ARTICLE

Check for updates

Seasonal forecasting of subsurface marine heatwaves

Ronan McAdam [1✉], Simona Masina [1] & Silvio Gualdi[2]

Marine heatwaves damage marine ecosystems and services, with effects identified mostly below the ocean surface. To create a truly user-relevant detection system, it is necessary to provide subsurface forecasts. Here, we demonstrate the feasibility of seasonal forecasting of subsurface marine heatwaves by using upper ocean heat content. We validate surface and subsurface events forecast by an operational dynamical seasonal forecasting system against satellite observations and an ocean reanalysis, respectively. We show that indicators of summer events (number of days, strongest intensity, and number of events) are predicted with greater skill than surface equivalents across much of the global ocean. We identify regions which do not display significant surface skill but could still benefit from accurate subsurface early warning tools (e.g., the mid-latitudes). The dynamical system used here outperforms a persistence model and is not widely influenced by warming trends, demonstrating the ability of the system to capture relevant subseasonal variability.

[1] Ocean modeling and Data Assimilation Division, Fondazione Centro Euro-Mediterraneo sui Cambiamenti Climatici—CMCC, Bologna, Italy. [2] Climate Simulations and Predictions Division, Fondazione Centro Euro-Mediterraneo sui Cambiamenti Climatici—CMCC, Bologna, Italy. ✉email: ronan.mcadam@cmcc.it

Due to the widespread availability of sea surface temperature (SST) fields from satellites, marine heatwaves are typically tracked and forecast at the surface[1–4]. However, many ecological impacts associated with MHWs occur below the surface, including coral bleaching, species displacement and losses for fisheries and aquaculture[5–11]. In the Mediterranean Sea, for example, 81% of the 582 mortality events caused by temperature anomalies occurred within the upper 40 m (using mass mortality data from 1979 to 2017 available at https://digital.csic.es/handle/10261/171445); only 20% of the total, on the other hand, are defined as surface-only events (listed as occurring at 0 m)[12]. Daily vertical migration across tens of metres in the near-surface is a common behaviour for a range of marine species, from plankton and zooplankton[13,14] to larger predatory species[15,16]. As a result, the creation of marine ecosystem protection and monitoring programs is now treated as a four-dimensional problem, in which the subsurface is key[17]. Thus, a single-layer indicator for extreme events, such as MHWs defined by SST, may not accurately determine the extreme conditions to which marine wildlife is exposed.

Meanwhile, new MHW forecasting tools are being developed to work with and for marine stakeholders[18–21]. Just as there are many examples of the energy, agriculture and tourism industries benefiting from seasonal forecasts of atmospheric conditions[22], the potential for marine seasonal forecasting to aid marine stakeholders is also being realised. Seasonal forecasts of ocean variables, in particular temperature, have aided various activities, from marine wildlife conservation to industrial activities such as fishing and aquaculture farming, in making resource management decisions and boosting resilience to climate variability[3,18,20,23–26]. Many studies and applications, however, typically focus on surface variables (e.g., SST), mainly due to the greater availability and quality of surface forecasts and validation datasets compared to the subsurface. This is despite the fact that their target applications (e.g., disease outbreaks[10], coral reef bleaching[19], fishery stock changes[27]) are phenomena which occur below the surface. Caged fish have been observed to avoid the surface during warmer surface conditions, implying full-cage MHWs are of more concern to aquaculture farms[28]. An increase in subsurface marine seasonal forecasting validation is the first step in the development of new and relevant forecasting tools which could eventually bring benefits for marine stakeholders. Given the evidence that MHWs manifest differently at depth than at the surface, MHW occurrence cannot be forecast by SST alone. Subsurface MHW indicators are necessary.

Here, we study subsurface MHWs in an operational, fully-coupled and high-resolution seasonal forecast system: the Seasonal Prediction System from the Euro-Mediterranean Centre on Climate Change (CMCC-SPS3.5)[29]. We compare re-forecasts (forecasts of past conditions) against European Space Agency Climate Change Initiative (ESA CCI) satellite observations[30,31] and the Global ocean Reanalysis Ensemble Product (GREP)[32], which provide long-term global records of surface and subsurface events respectively. Primarily, we compare subsurface MHW prediction skill to that of the surface, to find that subsurface MHW characteristics are predicted with greater accuracy across most of the global ocean. Then, we further explore the capabilities of the seasonal forecast system by comparing it to a computationally cheaper persistence model. Finally, we study the impact of warming trends on forecast skill.

Results

Comparison of forecast skill for surface and subsurface events.
In this study, subsurface MHWs are defined with ocean heat content in the upper 40 m. The main motivation for using OHC

Table 1 Summary of subsurface MHW indicator skill scores.

Correlation skill score	Global	Tropics	Mid-latitudes
Number of days	0.51	0.62	0.36
Strongest Intensity	0.48	0.59	0.32
Number of events	0.39	0.48	0.27
Area of significant skill (%)			
Number of days	70.5	81.7	55.0
Strongest Intensity	69.4	81.3	52.9
Number of events	62.9	74.6	49.5
Improvement over surface skill			
Number of days	0.11	0.09	0.14
Strongest Intensity	0.19	0.18	0.21
Number of events	0.12	0.10	0.17

The regions are defined as follows: global (60S-60N); tropics (30S-30N); midlatitudes (60S-30S, 30N-60N). Top three rows: area average of correlation skill scores. Middle three rows: area of significant skill scores, as a percentage of total regional area. Bottom three rows: average of improvement in correlation scores over surface skill, where subsurface skill is significant.

here is that species habitat varies greatly, meaning effective forecasting for a range of applications will require several target depth ranges[33–35]. In aquaculture farms, for example, cages in both coastal and open ocean settings extend to 5 m, 20 m and deeper[36]. Species in the wild, meanwhile, migrate vertically throughout the day; the range 0–40 m covers depths frequently passed through by many key components of the marine ecosystem, from plankton to foundation species to apex predators[13–16]. In practice, each potential user of MHW forecasts may be interested in different depth levels, as there is no one specific depth range relevant to all species or applications. The depth of 40 m is also deeper than the summer mixed layer in most of the ocean, meaning we can capture more than atmospheric-driven signals alone[37], and thus expect to see different characteristics from surface-defined events[38–40]. To confirm this, a 4D reconstruction of the ocean (reanalyses), known as the Global ocean Reanalysis Ensemble Product (GREP), is used to identify subsurface MHW events. There are clear differences between the average surface and subsurface MHW characteristics (number of events, duration and intensity); on average there are fewer, longer-lasting and less intense MHWs in the subsurface than at the surface (Fig. S1).

We use the GREP reanalysis product as a benchmark against which we quantify the skill of the forecasts. In the Methods section, we justify the use of reanalyses as a validation dataset for seasonal forecast data, by displaying good agreement with MHW characteristics at the surface derived from satellite-derived SST (Fig. S2). Across the global ocean, forecast skill of three MHW indicators (number of MHW days, strongest intensity and number of MHW events) is encouragingly high for the subsurface events (Fig. 1a–c). Significant positive correlation between the forecasts and validation dataset, for all indicators, is found in between 60 % and 70% of the ocean between 60°S and 60°N (Table 1). The tropics (between 30°S and 30°N) are the most skilfully predicted, with an average skill score around 0.6 for the number of days and strongest intensity. Overall, global patterns of skill resemble those of OHC anomaly forecast skill, although the skill of predicting extremes (i.e., MHWs) is not as widely significant as the skill of predicting anomalies (Fig. S3a). While the patterns and magnitude of skill is similar between the number of days and the strongest intensity, the number of events is predicted with less skill, and the extent of significant skill is smaller (Table 1).

Subsurface MHW skill also resembles global patterns of surface MHW skill[2] but is generally higher than or statistically indifferent to surface event skill for all indicators (Fig. 1d–f; Fig. S4; Table 1).
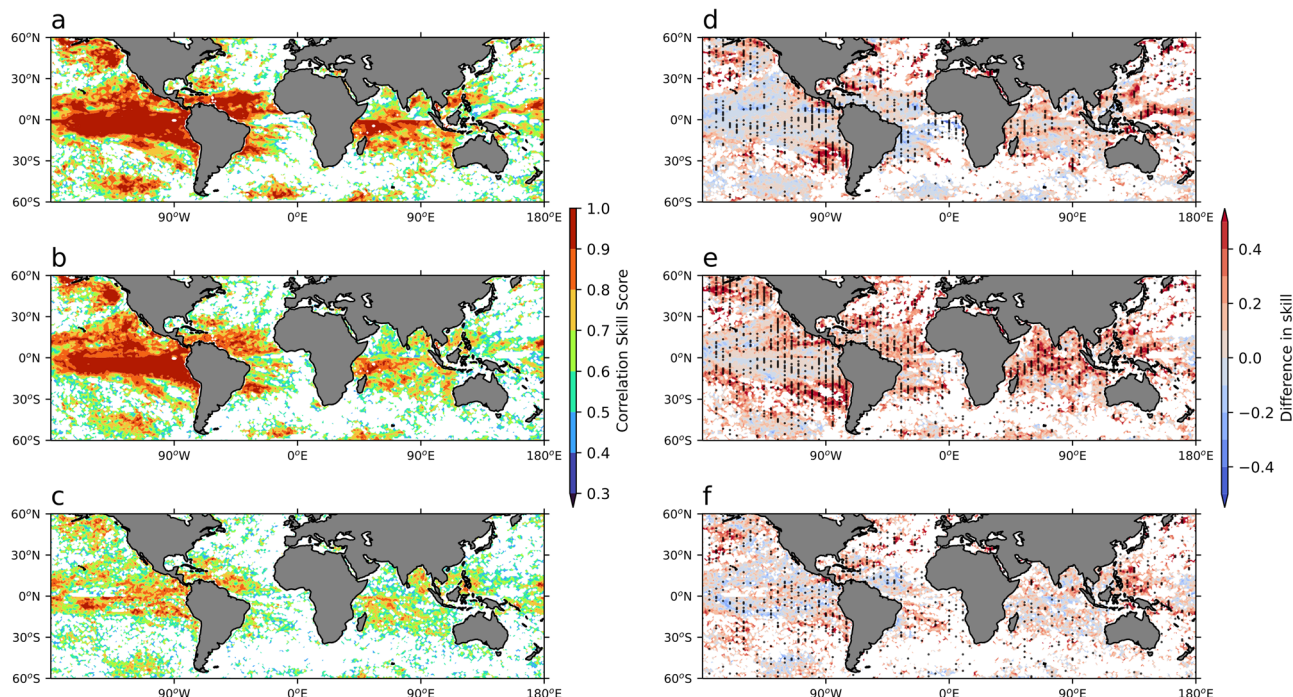
Fig. 1 Improved skill of subsurface MHW Indicators. Correlation skill score of subsurface MHW indicators for (**a**) number of MHW days, (**b**) strongest Intensity, (**c**) number of events. Difference between subsurface and surface MHW indicator skill (**d**) number of MHW days, (**e**) strongest Intensity, and (**f**) number of events. Positive values in (**d–f**) represent improvement in subsurface event detection over surface event detection. All scores correspond to the hemisphere-specific summer season and the 1993–2016 period. Regions in which skill scores in both models are insignificant are masked out (white). Black stippling indicates statistically significant differences in correlation (**d–f**).

Scores between surface and subsurface are similar in regions where surface skill is already high, so the improvements in subsurface skill are clearest away from the tropics. In the tropics, the average difference in skill score (where surface skill is significant) is 0.09 for the number of MHW days, while in the midlatitudes the value is 0.14 (Table 1). While these average values appear small, they represent averages of regions in which skill gain can reach 0.4 and higher (Fig. 1). Strongest intensity is the indicator which displays the greatest increase in skill in the subsurface, globally.

In the tropics, much of the significant skill difference corresponds to a decrease with respect to the surface skill (in the number of days and events). These are regions in which surface MHWs are well predicted, so the decrease in skill does not lead to insignificant subsurface skill (Fig. 1). The most important changes occur in the mid-latitudes, such as in the North Atlantic and the Mediterranean Sea, where subsurface skill is high and surface skill is poor. Thus, as we will discuss later, the use of subsurface indicators opens up new regions to potential benefits of forecast systems.

Explaining why subsurface MHWs are more predictable would involve regional-scale studies on drivers[38] and, specifically, changes in the ocean heat budget (during MHWs) throughout the target layer[33,41]. However, given that there is a widespread improvement across the ocean, we highlight here a general difference between SST and OHC. The inherent persistence of OHC implies it is generally more (theoretically) predictable than surface temperature on seasonal timescales[42,43]. Generally, OHC has lower variability and greater inertia (also known as memory) than SST (Fig. 2a, S5a)[44]; in other words, subsurface conditions tend to persist over longer timescales than surface conditions. The characteristic timescale for OHC 0–40 m ("Methods") falls within the subseasonal time scale (<45 days) only in the most powerful ocean currents and parts of the tropics, while elsewhere the

timescale is seasonal to annual (Fig. 2a, S5a). The upper 40 m is not necessarily shielded from atmospheric variability but has a slower response to it than SST would. Consequently, extremes of OHC are likely to occur less frequently and last longer than extremes of SST, as we show in this study (Fig. S1). A previous study showed there is correlation between forecast skill and event length (at the surface)[2], because longer-lasting events are likely to be picked up by initial conditions. There is indeed evidence to suggest a preconditioned ocean makes it easier to predict MHWs[45]. Here, we show that subsurface events are longer-lasting and predicted with greater skill. It is unclear why the different indicators experience different changes in skill, as all would benefit from the inherent subsurface persistence; the implication is that the strongest intensity of events (which experiences the greater increase in skill from surface to subsurface) is modulated more by inter-annual processes and is less susceptible to shorter-term variability than the other indicators.

To demonstrate how forecasts of subsurface MHWs can complement those of surface MHWs, we now study two regions which have experienced some of the most long-lasting or intense MHWs (with noted ecological impacts) on record[1]: the North Eastern Pacific and the Eastern Mediterranean Sea. These two regions also display significant improvement in subsurface over surface MHWs forecast skill (Fig. 1). In the North-Eastern Pacific, both the surface and subsurface experienced peaks of MHW days between 2013 and 2015, during the multi-year warm anomaly known as the "Blob"[5,6] (FigS. 3a, b, 4). Using SST alone, however, would have underestimated the amount of MHW days in those summers by roughly 50% (Fig. 3a, b, 4). Differences in MHW activity are also reflected in the spatial structure; in this region, the area which experiences MHW activity in the summer is typically lower for the subsurface (Fig. 4, Fig, S6a, b).

In the Eastern Mediterranean Sea, however, the surface and subsurface MHW records do not resemble each other (Fig. 3c, d).
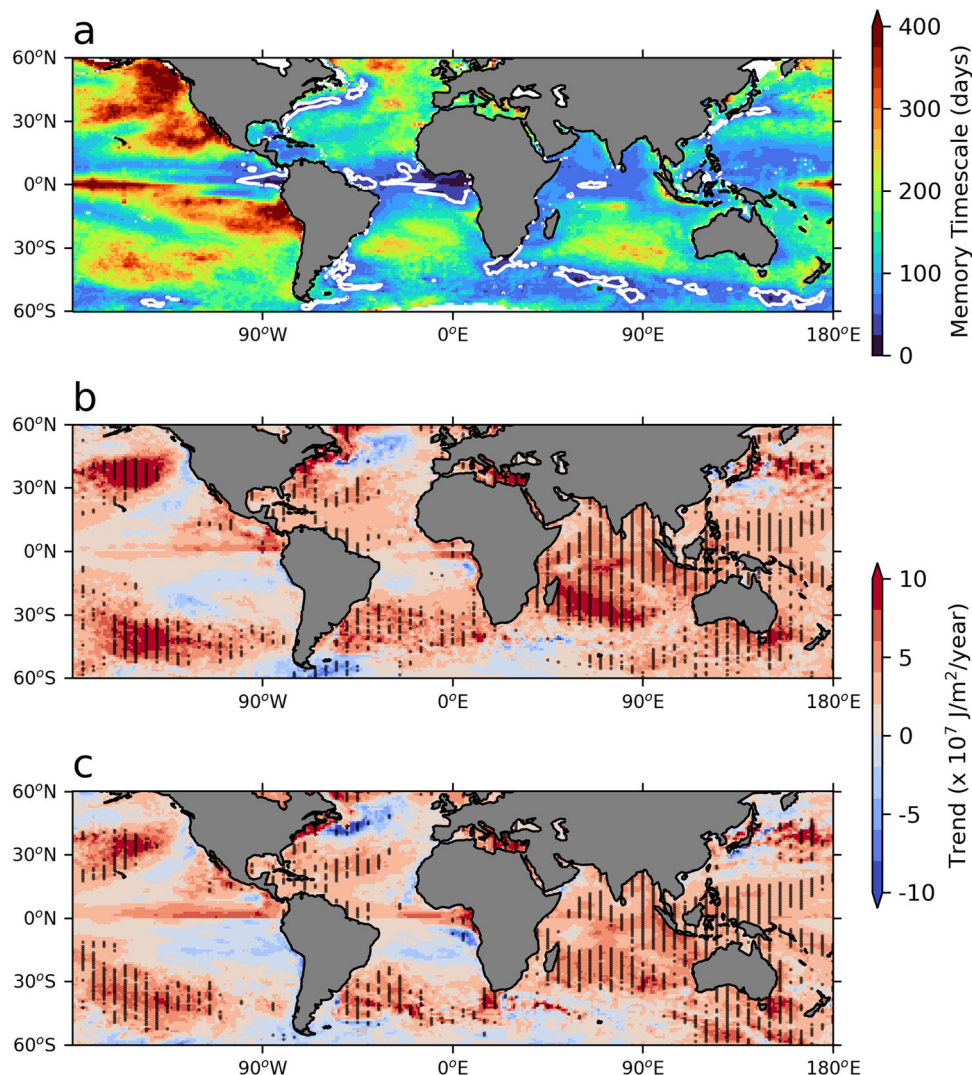
Fig. 2 Characteristic timescales and trends of OHC 0–40 m. (a) Memory timescale (calculated from autocorrelation of OHC; "Methods"). The white contour corresponds to 45 days. Trends of summer average anomalies over 1993–2016 are shown for GREP (**b**) and CMCC-SPS3.5 forecasts (**c**). Black stippling highlights statistically significant trends (using a Mann Kendall test and a p value threshold of 0.05; "Methods").

In 2003 there was a peak in the number of surface MHW days relating to the infamous summer-long event[46,47]. We find it left no signal on the subsurface indicator in the Eastern Mediterra-nean Sea; the subsurface signal of this event is restricted to the North-Western part of the Mediterranean Sea (Fig. 5). Instead, subsurface MHWs activity has been greater in recent years, as evidenced by the record of number of days and the area afflicted by MHWs (Fig. 3c, d, Fig. S6c, d). The notable difference between surface and subsurface extremes in the Eastern Mediterranean Sea is yet to be explored but implies that the influence of the recent warming trend dominates over the natural variability more in the subsurface than at the surface. Thus, not only can the characteristics of events, such as duration, differ between surface and subsurface events, but the occurrence can too. Surface MHW records, therefore, may fail to capture events which impact the subsurface alone.

In the two example regions, forecast skill in CMCC-SPS3.5 is higher for subsurface events than for surface events; the interannual variability of MHW occurrence (denoted by the higher correlation values) is better captured, as is the precision in the value of the number of days (denoted by the reduced error) (Fig. 3). The area affected by MHWs is also captured with greater accuracy in the subsurface (Fig. S6).

Seasonal forecasting systems typically provide an ensemble of forecasts as a means of quantifying the uncertainty of the forecasts, which arises due to error in the initial conditions or in the forecast system ("Methods"). A large ensemble spread (uncertainty) may indicate model failings or the unpredictable nature of particular phenomena; here, we note wide ensemble ranges in particular years (e.g., 2016 in the subsurface Eastern Mediterranean; Fig. 3d). However, the spread of the interquartile range is typically much lower than the range, implying there are often few extreme outliers which contribute to this large range. Reassuringly, the validation dataset values more often than not fall within the interquartile range.

Although subsurface forecast skill is similarly high in both regions, the Eastern Mediterranean Sea surface skill is consider-ably lower than in the North-Eastern Pacific. This is at least partially due to difference in the time scales of dominant drivers of variability in each region. In the North-Eastern Pacific, several inter-annual climate modes, such as ENSO, are known to provide predictability[48], and play a role in MHW development[49]. In the Mediterranean region, summer SST prediction skill on seasonal timescales is low compared to the skill in winter[50]. Surface conditions in the winter are modulated by the North Atlantic Oscillation, providing a source of predictability[51,52],
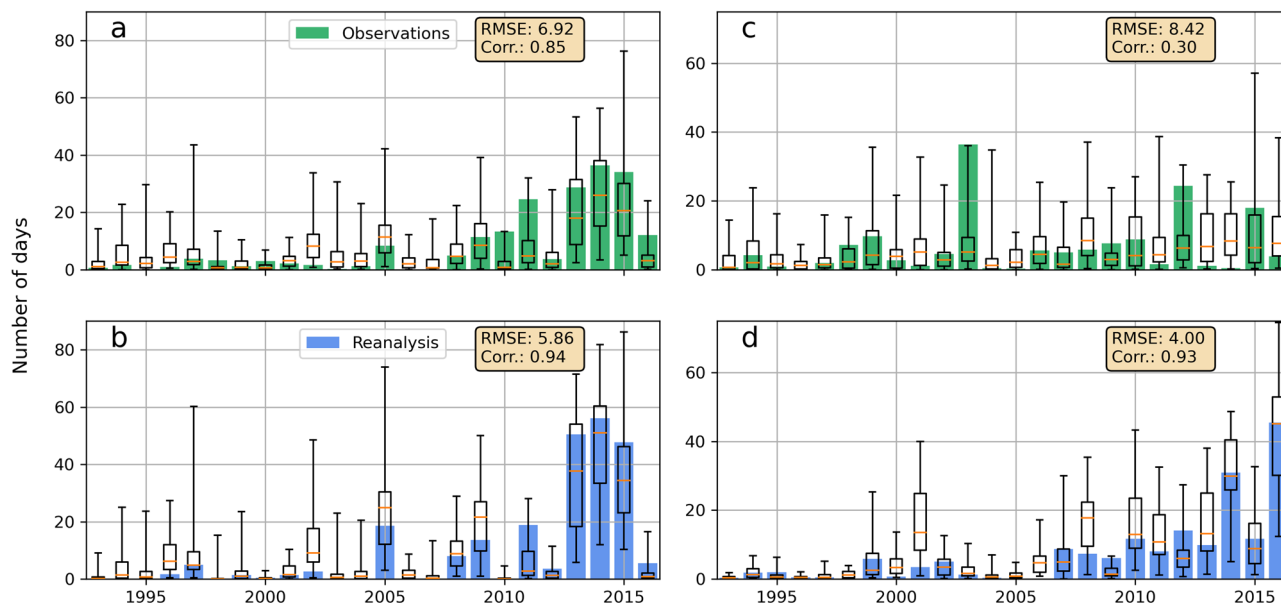
Fig. 3 Record of surface and subsurface MHWs in two regions of high MHW activity. The number of MHW days in the summer season for (**a**) surface North Eastern Pacific, (**b**) subsurface North-Eastern Pacific, (**c**) surface Eastern Mediterranean Sea and (**d**) subsurface Eastern Mediterranean Sea. Box plots represent the median, interquartile range, and range of the 40-member forecast ensemble. Bar charts represent either the surface MHWs in the observations (green) or the subsurface MHWs in the reanalysis (blue). The number of days is calculated as the average number of MHW days in all grid cells within the region, including cells with no MHWs (zero values). Root-mean square error (RMSE) and correlations between forecast ensemble medians and observation/reanalysis values are shown inset. See Figs. 4 and 5 for the definition of the areas used.
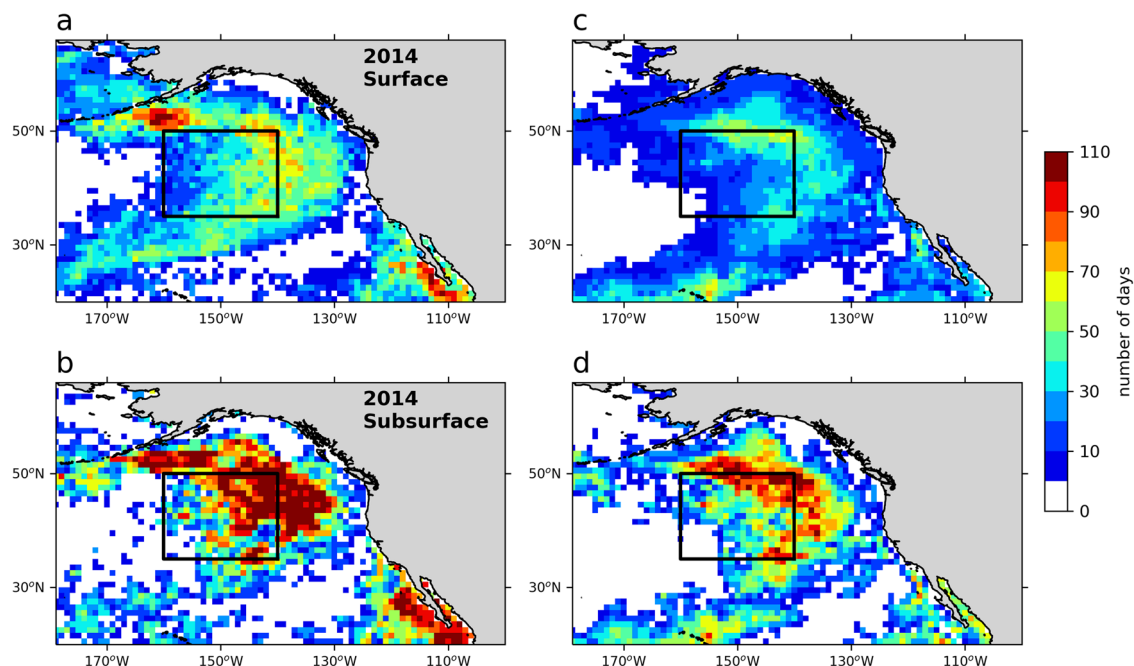

Fig. 4 Observations and forecasts of MHW activity in the North Eastern Pacific in the summer of 2014. Number of MHW days between mid-May and September for (**a**) surface events (ESA CCI satellite-derived SST), (**b**) subsurface events (GREP reanalysis), (**c**) forecast surface events and (**d**) forecast subsurface events (both from CMCC-SPS3.5). The area within the black box is used to calculate the time series in Fig. 1 and Extended Data Fig. S1.

while in the summer variability is typically driven by short-lived, small-scale convective phenomena[53]. It should be pointed out also that the quality and quantity of observations assimilated into the initial conditions in both regions may differ, impacting the forecast skill[54]. However, the skill of SST and OHC anomalies has been found to be similar in both regions (Fig. S3a), where also the reanalysis used here as initial conditions agrees well with other reanalyses and an observation-based product[43]. While there are fundamental differences in drivers of surface variability between the two regions, the subsurface MHW indicator is more accurately predicted in both. In particular, the Eastern Mediterranean Sea is an example of a region of low surface predictability which displays a significant gain in forecast skill with depth.
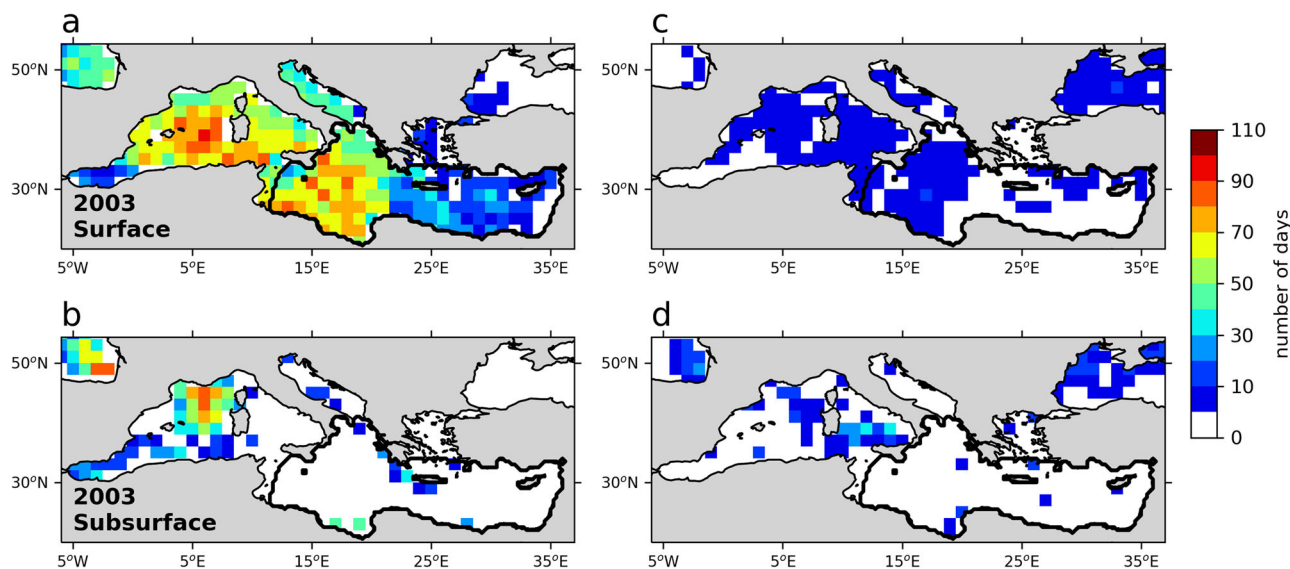
Fig. 5 Observations and forecasts of MHW activity in the Mediterranean Sea in the summer of 2003. Number of MHW days between mid-May and September for (**a**) surface events (ESA CCI satellite-derived SST), (**b**) subsurface events (GREP reanalysis), (**c**) forecast surface events and (**d**) forecast subsurface events (both from CMCC-SPS3.5). The area within the black contour is used to calculate the time series in Fig. 1 and Extended Data Fig. S1.

Identifying sources of forecast skill. We further explore the sources of MHW forecast skill on seasonal timescales while providing suggestions for future validation efforts. We have so far used a "dynamical" system—one which numerically solves equations of ocean state and motion, and therefore requires considerable computational resources. First, by comparing the dynamical system to a simple forecast model which assumes anomalous conditions persist in time, we justify the dynamical approach. Although the precise timing of MHW onset and decay is not explicitly covered in this work, improvement over persistence shows that the system is able to capture the subseasonal variability which determines MHW characteristics. Then, we quantify the extent to which warming trends in ocean heat provide forecast skill on seasonal timescales.

As it is common to test dynamical forecasts systems against computationally cheaper alternatives, we compare them to forecasts based on a persistence model ("Method"). Anomalies at the start of the forecast period are assumed to propagate forward (persist) in time and decay at a rate dependent on the characteristic timescales of OHC 0–40 m decay (see "Methods", Fig. 2a). Previous work on the upper 300 m assumed subsurface heat anomalies persist (with no decay) over seasonal time scales[43]; given that the heat content used in this work is shallower and that subsurface MHWs typically begin and end within seasonal timescales, the use of a decaying persistence model is more appropriate here. Overall, we find that the persistence model skill of MHW indicators is significant in only half of the ocean for the number of MHW days and the strongest intensity and in even less for the number of events (35%; non-masked regions in Fig. 6). In years in which there are several events (i.e., greater variability), persistence of anomalies will more likely underestimate the number of events. The number of MHW days, on the other hand, may be reasonably well predicted because subsurface events are long lasting, explaining why persistence skill is globally more skilful for that indicator (Fig. 6).

In much of the ocean, therefore, the persistence model is not a reasonable means of MHW prediction. In other words, prediction of subsurface MHWs requires forecasts systems to capture variability on subseasonal timescales, despite their longer-lasting nature than surface MHWs (Fig. S1). Globally, the improvement of our dynamical over the decaying persistence model is large

(correlation skill score increases of roughly 0.5, Fig. 6). An increase in skill over persistence is more common in the tropics than in the midlatitudes, as shown also in previous work on deeper OHC anomalies[43], but also greatly dependent on the region. The (significant) increase in skill of the dynamical system occurs in 62%, 80% and 84% of the regions in which there is significant difference between the systems, for the number of MHW days, strongest intensity and number of events respectively (Fig. 6., red regions).

In the equatorial Pacific, skill difference displays abrupt changes across the equator. Recall that each hemisphere covers a different time period, and that the El-Nino Southern Oscillation (ENSO) phenomenon (and therefore the ocean dynamics) are in a different phase in each period. In boreal summer (May-August), ENSO is typically in a decay phase and thus upper ocean heat is not well predicted by persistence. As a result, the dynamical system outperforms the MHW persistence, reaching global maxima in skill difference between the two (Fig. 6a, b). On the other side of the equator, in the austral summer (November-February), ENSO is in a mature and more stable phase. The upper ocean heat content anomalies are relatively stable during this period, explaining why persistence models show weak (but significant) improvements over the dynamical system.

Where anomalies decay slowly, such as in the North-Eastern Pacific and many parts of the tropics (Fig. 2a), and where long-lasting MHWs have been known to occur, it is reasonable to expect a persistence model to perform well on seasonal timescales. Nonetheless, in these regions, we highlight significant increase in skill in the dynamical system, thereby justifying its use. Elsewhere, we find that the dynamical system, in regions of relatively low skill (e.g., the North-Eastern Atlantic) does not outperform persistence. The Eastern Mediterranean Sea, for example, shows nearly no statistically significant difference in skill. Similarly poor increase in skill over persistence is found for other oceanic variables in other studies, suggesting this is not a problem related specifically to the representation of MHWs but instead indicates more general issues with seasonal forecasting of the ocean (discussed in the next section)[55].

A key influence on the occurrence and characteristics of MHWs is trends in temperature and heat in the ocean[56]. Such trends are introduced to forecast systems mainly through the initial
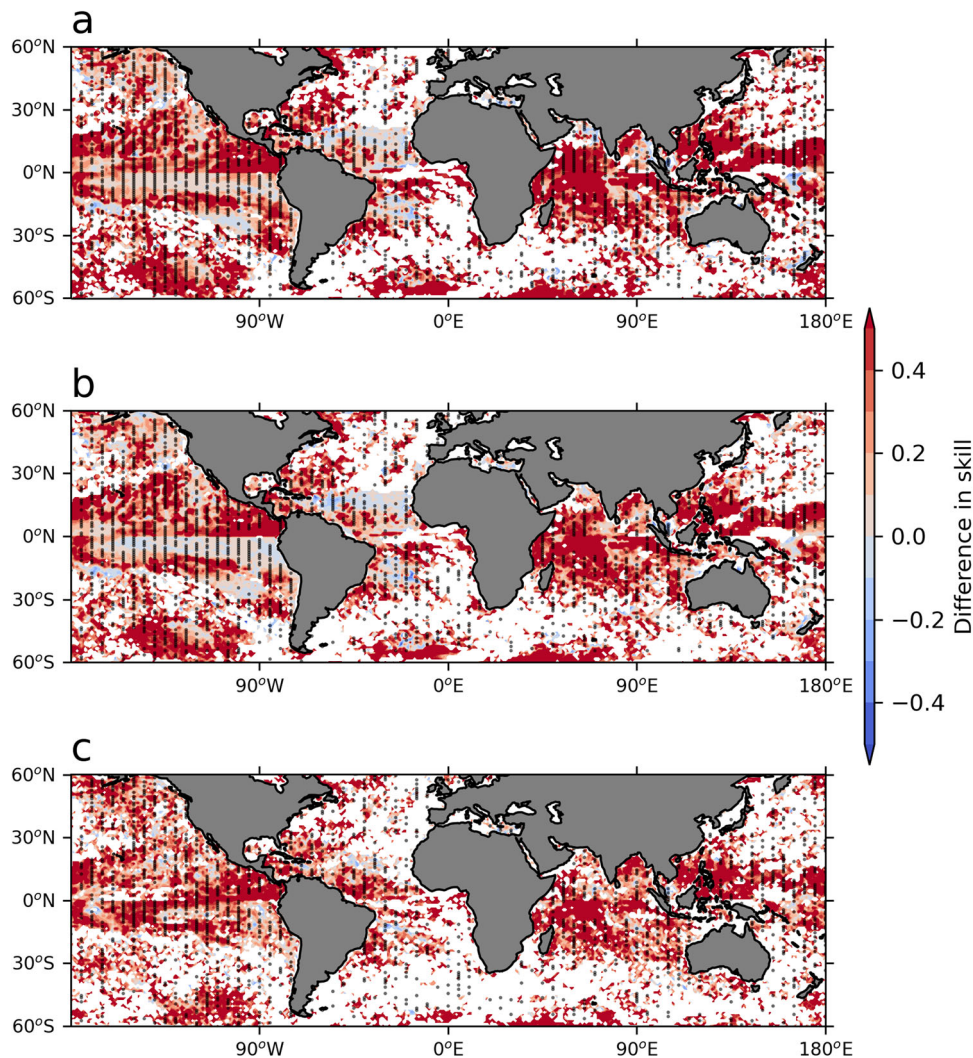
Fig. 6 Improvement of dynamical system subsurface MHW prediction over a persistence model. Difference between dynamical system skill scores (as seen in Fig. 4a–c) and persistence model skill scores in (**a**) Number of MHW days, (**b**) strongest intensity and (**c**) number of events. Positive values represent higher skill in the dynamical system. All scores correspond to the hemisphere-specific summer season and the 1993–2016 period. Regions in which skill scores in both models are insignificant are masked out (white). Black stippling indicates statistically significant differences in correlation.

conditions and also act as a source of forecast skill. Trends (and interannual variability) in SST and subsurface heat of the reanalysis used to initialise CMCC-SPS3.5, as well as the GREP reanalysis used as a validation dataset in this study, have been compared to other reanalyses and in-situ data, and are assumed to be accurate[32]. It is important to note the heterogeneity in summertime trends in OHC 0–40 m over the period used here (1993–2016) (Fig. 2b, c). Negative values are found in the transition between subtropical and subpolar gyres in the North Atlantic and the southern Pacific and Atlantic tropics. Positive trends are only significant in the gyre centres, the Eastern Mediterranean Sea and across the Indian Ocean (Fig. 2b, c). The influence of trends on forecast skill (over 1993–2016) is therefore expected to be confined to these specific regions. To confirm this, we removed the linear trend in OHC from both validation and forecast data ("Methods"), then recalculated MHWs with the detrended time series. The skill which arises from the inclusion of trends is shown in Fig. 7. The largest significant improvements in skill are found where the trends themselves are largest and significant: the Indian Ocean, North Eastern Pacific, Eastern Mediterranean and gyre centres (Fig. 2b, c, 7). A curious exception is the Equatorial Pacific in the Southern Hemisphere, where trends

are weak in both the validation and forecast data but also where there is a significant increase in skill when including trends. The austral summer period corresponds to the mature phase of ENSO, which is apparently sensitive to slight trends.

The influence of a trend on MHW forecast skill is clear at a few specific regions globally and is similar in the surface and subsurface (Fig. S6, S7). We find that the significant differences in trend correspond to increased skill in the detrended data. Finally, we note that by applying the same detrending analysis to subsurface anomalies, instead of MHW indicators, we find that the inclusion of a trend makes little difference to forecast skill even in regions of significant trends (Fig. S3b, c). The global patterns of skill difference for anomalies are similar to the global patterns of skill difference for MHW indicators, but the values are not significant. Here, therefore, we confirm where extreme event forecast skill appears to be particularly sensitive to warming trends.

Considerations for future forecasting efforts. In this study, we demonstrate that seasonal indicators of subsurface MHWs are predicted with greater accuracy than the more commonly-used surface-focused indicators. Economic and ecological impacts of MHWs, which could be mitigated with early warning, mostly
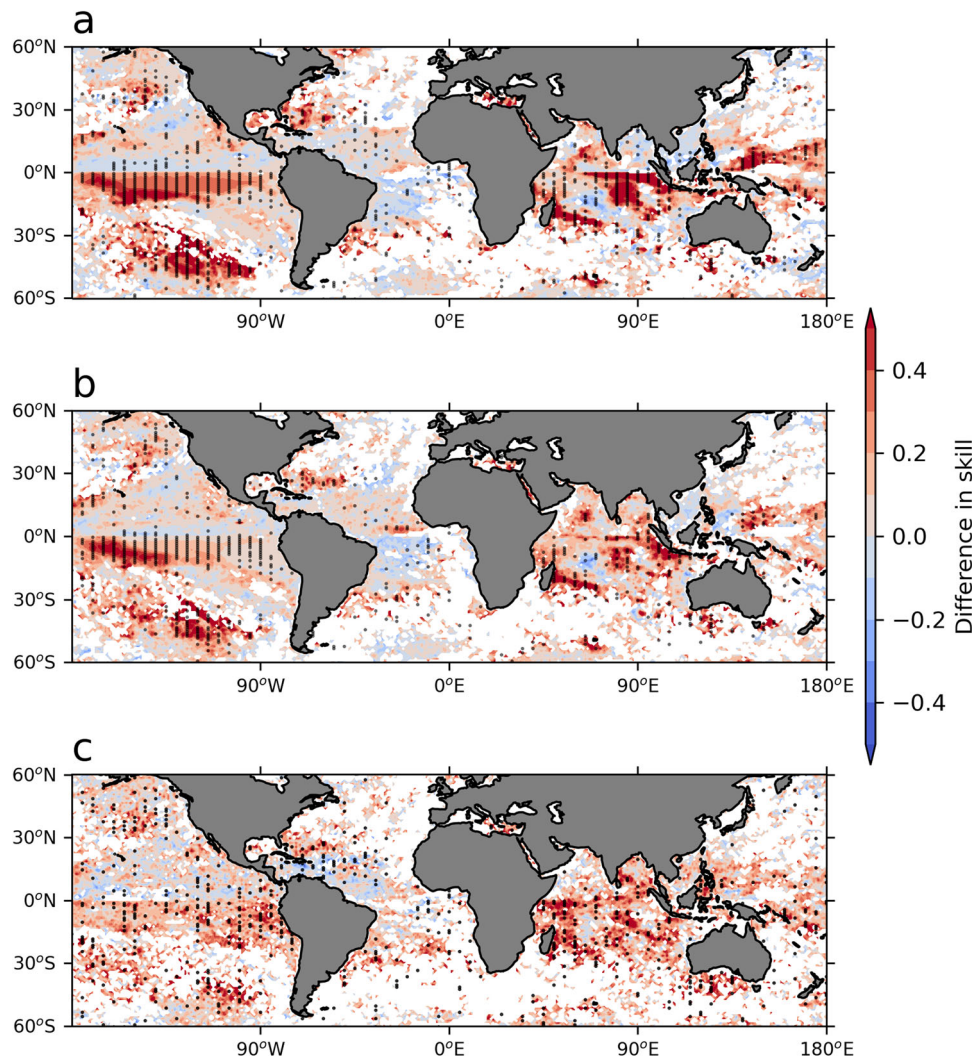
Fig. 7 Contribution of linear trends to subsurface MHW forecast skill. Difference between skill scores of original forecast output (Fig. 4a–c) and detrended forecast output for (**a**) Number of MHW days, (**b**) strongest intensity (**c**) number of events. Positive values represent higher skill in the MHW reforecasts when the linear trend is included. All scores correspond to the hemisphere-specific summer season and the 1993–2016 period. Regions in which skill scores in both models are insignificant are masked out (white). Black stippling indicates statistically significant differences in correlation.

occur below the surface, so we believe this work highlights the great potential for subsurface seasonal forecasting. During the process of creating forecast tools for marine stakeholders, regions in which surface skill is inadequate may be ruled out; however, using subsurface indicators leads to an increase in the number of regions, and therefore the number of potential stakeholders, which could benefit from tools based on seasonal forecasts.

On the time scales studied here, subsurface skill resembles surface skill which has previously been shown to be strongly driven by the largest modes of climate variability (e.g., ENSO)[2]. Here, we show that the greater predictability of subsurface extremes is found to be partly due to an increase in persistence of anomalies with depth, with warming trends playing a role in few specific regions. Nonetheless, the dynamical system used here outperforms a cheaper persistence model in large parts of the ocean. Where the dynamical system underperforms indicates where there is a need for fundamental improvements of forecasting systems, such as an increase in resolution and the increase of subsurface data to be assimilated into initial conditions. Such improvements might increase the validity of the current state-of-the-art of seasonal forecasts to be used for early-warning systems of extremes. A choice can be made by

providers and users of forecasts on whether to use detrended time series or not; doing so may increase skill in some regions but will not provide the "true" conditions in a warming ocean. Here, we highlight that the trade-off between the two choices is similar for surface and subsurface MHW indicators.

This study quantifies the skill of three user-relevant seasonal-aggregate indicators: the number of MHW days, the strongest intensity experienced and the number of MHW events. The first two are relevant for understanding whether MHW conditions push species beyond their thermal resilience, while the latter determines whether species have a suitable recovery time between events. Resilience and recovery time are of course species-dependent and, in practice, the combination of these indicators determines species response; for example, "degree heating days" (a product of intensity and duration) is used to indicate the extent of coral reef bleaching[38]. Here we show forecasts of MHW characteristics made 4 months in advance. There are many potential early warning systems that could be based on these indicators[10] in a co-development process with stakeholders (specifically linking physical conditions to impacts). For example, seasonal forecasts of MHW activity could provide fisheries with early warnings of reduced stocks, providing them time to cut

costs by reducing or redirecting fishing efforts on a season-by-season basis[57]. Early warning of MHWs could provide a proxy for disease outbreaks for aquaculture farms, which can help them decide whether to harvest early before a mass die-off[58]. Crucially, any such system should include subsurface indicators, in order to better represent the conditions experienced by much of the marine ecosystem (which is not constrained to the sea surface).

There are still several scientific and practical steps to take before making broadly useful operational MHW forecasts. First, the impact of MHWs on species depends also on their thermal resilience, which may change in response to extreme heat events and gradual climate change[8] – in other words, with a moving climatology[1]. MHWs can be defined with respect to a moving climatology if the data used covers a period of several decades, longer than the historical re-forecast output of most seasonal forecasting systems. CMCC-SPS3.5, for example, is constrained to a 24 -year baseline period, shorter than recommended although shorter periods have been used where necessary i.e., for observations with short temporal coverage[1]. Indeed, another major caveat to this study is that it uses only one forecast system. We wish to promote making subsurface ocean forecast data available, in order to foster the development of stakeholder-relevant applications. For surface extremes, the seasonal forecast data required to produce these indicators is already made freely available for a range of systems (e.g., through the Copernicus Climate Change Service). Unfortunately, ocean heat content and subsurface temperatures are not yet provided by seasonal forecasting centres, at least not openly. By showing that subsurface MHWs are predicted with greater skill than surface MHWs and that a dynamical system outperforms a simple persistence model, this study further motivates the use for subsurface indicators of MHW for seasonal forecasting.

Methods

We focus on summer marine heatwaves (MHWs) defined with hemisphere-specific definitions; those which take place over a 3.5-month period from mid-May to end of August in the Northern Hemisphere or from mid-November to the end of February in the Southern Hemisphere. All figures pertaining to forecast skill and trends (e.g., Fig. 1) therefore display information on different seasons in each hemisphere. Forecast start times are May 1st and November 1st of each year from 1993 to 2016. We remove the first two weeks of the forecast period, in line with previous studies on land-based and marine heatwaves[59], to focus on the forecast skill on seasonal timescales. For time scales shorter than 2 weeks, there exist dynamical forecasting systems designed specifically for the task, which benefit from a higher spatial resolution and daily initialisation.

The seasonal forecast system used is the CMCC Seasonal Prediction System version 3.5 (CMCC-SPS3.5)[29]. CMCC-SPS3.5 is a fully-coupled ocean-atmosphere-land-river-sea ice model which provides operational forecasts to the Copernicus Climate Change Service (C3S). The ocean component is a configuration of NEMO with an eddy-permitting horizontal resolution of 0.25°, and 50 vertical levels, but the version made available on C3S is an interpolation from the native NEMO grid to a regular 1° horizontal grid. Ocean-atmosphere coupling occurs every 90 min (three model time steps). Initial conditions are drawn from a version of the C-GLORS ocean reanalysis[60]. 40 ensemble members are produced for the hindcast output, representing error within the initial conditions.

To validate surface MHWs in CMCC-SPS3.5, we employ version 2.1 of the European Space Agency Climate Change Initiative (ESA CCI) satellite-derived L4 sea surface temperature (SST) record. It is constructed of data from several radiometers orbiting (not simultaneously) since 1981 and is provided at global 0.05° resolution[30,31]. To validate the subsurface MHWs, the Global ocean Reanalyses Product (GREP) is used[32]. GREP is a four-member ensemble of 4D NEMO-based reanalyses with an eddy-permitting horizontal resolution of 0.25°, which differ in their data assimilation methods and in the data assimilated. Although higher resolution reanalysis exist[61], the quality of GREP is comparable to that of observation-only products and it also has the advantage of providing the ensemble mean which, due to the cancellation of systematic errors, is more accurate than the individual members[32]. Moreover, GREP and CMCC-SP3.5 share the same horizontal resolution and therefore aim to simulate the same spatial variability, meaning GREP is a fairer benchmark. GREP has already been used in validation of seasonal forecasts of OHC 0–300 m; ensemble members are known to disagree most in near-coastal regions and around sharp fronts, thereby indicating where GREP's use as a validation tool is less reliable[43]. The suitability of GREP for studying MHWs is discussed below. Both validation datasets are freely available on

the Copernicus Marine Service, and in this study are interpolated to a regular 1° grid to be consistent with the re-forecasts provided by the C3S.

We computed the daily ocean heat content (OHC) of the upper 40 m using the following calculation:

$$c_p \rho \int_{z1=0\,m}^{z2=40\,m} T(z)dz \qquad (1)$$

in which c_p is the specific heat capacity of seawater (3996 J/(kg. C)), ρ is seawater density (1026 kg/m³), and T is the temperature at depth z. In the validation reanalysis and the forecasts system, there are 37 and 17 vertical levels in the upper 40 m, respectively. The respective depths of the uppermost level are 1.02 m and 0.49 m.

To our knowledge, the CMCC is the only seasonal forecasting centre which outputs shallow, near-surface OHC (e.g., 0–40 m). Here, the data availability covers 1993–2016. In many MHW studies, a period of at least 30 years is favoured. As marine variables in seasonal forecasting systems become more widely used and appreciated, we hope in the future to provide a more rigorous validation using a longer time period and re-forecasts from other seasonal forecasting centres.

MHWs are defined here based on the widely used statistical definition[62]. In this framework, MHWs occur when temperature or ocean heat content exceeds the 90th percentile for 5 or more days, allowing for gaps of less than 2 days. Here, climatologies and the 90th percentiles are calculated in two steps. First, an 11-day moving window is used to calculate the daily averages. Then, a polynomial fit is required to avoid discontinuities at the edge of the forecast period[63]. For the forecast data, the MHW detection algorithm is applied individually to all forecast ensemble members over the 6-month forecast period each year. The algorithm is applied to daily time series of SST and OHC 0–40 m. A list of MHW start, peak and end dates, intensities and durations is provided for each year and, in the case of forecast data, each ensemble member.

The intensity of events defined by SST and OHC will have different units and different background variability. Here, we normalise the intensity by the width between the (daily) average value and 90th percentile:

$$I_n = \frac{T - T_{90}}{T_{90} - T_c} \qquad (2)$$

in which T is the daily SST or OHC time series, and T_c and T_{90} are the corresponding daily climatology and 90th-percentile respectively. Since $T_{90} - T_c$ is dependent on position and time, a fixed anomaly (e.g., 2ºC) can lead to a different MHW intensity in different places and times. By using a normalised intensity, we not only remove the units but we also benefit from a fairer means of comparing the exceptionality between surface and subsurface events. This method is a variation on the MHW "category" metric[64]. In this study, "intensity" refers to the normalised intensity. MHW indicators for each grid cell are then built from duration and intensity and serve as the validation metrics. The event with the maximum intensity in the summer period provides the "Strongest Intensity", while the total number of MHW days in the summer period is added together to provide the "Number of MHW days". In the forecast data, MHWs characteristics are calculated for all ensemble members. Then, the ensemble median is calculated and used for the correlation scoring (e.g., Fig. 1).

To justify the use of GREP as a validation tool for MHW forecasts, we compare MHW characteristics at the surface in the satellite observations and GREP. At the surface, GREP generally underestimates the number and intensity of summer MHWs while overestimating the duration when compared to the satellite data (Fig. S1). These results agree with other works on inter-comparisons between model resolutions[45,65]. The average values between 60ºS and 60ºN for satellite observations, GREP SST and GREP OHC are, respectively, as follows: 17.5, 15.7 and 13.2 number of events; 10.9, 13.3, 17.2 days duration; 0.28, 0.25, 0.19 intensity. Regardless of which SST is used as a reference point, subsurface event characteristics are noticeably different to those at the surface (Fig. S1).

GREP's ability to capture interannual variability of MHW activity must also be compared, because the correlation score used in this study (e.g., in Fig. 1) quantifies the ability of the forecasts to capture the year-to-year variability of MHW characteristics. The correlation between GREP and satellite data is significantly high for all three MHW characteristics used in this study, in nearly all points in the global ocean (Fig. S2). In particular, the variability in the number of MHW days in summer is almost perfectly captured across the ocean. Therefore, we find that GREP can be considered a reliable validation dataset for interannual variability of MHW characteristics.

Here, we follow the approach of using seasonal forecasting indicators of MHW propensity i.e., how predisposed a season is to MHW activity[59]. Validation of the ability to predict MHW occurrence (i.e., whether a MHW happens or not) has previously been performed for surface MHWs[2]. Here, we test indicators which can provide further details on MHWs than their occurrence (number of MHW days, strongest intensity and number of events). We believe these indicators provide further detail and therefore may be of more benefit to potential stakeholders, as well as to those who wish to understand the dynamical capability of forecasting systems (see Discussion section). Skill is quantified here with the Pearson correlation coefficient of the validation datasets and the forecast system (see Statistics section for further detail).

The persistence model is built from anomalies in the validation datasets and the characteristic timescales of OHC 0–40 m. First, the average OHC 0–40 m anomaly

of the 5-day period prior to forecast start time ($t = -5$ to $t = -1$) is calculated (on a cell-by-cell basis). 5-days is chosen as it corresponds to the minimum duration of a MHW. Then, this anomaly is propagated forward throughout the forecast period, and an exponential decay is applied with a decay factor equal to the characteristic timescale. The timescale is estimated to first-order as follows:

$$\tau = \frac{-1}{\ln a} \tag{3}$$

where a is the autocorrelation of daily OHC 0–40 m timeseries, and τ is often referred to as ocean memory[1,44]. As in other seasonal forecast validation efforts, the characteristic timescale is assumed to be the typical decay time for anomalies[55]. Crucially, the decay factor is location-dependent; for example, anomalies will decay more slowly in the North East Pacific than in the tropics as the memory in the former is much longer (Fig. 2a, Fig. S6a). MHWs in the persistence model are then calculated using the previously-calculated climatology and 90th-percentile from the corresponding validation dataset. In practice, due to the timescales of OHC memory typically extending beyond seasonal timescales (Fig. 2a), results are similar when no decay is used (Fig. S8).

Lastly, a detrended time series is created for SST and OHC 0–40 m, in both the validation and forecast datasets. The 1993–2016 trend is calculated on a cell-by-cell level for each day of the forecast period and removed from the data. The trends removed are those corresponding to the dataset (i.e., there are different trends for forecast and validation data). Trends for the summer average anomalies are shown in Fig. 5 and Supplementary Fig. S6, to give an indication of the trends removed from the daily time series. MHWs are then recalculated with the detrended data (Fig. 7, Fig. S7). Given the short length of time period used here, the trends do not accurately represent long-term changes (e.g., anthropogenic climate change); instead, this analysis serves the purposes of identifying sources of forecast skill.

Statistics. Statistical significance of correlations is calculated using the two-sided test included in the *stats.pearsonr* function from the Python module *scipy*. Appropriate tests for statistical significance of differences between correlation skill scores (e.g., Fig. 1) are taken from Siegert et al (2017)[66]. We use tests for overlapping correlations when comparing the dynamical systems and the persistence model (because they are both compared to the same validation dataset), and tests for independent correlations when comparing surface and subsurface skill or when comparing MHW indicators defined by fixed and detrended climatologies (because they have different and independent validation datasets). The p value threshold for significance is 0.05 throughout the paper. In all cases, the samples size equals 24, the number of available re-forecast years.

Data availability

Code availability

References

1. Oliver, E. C. J. et al. Marine heatwaves. *Annu. Rev. Mar. Sci.* **13**, 313–342 (2021).
2. Jacox, M. G. et al. Global seasonal forecasts of marine heatwaves. *Nature* **604**, 486–490 (2022).
3. Benthuysen, J. A., Smith, G. A., Spillman, C. M. & Steinberg, C. R. Subseasonal prediction of the 2020 Great Barrier Reef and Coral Sea marine heatwave. *Environ. Res. Lett.* **16**, 124050 (2021).
4. Spillman, C. M., Smith, G. A., Hobday, A. J. & Hartog, J. R. Onset and decline rates of marine heatwaves: global trends, seasonal forecasts and marine management. *Front. Clim.* **3**, (2021).
5. Bond, N. A., Cronin, M. F., Freeland, H., & Mantua, N. Causes and impacts of the 2014 warm anomaly in the NE Pacific. *Geophys Res. Lett.* **42**, 3414–3420 (2015).
6. Cavole, L. M. et al. Biological impacts of the 2013–2015 warm-water anomaly in the Northeast Pacific: Winners, losers, and the future. *Oceanography* **29**, 273–285 (2016).
7. Wernberg, T. et al. An extreme climatic event alters marine ecosystem structure in a global biodiversity hotspot. *Nat. Clim. Chang.* **3**, 78–82 (2013).
8. Garrabou, J. et al. Marine heatwaves drive recurrent mass mortalities in the Mediterranean Sea. *Glob. Change Biol.* **28**, 5708–5725 (2022).
9. Caputi, N. et al. Management adaptation of invertebrate fisheries to an extreme marine heat wave event at a global warming hot spot. *Ecol. Evol.* **6**, 3583–3593 (2016).
10. Smith, K. E. et al. Socioeconomic impacts of marine heatwaves: Global issues and opportunities. *Science (1979)* **374**, eabj3593 (2021).
11. Smith, K. E. et al. Biological impacts of marine heatwaves. *Annu. Rev. Mar. Sci.* (2022) https://doi.org/10.1146/annurev-marine-032122-121437.
12. Garrabou, J. et al. Collaborative database to track mass mortality events in the Mediterranean sea. *Front. Mar. Sci.* **6**, 707 (2019).
13. van Haren, H. & Compton, T. J. Diel vertical migration in deep sea plankton is finely tuned to latitudinal and seasonal day length. *PLoS One* **8**, e64435 (2013).
14. Hays, G. C. A review of the adaptive significance and ecosystem consequences of zooplankton diel vertical migrations. in *Migrations and Dispersal of Marine Organisms* (eds. Jones, M. B. et al.) 163–170 (Springer Netherlands, 2003).
15. Andrzejaczek, S., Gleiss, A. C., Pattiaratchi, C. B. & Meekan, M. G. Patterns and drivers of vertical movements of the large fishes of the epipelagic. *Rev. Fish Biol. Fish* **29**, 335–354 (2019).
16. Lavender, E. et al. Environmental cycles and individual variation in the vertical movements of a benthic elasmobranch. *Mar. Biol.* **168**, 164 (2021).
17. Brito-Morales, I. et al. Towards climate-smart, three-dimensional protected areas for biodiversity conservation in the high seas. *Nat. Clim. Chang.* **12**, 402–407 (2022).
18. Jacox, M. G. et al. Seasonal-to-interannual prediction of North American coastal marine ecosystems: Forecast methods, mechanisms of predictability, and priority developments. *Prog. Oceanogr.* **183**, 102307 (2020).
19. Liu, G. et al. Predicting heat stress to inform reef management: NOAA coral reef watch's 4-month coral bleaching outlook. *Front Mar Sci* **5**, 57 (2018).
20. Stevens, C. L. et al. Horizon scan on the benefits of ocean seasonal forecasting in a future of increasing marine heatwaves for Aotearoa New Zealand. *Front. Clim.* **4**, (2022).
21. Hobday, A. J., Spillman, C. M., Paige Eveson, J. & Hartog, J. R. Seasonal forecasting for decision support in marine fisheries and aquaculture. *Fish. Oceanogr.* **25**, 45–56 (2016).
22. Goodess, C. M. et al. The value-add of tailored seasonal forecast information for industry decision making. *Climate* **10**, 152 (2022).
23. Widlansky, M. J. et al. Multimodel ensemble sea level forecasts for tropical pacific islands. *J. Appl. Meteorol. Clim.* **56**, 849–862 (2017).
24. Brodie, S. et al. Seasonal forecasting of dolphinfish distribution in eastern Australia to aid recreational fishers and managers. *Deep Sea Res. Part II: Topical Stud. Oceanogr.* **140**, 222–229 (2017).
25. Tommasi, D. et al. Managing living marine resources in a dynamic environment: The role of seasonal to decadal climate forecasts. *Prog. Oceanogr.* **152**, 15–49 (2017).
26. Spillman, C. M. Operational real-time seasonal forecasts for coral reef management. *J. Oper. Oceanogr.* **4**, 13–22 (2011).
27. Cheung, W. W. L. & Frölicher, T. L. Marine heatwaves exacerbate climate change impacts for fisheries in the northeast Pacific. *Sci. Rep.* **10**, 6678 (2020).
28. Gamperl, A. K., Zrini, Z. A. & Sandrelli, R. M. Atlantic Salmon (Salmo salar) cage-site distribution, behavior, and physiology during a newfoundland heat wave. *Front Physiol* **12**, 719594 (2021).
29. Gualdi, S. et al. The new CMCC Operational seasonal prediction system. *CentroEuro-Mediterraneo sui Cambiamenti Climatici. CMCC Tech. Rep.* 61pp (2020) https://doi.org/10.25424/CMCC/SPS3.5.
30. Good S. A.; Embury, O.; B. C. E.; M. J. ESA Sea Surface Temperature Climate Change Initiative (SST_cci): Level 4 Analysis Climate Data Record, version 2.1. *Centre for Environmental Data Analysis, 22 August 2019* (2019) https://doi.org/10.5285/62c0f97b1eac4e0197a674870afe1ee6.
31. Merchant, C. J. et al. Satellite-based time-series of sea-surface temperature since 1981 for climate applications. *Sci. Data* **6**, 223 (2019).
32. Storto, A. et al. 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. *Clim. Dyn.* **53**, 287–312 (2019).
33. Marin, M., Feng, M., Bindoff, N. L. & Phillips, H. E. Local drivers of extreme upper ocean marine heatwaves assessed using a global ocean circulation model. *Front. Clim.* **4**, 788390 (2022).
34. Galli, G., Solidoro, C. & Lovato, T. Marine heat waves hazard 3D maps and the risk for low motility organisms in a warming Mediterranean Sea. *Front. Mar. Sci.* **4**, 136 (2017).
35. Kerry, C., Roughan, M. & Azevedo Correia de Souza, J. M. Drivers of upper ocean heat content extremes around New Zealand revealed by Adjoint Sensitivity Analysis. *Front. Clim.* **4**, (2022).
36. Tacon, A. G. J. & Halwart, M. Cage aquaculture: a global overview. *FAO Fish. Tech. Pap.* **498**, 3 (2007).

37. Holte, J., Talley, L. D., Gilson, J. & Roemmich, D. An Argo mixed layer climatology and database. *Geophys. Res. Lett.* **44**, 5618–5626 (2017).

38. Wyatt, A. S. J. et al. Hidden heatwaves and severe coral bleaching linked to mesoscale eddies and thermocline dynamics. *Nat. Commun* **14**, 25 (2023).

39. Großelindemann, H., Ryan, S., Ummenhofer, C. C., Martin, T. & Biastoch, A. Marine heatwaves and their depth structures on the Northeast U.S. Continental Shelf. *Front. Clim.* **4**, (2022).

40. Juza, M., Fernández-Mora, À. & Tintoré, J. Sub-regional marine heat waves in the Mediterranean sea from observations: Long-term surface changes, sub-surface and coastal responses. *Front. Mar. Sci.* **9**, 785771 (2022).

41. Sen Gupta, A. et al. Drivers and impacts of the most extreme marine heatwave events. *Sci. Rep.* **10**, 19359 (2020).

42. Behrens, E., Fernandez, D. & Sutton, P. Meridional oceanic heat transport influences marine heatwaves in the tasman sea on interannual to decadal timescales. *Front. Mar. Sci.* **6**, 228 (2019).

43. McAdam, R. et al. Seasonal forecast skill of upper-ocean heat content in coupled high-resolution systems. *Clim. Dyn.* **58**, 3335–3350 (2022).

44. Shi, H. et al. Global decline in ocean memory over the 21st century. *Sci. Adv.* **8**, eabm3468 (2023).

45. Von Schuckmann, K. et al. Copernicus Marine service ocean state report, issue 6. *J. Operational Oceanography* **15**, 1–120 (2022).

46. Ibrahim, O., Mohamed, B. & Nagy, H. Spatial variability and trends of marine heat waves in the Eastern Mediterranean Sea over 39 years. *J. Marine Sci. Eng.* **9**, 643 (2021).

47. Darmaraki, S., Somot, S., Sevault, F. & Nabat, P. Past variability of Mediterranean sea marine heatwaves. *Geophys. Res. Lett.* **46**, 9813–9823 (2019).

48. Emery, W. J. & Hamilton, K. Atmospheric forcing of interannual variability in the northeast Pacific Ocean: Connections with El Niño. *J. Geophys. Res. Oceans* **90**, 857–868 (1985).

49. Holbrook, N. J. et al. A global assessment of marine heatwaves and their drivers. *Nat. Commun* **10**, 2624 (2019).

50. Calì Quaglia, F., Terzago, S. & von Hardenberg, J. Temperature and precipitation seasonal forecasts over the Mediterranean region: added value compared to simple forecasting methods. *Clim. Dyn.* **58**, 2167–2191 (2022).

51. Dunstone, N. et al. Skilful predictions of the winter North Atlantic Oscillation one year ahead. *Nat. Geosci.* **9**, 809–814 (2016).

52. Scaife, A. A. et al. Skillful long-range prediction of European and North American winters. *Geophys. Res. Lett.* **41**, 2514–2519 (2014).

53. Ardilouze, C. et al. Multi-model assessment of the impact of soil moisture initialization on mid-latitude summer predictability. *Clim. Dyn.* **49**, 3959–3974 (2017).

54. Balmaseda, M. A. et al. Ocean initialization for seasonal forecasts. *Oceanography* **22**, 154–159 (2009).

55. Long, X. et al. Seasonal forecasting skill of sea-level anomalies in a multi-model prediction framework. *J. Geophys. Res. Oceans* **126**, e2020JC017060 (2021).

56. Barkhordarian, A., Nielsen, D. M. & Baehr, J. Recent marine heatwaves in the North Pacific warming pool can be attributed to rising atmospheric levels of greenhouse gases. *Commun Earth Environ.* **3**, 131 (2022).

57. Mills, K. E. et al. Lessons from the 2012 Ocean Heat Wave in the Northwest Atlantic. *Oceanography* **26**, 191–195 (2013).

58. Lebel, L. et al. Innovation, Practice, and Adaptation to Climate in the Aquaculture Sector. *Rev. Fisher. Sci. Aquacult.* **29**, 721–738 (2021).

59. Prodhomme, C. et al. Seasonal prediction of European summer heatwaves. *Clim. Dyn.* **58**, 2149–2166 (2022).

60. Storto, A. & Masina, S. C-GLORSv5: an improved multipurpose global ocean eddy-permitting physical reanalysis. *Earth Syst. Sci. Data* **8**, 679–696 (2016).

61. Lellouche, JL et al. The Copernicus Global 1/12° Oceanic and Sea Ice GLORYS12 Reanalysis. *Front. Earth Sci. (Lausanne)* **9**, 698876 (2021).

62. Hobday, A. J. et al. A hierarchical approach to defining marine heatwaves. *Prog Oceanogr* **141**, 227–238 (2016).

63. Mahlstein, I., Spirig, C., Liniger, M. A. & Appenzeller, C. Estimating daily climatologies for climate indices derived from climate model data and observations. *J. Geophys. Res. Atmos.* **120**, 2808–2818 (2015).

64. Hobday, A. J. et al. Categorizing and naming marine heatwaves. *Oceanography* **31**, 162–173 (2018).

65. Pilo, G. S., Holbrook, N. J., Kiss, A. E. & Hogg, A.M. Sensitivity of marine heatwave metrics to ocean model resolution. *Geophys. Res. Lett.* **46**, 14604–14612 (2019).

66. Siegert, S., Bellprat, O., Ménégoz, M., Stephenson, D. B. & Doblas-Reyes, F. J. Detecting improvements in forecast correlation skill: Statistical testing and power analysis. *Mon. Weather Rev.* **145**, 437–450 (2017).

Author contributions

R.M., S.M. & S.G. contributed to conception and design of the study. R.M. defined the methodology and performed the analysis and design of the figures. RM wrote the first draft of the manuscript. R.M., S.M. & S.G. contributed to the development of the manuscript and approved the submitted version.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s43247-023-00892-5.

Correspondence and requests for materials should be addressed to Ronan McAdam.

Peer review information *Communications Earth & Environment* thanks the anonymous reviewers for their contribution to the peer review of this work. Primary Handling Editors: Heike Langenberg. A peer review file is available.

Reprints and permission information is available at http://www.nature.com/reprints

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.