(Mohamed et al., 2023)The paper by Lin and coauthors deals with marine heatwaves in the North Sea. The contents and quality of the presentation need improvement. I have some doubts on the methodology and clarity of the diagnosed mechanisms. The Authors should consider the comments provided below.

Response: Thank you very much. We carefully considered each of your comments and suggestions, and we would like to respond to them in the following content.

I have also noticed this paper https://www.researchsquare.com/article/rs-6503093/v1 from the same authors, but with a different first author. Contents are similar but not identical. Could you please clarify the situation?

Response: Thank you for raising this point. The manuscript you referred to is a preprint that was previously submitted to another journal and subsequently posted online. That submission was unfortunately not accepted for publication. We would like to clarify that the first author of the preprint is also the first author of the present manuscript. The corresponding author's name appeared first in the header of that preprint, but in the manuscript the order of authors is consistent in both versions.

The present manuscript in Ocean Science represents a substantially improved and revised version of that earlier work. In preparing the current submission to Ocean Science, we have reanalyzed the results, clarified the scientific focus, revised the methodology, and reorganized the manuscript to better address the scope of Ocean Science. While the two manuscripts share a common scientific background, the current version contains significant modifications and improvements and is not under consideration elsewhere.

General comments

While there may be implication for prediction, mentioning forecasting in the abstract can be misleading. I suggest to clarify that is a diagnostic analysis, instead.

Response: We agree that our study is based on diagnostic and mechanistic analyses rather than forecasting. To avoid potential misunderstanding, we have revised the abstract by clarifying the diagnostic nature of the study and by removing wording that could be interpreted as implying direct forecasting:

Revised Abstract:
"Global shelf seas have experienced unprecedented marine heatwaves (MHWs) in recent decades. Although MHWs have been extensively studied at the global scale, their regional variability and underlying mechanisms remain poorly understood, particularly in shelf seas influenced by multiple climate modes. Here, we examine MHW variability in the Northeastern Atlantic shelf using a correlation-based k-means clustering approach. Two distinct subregions with contrasting seasonal patterns are identified. In winter, the southern North Sea experiences increased MHW frequency, intensity, and duration. This enhancement is linked to a positive East Atlantic Pattern, which intensifies westerly winds and enhances warm Atlantic inflow through both atmospheric and oceanic pathways. In contrast, the

northern North Sea shows enhanced MHW frequency and duration in summer, while MHW intensity weakens. This summer response is modulated by Atlantic Multidecadal Variability, with its positive phase strengthening Pacific-Atlantic connections via Rossby wave propagation, altering cloud cover and surface radiative forcing. A shallow mixed layer, enhanced stratification, and circulation-induced upwelling favor frequent and persistent but less intense summer MHWs. This north-south contrast demonstrates that different combinations of atmospheric and oceanic processes shape MHW variability across the shelf, providing a diagnostic and mechanistic framework for understanding regional MHW variability and its potential predictability."

"variabilities" in the plural form sounds strange. Suggest using singular or another noun, e.g. "modes" depending on what you want to convey.

Response: We agree that the plural form "variabilities" is confusing in this context. We have revised the manuscript by replacing "climate variabilities" with "climate modes" to better reflect the physical meaning.

Nonlinear quantities are computed for both atmospheric and ocean reanalysis data; can you estimate the error made with this compared to using daily data?

Response: Thank you for this important comment. We agree that nonlinear terms calculated from monthly mean reanalysis fields may differ from the values calculated at daily resolution. We would like to clarify that our analysis focuses on seasonal to interannual variability, for which monthly-mean fields are commonly used in existing studies (Liu et al., 2022; Mohamed et al., 2023). At these timescales, high-frequency synoptic fluctuations tend to be smoothed when they average. As a result, our estimates should be interpreted as diagnostics of low-frequency, large-scale contributions, rather than an exact closure of the instantaneous nonlinear budget.

L135-L147: "Atmospheric conditions were characterized using the ERA5 reanalysis dataset (Hersbach et al., 2020) produced by the European Centre for Medium-Range Weather Forecasts (ECMWF). ERA5 provides monthly variables at a spatial at resolution of 0.25°×0.25°. The variables used in this study include 10-m wind components, 500-hPa geopotential height, precipitation, and surface heat fluxes. Although ERA5 is available from 1940 onward, only data from 1982-2021 were used in this study to ensure consistency with the satellite-based SST record used for marine heatwave detection.

Oceanic conditions were analyzed using monthly data from the ECMWF Ocean Reanalysis System 5 (ORAS5) (Zuo et al., 2019), which provides global ocean reanalysis data at a horizontal resolution of 0.25°. Variables of interest include potential temperature and ocean current velocity, which were used to examine subsurface thermal structure and oceanic heat transport associated with MHW variability. Our analysis focuses on seasonal to interannual variability, for which monthly-mean fields are commonly used in existing studies (Liu et al., 2022; Mohamed et al., 2023)."

I am not sure to understand your use of K-means in this work. In other works (e.g. Vogt et al. 2022 10.3389/fclim.2022.847995, Wong et al. 2024

Response: We thank the reviewer for pointing out these relevant studies. Indeed, in previous works such as Vogt et al. (2022) and Wong et al. (2024), k-means clustering was primarily applied to classify MHWs or compound extreme events into different types, typically within predefined regions, with the clustering performed on event-based characteristics (e.g., duration, intensity, or associated driver anomalies).

Our study addresses a distinct methodological objective. Rather than clustering individual MHW events, each spatial grid point is treated as a data point in our study, characterized by its seasonal pattern of interannual MHWCI variability. The k-means clustering is therefore used to group grid points that exhibit similar seasonal variability structures, and the two regions emerge as the spatial manifestation of these clusters, rather than being prescribed a priori.

Conceptually, our approach is closer to clustering frameworks used to identify dynamic states or regimes based on temporal variability patterns, such as those applied in studies of brain connectivity dynamics (Allen et al., 2014). In that context, spatial units are grouped according to similarities in their temporal variability behavior, and the resulting clusters are interpreted as distinct connectivity states. Similarly, our clustering aims to identify spatially coherent regions with distinct seasonal responses to large-scale climate modes, rather than to define types of individual MHW events.

We have clarified this distinction in the revised manuscript to avoid potential confusion:

L167-L173: "Distinct regional patterns of MHWCI variability in the North Sea were identified using a correlation-based K-means clustering approach (Jain, 2010; Lloyd, 1982). Unlike conventional K-means clustering, which minimizes Euclidean distance, the correlation-based version groups spatial grid points according to the similarity in the shape of their normalized feature vectors, thereby emphasizing pattern similarity rather than absolute amplitude. Such correlation-based K-means clustering has been successfully applied in gene expression data analyses (Loganantharaj et al., 2006) and brain connectivity dynamics (Allen et al., 2014), where spatial units are grouped based on similarities in their temporal or spatial variability structures to identify coherent regions. "

Given the pronounced seasonal variability of MHW in this region, this approach classifies grid points according to their seasonal patterns of MHWCI interannual variability, thereby revealing spatially coherent regions with distinct seasonal responses to large-scale climate forcing.

Response: We agree that Fig. 3 highlights contrasting seasonal dominance between the two clusters, which may at first glance appear as a seasonal distinction. However, Fig. 3 is intended to identify the seasonal window during which interannual MHW variability is strongest in each domain, rather than to directly explain MHW occurrence mechanisms. The relevance to MHW occurrence is established in the subsequent analyses, where we show that the season of dominant variability determines which physical

processes and climate drivers can effectively modulate MHW characteristics. Specifically, winter-dominated variability in the southern domain is associated with enhanced atmospheric circulation anomalies and oceanic heat advection, leading to increases in MHW frequency, intensity, and duration, whereas summer-dominated variability in the northern domain is primarily regulated by surface heat fluxes and shallow mixed layers, affecting MHW frequency and duration. Therefore, the two domains are not defined by seasonality alone, but by distinct seasonal sensitivity to climate forcing, which results in different formation mechanisms of MHW. We believe this linkage is demonstrated by the combined interpretation of Figs. 3-10. We have revised the manuscript to avoid potential confusion:

L268-L292: "Based on the seasonal patterns of interannual MHWCI variability, the correlation-based K-means clustering identifies two spatially coherent subregions (named as Cluster 1 and 2, respectively) in the North Sea (Fig. 2a). Cluster 1 mainly covers the central and southern parts of the North Sea, while Cluster 2 includes the deeper northern region. The two clusters are geographically contiguous and broadly aligned with the contrast between the shallow southern shelf and the deeper northern basin.

Seasonal differences between the two clusters were quantified by calculating the relative contribution of each season, expressed as the percentage of its spatially averaged interannual variability intensity (VI) relative to all four seasons within each cluster. Positive values of VI indicate seasons that contribute more strongly to the total variability of that cluster, while negative values indicate weaker contributions. In Cluster 1, variability is dominated by winter (VI = 0.66) and autumn (VI = 0.41), with substantially weaker variability in summer (VI = -0.28) and spring (VI = 0.07) (Fig. 2b). Cluster 2, by contrast, shows its strongest variability in summer (VI = 0.57) and moderate variability in spring (VI = 0.24), while variability in autumn (VI = -0.08) and winter (VI = -0.30) is relatively weak. These results indicate a clear seasonal asymmetry between the central-southern and northern North Sea, implying that seasons with the most active MHW-related variability differ between the two regions.

In the central-southern North Sea (Cluster 1), variability is dominated by winter, whereas in the northern deeper region (Cluster 2) it is primarily expressed in summer. These seasonal contrasts are derived from the relative contributions of each season to the standardized interannual variability of MHWCI and therefore reflect when MHW-related variability is most active in each region. Based on this diagnosis, the subsequent analyses focus on winter MHWCI for Cluster 1 and summer MHWCI for Cluster 2 in order to examine large-scale climate influences during the season of maximum variability for each region. The relationship between regional differences and large-scale climate modes was examined using the time series of domain-average MHWCI extracted for each cluster. Consistent with the identified seasonal dominance, the winter series was analyzed for Cluster 1 and the summer series for Cluster 2."

From line 201 and following I understand MHWs may happen predominantly in one or the other region depending on the season, but more explanation is due.

Response: We agree that additional explanation is needed. In the revised version, we have added an explanatory paragraph to explicitly link the seasonal diagnosis shown in Fig. 2b with the subsequent analysis strategy (See above: **L268-L292**).

Various climate indices are shown; sources for the associated data is missing.

Response: We thank the reviewer for pointing this out. We have revised the manuscript to explicitly describe the data sources and calculation methods of all climate indices used in this study. A new subsection (Sect. 2.4) has been added to the Methods section, detailing the definitions, computational approaches, and data sources for the NAO, EAP, and AMV indices.

L142-L157:
"2.1.3 Climate modes

The influence of large-scale climate modes is investigated using indices of the NAO, the EAP, and the AMV, which are known to modulate atmospheric circulation over the North Atlantic and adjacent shelf seas. The NAO and EAP indices are obtained from the National Oceanic and Atmospheric Administration (NOAA) Climate Prediction Center and are derived from the first and second modes of a rotated empirical orthogonal function (EOF) analysis of monthly mean 500-hPa geopotential height anomalies over the North Atlantic.

The AMV index is derived from basin-averaged SST anomalies over the North Atlantic. To isolate internal multidecadal variability, the AMV index is calculated by removing the global-mean SST anomaly from the North Atlantic mean SST anomaly (Trenberth & Shea, 2006). Positive and negative phases of each climate mode are defined using a fixed threshold at zero, with values greater (less) than zero indicating positive (negative) phases. Seasonal indices are computed by averaging the corresponding monthly values over the relevant seasons.

Previous studies have shown that positive phases of the AMV and EAP are associated with increased annual MHW cumulative intensity in the North Sea, whereas positive NAO phases primarily enhance wintertime MHW occurrence in the southern North Sea (Mohamed et al., 2023; Mohamed et al., 2025)."

The discussion on the mechanisms is quite poor, as results for a much larger domain is presented, so local mechanisms are not discussed. More details should be provided on the stratification procedure;

Response: We agree that the discussion of mechanisms would benefit from a clearer emphasis on local upper-ocean processes and stratification. Following this suggestion, we have expanded the analysis of mixed-layer dynamics and vertical stability to better link large-scale forcing with local physical responses in the North Sea. Corresponding revisions are shown below.

L208-L222: "2.3 Quantification of upper-ocean stratification

To characterize the upper-ocean stratification associated with MHWs, mixed layer depth (MLD) is diagnosed following a density-based criterion. Specifically, MLD is defined as the depth at which the potential density first exceeds the potential density at 10 m by 0.03 kg m^{-3} (England et al., 2025), which has been shown to be suitable for shallow and seasonally stratified shelf seas. Vertical

stratification and dynamical stability are further quantified using the gradient Richardson number, Ri, defined as:

$$Ri = \frac{N^2}{S^2}, \qquad (3)$$

where N is the buoyancy frequency,

$$N = \left(-\frac{g}{\rho}\frac{\partial \rho}{\partial z}\right)^{1/2}, \quad (4)$$

and S is the vertical shear

$$S = \left[\left(\frac{\partial u}{\partial z}\right)^2 + \left(\frac{\partial v}{\partial z}\right)^2\right]^{1/2} \quad (5).$$

Here, g denotes gravitational acceleration, ρ is seawater density, and u and v are the zonal and meridional velocity components, respectively. All variables are derived from the ORAS5 reanalysis and are computed over the same temporal intervals as the temperature fields to ensure consistency among stratification and heat budget diagnostics."

L407-L435:

"In winter, the Cluster 1 region is predominantly well mixed due to strong wind forcing and shallow bathymetry, as reflected by weak climatological stratification (Fig. S4a) with low Ri (Fig. S3). Consistent with this background state, composite MLD anomalies (Fig. 7b) are negligible across most of the Cluster 1 region, with negative anomalies confined to very shallow shelf regions.

The Richardson number anomalies displayed in Fig. 7c exhibit a similar spatial structure. While Ri anomalies are weak or near zero over the well-mixed central basin, pronounced negative anomalies appear in the same shallow shelf regions where MLD shoaling is observed. This spatial correspondence indicates that changes in stratification and velocity shear are spatially collocated over shallow shelf areas during enhanced MHWCI winters.

These localized structural changes occur together with coherent atmospheric and oceanic anomalies. Enhanced Atlantic inflow provides a positive contribution to the mixed-layer heat budget, while strengthened southwesterly winds are associated with increased surface heat loss. Additionally, shoaling of the MLD reduces the effective heat capacity of the upper ocean, allowing advected heat to produce a stronger temperature response. Meanwhile, reduced Ri indicates enhanced shear relative to stratification, which modulates vertical mixing without promoting deep redistribution of heat. Consequently, heat transported by horizontal advection can be preferentially retained in the near-surface layer, favoring the development and persistence of winter MHWs despite concurrent surface heat loss.

To further assess the dynamical response associated with these coupled structural and heat transport anomalies, the corresponding overturning circulation anomalies are investigated. Figure 7d presents composite anomalies of the meridional overturning circulation (MOC) during enhanced MHWCI winters.

The MOC is strengthened, with increased northward and downward transport in the southern and central North Sea, respectively. These overturning features align with regions of enhanced Atlantic inflow and current convergence. The enhanced downward transport in the central North Sea occurs within a water column that is already well mixed during winter. As a result, the intensified overturning circulation does not produce a discernible mixed-layer deepening, consistent with the absent MLD and Ri anomalies observed in the central basin. Instead, the enhanced overturning facilitates the redistribution and retention of advected heat within the water column, contributing to elevated winter temperatures and enhanced MHWCI in the central North Sea."

L496-L514:

"The effect of surface radiative forcing in summer is closely linked to the upper-ocean structure in Cluster 2. During summer, the climatological MLD is quite shallow (Fig. S4b). Composite anomalies further show a basin-wide shoaling of the MLD in Cluster 2 (Fig. 7f), implying a further shoaling of the near-surface layer during enhanced MHWCI summer. Fig 7g also shows predominantly positive Ri anomalies over the same region, indicating enhanced stratification stability and a reduced susceptibility to shear-driven instabilities. The spatial co-location of MLD shoaling and increased Ri suggests that buoyancy-driven stratification strengthens sufficiently to offset any increase in velocity shear, resulting in a more stably stratified upper ocean during these summers. Together, these changes imply that radiative and surface heat flux anomalies are confined to a thin surface layer, allowing relatively short-lived atmospheric forcing anomalies to produce a pronounced near-surface thermal response.

In addition to surface forcing, the regression patterns indicate a dynamical response in the regional circulation. The easterly wind anomalies counteract the climatological southwesterlies, leading to weakened surface winds over the northern North Sea. These wind anomalies are accompanied by negative MOC anomalies (Fig. 7i), indicating an anomalous anticlockwise overturning circulation with enhanced upwelling over the northern North Sea. When combined with the summer upper-ocean structure (shallow MLD and increased Ri), this circulation configuration provides a consistent explanation for the observed MHW characteristics in Cluster 2: a shallow, stably stratified surface layer increases the sensitivity of SST to radiative anomalies, favoring more frequent and longer-lasting warm events, whereas the upwelling-related cooling tendency limits the buildup of extreme surface warming, consistent with reduced MHW intensity."

are you using fixed or dynamic thresholds? This should be consistent across indices, e.g. if they are all standardized already.

Response: We would like to clarify that fixed thresholds are used throughout the analysis, but they are applied differently depending on the type of variable. Details are provided below.

(1) Climate indices: The large-scale climate indices (NAO, EAP, and AMV) are obtained directly from publicly available datasets. These indices are anomaly-based indices derived from atmospheric or oceanic fields, and we use them in their original form. Positive and negative phases shown in Fig. 2 and

Fig. 9 are therefore defined using a fixed threshold at zero, with values greater (less) than zero indicating positive (negative) phases. We have revised the Methods sections:

L158-L159: "Positive and negative phases of each climate mode are defined using a fixed threshold at zero, with values greater (less) than zero indicating positive (negative) phases."

(2) Cluster-specific MHWCI time series: The cluster-specific MHWCI time series are constructed in this study and are explicitly standardized to zero mean and unit variance. Periods of enhanced MHWCI are defined using a fixed threshold of +1 standard deviation, corresponding to values greater than 1 in the standardized time series. This threshold is applied consistently in all composite and regression analyses. We have revised the Results sections to explicitly clarify these definitions and to avoid any ambiguity regarding the thresholds used.

L339-L343: "The atmospheric and oceanic conditions associated with enhanced winter MHWCI in Cluster 1 were examined using composite anomalies during periods when the standardized cluster-mean MHWCI exceeds +1 standard deviation (black line in Fig. 2c). The analysis focuses on large-scale atmospheric circulation, surface forcing, and oceanic circulation and heat transport during the months from October of the preceding year to January of the MHW winter. "

The level of discussion when presenting figures (which are often hard to read) is insufficient, and the various domains used (e.g., larger in Fig. 10) complicates comparisons.

Response: We appreciate this comment and have revised the manuscript accordingly. The presentation of figures in the Results section has been expanded and clarified to guide the reader more explicitly through each diagnostic. For figures with complex spatial structure or larger domains (e.g., Fig. 10), we now clearly state the purpose of the domain choice and explicitly identify the subregions relevant to the North Sea analysis. The text now explains which parts of the broader domain are being interpreted and how they relate to the cluster-based regions defined earlier, thereby reducing ambiguity and facilitating comparison with other figures. Moreover, for all key figures we added more detailed figure-by-figure descriptions, explicitly outlining spatial contrasts, temporal differences between phases, and consistency across variables. These revisions improve comprehensibility when comparing figures that use different spatial extents.

I would suggest to first analyse non-MHW variables and understand how they change over time and due to teleconnections, and then focus on the influence of these indices on MHWs.

Response: We have revised the Results section to follow the suggested analytical sequence. In the revised manuscript, we first examine non-MHW variables, including large-scale atmospheric circulation, teleconnection patterns, surface wind, cloud cover, radiative fluxes, upper-ocean structure (MLD and Richardson number), and circulation diagnostics (e.g., overturning). These variables are analyzed in detail to establish their temporal evolution and phase dependence associated with large-scale climate modes.

Based on this background understanding, we then discuss how these non-MHW processes jointly influence MHW characteristics. In particular, we now explicitly distinguish between different aspects of MHW (frequency, duration, and intensity) and show how they respond to the identified atmospheric and oceanic conditions. For example, in summer, over the northern North Sea, shallow mixed layers and enhanced stratification favor frequent and long-lasting MHWs, while circulation-induced upwelling limits extreme warming, explaining the reduced MHW intensity.

This revised structure ensures that MHW variability is interpreted as a consequence of previously diagnosed physical processes, strengthening the mechanistic clarity of the study.

I feel like the summary of Fig. 11 is actually not aiding interpretation, as new concepts and indices are added. This should be simplified.

Response: We agree that the current summary figure (Fig. 11) introduces additional concepts and climate indices, which may complicate the interpretation. In the revised manuscript, we have simplified Fig. 11 (shown below) so that it functions purely as a conceptual summary of the key mechanisms.
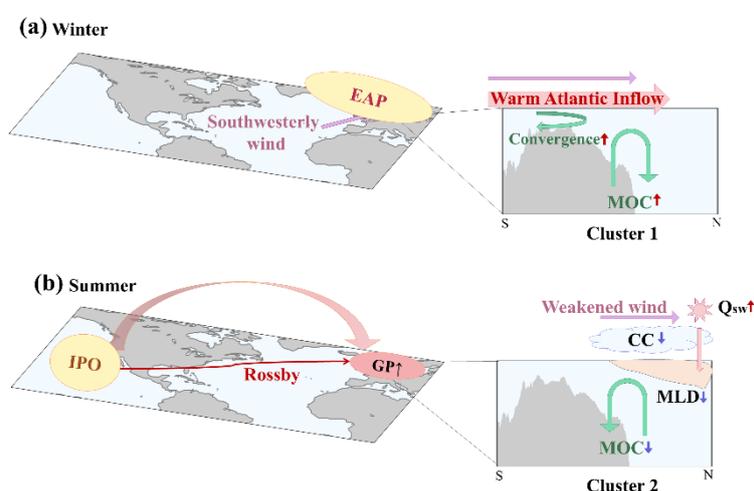
Updated Figure 11. Schematic illustration of two distinct cluster patterns showing the synergistic impacts of climate modes on MHWCI in the Northwestern European shelf. (a) The winter pattern characterized by a combined negative NAO phase in late autumn and a positive EAP phase in winter. (b) The summer pattern characterized by simultaneous positive phases of AMV and IPO. Purple arrows indicate wind anomalies. Red upward and blue downward arrows represent increases and decreases in the associated variables. Blue cloud shapes denote cloud cover (CC), while red circles denote geopotential (GP). Q_{sw} indicates the downward shortwave heat flux. Green arrows mark horizontal flow convergence or the Meridional Overturning Circulation (MOC).

Appendix A seems unnecessary, as this has likely been done by the producers in more detail.

Response: We agree that the validation of the SST product has been extensively documented by the data producers. Following this recommendation, we have removed Appendix A from the manuscript. The validation has been mentioned in the Data section as supplementary information.

L106-L108: "An independent observation-based SST dataset from the FerryBox system (V. Macovei et al., 2021), developed by Helmholtz-Zentrum Hereon, is used for validation purposes (Fig. S1)."

The work needs careful proofreading, as there are errors in the titles of most figures and references. Only some examples are given below.

Response: Thank you. We have carefully proofread the entire manuscript and corrected typos and inconsistencies in figure titles, references, and text throughout the paper.

Comments by line
l21 why speaking about "prediction"?

Response: We agree that the reference to "prediction" in the abstract is confusing. As the present study is based on diagnostic and mechanistic analyses rather than forecasting, we have revised the abstract to clarify its diagnostic nature and remove confusing wording.

L20-L33: "Global shelf seas have experienced unprecedented marine heatwaves (MHWs) in recent decades. Although MHWs have been extensively studied at the global scale, their regional variability and underlying mechanisms remain poorly understood, particularly in shelf seas influenced by multiple climate modes. Here, we examine MHW variability in the Northeastern Atlantic shelf using a correlation-based k-means clustering approach. Two distinct subregions with contrasting seasonal patterns are identified. In winter, the southern North Sea experiences increased MHW frequency, intensity, and duration. This enhancement is linked to a positive East Atlantic Pattern, which intensifies westerly winds and enhances warm Atlantic inflow through both atmospheric and oceanic pathways. In contrast, the northern North Sea shows enhanced MHW frequency and duration in summer, while MHW intensity weakens. This summer response is modulated by Atlantic Multidecadal Variability, with its positive phase strengthening Pacific-Atlantic connections via Rossby wave propagation, altering cloud cover and surface radiative forcing. A shallow mixed layer, enhanced stratification, and circulation-induced upwelling favor frequent and persistent but less intense summer MHWs. This north-south contrast demonstrates that different combinations of atmospheric and oceanic processes shape MHW variability across the shelf, providing a diagnostic and mechanistic framework for understanding regional MHW variability and its potential predictability."

l69 same comment as for the title

Response: We have revised the manuscript at line 69, replacing "climate variabilities" with "climate modes".

Fig. 1 typo in panel b title

Response: The typo in the title of Fig. 1b has been corrected from "MHWTI" to "MHWCI".

l88/184 as far as I understand, cumulative intensity is a per-event quantity, and events are discrete. How do you compute the Fourier spectrum then? Is it just SST?

Response: We would like to clarify that the Fourier spectrum is not computed from individual MHW events. While cumulative intensity is defined at the event level, we constructed a continuous time series by aggregating the cumulative intensity of all MHW events occurring within each month. MHWCI is set to zero when no marine heatwaves occur during a given month. This results in a monthly MHWCI time series, which is then used for the spectral analysis shown in Fig. 1b. The analysis is therefore based on monthly MHWCI, not on SST directly. We have revised the manuscript to avoid potential confusion:

L261-L266: "Power spectrum analysis of the detrended monthly MHWCI (Fig. 1b), constructed as a continuous time series by summing the cumulative intensity of all MHW events occurring within each month (and set to zero when no events occur), shows that, after removing the dominant seasonal pattern, substantial variability remains at interannual timescales. Several spectral peaks exceed the 90% confidence level, suggesting that MHWCI variability in the North Sea is organized on multi-year timescales. This provides a basis for exploring whether such variability exhibits coherent spatial and seasonal structures across the region."

l102 The data record starts in 1940. Are you using 1982-2021 as for the SSTs? Please clarify

Response: The satellite-based OSTIA SST product used to identify MHW is available only from 1982 onwards. Although the atmospheric datasets (e.g. ERA5) extend back to 1940, our analysis is restricted to the common period 1982-2021, consistent with the SST record. Corresponding revisions are shown below.

L138-L140: "Although ERA5 is available from 1940 onward, only data from 1982-2021 were used in this study to ensure consistency with the satellite-based SST record used for marine heatwave detection."

l104 A reference and some more details, both on ERA5 and ORAS5, should be given

Response: We have revised the manuscript to include appropriate references and additional details for both ERA5 and ORAS5 in the Data section, including a brief description of the datasets and their temporal and spatial resolutions.

L135-L147:

   "Atmospheric conditions were characterized using the ERA5 reanalysis dataset (Hersbach et al., 2020) produced by the European Centre for Medium-Range Weather Forecasts (ECMWF). ERA5 provides monthly variables at a spatial at resolution of 0.25°×0.25°. The variables used in this study include 10-m wind components, 500-hPa geopotential height, precipitation, and surface heat fluxes. Although ERA5 is available from 1940 onward, only data from 1982-2021 were used in this study to ensure consistency with the satellite-based SST record used for marine heatwave detection.

   Oceanic conditions were analyzed using monthly data from the ECMWF Ocean Reanalysis System 5 (ORAS5) (Zuo et al., 2019), which provides global ocean reanalysis data at a horizontal resolution of 0.25°. Variables of interest include potential temperature and ocean current velocity, which were used to examine subsurface thermal structure and oceanic heat transport associated with MHW variability. Our

analysis focuses on seasonal to interannual variability, for which monthly-mean fields are commonly used in existing studies (Liu et al., 2022; Mohamed et al., 2023)."

Response: Thank you for this comment. We constructed monthly MHWCI by aggregating event cumulative intensity within each calendar month. Specifically, monthly MHWCI is obtained by summing the cumulative intensity of all events whose onset date falls in that month. Therefore, an event starting on 27 January and ending on 10 February is assigned to January in our monthly aggregation. We additionally tested the sensitivity of the k-means classification to the event-to-month assignment by assigning events to their termination month instead of their initiation month. The resulting spatial clustering remains very similar to the original classification. The main difference is a slight seasonal shift in the relative contributions within each cluster: in Cluster 1, which is primarily winter-dominated, the contribution from spring becomes somewhat larger, while in Cluster 2, which is primarily summer-dominated, the contribution from autumn increases. Importantly, however, the dominant seasons associated with MHW occurrence remain winter and summer, respectively. Therefore, this sensitivity does not affect the main physical interpretation or conclusions of the study. Corresponding revisions are shown below.

L128-L132: "Specifically, the cumulative intensity of all MHW events whose onset date falls within a given month was summed to form the monthly MHWCI, while months without any MHW events were assigned a value of zero. This results in a continuous monthly MHWCI time series, which was subsequently expressed as anomalies relative to its long-term mean to characterize interannual variability"

Furthermore, the MHW threshold (90th percentile) is computed from the detrended SST time series using the baseline period 1982-2021. Explanations have been provided in the revision:

L112-L116: "The climatological mean and percentile-based threshold used to detect MHW events are computed over the period 1982-2021, which corresponds to the full temporal coverage of the SST dataset used in this study. To ensure internal consistency across all analyses, a unified climatological baseline and detrending strategy is applied throughout this study. All anomalies are computed relative to the same reference period, and long-term trends are removed before analyses that focus on variability and climate-mode relationships."

Response: The mixed-layer heat tendency equation (Eq. 3) follows the standard formulation commonly used in previous studies of upper-ocean heat budgets (Liu et al., 2014; Tan et al., 2016). We have added the relevant references in the manuscript.

ORAS5 provides ocean variables on 75 vertical levels (level spacing increasing from 1 m at the surface to 200 m in the deep ocean), allowing a physically consistent estimation of mixed-layer depth and associated heat budget terms at seasonal to interannual timescales. All heat budget terms are computed using monthly mean data. We have clarified these points in the Methods section.

L238-L241: "ORAS5 provides ocean variables on 75 vertical levels, with vertical spacing increasing from approximately 1 m near the surface, allowing a physically consistent estimation of MLD and mixed-layer heat budget terms at seasonal to interannual timescales."

L245-L246: "This heat tendency equation (Eq. 6) follows the standard formulation commonly used in upper-ocean heat budget studies. All terms are computed using monthly mean fields."

l189 so you apply K-means clustering over the whole year or individual seasons?

Response: The K-means clustering is not applied separately to individual seasons, nor directly to the raw monthly time series. Instead, it is performed using seasonal variability characteristics derived from the full monthly record. Specifically, for each grid point, we compute the interannual variability (standard deviation) of MHWCI anomalies separately for DJF, MAM, JJA, and SON (with DJF treated across calendar years). These four seasonal variability measures form a feature vector that is used as input to the K-means clustering. Therefore, the clustering is based on information from all seasons, while explicitly accounting for their distinct variability characteristics. Corresponding revisions are shown below.

L175-L179: "Given the pronounced seasonal variability of MHW in this region, this approach classifies grid points according to their seasonally resolved interannual variability characteristics, rather than applying the clustering separately to individual seasons. Specifically, a K-means clustering is performed using seasonal variability information derived from the full monthly record, thereby revealing spatially coherent regions with distinct seasonal responses to large-scale climate forcing. "

Fig. 2 typos in titles, and no units in Fig c and d; how are time series normalized?

Response: Thank you for pointing this out. As response,
(1) We have corrected the typos in the titles of Fig. 2c-d.
(2) We have clarified in the figure labels and captions that the time series shown in Fig. 2c and d are standardized MHWCI, which are therefore dimensionless.
(3) The time series are normalized using z-score standardization, defined as

$$z(t) = \frac{x(t) - \mu}{\sigma},$$

where μ and σ denote the mean and standard deviation of the original time series, respectively. This procedure results in dimensionless time series with zero mean and unit variance.

L330-L331 in Fig.2 caption: "All time series in (c) and (d) are standardized using z-score normalization and are therefore dimensionless."

l247 anomalies from what? Is this is stratified according to some indices? How?

Response: The geopotential anomalies shown in Fig. 4 are defined relative to the long-term monthly climatology. They are obtained using composite analysis during periods of enhanced MHWCI, defined as months when the standardized cluster-specific MHWCI time series exceeds +1 standard deviation. We have clarified the anomaly reference state and the composite criteria in the revised manuscript.

L339-L343: "The atmospheric and oceanic conditions associated with enhanced winter MHWCI in Cluster 1 were examined using composite anomalies during periods when the standardized cluster-mean MHWCI exceeds +1 standard deviation (black line in Fig. 2c). The analysis focuses on large-scale atmospheric circulation, surface forcing, and oceanic circulation and heat transport during the months from October of the preceding year to January of the MHW winter."

l266 with units m^2/s^2, this is geopotential (not height)

Response: Thank you for pointing this out. We agree that the variable has units of $m^2\,s^{-2}$ and therefore represents geopotential rather than geopotential height. We have revised the text and figure labels throughout the manuscript accordingly to ensure consistent terminology.

l269 why this value? Why not using some significance threshold?

Response: The composite anomalies shown are tested for statistical significance at the 90% confidence level. For the wave activity flux vectors, in addition to significance testing, we apply a display threshold of $0.1\ m^2\,s^{-2}$ in the figure to emphasize dynamically meaningful flux patterns. Similar display thresholds have been used in previous studies of atmospheric wave activity flux diagnostics (Hou et al., 2023). Corresponding revisions are shown below.

L346-L347: "The anomalous WAF flux is shown only when its magnitude is larger than $0.1\ m^2/s^2$ (Hou et al., 2023). "

Fig 7 cannot be understood. Why are arrows colored? They are hard to see. What are the black arrows on the right side? Missing labels on axes

Response: We agree that Fig. 7 currently contains too many elements, which make it difficult to interpret. In the revised version (shown below), we have simplified the presentation of the vector fields by representing current speed through arrow length rather than color, while using a uniform color for all vectors. In addition, we have increased arrow size and spacing to further enhance readability. The black arrows on the right-hand side indicate the direction of the meridional overturning circulation. We agree that this was not sufficiently clear in the original version and this is now explicitly clarified in both the caption and the main text. In addition, we have added missing axis labels to ensure that all plotted elements are clearly defined.
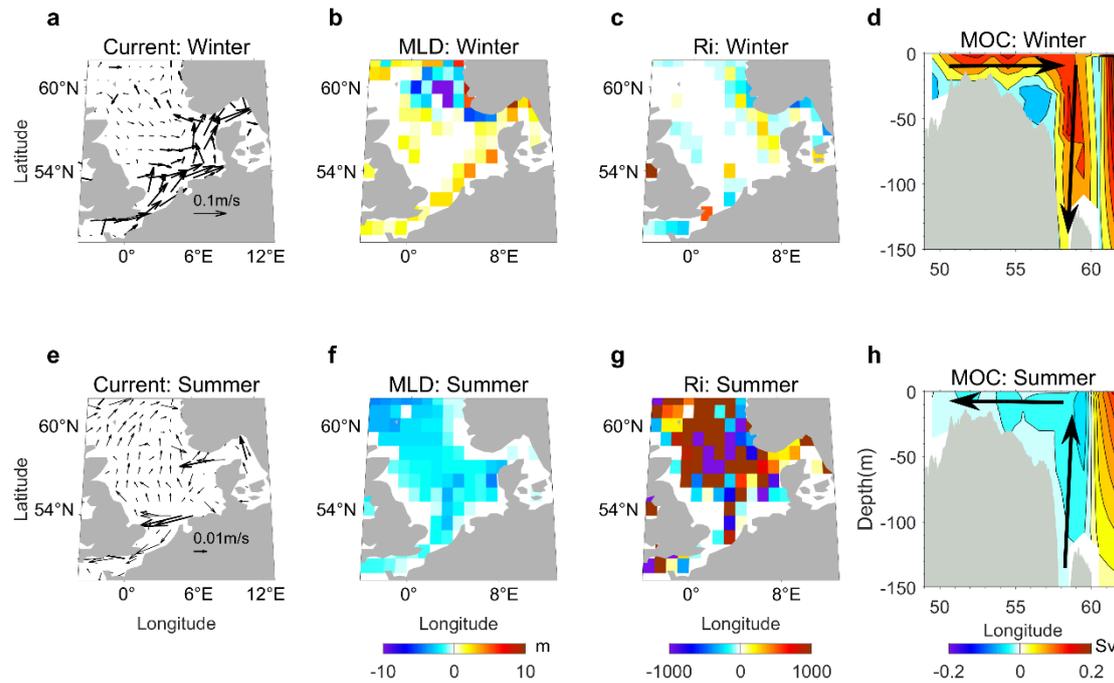
Updated Figure 7. Composite anomalies of oceanic variables during positive phases of the cluster-specific time series in winter (a-d) and summer (e-h). (a), (e) Oceanic current (vectors, m/s), (b), (f) mixed layer depth (MLD, m), (c), (g) Richardson number (Ri), (d), (h) meridional overturning circulations (MOC, Sv). The MOC is domain-integrated overturning circulations, obtained by integrating the velocity field over the entire zonal and meridional extent of the study domain, respectively. Negative stream function values correspond to clockwise circulation, while positive values indicate counterclockwise circulation. Black arrow-headed lines in (d) and (h) indicate the direction of the overturning circulation.

l277 is this some sort on average on some sub-domain?

Response: The contributions shown in Fig. 6 are computed as spatial averages over the Cluster 1 region during periods of enhanced MHWCI. We have clarified this in the revised manuscript.

L386-L387: "Figure 6 presents composite anomalies of the net surface heat flux and its individual components averaged over the Cluster 1 region during winters with enhanced MHWCI."

l318 "telecommunications"?

Response: The term "telecommunications" was an error and has been corrected to "teleconnections" in the revised manuscript.

l340 "predominated"

Response: The term "predominated" has been replaced with "dominated" in the revised manuscript.

l342 and otherwise? Some climatological maps should be presented.

Response: We understand that the concern refers to the description of mixed layer depth (MLD), where only regional mean values are reported, and more generally to the representation of stratification in the manuscript. In the revised manuscript, we have included climatological maps of MLD (Fig. S4, see below) and discuss stratification together with MLD, linking this discussion to the heat budget analysis and to the contrasting summer and winter mechanisms identified in the two clusters.
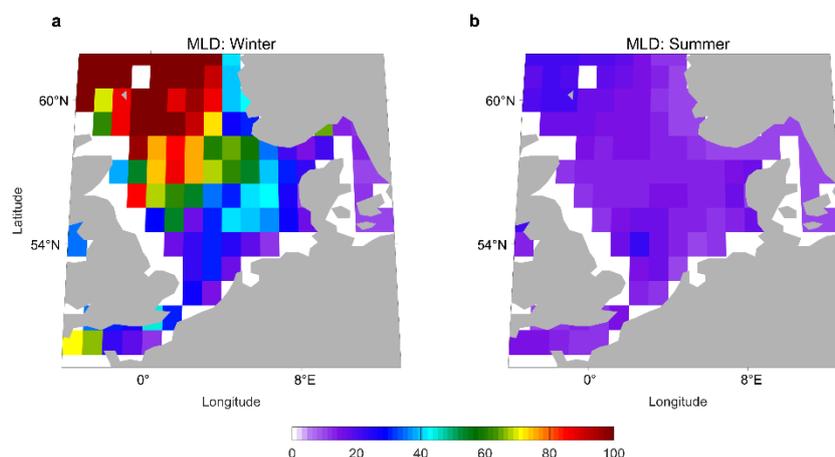
Supplementary Figure S4. Climatological mixed layer depth (MLD, m) in the North Sea during winter (a) and (b) summer.

L407-L415:" In winter, the Cluster 1 region is predominantly well mixed due to strong wind forcing and shallow bathymetry, as reflected by weak climatological stratification (Fig. S4a) with low Ri (Fig. S3). Consistent with this background state, composite MLD anomalies (Fig. 7b) are negligible across most of the Cluster 1 region, with negative anomalies confined to very shallow shelf regions.

The Richardson number anomalies displayed in Fig. 7c exhibit a similar spatial structure. While Ri anomalies are weak or near zero over the well-mixed central basin, pronounced negative anomalies appear in the same shallow shelf regions where MLD shoaling is observed. This spatial correspondence indicates that changes in stratification and velocity shear are spatially collocated over shallow shelf areas during enhanced MHWCI winters."

L496-L504: "The effect of surface radiative forcing in summer is closely linked to the upper-ocean structure in Cluster 2. During summer, the climatological MLD is quite shallow (Fig. S4b). Composite anomalies further show a basin-wide shoaling of the MLD in Cluster 2 (Fig. 7f), implying a further shoaling of the near-surface layer during enhanced MHWCI summer. Fig 7g also shows predominantly positive Ri anomalies over the same region, indicating enhanced stratification stability and a reduced susceptibility to shear-driven instabilities. The spatial co-location of MLD shoaling and increased Ri suggests that buoyancy-driven stratification strengthens sufficiently to offset any increase in velocity shear, resulting in a more stably stratified upper ocean during these summers. Together, these changes imply that radiative and surface heat flux anomalies are confined to a thin surface layer, allowing relatively short-lived atmospheric forcing anomalies to produce a pronounced near-surface thermal response."

Fig. 9 why do arrows look quite different? And what is the box in the maps?

Response: The differences in the WAF vectors among the three panels reflect physically distinct wave propagation patterns associated with different AMV phases and background climate states, as all fields are consistently regressed onto the IPO index. In addition, the rectangular box shown in Fig. 9 indicates the North Sea region, which is the focus of the present study. This has now been stated in the caption.

l400 is it because some versions of the AMO index are influenced by trends?

Response: Thank you for this insightful question. We agree that some versions of the AMO index can be influenced by long-term trends. However, the AMO index used in Mohamed et al. (2023) is based on the NOAA index definition in which the North Atlantic SST time series is detrended and thus represents multidecadal internal variability rather than the long-term warming signal.

In our study, we detect marine heatwaves using detrended SST. Therefore, the relationships discussed here are unlikely to arise from shared long-term trends but instead reflect the influence of low-frequency Atlantic variability and associated atmosphere-ocean processes.

l425 why the West Med now?

Response: We agree that the reference to the western Mediterranean is not essential in this context and may appear abrupt. We have removed this part of the sentence in the revised manuscript to improve the focus and clarity of the discussion.

l662 incomplete citation, also l681, l700...

Response: We have carefully checked the entire reference list and correct all incomplete citations, including those at lines 662, 681, and 700.

Fig. B1 what are the units of the ordinate?

Response: The ordinate in Fig. B1 represents the within-cluster sum of squared correlation distances (WCSS), which is a dimensionless clustering metric. We have clarified this in the figure caption:

L657-L660: "**Figure. A1**. Figure. A1. Elbow method validation of K-means clustering of marine heatwave cumulative intensity (MHWCI) in the North Sea. The within-cluster sum of squared correlation distances (WCSS) is shown as a function of cluster numbers. The red circle indicates the optimal number of clusters determined by the Elbow method. WCSS denotes the within-cluster sum of squared correlation distances and is a dimensionless clustering metric."

Reference used in the response

Allen, E. A., Damaraju, E., Plis, S. M., Erhardt, E. B., Eichele, T., & Calhoun, V. D. (2014). Tracking whole-brain connectivity dynamics in the resting state. *Cereb Cortex*, *24*(3), 663-676. https://doi.org/10.1093/cercor/bhs352

England, M. H., Li, Z., Huguenin, M. F., Kiss, A. E., Sen Gupta, A., Holmes, R. M., & Rahmstorf, S. (2025). Drivers of the extreme North Atlantic marine heatwave during 2023. *Nature*, *642*(8068), 636-643. https://doi.org/10.1038/s41586-025-08903-5

Hou, J., Fang, Z., & Geng, X. (2023). Recent Strengthening of the ENSO Influence on the Early Winter East Atlantic Pattern. *Atmosphere*, *14*(12), 1809. https://www.mdpi.com/2073-4433/14/12/1809

Liu, K., Xu, K., Zhu, C., & Liu, B. (2022). Diversity of Marine Heatwaves in the South China Sea Regulated by ENSO Phase. *Journal of Climate*, *35*(2), 877-893. https://doi.org/https://doi.org/10.1175/JCLI-D-21-0309.1

Liu, Q. Y., Wang, D., Wang, X., Shu, Y., Xie, Q., Chen, J., Liu, Q. Y., Wang, D., Wang, X., & Shu, Y. (2014). Thermal variations in the South China Sea associated with the eastern and central Pacific El Nino events and their mechanisms. *Journal of Geophysical Research Oceans*, *119*(12), 8955-8972.

Mohamed, B., Barth, A., & Alvera-Azcárate, A. (2023). Extreme marine heatwaves and cold-spells events in the Southern North Sea: classifications, patterns, and trends. *Frontiers in Marine Science*, *10*, Article 1258117. https://doi.org/10.3389/fmars.2023.1258117

Tan, W., Wang, X., Wang, W., Wang, C., & Zuo, J. (2016). Different Responses of Sea Surface Temperature in the South China Sea to Various El Niño Events during Boreal Autumn. *Journal of Climate*, *29*(3), 1127-1142.