Sea	0.85 (0.81)	0.83 (0.81)	0.83 (0.79)	<i>0.89 (0.86)</i>
Great Barrier Reef	0.92 (0.90)	0.87 (0.82)	0.89 (0.83)	<i>0.92 (0.89)</i>
Santa Barbara	0.94 (0.85)	0.86 (0.70)	<i>0.96 (0.92)</i>	0.95 (0.88)
Northwest Atlantic	0.74 (0.69)	0.72 (0.65)	0.74 (0.69)	<i>0.79 (0.78)</i>

Table 3. Spatial correlations on peak anomaly maps for each region, non-detrended data in black and detrended data in parentheses (marked in bold regions with higher correlation for a given model, and in italic characters the best performing product per region).

Region\model (detrended)	cglo	glor	oras	ensemble mean
Western Australia	0.97 (0.97)	0.95 (0.95)	<i>0.98 (0.98)</i>	0.97 (0.97)
Northeast Pacific Blob	0.96 (0.96)	0.96 (0.96)	<i>0.97 (0.97)</i>	<i>0.97 (0.97)</i>
Mediterranean Sea	<i>0.97 (0.97)</i>	0.95 (0.95)	0.97 (0.97)	0.97 (0.97)
Tasman Sea	0.90 (0.90)	0.90 (0.89)	0.90 (0.90)	<i>0.93 (0.93)</i>
Great Barrier Reef	0.94 (0.93)	0.922 (0.91)	<i>0.97 (0.96)</i>	0.97 (0.95)
Santa Barbara	0.94 (0.92)	0.91 (0.89)	<i>0.95 (0.94)</i>	0.94 (0.92)
Northwest Atlantic	0.85 (0.83)	0.90 (0.87)	0.84 (0.82)	<i>0.90 (0.89)</i>

Table 4. Temporal correlations between time series of spatially averaged anomalies for each region, non-detrended data in black and detrended data in parentheses (marked in bold regions with higher correlation for a given model, and in italic characters the best performing product per region).

The most extreme event among AOIs listed in Table 2 (Western Australia 2011) is captured well and is better resolved than others (together with the Mediterranean Sea and Northeast Pacific Blob events is among the ones that show the highest



spatial and temporal correlation with ESA CCI SST data, see Table 3 and Table 4 - the ensemble mean outperform all the members in having the highest spatial correlations for the peak of the events in every region exception made for Santa Barbara). Figures 4-5 focus on the Northeast Pacific Blob, showing the large spatial coherence and persistence of positive SST anomalies, together with areas classified as strong and extreme MHW categories. The temporal evolution highlights a progressive intensification and broadening of the anomaly field, while differences among reanalysis products remain comparatively small, particularly during the mature phase of the event. Figures 6-7 describe the Mediterranean Sea event, characterised by a more confined spatial structure and stronger regional heterogeneity, with substantial variability in the timing and peak intensity among products, reflecting the influence of semi-enclosed basin dynamics and mesoscale variability. In supplementary material (figures S2-S9), we show the rest of the AOIs, which have smaller correlation with ESA CCI SST data, but still show a good consistency in representing events across models.

The set of surface regional figures (2-7, S2-S9) also illustrates the effect of detrending the SST time series before MHW detection. Detrending generally reduces the magnitude, duration, and spatial extent of the detected events, particularly in regions strongly affected by long-term ocean warming, such as the Mediterranean Sea (Capotondi et al., 2024, Pisano et al., 2026). However, the main spatial structures, timing of peak anomalies, and appearance of positive anomalies at depth (not shown) remain largely preserved after detrending, indicating that the identified events are not solely artefacts of long-term warming trends but are also associated with substantial interannual-to-decadal variability. For this reason, and to maintain consistency with definitions of marine heatwaves that may be more appropriate for assessing MHW impacts on marine ecosystems under actual environmental conditions, the main analyses presented in the paper were conducted using non-detrended input data. Retaining the long-term warming signal allows the cumulative effect of climate change and natural variability to be considered together, thereby providing a more realistic assessment of present-day MHW intensity and persistence, which may have a larger ecological relevance.

3.3 Subsurface Properties – Global

In this section, we examine global characterisation of subsurface temperature extremes, across both shallower than 50 meters and deeper than 50 meters ocean layers. Specifically, we consider mean duration, onset and decline rates, maximum category, cumulative intensity and variance for the shallower and deeper groups. These quantities can be expected to be very relevant for characterizing the impact of MHW events, thus supporting decision-making. These characteristics can also provide insights on the physical mechanisms involved in MHW development (e.g., Zhao et al. 2026).

Figure 8 reveals that MHW events are longer-lasting the near-surface waters of the tropical Pacific, with some regions experiencing durations of up to 28 days. These longer durations are likely associated with the presence of the El Niño Southern Oscillation variability. Below 50 m, durations increase to values up to 35-38 days, except in the tropical regions where they remain similar to the surface values and in the Southern Ocean poleward of about 60S. The ensemble standard deviation range is confined well below those maximum values, reaching up to about 4.86 days for shallow events and 8.75



days for deep ones. Figure 9 shows that deep events exhibit a faster emergence of thermal anomalies in equatorial and western boundary current regions, where advective processes dominate (Zhao et al. 2026), while surface events have, in general, a lower mean rate of onset in those regions, as air-sea interactions may also play a role. The ensemble spread variations again exceed typical values for the onset rate (0.56 for events shallower, 0.81 for the ones deeper than 50 meters). Across all these figures, the three reanalyses (CGLO, GLOR, ORAS) show broadly consistent spatial patterns, lending robustness to the patterns in each product. However, notable inter-model discrepancies emerge, especially in high-latitude regions and in areas of complex oceanographic dynamics such as the Southern Ocean and the North Atlantic. For instance, the GLOR experiment tends to produce higher onset and decline rates of both shallow and deep events compared to CGLO and ORAS, which may be attributed to differences in horizontal and vertical resolutions, parameterisation schemes, and even data assimilation techniques.

The rate of thermal decline (Figure 10), indicates that surface waters tend to cool more rapidly than deep ones especially within the mid-latitudinal [30, 60] and [-30,-60] bands—likely due to stronger turbulent air-sea fluxes and vertical mixing. Deeper layers, on the other hand, exhibit slower decay rates (except for the intertropical convergence zone area and along major western boundary currents), reflecting greater thermal inertia due to reduced exposure to atmospheric forcing, thus partially explaining the longer persistence generally seen for these events in Fig. 8. Deep events have a relatively higher decline rate than shallow ones in the equatorial band and along the major western boundary currents, where they are affected by deeper reaching dynamical processes. In these regions the ensemble standard deviation ranges are respectively a bit higher than typical decline rate values: 0.39 in the shallow events, more than doubled in deep ones (0.80).

The maximum categories across all the events highlight that extreme-category events cover a larger extent in deeper than 50m waters than in shallower than 50m ones, almost everywhere, but especially in the Tropical Pacific and, to a lesser degree, in the other tropical basins. The ensemble spread has a maximum range of 1.41 for both shallow and deep events, indicating that inter-model discrepancies across categories rarely go beyond a difference of one category between any two members.

Figure 12 presents global distributions of the mean cumulative intensity (see Appendix II: Metrics definitions) - a metric which is important for the quantification of the heat stress on ecosystems. In shallow waters (<50 m), higher intensities are concentrated in equatorial and coastal regions, with notable variation across models. Deeper layers (>50 m), on the other hand, reveal broader areas of relatively high intensity with generally reduced spatial variability. Regional differences of cumulative intensity between the surface and the deeper ocean layers can be expected to reflect differences of mean duration, onset, and decline rate discussed earlier.

The mean intensity variance, or mean heterogeneity (Figure 13), shows a pattern which is very similar to that of the onset-decline rate: in particular, deep events are more heterogeneous than surface ones, especially in the equatorial region and along major western boundary currents. The ensemble standard deviation range for this quantity is below typical values for the heterogeneity (shallow 0.33, deep 0.41), indicating consistency among reanalysis products.



In general, we observe that when the ensemble spread range for a given variable is higher than typical values for that variable in shallow events, the same happens also for the deep events, suggesting that variable inter-model discrepancies may be the predominant source of differences between shallow and deep events.

270 Regardless of the inter-model discrepancies and focusing on the shallow/deep categorisation, the global landscape offered by the above analyses leads to the conclusion that deep events tend to appear more suddenly and decay more slowly, therefore being generally more persistent and have a higher level of heterogeneity than shallow events, especially in the tropical oceans, along major Western Boundary Currents, in the Agulhas, along the Antarctic Circumpolar Current (ACC) and in most of the major hotspots of MHW activity (see historical MHWs in Capotondi et al., 2024).

275 **3.4 Subsurface Properties – Regional**

We focus on Western Australia for comparing MHW subsurface characteristics across the GREP ensemble. We know that MHWs in this region are significantly affected by advective processes, as well as local air-sea interactions, providing an opportunity to assess possible differences in how different reanalyses represent these processes. The spatially averaged anomalies highlight the consistency of GREP members in reproducing the mean state over this region, particularly in the uppermost 100m, with anomalies of the same magnitude as ARMOR3D data (see Supplementary Figure S21). This gives us confidence in the ability to represent MHW in this region, and Figure 14 illustrates the maximum MHW anomaly reached over the selected region in Western Australia between November 2010 and July 2011 as a function of depth, based on each GREP Member. All models consistently indicate the occurrence of intense and persistent MHWs during early to mid-2011, coinciding with the peak of the well-documented 2011 Western Australian marine heatwave. The strongest anomalies are concentrated in the upper 50 m of the water column, although more moderate anomalies are also seen at depths below ~80 m. Looking at the ensemble mean panel, we can notice that for nearly the whole event duration, the highest ensemble spread values are concentrated below the mixed layer depth, indicating some spread in the appearance of anomalies at depth across products (i.e., products do not completely agree on the depth at which an anomaly can take place). The ensemble spread values (gray contours in the bottom-right panel of Figure 14) remain, however, below 1°C on average, indicating a relatively good agreement among the three reanalyses.

Differences include higher intensities in CGLORS and GLORYS relative to ORAS, especially below the mixed layer, with widespread severe to extreme anomalies observed during March–May 2011. The evolution of the MLD (blue line) provides further insight into the vertical distribution of MHWs. The peak anomalies are predominantly confined within, or slightly below, the mixed layer, at a time when the ML is relatively shallow (during austral summer) while the observed deepening of the MLD during mid-2011 appears to coincide with the presence of subsurface anomalies down to 200m. Notice, however, that the largest subsurface anomalies in GLOR occur concurrently with the surface event, suggesting the presence of independent processes operating at depth, e.g., advection or vertical thermocline displacements related to eddy activity, processes that may not be properly captured by reanalyses run at coarser resolution.



Similar time–depth sections of maximum MHW anomalies are also examined. for the Northeast Pacific Blob (Figure 15) and
300 Northwest Atlantic event (Figure S13). Mixed layer depth is also shown in both figures. Both figures display deep anomaly
penetration and long persistence below the mixed layer, indicating enhanced upper-ocean heat storage. In contrast, the
Mediterranean Sea (Figure 16) and the Santa Barbara region (Figure S12) events show shallower vertical structures and
more intermittent subsurface anomalies. The Tasman Sea (Figure S10) and Great Barrier Reef (Figure S11) events exhibit
intermediate behaviour, with anomalies extending below the mixed layer during peak phases but remaining strongly
305 modulated by seasonal stratification and mixed-layer variability. Across all regions, the comparison between CGLO and
GLOR products demonstrates substantial consistency in the timing and vertical extent of the strongest anomalies, despite
differences in anomaly magnitude and mixed-layer depth representation.

Violin-plot distributions in Figure 17 for the Western Australia region (other AOI in figures S15–S20) summarize the results
for each metric, for each member of the GREP, and for the ensemble mean, distinguishing between shallow and deep events.
310 The main source of differences in the parameters of the distributions (mean, interquartile ranges, shapes, tails) is the choice
of a given member, rather than the classification between shallow-deep events - exception made for the maximum and mean
absolute-value intensities of the events (more or less true for all the areas of interest chosen). These variations underscore the
importance of using ensemble approaches in climate diagnostics, as single-model outputs may under- or overestimate key
metrics, such as duration, intensity, heterogeneity or cumulative intensity, depending on structural biases, differences in
315 parameterisations of vertical mixing, data assimilation background-error structures, etc., which may influence the
propagation of temperature anomalies over the water column.

Figures S14–S20 compare MHW characteristics in waters shallower than 50 m and deeper than 50 meters. Figure S14 shows
the global distribution of MHW-day percentages, highlighting a systematic increase in persistence and occurrence in shallow
coastal environments, while deeper regions exhibit lower occurrence but often larger temporal continuity. The subsequent
320 regional comparisons (Figures S15–S20) reveal that shallow waters generally experience shorter but more intense events,
whereas deeper waters tend to exhibit longer persistence and enhanced cumulative intensity due to subsurface heat retention.
These differences are particularly evident in the Northeast Pacific Blob and Northwest Atlantic events, where deep-water
anomalies remain detectable long after surface anomalies weaken. A common feature emerging from all supplementary
figures is the overall consistency of the MHW spatial patterns and temporal evolution among the different reanalysis
325 products, despite regional differences in intensity and duration.

4 Summary and Discussions

This study investigates the subsurface characteristics of Marine Heatwaves (MHWs) using the ESA Climate Change
Initiative SST dataset (Merchant et al., 2014; Merchant et al., 2019) and the Copernicus Marine Environmental Monitoring
Service (CMEMS) Global Ocean Reanalysis Ensemble Products (GREP; Donlon et al., 2012). By analysing multiple depths



330 over the range 0.5–200 m across the three different reanalysis members, we used the traditional surface-based definitions of MHWs (Hobday et al., 2016) to examine MHWs vertical extent and persistence.

Our results confirm that well-documented events such as the 2011 Western Australia heatwave (Wernberg et al., 2013; Thomson et al., 2015) and the 2015 Northeast Pacific Blob (Bond et al., 2015; Di Lorenzo & Mantua, 2016; Cavole et al., 2016) were successfully detected across GREP members, with anomalies also showing up below the mixed layer to depths
335 of 100–150 m. While all reanalyses captured the timing and general structure of these events, differences emerged in intensity and duration. ORAS and CGLORS produced slightly more intense and persistent signals than GLORYS, resulting in a greater extent of the extreme category area on the date of peak values. These discrepancies highlight the sensitivity of subsurface MHW detection to model configurations and assimilation schemes (Holbrook et al., 2019; Sun et al., 2023; Zhang et al., 2023).

340 Some studies in the literature highlighted how the coupling between vertical anomaly distribution and mixed-layer dynamics underscores the role of vertical mixing in modulating subsurface thermal extremes (Chen et al., 2024; Yuan et al., 2024). Here, the assessment of the ensemble product revealed consistent spatial patterns in MHW frequency, onset, and decline rates, but also exposed inter-model differences, particularly in high-latitude regions and dynamically complex basins such as the Southern Ocean and North Atlantic (Frölicher et al., 2018; Sen Gupta et al., 2020). Furthermore, the inspection of MHW
345 characteristics globally revealed that, despite marked inter-model differences, all the GREP members show overall agreement in the representation of deep events, which have a higher duration, intensity, category level, heterogeneity, and also have sharper onset and decline rates than near surface events.

From an ecological and socio-economic perspective, the subsurface expression of MHWs is critical, as confirmed by the analyses in the present manuscript. Many marine organisms and ecosystems—coral reefs, kelp forests, pelagic fisheries—are
350 sensitive to subsurface thermal conditions (Hughes et al., 2017; Smale et al., 2019). Events that extend below the surface layer may have longer-lasting and more severe impacts than those inferred from SST alone (Oliver et al., 2017; Garrabou et al., 2009; Mills et al., 2013). This has direct implications for fisheries management, aquaculture, and climate adaptation strategies, which must incorporate subsurface diagnostics to anticipate ecosystem vulnerability, highlighting the need for reliable forecasting systems that extend below the surface (McAdam et al., 2023); since such systems are typically initialised
355 with ocean reanalyses, their skill in detecting these extremes is itself dependent on reanalysis quality, reinforcing the validation carried out in this paper.

In conclusion, our findings demonstrate that subsurface reanalysis data provide valuable insights into the vertical structure of MHWs, complementing satellite-based SST observations. However, differences among reanalysis products emphasise the need for multi-model ensemble approaches to characterise uncertainty. Future analyses should better integrate in situ
360 observations (e.g., Argo floats) by co-locating in-situ profiles with reanalyses during MHW events for an enhanced comparison, use higher resolution products, or refine the vertical classification framework to account for dynamically relevant subsurface levels, such as mixed-layer or thermocline depths (Malan et al. 2025). Future studies should also extend surface MHWs tracking algorithms (e.g., Scannell et al., 2024; Zhao et al., 2026) to a 4-dimensional domain to more fully



understand vertical propagation of anomalies. Such efforts will inform predictive capacity and support more effective
365 climate resilience planning in marine systems.

Figures

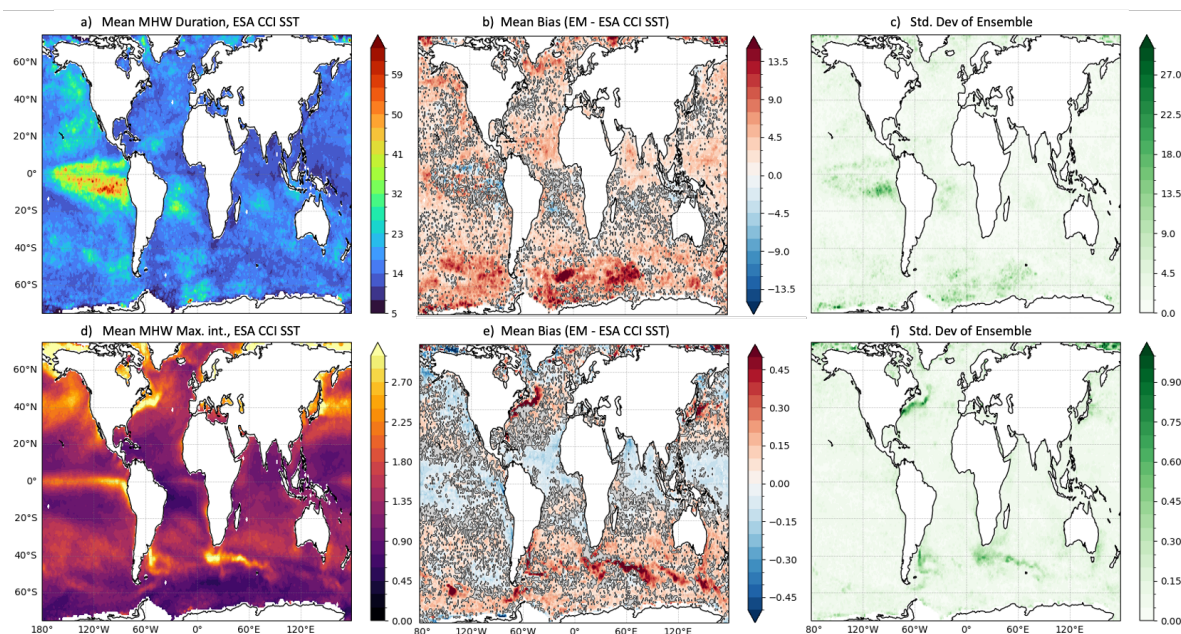


Figure 1. Mean MHW duration and Maximum intensity across all surface events: panel a), b), c) show the mean MHW
370 duration for ESACCI SST, the bias in mean duration of the ensemble mean with respect to satellite data (shaded in grey
regions where the bias is higher than two times the ensemble standard deviation), and the ensemble standard deviation of
mean duration expressed in days. Panels d), e) and f) show the same, but for the maximum intensity above the threshold
reached across all the events, expressed in Celsius degrees.

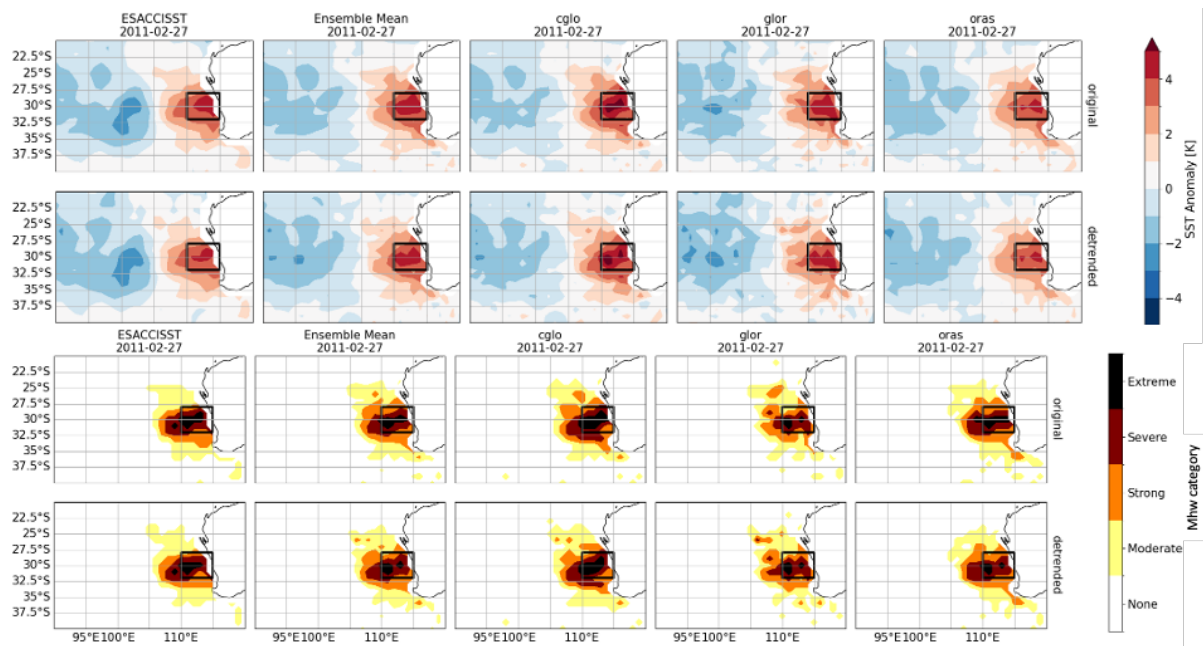


Figure 2. Anomalies at the peak of the Western Australia 2011 event (top two rows, with the first row indicating results for not-detrended data, and second row results for detrended data) and categories (bottom rows, with third and fourth rows showing results for non-detrended and detrended data, respectively), based on the time series inside the subregion defined in Table 1.

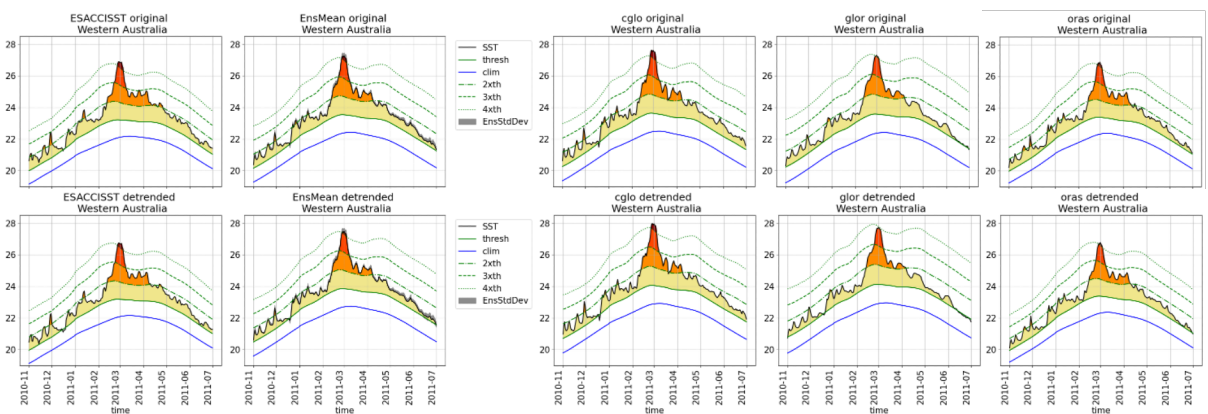
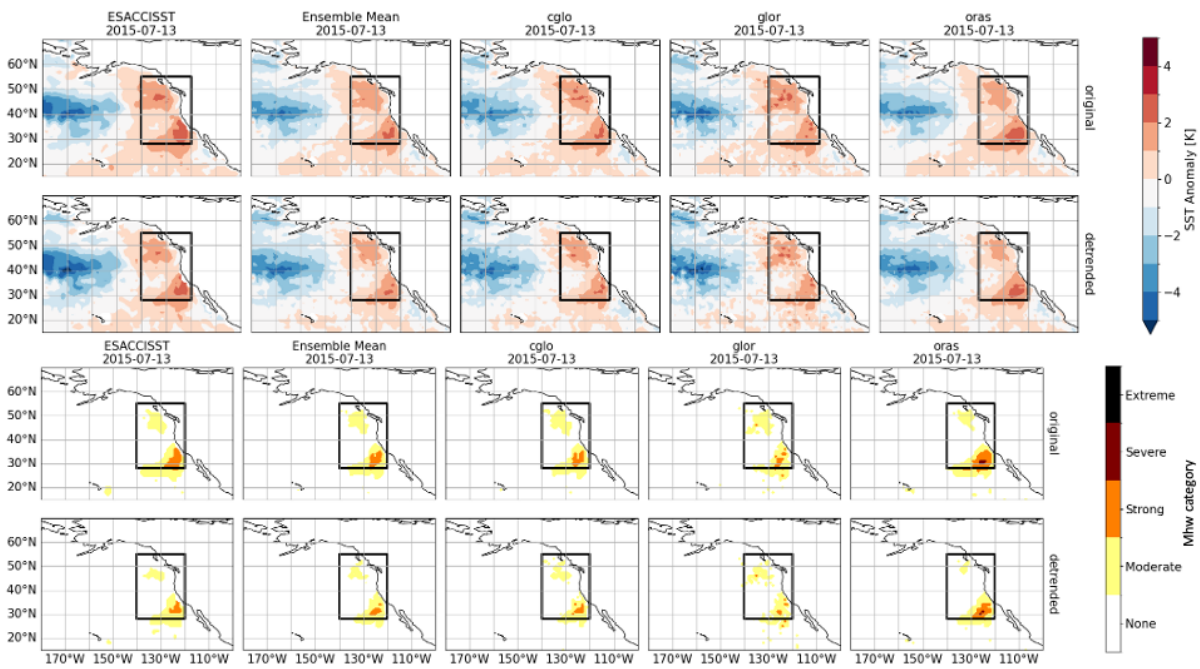


Figure 3: Time series of SST anomalies in the subregion chosen for the Western Australia 2011 event. 2xth, 3xth and 4xth are respectively the indication of the thresholds for category definition (two, three and four times the difference between the climatology and threshold). In the second column (ensemble mean), shaded regions indicate the



ensemble standard deviation along the time series. Top row shows results for non-detrended data, and the bottom row for detrended data.



385 **Figure 4.** Same as Figure 2, but for the Northeast Pacific Blob region.

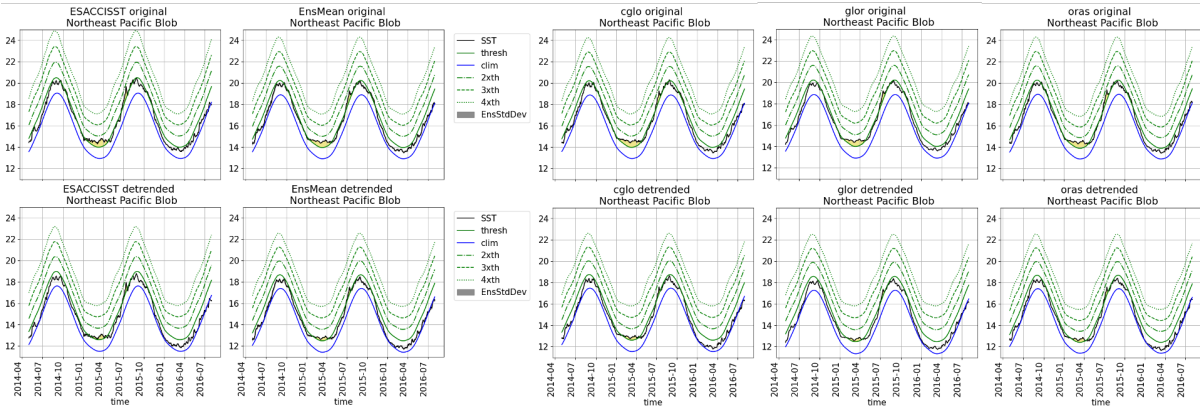


Figure 5.

Same as Figure 3, but for the Northeast Pacific Blob region.

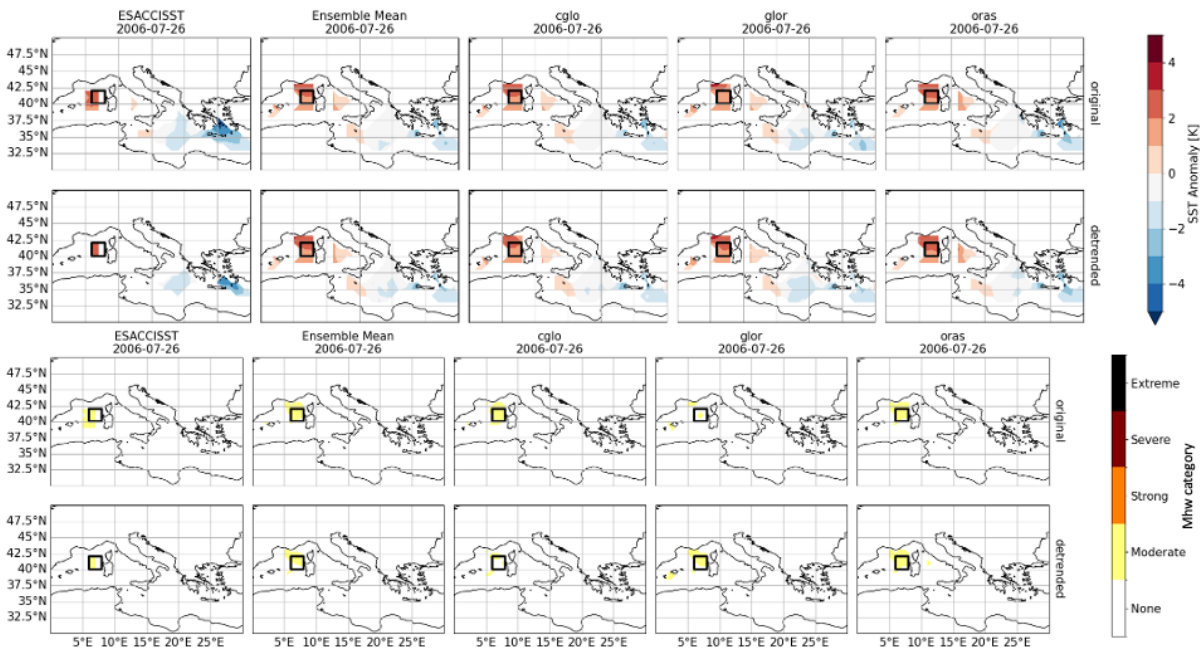


Figure 6. Same as Figure 2, but for the Mediterranean Sea region.

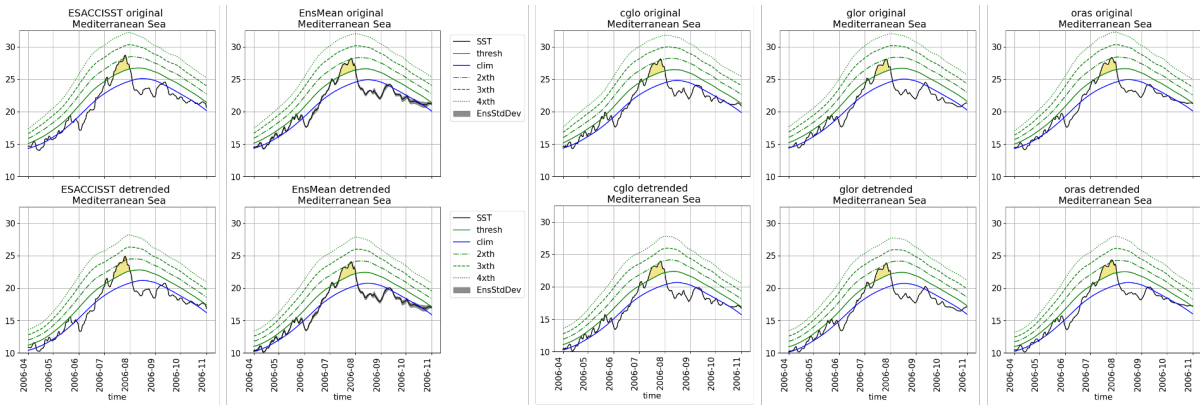


Figure 7. Same as Figure 3, but for the Mediterranean Sea AOI.

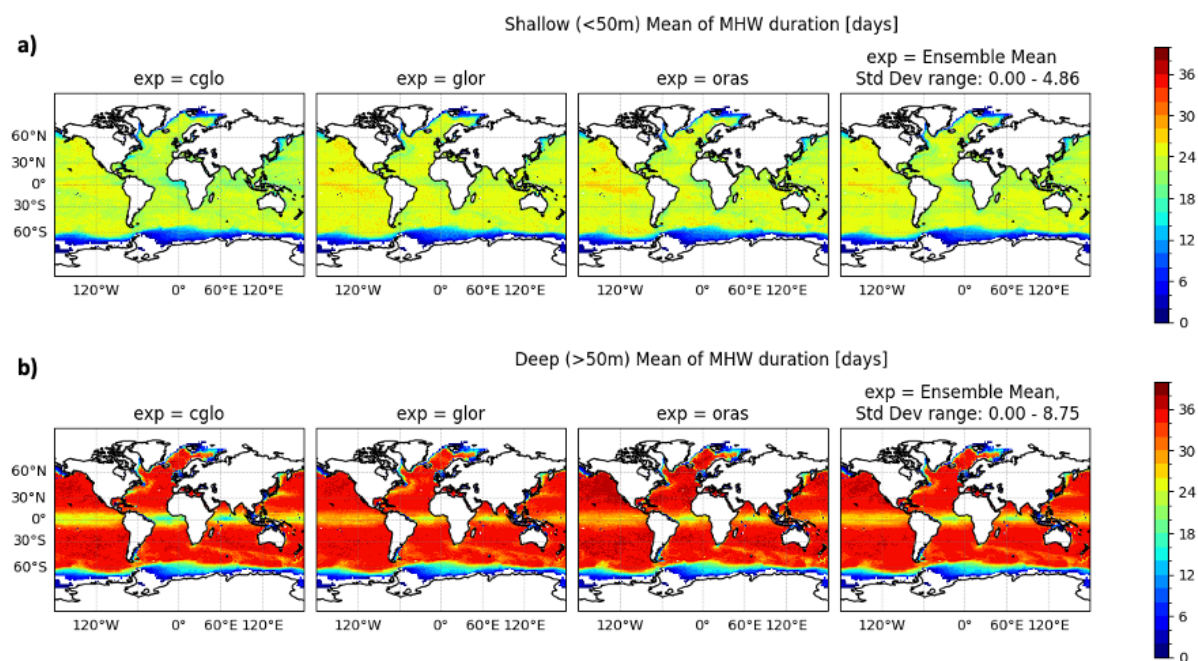


Figure 8.

Mean duration of shallow and deep (panels a and b, respectively) events from each GREP member (columns - ensemble mean shown in the last column).

400

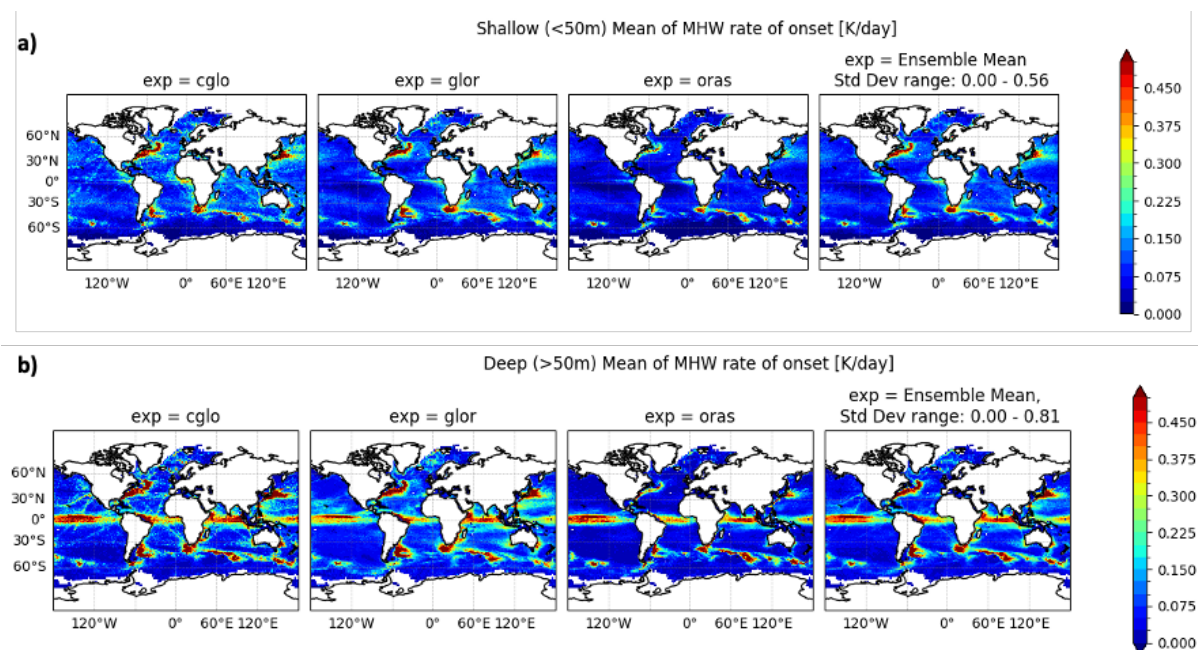
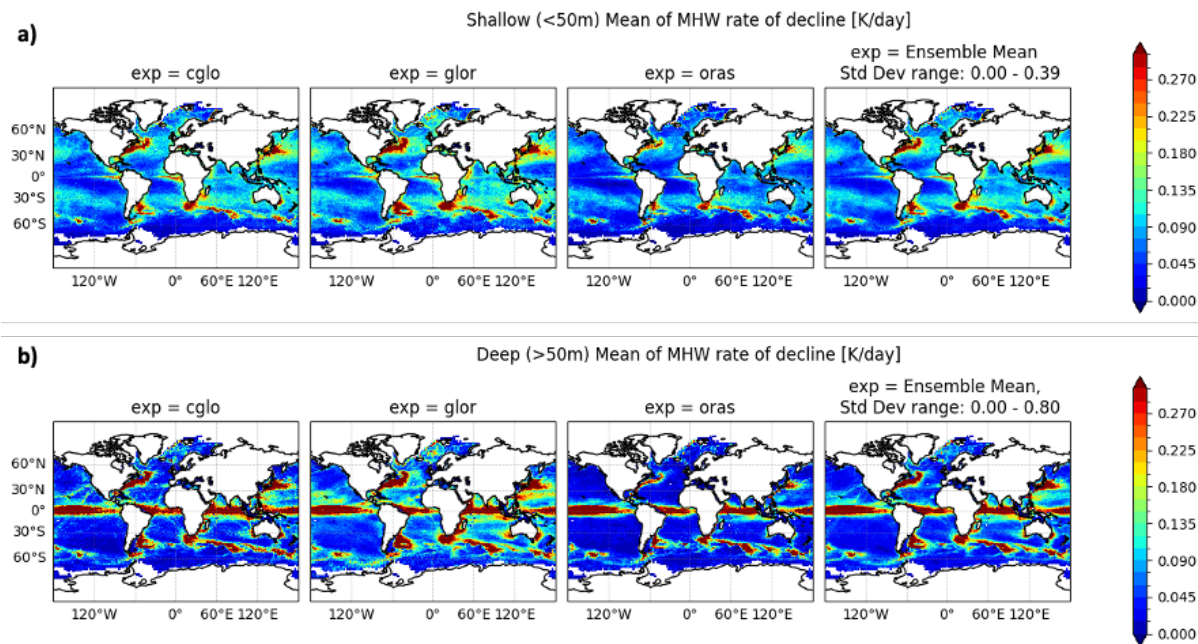




Figure 9. Mean onset rate in degrees/day of shallow and deep (panels a and b, respectively) events from each GREP member (columns - ensemble mean shown in the last column).



405 **Figure 10.** Mean decline rate in degrees/day of shallow and deep (panels a and b, respectively) events from each GREP member (columns - ensemble mean shown in in the last column).

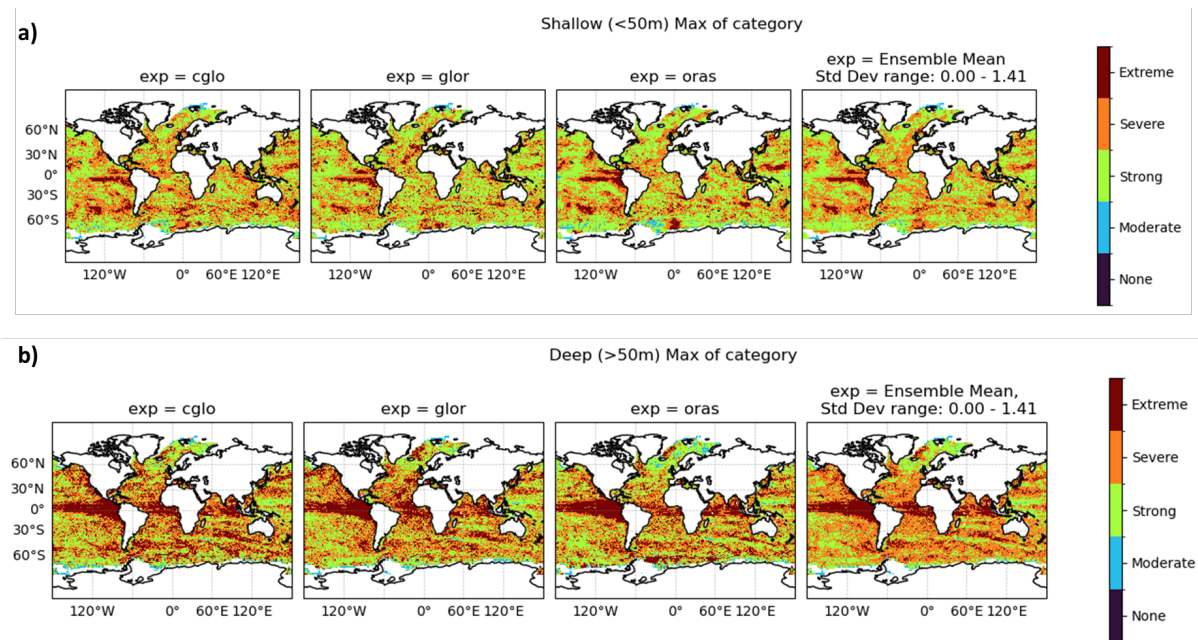




Figure 11. Maximum Category reached for Shallow and Deep events (panels a, b respectively) from each GREP member
410 (columns - ensemble mean in the last column).

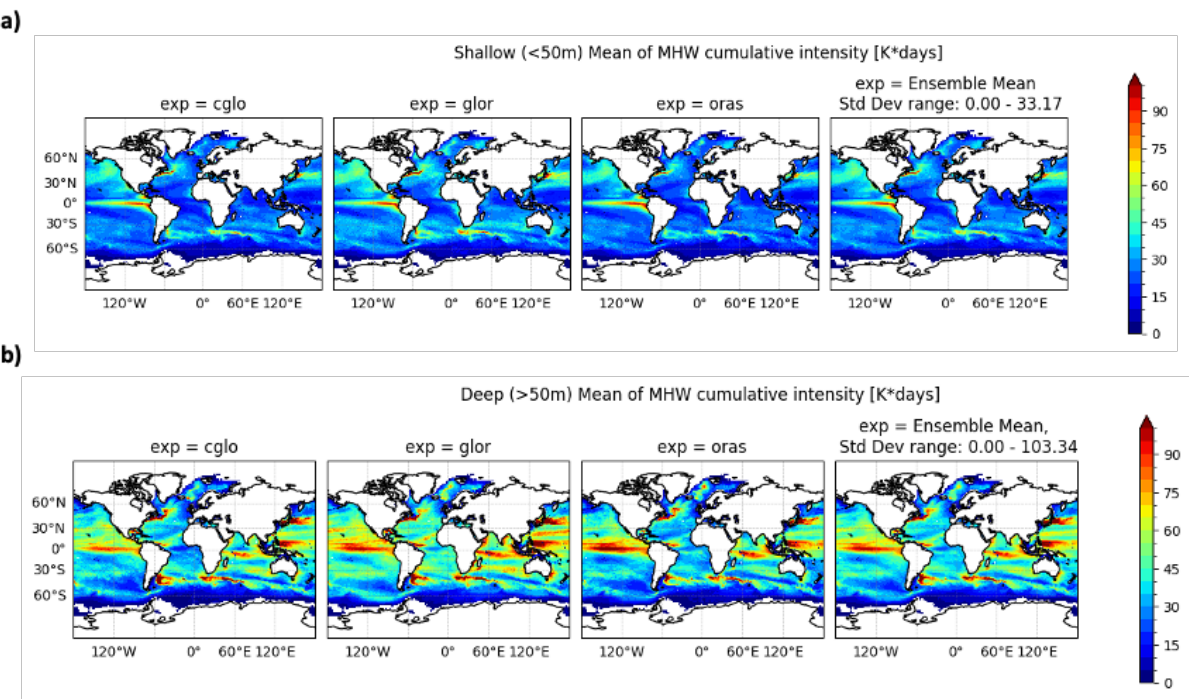


Figure 12. Mean of cumulative intensity in degrees day during all events in Shallow (a) and Deep (b) events from each
415 GREP member (columns - ensemble mean in the last column).

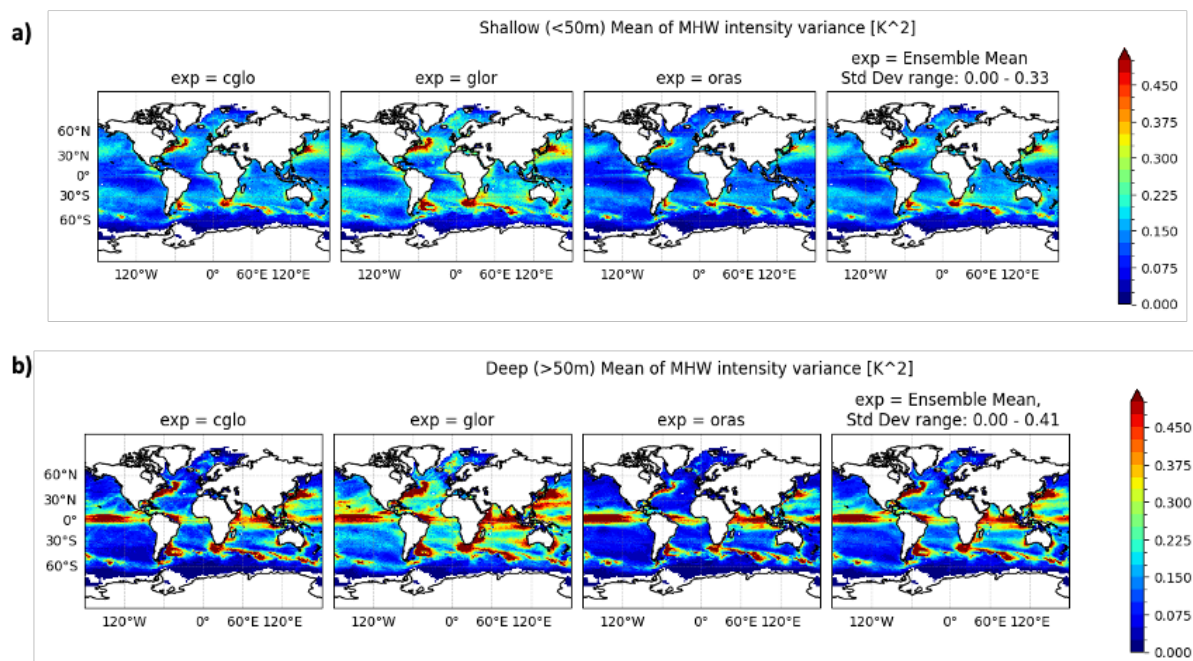


Figure 13. Mean Variance (i.e. heterogeneity) of the intensity in degrees during all events in Shallow (a) and Deep (b) events from each GREP member (columns - ensemble mean in the last column).

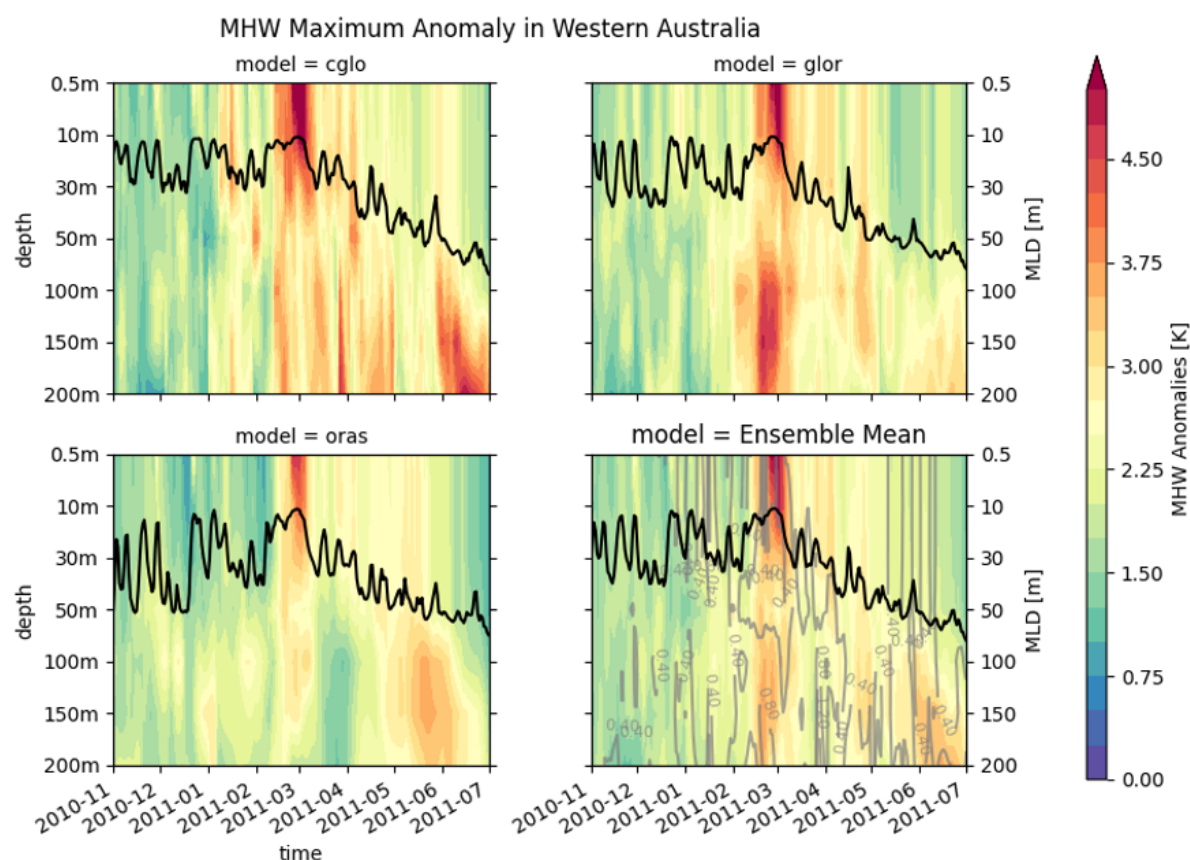


Figure 14. Spatial Maximum of the MHW intensity within its subregion (defined in Table 1) of the Western Australia event, depth vs time plots. Superimposed in black is the time variability of the Mixed Layer Depth (MLD) from each reanalysis. Grey contours in the bottom-right panel show the ensemble spread (contour interval is such that there are 5 levels in total, equally distributed on ensemble spread range)

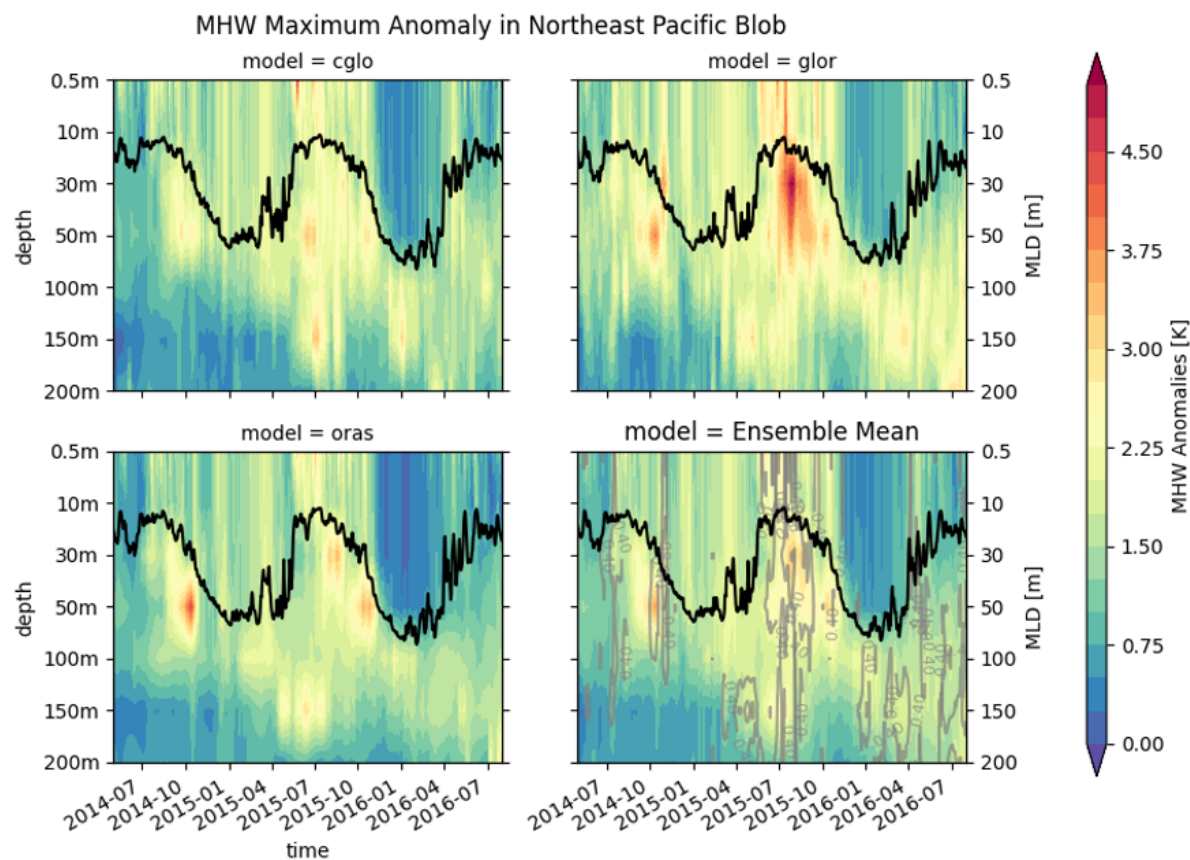


Figure 15.

425 Spatial Maximum of the anomaly above the threshold within its subregion (defined in Table 1) of the Northeast Pacific Blob event, depth vs time plots. Superimposed in black is the time variability of the Mixed Layer Depth (MLD) from each reanalysis.

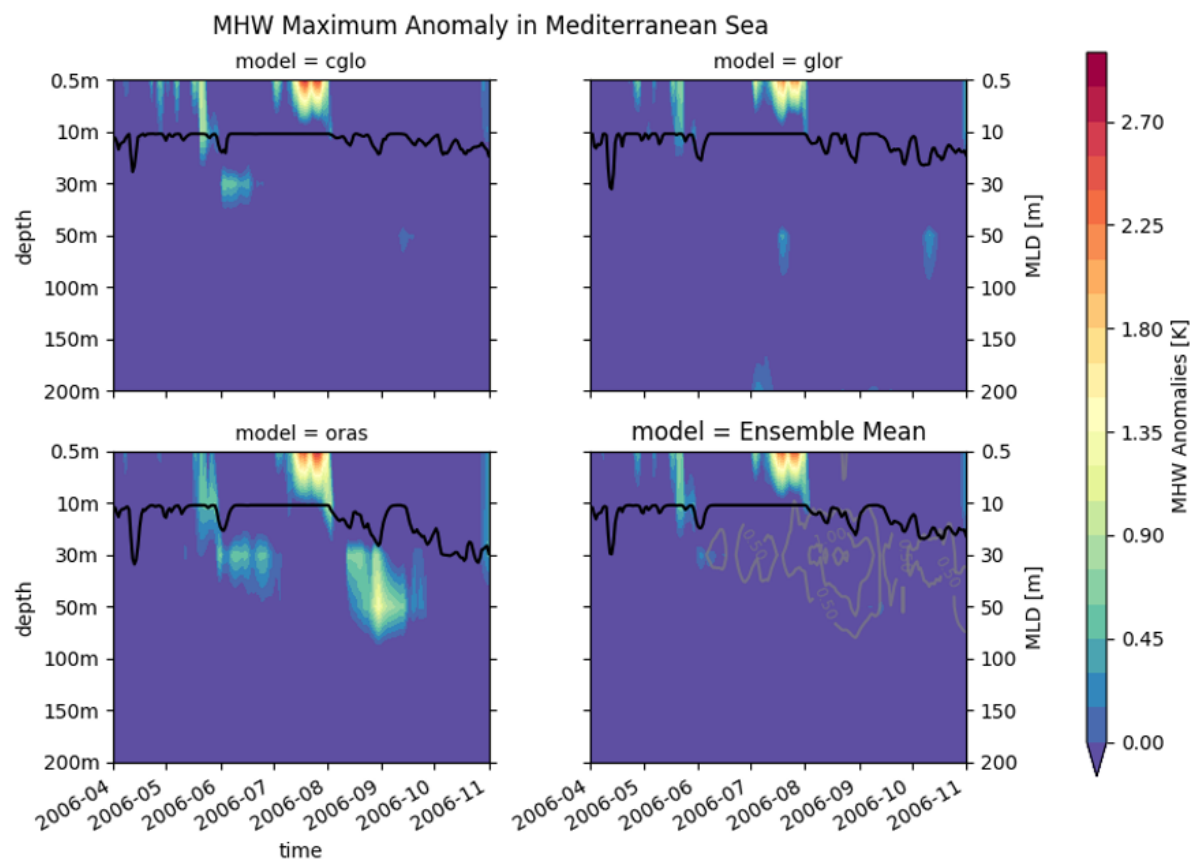


Figure 16.

Spatial Maximum of the anomaly above the threshold within its subregion (defined in Table 1) of the Mediterranean Sea event, depth vs time plots. Superimposed in black is the time variability of the Mixed Layer Depth (MLD) from each reanalysis.

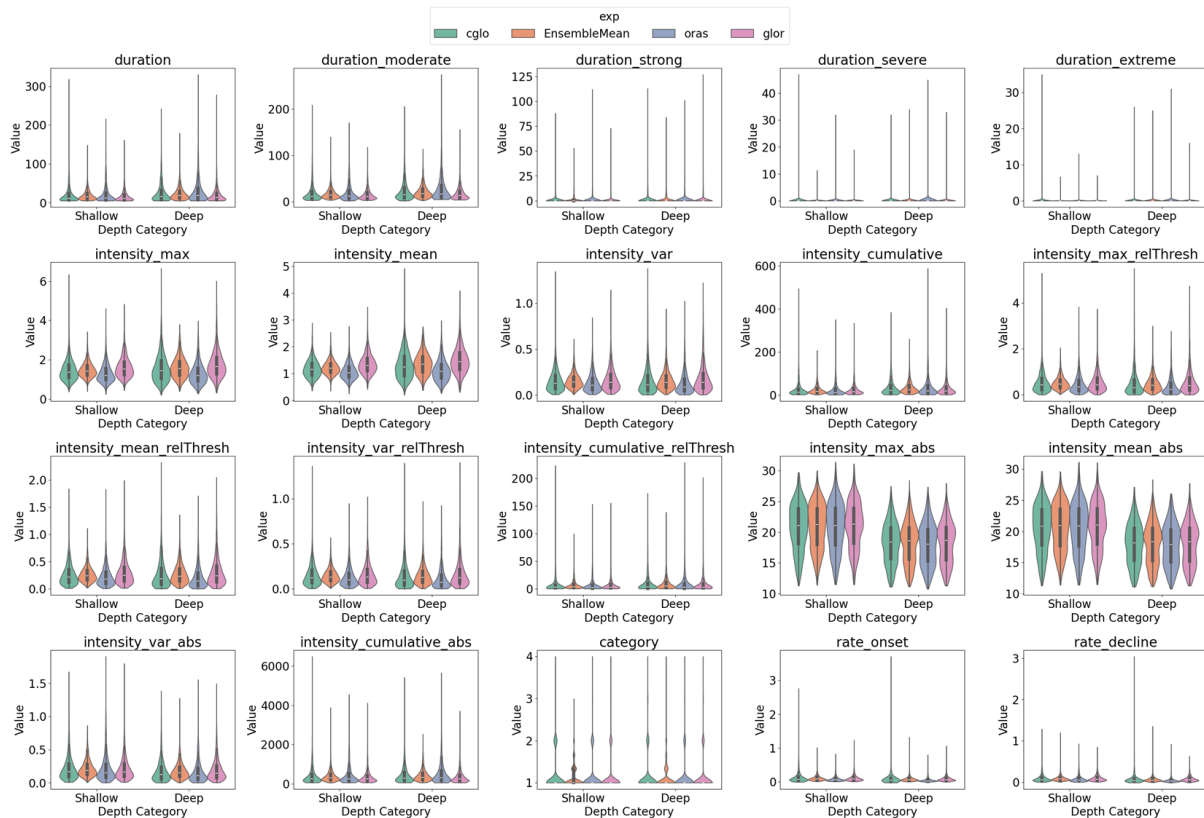


Figure 17. Violin-Plot of the distribution of variables in the Western Australia region, for Surface and Deep Variables, for each of the GREP members, over the entire range of events detected.

Appendix A

In this paper, we followed two approaches for MHW detection. The first implementation has been done by rewriting from scratch the code for MHW detection, following the recipe of Hobday et al. 2016: as an outcome of the detection, we have a matrix with the same dimension as the input data (i.e., time, latitude and longitude for each selected depth level). The main limitation of this approach (i.e., the output matrix having the same shape and dimension as the input data) is the difficulty in getting single events, and also the metrics provided by Oliver's code.

Therefore, another code version was developed and used for global detection results in the following section, i.e. an xarray wrapping of the Python version used in Hobday et al., 2016 (Oliver's code). The disadvantage of this approach is that the index along the event dimension is local in space, i.e. the starting and ending dates for an event index at a given location do not coincide with the ones for the same event index at another location. But as long as we look at quantities which are averaged on the event dimension for each point in space, this issue is practically irrelevant.



Appendix B

In this section, we define the metrics used throughout the paper. See Figure B1 for a resume with mathematical details.

Marine Heatwaves Global Metrics: Math

Let N be the total number of detected MHW events. For each event i , let t_i^{start} , t_i^{peak} , t_i^{end} be the start, peak and end dates, and $I_i(t)$ the intensity time series.

Mean Duration:

$$\bar{D} = \frac{1}{N} \sum_{i=1}^N (t_i^{\text{end}} - t_i^{\text{start}} + 1)$$

Mean Onset Rate:

$$\bar{R}_{\text{onset}} = \frac{1}{N} \sum_{i=1}^N \frac{I_i^{\text{max}} - I_i(t_i^{\text{start}})}{t_i^{\text{peak}} - t_i^{\text{start}}}$$

Maximum Category Reached:

$$C_{\text{max}} = \max_{i=1, \dots, N} \left(\max_{t \in [t_i^{\text{start}}, t_i^{\text{end}}]} C_i(t) \right)$$

Mean Decline Rate:

$$\bar{R}_{\text{decline}} = \frac{1}{N} \sum_{i=1}^N \frac{I_i^{\text{max}} - I_i(t_i^{\text{end}})}{t_i^{\text{end}} - t_i^{\text{peak}}}$$

Mean Cumulative Intensity:

$$\bar{CI} = \frac{1}{N} \sum_{i=1}^N \sum_{t=t_i^{\text{start}}}^{t_i^{\text{end}}} I_i(t)$$

450 **Figure B1.** Mathematical definitions for the Mean Duration, Mean Onset Rate, Maximum Category Reached, Mean Decline Rate and Mean Cumulative Intensity.

Code and data availability

All the codes used in this paper, together with the ones used to produce results and figures within this manuscript, are available in a public GitHub repository (<https://github.com/vincenzodetoma/ROHAN>).

455 Data used to carry out the analysis and produce the figures are freely available upon registration on CMEMS portal.

Author contributions

VdT and AS made the conceptualisation, and VdT wrote the code and performed the analysis, providing the initial draft. AC, RMA and CY contributed to the discussion on diagnostics to include and to the review of the initial draft.



Competing interests

460 At least one of the (co-)authors serves as editor for the special issue to which this paper belongs.

Acknowledgements

The authors acknowledge Vega Forneris, Flavio La Padula and Filippo Manfredonia for the IT support provided during the different phases of code development, and Gael Forget and Katie Smith for fruitful discussions at the Ocean Science Meeting in Glasgow.

465 References

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. <https://doi.org/10.1038/s41467-023-36567-0>

470 Artana, C., Kaplan, A., Ramírez, F., Ortega, M., Steenbeek, J. G., & Coll, M. (2026). Marine heatwaves are transforming Western mediterranean marine ecosystems. *Scientific Reports*. <https://doi.org/10.1038/s41598-026-35813-x>

Bian, C., Jing, Z., Wang, H., Wu, L., Chen, Z., Gan, B., & Yang, H. (2023). Oceanic mesoscale eddies as crucial drivers of global marine heatwaves. *Nature Communications*, 14(1), 2970. <https://doi.org/10.1038/s41467-023-38811-z>

475

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. <https://doi.org/10.1002/2015GL063306>

Bourdallé-Badie, R., Karioua, S., Carabin, F., Cipollone, A., Zuo, H., Parent, L., ... & Drévillon, M. (2024). Global
480 Production Centre GLOBAL_MULTIYEAR_PHY_ENS_001_031. <https://documentation.marine.copernicus.eu/QUID/CMEMS-GLO-QUID-001-031.pdf>

Capotondi, A., Rodrigues, R. R., Sen Gupta, A., Benthuyssen, J. A., Deser, C., Frölicher, T. L., ... Wang, C. (2024). A global overview of marine heatwaves in a changing climate. *Communications Earth & Environment*, 5(1), Article 701.
485 <https://doi.org/10.1038/s43247-024-01806-9>



- 490 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. <https://doi.org/10.5670/oceanog.2016.32>
- Chen, K., Gawarkiewicz, G. G., Lentz, S. J., & Bane, J. M. (2014). Diagnosing the warming of the northeastern U.S. coastal ocean in 2012: A linkage between the atmospheric jet stream variability and ocean response. *Journal of Geophysical Research: Oceans*, 119(1), 218–227. <https://doi.org/10.1002/2013JC009393>
- 495 Chen, W., He, Q., & Staneva, J. (2024). Impact of anticyclonic mesoscale eddies on the vertical structures of marine heatwaves in the South China Sea. *Remote Sensing*, 17(3), 370. <https://doi.org/10.3390/rs17030370>
- Chen, H. H., Wang, Y., Xiu, P., Yu, Y., Ma, W., & Chai, F. (2023). Combined oceanic and atmospheric forcing of the 2013/14 marine heatwave in the northeast Pacific. *npj Climate and Atmospheric Science*, 6(1), 3. <https://doi.org/10.1038/s41612-023-00327-0>
- 500 <https://doi.org/10.1038/s41612-023-00327-0>
- Chevillard, C., Le Gendre, R., Menkes, C., Izumo, T., Pagli, B., Van Wynsberge, S., & Cravatte, S. (2026). Sensitivity of marine heatwaves metrics to SST products, focusing on the Tropical Pacific. *Ocean Science*, 22(2), 1213–1236. <https://doi.org/10.5194/os-22-1213-2026>
- 505 <https://doi.org/10.5194/os-22-1213-2026>
- Gounou, A., Dré villon, M., & Clavier, M. (2020). PRODUCT USER MANUAL For Global Ocean Reanalysis Products. EU Copernicus Mar. Serv, Inf.[Online]. <https://documentation.marine.copernicus.eu/PUM/CMEMS-GLO-PUM-001-031.pdf>
- 510 Di Lorenzo, E., & Mantua, N. (2016). Multi-year persistence of the 2014/15 North Pacific marine heatwave. *Nature Climate Change*, 6, 1042–1047. <https://doi.org/10.1038/nclimate3082>
- Donlon, C. J., Martin, M., Stark, J., Roberts-Jones, J., Fiedler, E., & Wimmer, W. (2012). The operational sea surface temperature and sea ice analysis (OSTIA) system. *Remote Sensing of Environment*, 116, 140–158. <https://doi.org/10.1016/j.rse.2010.10.017>
- 515 <https://doi.org/10.1016/j.rse.2010.10.017>
- Dayan, H., McAdam, R., Juza, M., Masina, S., & Speich, S. (2023). Marine heat waves in the Mediterranean Sea: An assessment from the surface to the subsurface to meet national needs. *Frontiers in Marine Science*, 10, 1045138. <https://doi.org/10.3389/fmars.2023.1045138>
- 520 <https://doi.org/10.3389/fmars.2023.1045138>



Easterling, D. R., Evans, J. L., Groisman, P. Y., Karl, T. R., Kunkel, K. E., & Ambenje, P. (2000). Observed variability and trends in extreme climate events: A brief review. *Bulletin of the American Meteorological Society*, 81(3), 417–426.

[https://doi.org/10.1175/1520-0477\(2000\)081<0417:OVATIE>2.3.CO;2](https://doi.org/10.1175/1520-0477(2000)081<0417:OVATIE>2.3.CO;2)

525

Frölicher, T. L., Fischer, E. M., & Gruber, N. (2018). Marine heatwaves under global warming. *Nature*, 560(7718), 360–364.

<https://doi.org/10.1038/s41586-018-0383-9>

530 Garrabou, J., Coma, R., Bensoussan, N., Bally, M., Chevaldonné, P., Cigliano, M., ... Cerrano, C. (2009). Mass mortality in northwestern Mediterranean rocky benthic communities: Effects of the 2003 heat wave. *Global Change Biology*, 15(5), 1090–1103. <https://doi.org/10.1111/j.1365-2486.2008.01823.x>

Guinehut, S., Dhomps, A. L., Larnicol, G., & Le Traon, P. Y. (2012). High resolution 3-D temperature and salinity fields derived from in situ and satellite observations. *Ocean Science*, 8(5), 845–857. <https://doi.org/10.5194/os-8-845-2012>

535

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.

<https://doi.org/10.1016/j.pocean.2015.12.014>

540 Hobday, A.J., E.C.J. Oliver, A. Sen Gupta, J.A. Benthuisen, M.T. Burrows, M.G. Donat,

N.J. Holbrook, P.J. Moore, M.S. Thomsen, T. Wernberg, and D.A. Smale. 2018.

Categorizing and naming marine heatwaves. *Oceanography* 31(2), <https://doi.org/10.5670/oceanog.2018.205>.

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

550 Hughes, T. P., Kerry, J. T., Álvarez-Noriega, M., Álvarez-Romero, J. G., Anderson, K. D., Baird, A. H., ... Wilson, S. K. (2017). Global warming and recurrent mass bleaching of corals. *Nature*, 543(7645), 373–377. <https://doi.org/10.1038/nature21707>

Lalli, C. (1997). *Biological oceanography: An introduction*. Elsevier.



- Malan, N., Gupta, A. S., Schaeffer, A., Zhang, S., Doblin, M. A., Pilo, G. S., ... & Spillman, C. M. (2025). Lifting the lid on
555 Marine Heatwaves. *Progress in Oceanography*, 103539. <https://doi.org/10.1016/j.pocean.2025.103539>
- McAdam, R., Masina, S., & Gualdi, S. (2023). Seasonal forecasting of subsurface marine heatwaves. *Communications Earth
& Environment*, 4(1), 225. <https://doi.org/10.1038/s43247-023-00892-5>
- 560 Merchant, C. J., Embury, O., Bulgin, C. E., Block, T., Corlett, G. K., Fiedler, E., et al. (2019). Satellite-based time-series of
sea-surface temperature since 1981 for climate applications. *Scientific Data*, 6(1), 1–18. <https://doi.org/10.1038/s41597-019-0236-x>
- Merchant, C. J., Embury, O., Roberts-Jones, J., Fiedler, E., Bulgin, C. E., Corlett, G. K., et al. (2014). Sea surface
565 temperature datasets for climate applications from phase 1 of the European Space Agency Climate Change Initiative (SST
CCI). *Geoscience Data Journal*, 1(2), 179–191. <https://doi.org/10.1002/gdj3.20>
- Mills, K. E., Pershing, A. J., Brown, C. J., Chen, Y., Chiang, F.-S., Holland, D. S., ... Wahle, R. A. (2013). Fisheries
management in a changing climate: Lessons from the 2012 ocean heat wave in the Northwest Atlantic. *Oceanography*, 26(2),
570 191–195. <https://doi.org/10.5670/oceanog.2013.27>
- Mishonov, Alexey V., Boyer, Tim P., Baranova, Olga K., Bouchard, Courtney N., Cross, Scott L., Garcia, Hernan E.,
Locarnini, Ricardo A., Paver, Christopher R., Reagan, James R., Wang, Zhankun, Seidov, Dan, Grodsky, Alexandra I.,
Beauchamp, James G. (2024). World Ocean Database 2023. <https://doi.org/10.25923/z885-h264>
- 575 Olita, A., Ribotti, A., Fazioli, L., Perilli, A., & Sorgente, R. (2007). Effects of the 2003 European heatwave on the central
Mediterranean Sea: Surface fluxes and the dynamical response. *Ocean Science*, 3(2), 273–289. <https://doi.org/10.5194/os-3-273-2007>
- 580 Oliver, E. C. J., Donat, M. G., Burrows, M. T., Moore, P. J., Smale, D. A., Alexander, L. V., ... Holbrook, N. J. (2018).
Longer and more frequent marine heatwaves over the past century. *Nature Communications*, 9, 1324.
<https://doi.org/10.1038/s41467-018-03732-9>
- Oliver, E. C. J., Holbrook, N. J., Sen Gupta, A., Benthuyssen, J. A., Donat, M. G., Hobday, A. J., ... Alexander, L. V. (2019).
585 Projected marine heatwaves in the 21st century and the potential for ecological impact. *Frontiers in Marine Science*, 6, 734.
<https://doi.org/10.3389/fmars.2019.00734>



Oliver, E. C. J., Wernberg, T., Smale, D. A., Walsh, J. C., & Smit, A. J. (2017). The unprecedented 2015/16 Tasman Sea marine heatwave. *Nature Communications*, 8, 16101. <https://doi.org/10.1038/ncomms16101>

590

Paul, A., Maheswaran, P. A., Satheesan, K., Kottayil, A. (2025). Assessing ocean reanalysis accuracy for marine extremes in the Indian Ocean using in-situ observations. *Climate Dynamics*, 63, Article 343. <https://doi.org/10.1007/s00382-025-07780-y>

595 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. <https://doi.org/10.1029/2019GL084928>

Pisano, A., de Toma, V., Avolio, E., Serva, F., Fanelli, C., Ciani, D., ... & Marullo, S. (2026). Observational evidence for a shift toward a warming-dominated phase of Mediterranean SST evolution. *Environmental Research Letters*.
600 <https://doi.org/10.1088/1748-9326/ae9ee2>

Scannell, H. A., Cai, C., Thompson, L., Whitt, D. B., Gagne, D. J., & Abernathey, R. P. (2024). Spatiotemporal Evolution of Marine Heatwaves Globally. *Journal of Atmospheric and Oceanic Technology*, 41(12), 1247–1263.
<https://doi.org/10.1175/JTECH-D-23-0126.1>

605

Schulzweida, U., Kornblueh, L., & Quast, R. (2019). CDO user guide. <https://code.mpimet.mpg.de/projects/cdo/wiki/Cdo#Documentation>

Sen Gupta, A., Thomsen, M., Benthuyssen, 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. <https://doi.org/10.1038/s41598-020-75445-3>

610

Smale, D. A., Wernberg, T., Oliver, E. C. J., Thomsen, M. S., 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(4), 306–312.
615 <https://doi.org/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. <https://doi.org/10.1146/annurev-marine-032122-121437>

620



- Storto, A., Masina, S., Simoncelli, S., Iovino, D., Cipollone, A., Drevillon, M., ... & Peterson, K. A. (2019). The added value of the multi-system spread information for ocean heat content and steric sea level investigations in the CMEMS GREP ensemble reanalysis product. *Climate dynamics*, 53(1), 287–312. <https://doi.org/10.1007/s00382-018-4585-5>
- 625 Sun, D., Jing, Z., Li, F., & Wu, L. (2023). Characterizing global marine heatwaves under a spatio-temporal framework. *Progress in Oceanography*, 211, 102947. <https://doi.org/10.1016/j.pocean.2022.102947>
- Sun, D., Li, F., Jing, Z., Hu, S., & Zhang, B. (2023). Frequent marine heatwaves hidden below the surface of the global ocean. *Nature Geoscience*, 16, 1099–1104. <https://doi.org/10.1038/s41561-023-01325-w>
- 630 Tan, H., & Cai, R. (2018). What caused the record-breaking warming in East China Seas during August 2016? *Atmospheric Science Letters*, 19(8), e853. <https://doi.org/10.1002/asl.853>
- Thomson, J. A., Burkholder, D. A., Heithaus, M. R., Fourqurean, J. W., Fraser, M. W., Statton, J., ... Kendrick, G. A. (2015). Extreme temperatures, foundation species, and abrupt ecosystem change: An example from an iconic seagrass ecosystem. *Global Change Biology*, 21(4), 1463–1474. <https://doi.org/10.1111/gcb.12835>
- 635 Wang, Q., Zhang, B., Zeng, L., He, Y., Wu, Z., & Chen, J. (2022). Properties and drivers of marine heat waves in the northern South China Sea. *Journal of Physical Oceanography*. <https://doi.org/10.1175/JPO-D-21-0236.1>
- 640 Wernberg, T., Smale, D. A., Tuya, F., Thomsen, M. S., Langlois, T. J., de Bettignies, T., ... Bennett, S. (2013). An extreme climatic event alters marine ecosystem structure in a global biodiversity hotspot. *Nature Climate Change*, 3, 78–82. <https://doi.org/10.1038/nclimate1627>
- 645 Yuan, T., Zhang, J., Yang, S., Su, Q., Zhao, Z., & Ren, K. (2024). Vertical structures of marine heatwaves in the South China Sea: Characteristics, drivers and impacts on chlorophyll concentration. *Journal of Geophysical Research: Oceans*, 129(12), e2024JC021091. <https://doi.org/10.1029/2024JC021091>
- Zhang, Y., Du, Y., Feng, M., & Hobday, A. J. (2023). Vertical structures of marine heatwaves. *Nature Communications*, 14, 6483. <https://doi.org/10.1038/s41467-023-42219-0>
- 650 Zhao, Z., Holbrook, N. J., Capotondi, A., Cravatte, S., Kajtar, J. B., Sen Gupta, A., ... & Spillman, C. M. (2026). Toward a mechanistic characterisation of marine heatwaves. *Scientific Reports*. <https://doi.org/10.1038/s41598-026-40354-4>