Author's response to Reviewer 1

General comments:

This paper presents an analysis of the factors that affect the occurrence of marine heatwaves in the North Sea. A correlation-based K-means clustering approach was used to divide the North Sea into two regions that have different variability in marine heatwave cumulative intensity, and detailed analysis carried out to determine the mechanisms that cause that variability. The paper is well written and presented, and the topic and results are of strong interest.

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.

Specific comments:

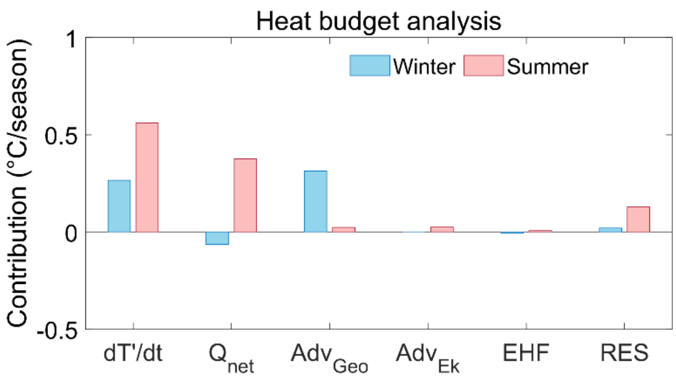
Line 97 – should Worsfold et al. (2024; https://doi.org/10.3390/rs16183358) be referenced here?

Response: Thank you for noting this relevant paper. We have read it through and referenced it at appropriate places in the revised texts.

Line 157 – a minor comment, but why include November and May in the calculation of tendency but not March or September? Shouldn't it be calculated either November - March and May – September if including the months surrounding the season, or December – February and June – August if not?

Response: We agree that our description of the SST anomaly tendency window could be confusing. To ensure temporal consistency between the left-hand-side tendency and the seasonally averaged budget terms, we have revised the tendency calculation, so that it is computed within the core season: (Feb-Dec) for winter (DJF) and (Aug-Jun) for summer (JJA) and normalized by the corresponding time interval. All right-hand-side terms are now averaged over the same DJF/JJA periods. The manuscript and figure have been updated accordingly.

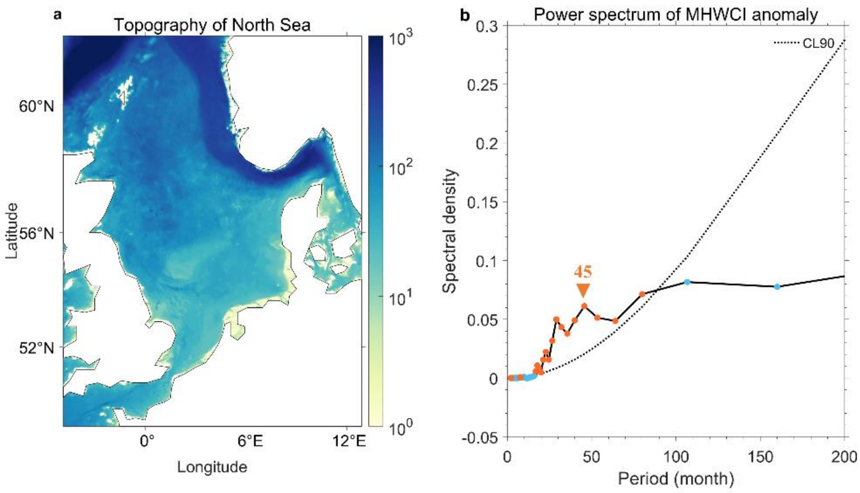
L231-L232: "We calculated this tendency as the temperature difference between February and December for winter and between August and June for summer, respectively. "



Updated Figure 8. Quantitative contribution (˚C/season) of atmospheric and oceanic processes to the upper ocean temperature anomaly tendency for two clusters identified by the K-means algorithm. The atmospheric and oceanic variables are shown as composite patterns during positive phases of the cluster-specific time series in winter (blue) and summer (red). The temperature tendency (*dT'/dt*) represents the difference in the sea surface temperature anomaly between February and December for winter and between August and June for summer, respectively. Contributions from the net heat flux anomaly (Q_{net}), geostrophic heat advection (Adv_{Geo}), Ekman heat advection (Adv_{Ek}), entrainment heat flux (*EHF*), and residual term (*RES*), averaged over the same periods, are also shown.

Line 184 – I struggle to relate the text here, which says two distinct temporal patterns, to Figure 1b, which does not seem to show a strong peak at ~6 month period. Perhaps a log scale on the y axis might help?

Response: Thank you for this important and constructive comment. We apologize for not clearly describing the preprocessing applied prior to the spectral analysis. The power spectrum shown in Figure 1b was computed from a 12-month running-mean MHWCI anomaly, which acts as a low-pass filter and was intended to emphasize interannual variability. As a result, seasonal variability is strongly dampened and is therefore not expected to appear as a pronounced spectral peak in this figure. To avoid confusion, we have revised this figure (see below) and corresponding texts in the manuscript.



Updated Figure 1. (a) The research domain of the North Sea. The bathymetry (m) is shown on a logarithmic scale. (b) Power spectrum analysis of 12-month running-mean marine heatwave cumulative intensity (MHWCI) anomalies in the North Sea derived from OSTIA. Orange dots indicate periods exceeding the 90% confidence level (black dashed line), while blue dots represent periods below this threshold. The 90% confidence level is calculated using a χ^2 test against a red-noise (AR(1)) background spectrum based on the lag-1 autocorrelation.

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.

Equation 3 – Is the left brace notation intended? The equivalent equation in the referenced paper uses square braces. Also, I'm not clear why the stationary version of the equation is used rather than the full one?

Response: You are correct regarding the notation. We have revised Eq. (3) to use square brackets for consistency with the original presentation in Takaya and Nakamura (2001). The wave activity flux formulation adopted here follows the stationary Rossby wave activity flux as defined by Takaya and Nakamura (2001), which is widely used to diagnose the propagation of quasi-stationary Rossby wave energy in seasonal-mean and large-scale circulation patterns. Since our analysis focuses on seasonally averaged atmospheric teleconnections rather than transient wave evolution, we argue that the stationary version of formulation is appropriate for the purpose of this study.

Updated Eq.3: $$WAF = \frac{p\cos\varphi}{2|U|}\begin{bmatrix} \frac{U}{a^2\cos^2\varphi}\left[\left(\frac{\partial\psi'}{\partial\lambda}\right)^2 - \psi'\frac{\partial^2\psi'}{\partial\lambda^2}\right] + \frac{V}{a^2\cos\phi}\left(\frac{\partial\psi'}{\partial\lambda}\frac{\partial\psi'}{\partial\varphi} - \psi'\frac{\partial^2\psi'}{\partial\lambda\partial\varphi}\right) \\ \frac{U}{a^2\cos\varphi}\left(\frac{\partial\psi'}{\partial\lambda}\frac{\partial\psi'}{\partial\varphi} - \psi'\frac{\partial^2\psi'}{\partial\lambda\partial\varphi}\right) + \frac{V}{a^2}\left[\left(\frac{\partial\psi'}{\partial\varphi}\right)^2 - \psi'\frac{\partial^2\psi'}{\partial\varphi^2}\right] \end{bmatrix}$$

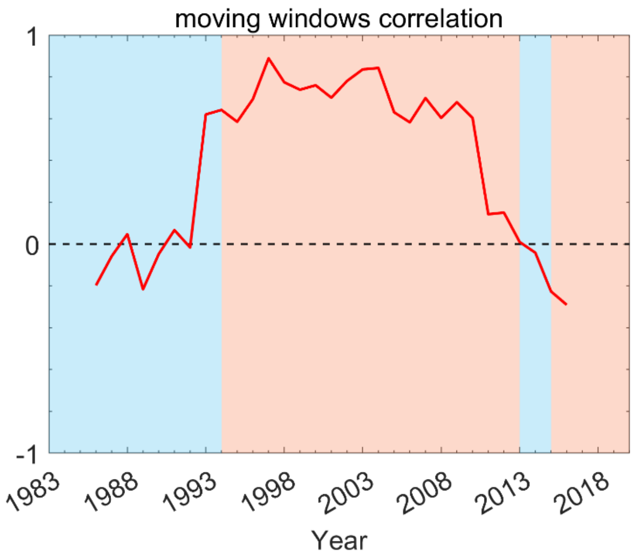
Line 185 – How was the 90% confidence interval calculated?

Response: The 90% confidence level in Fig. 1b was estimated using a chi-square significance test against a stochastic noise background. Specifically, we computed the power spectrum from the detrended monthly MHWCI anomalies, estimated the effective degrees of freedom, and then derived the 90% confidence threshold from the χ^2 distribution. When the lag-1 autocorrelation is positive (and larger than lag-2), we adopt a red-noise (AR(1)) background spectrum following the standard formulation; otherwise a white-noise background is used. Spectral peaks exceeding this 90% threshold are marked as significant in the figure. We have clarified this procedure in the updated caption of Figure 1.

L95-L99: "Figure 1. (a) The research domain of the North Sea. The bathymetry (m) is shown on a logarithmic scale. (b) Power spectrum analysis of 12-month running-mean marine heatwave cumulative intensity (MHWCI) anomalies in the North Sea derived from OSTIA. Orange dots indicate periods exceeding the 90% confidence level (black dashed line), while blue dots represent periods below this threshold. The 90% confidence level is calculated using a χ^2 test against a red-noise (AR(1)) background spectrum based on the lag-1 autocorrelation."

Line 211 – 2013 onwards seems a short period to calculate the correlation, even though it is statistically significant. Do any other similar length periods in the record have low correlation?

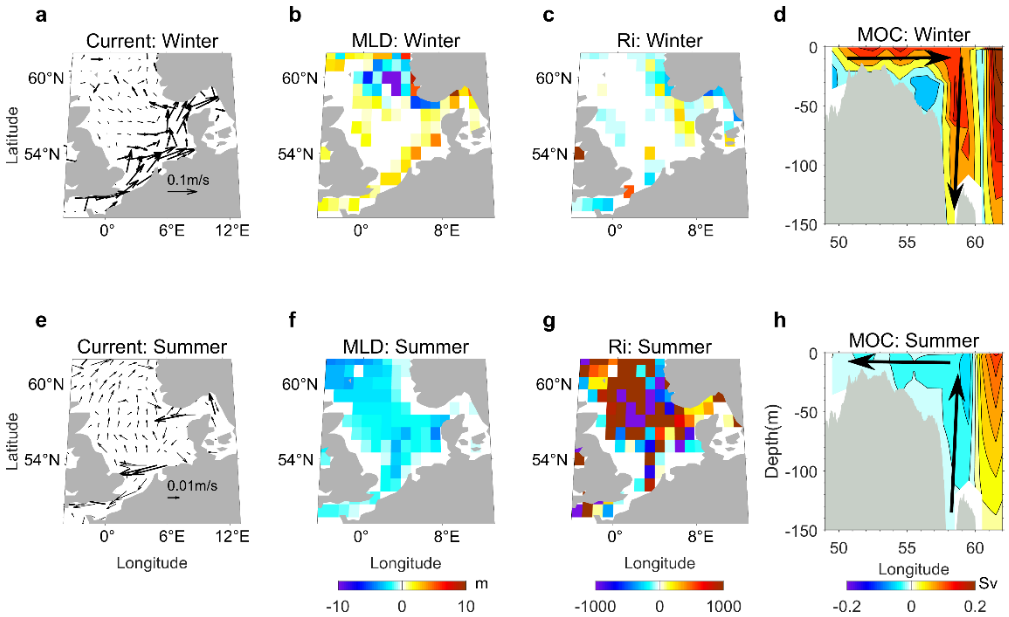
Response: To address this concern, we calculated correlations using moving windows of the same length as the 2013-2022 period across the full time series record. The sliding-window analysis shows that correlations remain consistently high during positive AMV phase (1994-2013), even when the same short window length is applied. In contrast, correlations are weak as those observed after 2013 (Fig. S2). This indicates that the 2013-2022 reduction is not a feature of short windows but instead reflects a distinct weakening of the relationship. We have added this figure as Supplementary Figure S2 in the revised manuscript.



Supplementary Figure S2. Sliding-window correlation between domain-averaged MHWCI for Cluster 2 during summer and Interdecadal Pacific Oscillation (IPO) index using a 9-yr window. blue and red shading indicate negative and positive phases of Atlantic Multidecadal Variability (AMV), respectively. The blue bar from 2013-2015 indicates the Atlantic "cold blob" event.

Figures using arrow vectors (particularly Figure 7) – Is it possible to make the arrows clearer as their direction is not easy to see?

Response: We have improved Figure 7 (see below) by increasing the arrow size to make the flow direction more easily discernible. In addition, we have simplified the vector representation by using arrow length to indicate current speed and a uniform color for all vectors.



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.

Figure 7 – Colour bar for panels a and f uses a diverging colour scale and it would be better to use a non-diverging scale as the scale goes from 0 to 0.02.

Response: Thank you for this helpful suggestion. In the revised Figure 7 (see above), we have further simplified the presentation of the velocity fields. For panels (a) and (f), the flow speed is no longer represented using a color scale. Instead, all vectors are plotted in black, and the current speed is indicated solely by arrow length.

Figure 7 – Are the MOC and ZOC calculated at a particular latitude / longitude?

Response: The MOC shown in Figure 7 are not calculated at a single fixed latitude or longitude. Instead, they are obtained by integrating the velocity field across the entire zonal or meridional extent of the study domain, respectively. Therefore, the resulting MOC represents basin-integrated overturning circulations rather than transports along a particular section. We have clarified this in the revised figure caption (see above).

Figure 7 convergence and overturning circulation colour scales - What direction do negative or positive numbers indicate?

Response: We apologize for not clearly stating this in the original caption. In Figure 7, "Negative stream function values correspond to clockwise circulation, while positive values indicate counterclockwise circulation.". This information has been added to the figure caption for clarity.

Technical comments:

Line 400 – Should it say AMV instead of AMO?

Response: You are right. We have revised the manuscript and consistently use AMV.

Fig 2c and d – Series is misspelt in the titles.

Response: Thank you. The spelling errors have been corrected in the revised figure.

Figure 7d, e, i, j – Suggest adding Latitude or Longitude labels on the x axis.

Response: Thank you for this helpful suggestion. Latitude/longitude labels have been added to the x axis.

Figure 8 – budget is spelt wrong in the title.

Response: The spelling error has been corrected in the revised figure.

Figure 9 – Colour bar should be labelled

Response: The colour bar in Figure 9 is now labelled with the unit (m^2/s^2).

Author's response to Reviewer 2

(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 https://agupubs.onlinelibrary.wiley.com/doi/10.1029/2023AV001059) the procedure is meant to identify groups of data points, while in your case two regions (Fig. 1) are identified.

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.

And from Fig. 3, I have the impression this just means seasonality. How can these two domains be related to MHW occurrence?

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,

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

and *S* is the vertical shear

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

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.

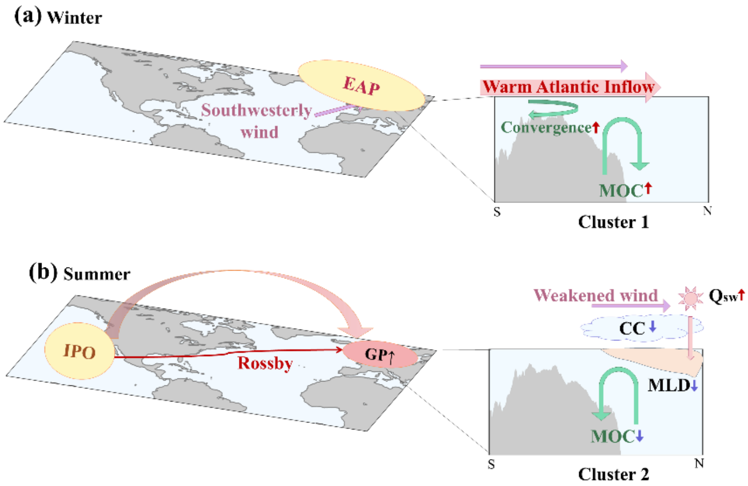
I would suggest to first analyse non-MHW variables and understand how they change over time and due to teleconnections, and then focus on the influence of these indices on MHWs.

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)."

l113 MHWCI definition ain't very clear to me. What happens say if an event starts in Jan 27 and ends on Feb 10? Which month gets the CI? Please specify also the baseline period used.

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."

l154 state the source of this version of the heat tendency equation. Information on the number of levels in ORAS5 should be added, to clarify which is the accuracy of the MLD estimated as such eq3 I imagine these equation is also calculated with monthly mean data?

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.

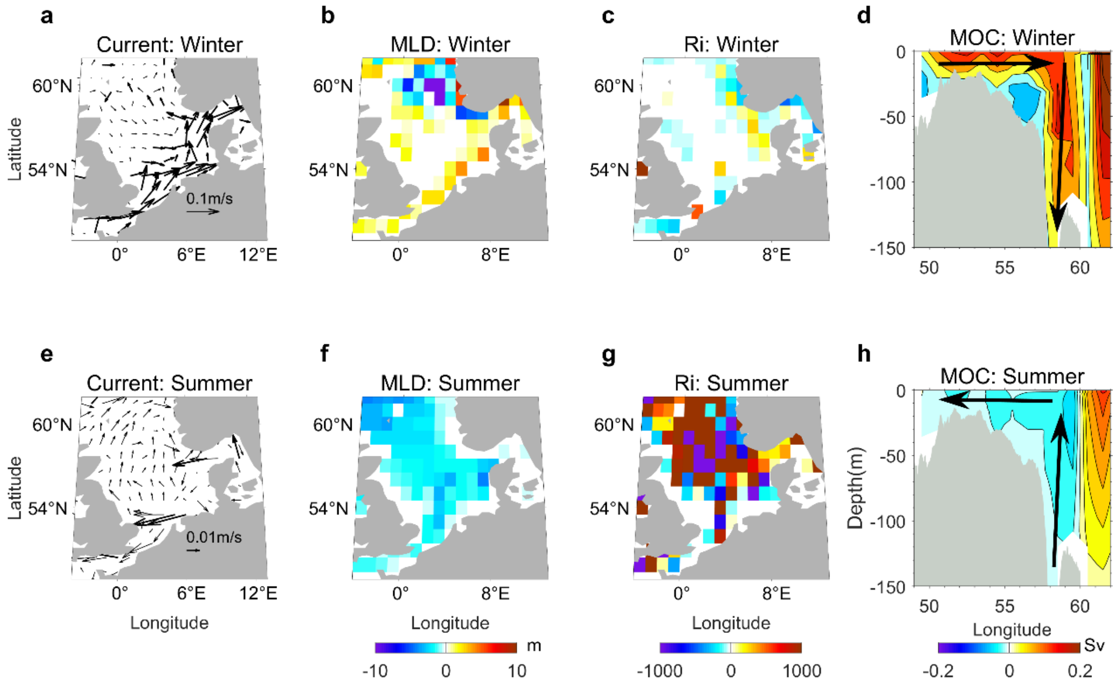
l269 why this value? Why not using some significance threshold?

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"?

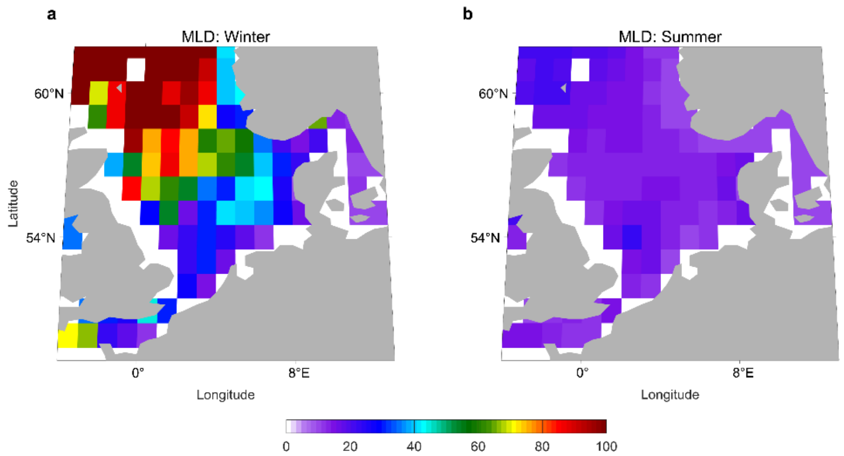
Response: The term "telecommunications" was an error and has been corrected to "teleconnections" in the revised manuscript.

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.

Author's response to Reviewer 3

The manuscript examines the spatio-temporal variability of marine heatwaves (MHWs) in the North Sea and explores how different large-scale climate variabilities interact to shape their occurrence and characteristics. The authors use long-term SST observations, identify two main regional patterns of MHW variability, and relate these patterns to different atmospheric and oceanic climate modes, with a primary focus on seasonal differences.

Overall, the topic is important and highly relevant, especially given the increasing frequency and impacts of marine heatwaves in shelf seas. The focus on the North Sea is well chosen, and the approach of examining the combined influence of climate modes, rather than treating them independently, is interesting. The study appears scientifically sound, and the results are clearly relevant to the community.

However, I believe the manuscript still requires major revisions. My main concern is the lack of several methodological details and the presentation, particularly in the results section, which makes it difficult to follow the main results and understand them. With significant improvements in clarity and structure, the paper could be much stronger.

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.

Introduction

1. The introduction provides useful background, but the aim of the study is not clearly stated. I strongly recommend rewriting the objectives part of the introduction to clearly state the goal(s) of the work and the main research questions.

Response: We agree that the original Introduction does not sufficiently highlight the specific aims of the study. We have revised the Introduction to clearly state the aim of the study and to explicitly formulate the main objectives, including the identification of seasonally distinct spatial domains of MHW variability and the investigation of their associated physical mechanisms.

L81-L91: "In this study, we address the knowledge gap in understanding synergistic impacts of climate modes on MHWs in shelf seas. We focus on the greater North Sea region, which is projected to warm as fast as global levels (Hobday & Pecl, 2014) but exhibits large climate variability caused by interactions between the Arctic and subtropical zones (Quante & Colijn, 2016). This makes it an ideal regional example for understanding how climate modes synergistically influence MHWs in complex shelf sea settings. Specifically, this study aims to (1) identify coherent spatial domains in the North Sea based on the seasonal patterns of interannual marine heatwave cumulative intensity (MHWCI) variability, (2) quantify how the frequency, intensity, and duration of MHWs differ between these domains during their seasonally dominant periods, and (3) examine the associated atmospheric and oceanic processes that modulate the observed regional contrasts. A correlation-based k-means clustering approach is used to characterize MHWCI variability, followed by an examination of the associated large-scale climate variability, air-sea heat fluxes, and oceanic heat transport."

2. Additionally, since the analysis relies heavily on different climate modes, it would be helpful to define these climate modes early on, either in the introduction or in the data/methodology section. And add a brief explanation of how the positive and negative phases of these modes could affect SST variability in the North Sea.

Response: We agree that a clear definition of the climate modes and their relevance to North Sea SST variability should be provided early in the manuscript. We have revised the manuscript to explicitly define all climate modes used in this study and to clarify their physical relevance. Specifically, a new subsection (Sect. 2.4) has been added to the Methods section, which details the definitions, data sources, and calculation methods of the NAO, EAP, and AMV indices. In addition, the section has included a brief explanation of how the positive and negative phases of these climate modes can influence SST variability and marine heatwaves in the North Sea.

L149-L164:

"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)."

3. One minor point: the manuscript sometimes refers to the Atlantic Multidecadal Oscillation as AMO and sometimes as AMV. Please use only one term throughout the manuscript, as switching between them is confusing.

Response: Thank you for pointing this out. We have revised the manuscript with a consistent use of AMV.

Data and Methods

This section requires the most attention, primarily because important details are missing or not clearly described.

1. The data section is not sufficiently clear. I suggest rewriting it in a more straightforward manner, listing the datasets used, their sources, temporal and spatial resolution, and how they are used in the analysis. Currently, it is difficult to understand exactly which products are used and why.

Response: We have rewritten and reorganized the Data section to improve clarity and transparency following your suggestion. The revised version now explicitly lists all datasets used in the study, including their sources, temporal and spatial resolutions, and their specific roles in the analysis.

Specifically, the Data section has been restructured into separate subsections (listed below) for (i) SST data and MHW identification, (ii) atmospheric and oceanic reanalysis data, and (iii) climate mode indices. This reorganization provides a clear description about which data products are used at each stage of the analysis and for what purpose.

L100-L164:

"2 Data and Methods

2.1 Data

2.1.1 SST and MHW Identification

Our analysis of MHWs in the North Sea (Fig. 2a) is based on high-resolution SST data (0.05°×0.05°, daily) obtained from the Copernicus Marine Environment Monitoring Service (CMEMS). This product is generated by the Operational Sea Surface Temperature and Sea Ice Analysis (OSTIA) , which integrates multiple satellite observations and in situ measurements, and covers the period 1982-2021. An independent observation-based SST dataset from the FerryBox system (Macovei et al., 2021), developed by Helmholtz-Zentrum Hereon, is used for validation purposes (Fig. S1).

MHWs, which are defined as individual thermal events during which their associated temperature is larger than the 90th percentile threshold for at least 5 days (Hobday et al., 2016), are detected by using the MATLAB Marine Heatwaves (M_MHW) toolbox (Zhao & Marin, 2019). 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.

The long-term warming trend was subtracted to better estimate the effects of climate modes (Liu et al., 2022). To isolate variability relative to the long-term warming background, SST was detrended before MHW detection. The detrending was performed by identifying a linear trend estimated from the global-mean SST time series over 1982-2021 and subtracting this common trend from the SST time series at each grid point. This procedure reduces the influence of externally forced warming on the identification of marine heatwaves and subsequent analyses.

Characteristics of MHWs, including frequency, duration, mean intensity, and cumulative intensity, were derived following event detection. Among these metrics, cumulative intensity integrates the effects of event duration and mean intensity and provides an effective measure of the overall impact of MHWs (Marin et al., 2021). In this study, the cumulative intensity was summed to derive the monthly MHWCI, representing the combined impact of all MHW events occurring within a given month (Gröger et al., 2024; Mohamed et al., 2023). 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

2.1.2 Atmospheric and oceanic reanalysis data

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).

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).”

2. The SST analysis is limited to 2021, but high-resolution SST products are available at least until the end of 2024. It is unclear why the study stops at 2021. If this limitation is due to other variables or datasets, it should be clearly stated. Otherwise, extending the analysis to include the most recent years would strengthen the study, especially since recent years include strong extremes that could affect the statistics.

Response: Thank you for noting the temporal coverage of the SST analysis. The original manuscript used the CMEMS OSTIA reprocessed analysis product, which is a long-term, homogeneous dataset specifically designed for climate studies and marine heatwave statistics. The reprocessed OSTIA product is generated retrospectively using the OSTIA system with re-processed satellite observations (including ESA SST CCI and C3S products, as well as EUMETSAT and REMSS data) and in situ measurements from the HadIOD dataset, all processed using a consistent system and algorithms. This ensures temporal homogeneity and provides a climate-quality estimate of foundation SST, which is free of diurnal variability.

At the time when the data were processed and the manuscript was prepared, this reprocessed dataset was available only up to 2021, which explains the end year of the original analysis. We note that near-real-time (NRT) OSTIA SST products extend to the present day. However, the NRT-ANALYSIS is a near-real-time hindcast produced with less optimal observational coverage and quality compared to the reprocessed best-analysis product, and may involve temporal inhomogeneities. For this reason, we did not merge the NRT product with the reprocessed dataset for long-term climatological and marine heatwave analyses.

3. Regarding the marine heatwave definition, the authors use the standard Hobday et al. method, but the climatology period used to define the MHW threshold is not mentioned. It must be included in the methodology. Also, please clarify whether the same climatology baseline period is used when computing other anomalies in the study, as this is not currently clear.

Response: The MHWs are identified using the standard Hobday et al. methodology, with the climatological mean and threshold computed over the 1982-2021 period, consistent with the temporal coverage of the SST dataset. In addition, all other anomalies analyzed in this study are calculated relative to the same 1982-2021 climatological baseline. We have clarified these points in the Methods section.

L110-L116: “MHWs, which are defined as individual thermal events during which their associated temperature is larger than the 90th percentile threshold for at least 5 days (Hobday et al., 2016), are detected by using the MATLAB Marine Heatwaves (M_MHW) toolbox (Zhao & Marin, 2019). 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.”

L118-L122: “The long-term warming trend was subtracted to better estimate the effects of climate modes (Liu et al., 2022). To isolate variability relative to the long-term warming background, SST was detrended before MHW detection. The detrending was performed by identifying a linear trend estimated from the global-mean SST time series over 1982-2021 and subtracting this common trend from the SST time series at each grid point. This procedure reduces the influence of externally forced warming on the identification of marine heatwaves and subsequent analyses.”

4. Another confusing aspect is the detrending or trend removal. In lines 109–110, the manuscript refers to trended and detrended data, but it is not clearly explained: how the detrending is performed? and how it affects the results? This needs a clearer explanation, as detrending choices can influence both clustering and correlations with climate modes.

Response: The detrending procedure has been clarified in the revised manuscript. Specifically, long-term warming is identified by estimating a linear trend of the global-mean SST over the study period (1982-2021). This global trend is then subtracted from the original SST field prior to marine heatwave detection. This approach follows common practice in MHW studies and is intended to remove the externally forced warming signal while retaining internal variability. As a result, the clustering and the relationships with climate modes reflect variability relative to the evolving background state, rather than being dominated by the long-term warming trend. We have added a description of this procedure and its rationale in the Methods section.

L118-L122: “The long-term warming trend was subtracted to better estimate the effects of climate modes (Liu et al., 2022). To isolate variability relative to the long-term warming background, SST was detrended before MHW detection. The detrending was performed by identifying a linear trend estimated from the global-mean SST time series over 1982-2021 and subtracting this common trend from the SST time series at each grid point. This procedure reduces the influence of externally forced warming on the identification of marine heatwaves and subsequent analyses.”

5. I also recommend a minor change in the subsection title in the methods (line 106). The subtitle is “Cumulative Intensity of Marine Heatwaves,” but the section actually describes the overall MHW calculation and detection method and includes multiple metrics. A title such as “Marine Heatwave Calculation” would be more accurate and clearer.

Response: We agree that the scope of the original subsection title was too narrow. The subsection title has been updated to more accurately reflect the content of the section.

6. I was confused by the calculation of the “cumulative intensity anomaly.” Since MHWs are already defined as anomalous events, it is not clear why an additional anomaly is needed. If the authors choose to keep this variable, they need to clearly explain its importance and what additional information it provides.

Response: Thank you for raising this point. The MHWCI anomaly is introduced for the purpose of our study. In this study, the cumulative intensity is used to quantify the overall impact of all MHW events occurring within a given year or season. The anomaly of this metric is then calculated to identify interannual variability, allowing a direct comparison of how active or inactive different years are relative to the long-term mean. We acknowledge that this distinction was not sufficiently explained in the current manuscript. In the revised version, we have clarified the rationale for using cumulative intensity anomalies and explained the additional information they provide beyond the event-based MHW definition.

L126-L132: "In this study, the cumulative intensity was summed to derive the monthly MHWCI, representing the combined impact of all MHW events occurring within a given month (Gröger et al., 2024; Mohamed et al., 2023). 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."

7. Since the analysis relies heavily on seasons, the seasonal divisions should be clearly stated and justified. Currently, there is no clear explanation of why these particular seasonal divisions are chosen or what they represent physically for the North Sea.

Response: Thank you for this comment. We agree that although the seasonal divisions (DJF, MAM, JJA, SON) are explicitly defined in the Methods section, their physical relevance to the North Sea was not sufficiently explained.

These seasonal divisions correspond to well-established physical regimes in the North Sea, including weakening of the winter thermal stratification and summer stratification with shallow mixed layers. Similar seasonal divisions have been adopted in previous studies on North Sea MHWs, which highlight pronounced contrasts between winter and summer mechanisms (Chen & Staneva, 2024; Mohamed et al., 2023). We have clarified this in the revision.

L182-L186: "The four seasons are defined as winter (December of the previous year and January-February of the current year, DJF), spring (March-May, MAM), summer (June-August, JJA), and autumn (September-November, SON), which correspond to well-established physical regimes in the North Sea (Chen & Staneva, 2024; Mohamed et al., 2023), including wintertime deep mixing and weak stratification, and summertime strong stratification with shallow mixed layers."

Results

For me, the Results section is currently difficult to follow, not because the figures are unclear, but because the text does not clearly describe the results shown in the figures. In many paragraphs, the text moves directly to interpretation and discussion without first explaining what is actually being observed. As a result, the reader must go back and forth between the figures and the text to understand what is being claimed.

1. I strongly recommend rewriting the Results section to be more descriptive: each figure should be clearly explained first (what pattern is shown, what changes occur, what differences appear), and only after that should the interpretation be introduced. This would make the manuscript much easier to read.

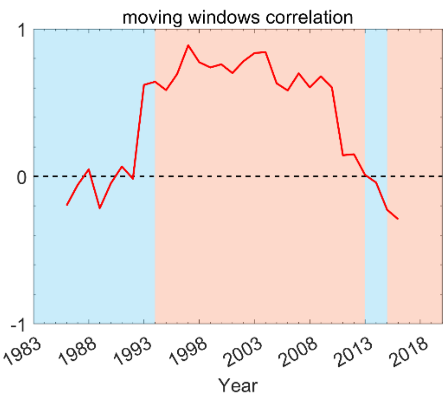
Response: Thank you for this constructive suggestion. Following this recommendation, we have substantially revised the entire Results section to adopt a more descriptive, step-by-step presentation. Specifically, we rewrote Sections 3.2 and 3.3 so that each figure is now introduced and described explicitly before further interpretation is provided. For each figure, we first describe the dominant spatial patterns, the sign and magnitude of anomalies, and the key differences between regions or phases. After explaining the basic elements of each figure, we introduce physical interpretations and linkages to MHW characteristics. In addition, we reorganized the Results to ensure a clear top-down progression from large-scale atmospheric circulation to surface forcing, upper-ocean structure (e.g., mixed layer depth and Richardson number), ocean circulation diagnostics (e.g., overturning), and finally MHW characteristics. This restructuring ensures that interpretations are based on clearly described observational evidence, improving readability and logical flow throughout the Results section.

2. Additionally, in lines 355-361, the manuscript separates the analysis into periods before and after 2013, but it is not clear why 2013 was chosen as a breakpoint. Where is this significant in the results? What evidence supports this split? This section needs more explanation.

Response: Thank you for raising this point. We would like to clarify that this choice is not arbitrary but is supported by both our results and independent evidence from existing literature.

First, our moving-window correlation analysis (Fig. S1, shown below) reveals a clear change in the relationship between summer MHWCI in Cluster 2 and the IPO around 2013, indicating a shift in the teleconnection structure at that time. We have added this figure as supplementary of revised manuscript to support this.

Second, previous studies have documented an abrupt cooling in the subpolar North Atlantic during 2013-2014, often referred to as the development of the “cold anomaly” or “cold blob,” which coincides with a weakening of the AMV signal during the early 2010s (Frajka-Williams et al., 2017; Josey & Sinha, 2022; Moat et al., 2020). In the revised manuscript, we have clarified this by explicitly linking the period separation to the moving-window correlation results and by citing relevant studies documenting the North Atlantic cooling and the contemporaneous weakening of the AMV.



Supplementary Figure S2. Sliding-window correlation between domain-averaged MHWCI for Cluster 2 during summer and Interdecadal Pacific Oscillation (IPO) index using a 9-yr window. blue and red shading indicate negative and positive phases of Atlantic Multidecadal Variability (AMV), respectively. The blue bar from 2013-2015 indicates the Atlantic "cold blob" event.

L309-L315: "In Cluster 2, the summer-averaged MHWCI displays pronounced decadal modulation (Fig. 2d). During the negative AMV phase (1982-1994, blue shading in Fig. 2d), MHWCI exhibits relatively weak intensity. In contrast, during the positive AMV phase (1994-2013, red shading in Fig. 2d), MHWCI shows notable enhancement, with its interannual variability strongly correlated with both ENSO and the Interdecadal Pacific Oscillation (IPO) ($r = 0.69$ and 0.70, respectively). However, this teleconnection substantially weakens after 2013 (Fig. S2), with correlations becoming statistically insignificant ($p > 0.1$), indicating a substantial reduction in Pacific influence on North Sea MHWs."

L516-L523: "After 2013, previous studies have documented an abrupt cooling in the subpolar North Atlantic (Mooney, 2015), often referred to as the development of the "cold anomaly" or "cold blob," which coincides with a weakening of the AMV signal during the early 2010s (Frajka-Williams et al., 2017; Josey & Sinha, 2022; Moat et al., 2020). During this period, the tropical and subtropical Pacific exhibit markedly diminished atmospheric responses (Fig. 10d-f). Tropical Pacific SST anomalies weaken substantially compared to 1994-2012, accompanied by the reductions in precipitation. Moreover, the geopotential anomalies in the North Pacific weaken significantly (Fig. 9b), and the dipole circulation pattern between the subpolar and subtropical Pacific substantially weakens. WAF analysis reveals a marked disruption of Rossby wave energy propagation from the Pacific to the Atlantic (Fig. 9b)."

Minor Comments

L26–27: The sentence is long and somewhat difficult to understand; consider splitting it for clarity.

Response: This sentence has been split to shorter sentences to improve clarity.

L24-L26: "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. "

L106–112: Please clarify the detrending approach and its spatial application.

Response: The detrending is based on a linear trend estimated from the global-mean SST time series over 1982-2021, which represents the externally forced warming signal. This common trend is then subtracted from the SST time series at each grid point prior to marine heatwave detection. We have clarified this spatial application of the detrending procedure in the revised manuscript.

L118-L122: "The long-term warming trend was subtracted to better estimate the effects of climate modes (Liu et al., 2022). To isolate variability relative to the long-term warming background, SST was detrended before MHW detection. The detrending was performed by identifying a linear trend estimated from the global-mean SST time series over 1982-2021 and subtracting this common trend from the SST time series at each grid point. This procedure reduces the influence of externally forced warming on the identification of marine heatwaves and subsequent analyses."

L118–124: Seasonal definitions should be explicitly stated.

Response: The seasonal definitions have been explicitly stated in the Methods section.

L182-L186: "The four seasons are defined as winter (December of the previous year and January-February of the current year, DJF), spring (March-May, MAM), summer (June-August, JJA), and autumn (September-November, SON), which correspond to well-established physical regimes in the North Sea (Chen & Staneva, 2024; Mohamed et al., 2023), including wintertime deep mixing and weak stratification, and summertime strong stratification with shallow mixed layers."

L194: The sentence structure is heavy; consider rephrasing.

Response: We have rephrased this sentence to improve clarity and readability.

L274-L282: "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."

Reference used in the response

Chen, W., & Staneva, J. (2024). Characteristics and trends of marine heatwaves in the northwest European Shelf and the impacts on density stratification. *8th edition of the Copernicus Ocean State Report (OSR8)*, *4-osr8*, 7. https://doi.org/10.5194/sp-4-osr8-7-2024

Frajka-Williams, E., Beaulieu, C., & Duchez, A. (2017). Emerging negative Atlantic Multidecadal Oscillation index in spite of warm subtropics. *Scientific Reports*, *7*(1), 11224. https://doi.org/10.1038/s41598-017-11046-x

Good, S., Fiedler, E., Mao, C., Martin, M. J., Maycock, A., Reid, R., Roberts-Jones, J., Searle, T., Waters, J., While, J., & Worsfold, M. (2020). The Current Configuration of the OSTIA System for Operational Production of Foundation Sea Surface Temperature and Ice Concentration Analyses. *Remote Sensing*, *12*(4), 720. https://www.mdpi.com/2072-4292/12/4/720

Gröger, M., Dutheil, C., Börgel, F., & Meier, M. H. E. (2024). Drivers of marine heatwaves in a stratified marginal sea. *Climate Dynamics*, *62*(5), 3231-3243. https://doi.org/10.1007/s00382-023-07062-5

Hersbach, H., Bell, B., Berrisford, P., Hirahara, S., Horányi, A., Muñoz Sabater, J., Nicolas, J., Peubey, C., Radu, R., Schepers, D., Simmons, A., Soci, C., Abdalla, S., Abellan, X., Balsamo, G., Bechtold, P., Biavati, G., Bidlot, J., Bonavita, M., & Thépaut, J. N. (2020). The ERA5 global reanalysis. *Quarterly Journal of the Royal Meteorological Society*, *146*. https://doi.org/10.1002/qj.3803

Hobday, A. J., Alexander, L. V., Perkins, S. E., Smale, D. A., Straub, S. C., Oliver, E. C. J., Benthuysen, J. A., Burrows, M. T., Donat, M. G., Feng, M., Holbrook, N. J., Moore, P. J., Scannell, H. A., Sen Gupta, A., & Wernberg, T. (2016). A hierarchical approach to defining marine heatwaves. *Progress in Oceanography*, *141*, 227-238. https://doi.org/https://doi.org/10.1016/j.pocean.2015.12.014

Hobday, A. J., & Pecl, G. T. (2014). Identification of global marine hotspots: sentinels for change and vanguards for adaptation action. *Reviews in Fish Biology and Fisheries*, *24*(2), 415-425. https://doi.org/10.1007/s11160-013-9326-6

Josey, S. A., & Sinha, B. (2022). Subpolar Atlantic Ocean mixed layer heat content variability is increasingly driven by an active ocean. *Communications Earth & Environment*, *3*(1), 111. https://doi.org/10.1038/s43247-022-00433-6

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

Macovei, V. A., Voynova, Y. G., Gehrung, M., & Petersen, W. (2021). Ship-of-Opportunity, FerryBox-integrated, membrane-based sensor pCO2, temperature and salinity measurements in the surface North Sea since 2013. In: PANGAEA.

Marin, M., Feng, M., Phillips, H. E., & Bindoff, N. L. (2021). A Global, Multiproduct Analysis of Coastal Marine Heatwaves: Distribution, Characteristics, and Long-Term Trends. *Journal of Geophysical Research: Oceans*, *126*(2), e2020JC016708. https://doi.org/https://doi.org/10.1029/2020JC016708

Moat, B. I., Smeed, D. A., Frajka-Williams, E., Desbruyères, D. G., Beaulieu, C., Johns, W. E., Rayner, D., Sanchez-Franks, A., Baringer, M. O., Volkov, D., Jackson, L. C., & Bryden, H. L. (2020). Pending recovery in the strength of the meridional overturning circulation at 26° N. *Ocean Sci.*, *16*(4), 863-874. https://doi.org/10.5194/os-16-863-2020

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

Mohamed, B., Barth, A., Van der Zande, D., & Alvera-Azcárate, A. (2025). Amplified warming and marine heatwaves in the North Sea under a warming climate and their impacts. *Ocean Sci.*, *21*(5), 2505-2525. https://doi.org/10.5194/os-21-2505-2025

Quante, M., & Colijn, F. (2016). *North Sea Region Climate Change Assessment*. Springer International Publishing : Imprint : Springer. https://doi.org/10.1007/978-3-319-39745-0

Trenberth, K. E., & Shea, D. J. (2006). Atlantic hurricanes and natural variability in 2005. *Geophysical Research Letters*, *33*(12). https://doi.org/https://doi.org/10.1029/2006GL026894

Worsfold, M., Good, S., Atkinson, C., & Embury, O. (2024). Presenting a Long-Term, Reprocessed Dataset of Global Sea Surface Temperature Produced Using the OSTIA System. *Remote Sensing*, *16*(18), 3358. https://www.mdpi.com/2072-4292/16/18/3358

Zhao, Z., & Marin, M. (2019). A MATLAB toolbox to detect and analyze marine heatwaves Software. *Journal of Open Source Software*, *4*. https://doi.org/10.21105/joss.01124

Zuo, H., Balmaseda, M. A., Tietsche, S., Mogensen, K., & Mayer, M. (2019). The ECMWF operational ensemble reanalysis–analysis system for ocean and sea ice: a description of the system and assessment. *Ocean Sci.*, *15*(3), 779-808. https://doi.org/10.5194/os-15-779-2019