



Seasonal forecasting of the European North-West shelf seas: limits of winter and summer sea surface temperature predictability

Jamie R. C. Atkins¹ · Jonathan Tinker² · Jennifer A. Graham^{3,4} · Adam A. Scaife^{1,2} · Paul R. Halloran¹

Received: 22 February 2024 / Accepted: 3 September 2024 / Published online: 18 September 2024
© The Author(s) 2024

Abstract

The European North-West shelf seas (NWS) support economic interests and provide environmental services to adjacent countries. Expansion of offshore activities, such as renewable energy infrastructure, aquaculture, and growth of international shipping, will place increasingly complex demands on the marine environment over the coming decades. Skilful forecasting of NWS properties on seasonal timescales will help to effectively manage these activities. Here we quantify the skill of an operational large-ensemble ocean-atmosphere coupled global forecasting system (GloSea), as well as benchmark persistence forecasts, for predictions of NWS sea surface temperature (SST) at 2–4 months lead time in winter and summer. We identify sources of and limits to SST predictability, considering what additional skill may be available in the future. We find that GloSea NWS SST skill is generally high in winter and low in summer. GloSea outperforms simple persistence forecasts by adding information about atmospheric variability, but only to a modest extent as persistence of anomalies in the initial conditions contributes substantially to predictability. Where persistence is low – for example in seasonally stratified regions – GloSea forecasts show lower skill. GloSea skill can be degraded by model deficiencies in the relatively coarse global ocean component, which lacks dynamic tides and subsequently fails to robustly represent local circulation and mixing. However, “atmospheric mode matched” tests show potential for improving prediction skill of currently low performing regions if atmospheric circulation forecasts can be improved. This underlines the importance of coupled atmosphere-ocean model development for NWS seasonal forecasting applications.

Keywords Seasonal forecasting · Predictability · Shelf seas · Sea surface temperature · Atmospheric circulations · Dynamical forecasting systems

1 Introduction

Near-term climate predictions are vital to both land (Ely et al. 2013; Palin et al. 2016; Clark et al. 2017; Turco et al. 2017; Bell et al. 2017; Ceglar et al. 2018) and marine sector applications (Stock et al. 2015; Hobday et al. 2016; Mills et al. 2017; Payne et al. 2017; Scott et al. 2021). Traditionally,

however, there has been less work on understanding and predicting seasonal variability of shelf sea environments, where shallow bathymetry creates a dynamic physical environment subject to high variability via energetic mixing processes as well as the influence of atmospheric variability (Otto et al. 1990; Huthnance 1991; Becker and Pauly 1996; Dippner 1997; Pätsch et al. 2017). This knowledge gap is especially pertinent given the disproportionately high importance of shelf seas to humans and ecosystems. For example, despite occupying just 7% of global ocean area they support over 90% of the world’s fisheries (Pauly et al. 2002; Kröger et al. 2018).

We identify the European North-West shelf seas (hereafter NWS; area within the 200 m isobath marked in Fig. 1) as a prime region of interest. The operations of many NWS industries are sensitive to variability of physical fields, especially temperatures, and could benefit from accurate forecasts months ahead. Such industries include fisheries,

✉ Jamie R. C. Atkins
ja661@exeter.ac.uk

¹ Faculty of Environment, Science and Economy, University of Exeter, Exeter, UK

² UK Met Office Hadley Centre, Exeter, UK

³ Centre for Environment, Fisheries and Aquaculture Science (Cefas), Lowestoft, UK

⁴ School of Environmental Sciences, University of East Anglia, Norwich, UK

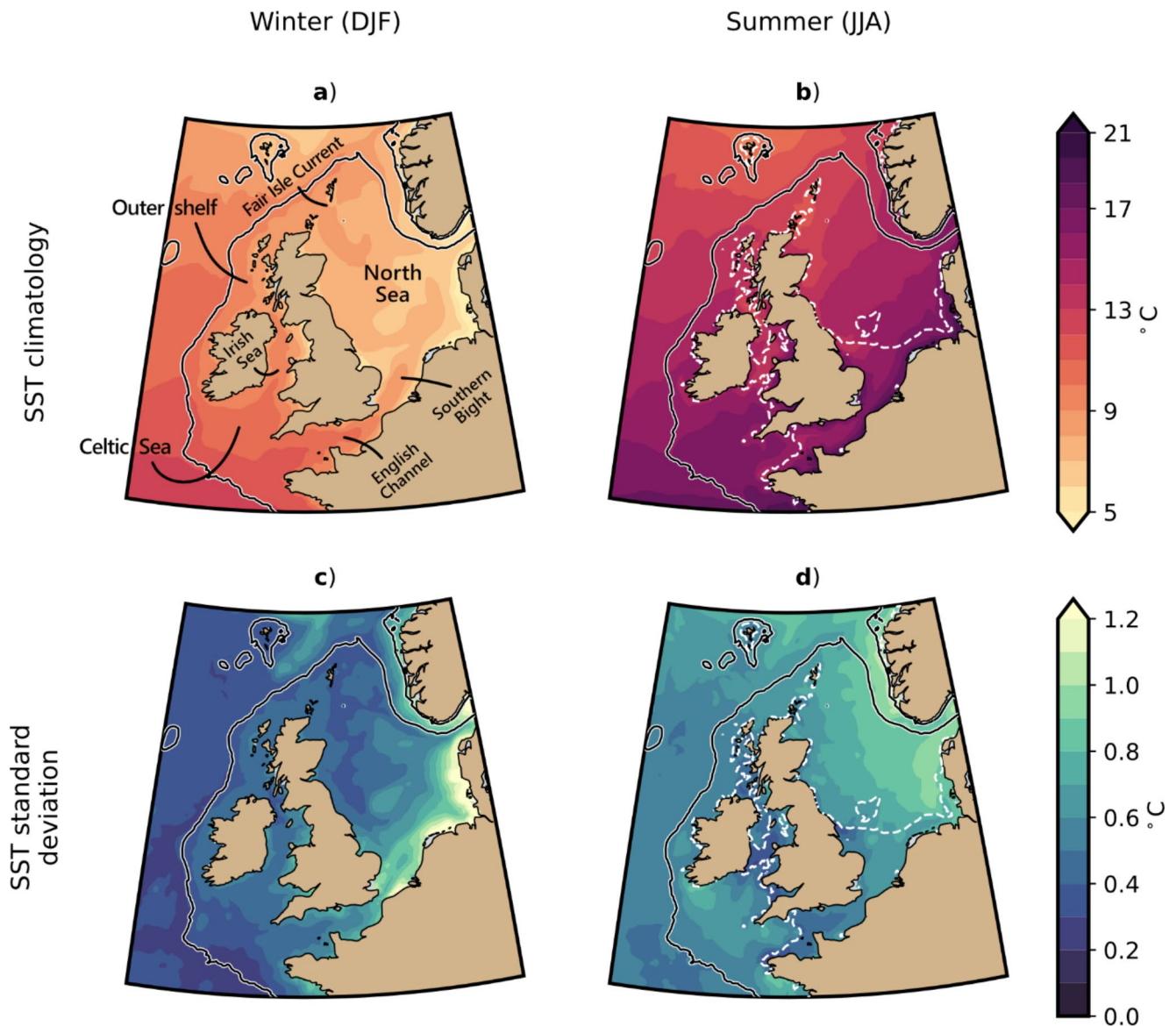


Fig. 1 Climatologies and variability of the European North-West shelf seas. CMEMS-v5 SST climatologies (**a**, **b**) and standard deviations (**c**, **d**) across the period 1993–2021 in winter (DJF; panel column **a**, **c**) and summer (JJA; column **b** – **d**). Black line marks the 200 m iso-

bath. Mixing fronts are shown by white dashed contours in **b**) and **d**). Key NWS regions are marked by text annotations in **a**)

renewable and non-renewable energy, transport and recreation (Payne et al. 2017, 2019). Interannual sea surface temperature (SST) variability can be on the order of degrees Celsius in large areas of the NWS (Fig. 1; Tinker et al. 2018) and seasonal climatologies will therefore not suffice for making predictions of upcoming conditions. Whilst 6-day operational forecasts (Tonani et al. 2019) and uninitialized end-of-century climate projections (Tinker et al. 2016) exist for the physical marine environment of the NWS, no operational NWS seasonal product currently exists. Tinker et al. (2018) and Tinker and Hermanson (2021) began to explore the prospect for NWS seasonal forecasting and suggest

there may be potential predictability of temperatures on the NWS in boreal winter (DJF). There is likely also scope for predictability in other seasons which have not yet been explored, for example boreal summer (JJA), albeit limited by internal predictability limits as well as ocean or atmospheric modelling deficiencies.

The shallow NWS environment has lower thermal inertia than the deep open ocean and is primarily driven by variability in the atmosphere (Becker and Pauly 1996; Dippner 1997; Sharples et al. 2006). For example, from a simple box model, Holt et al. (2010) suggested that surface fluxes could be expected to be twice as important as lateral heat transport

for determining the heat content. However, advection is likely to play a more substantial role around the shelf break. Here, whilst sub-mesoscale processes may occur, particularly sub-surface, on-shelf transport in the surface layer is primarily influenced by wind-driven Ekman transport (e.g. Huthnance et al. 2009; Graham et al. 2018).

North Atlantic – European domain-scale atmospheric circulation patterns, most notably the North Atlantic Oscillation (NAO), are well-known determinants of seasonal surface climate over Europe (Hurrell 1995; Hurrell and Deser 2009). Capturing changes in the dominant circulation patterns in operational seasonal forecasting systems remains a challenge (e.g. Kim et al. 2012), with the recent exception of winter (DJF) surface NAO (Riddle et al. 2013; Scaife et al. 2014; Palin et al. 2016; Clark et al. 2017; Athanasiadis et al. 2017; Baker et al. 2018; Thornton et al. 2023). The summer NAO is also known to impact European surface climate (Folland et al. 2009; Bladé et al. 2012) but is found to have a much weaker predictable signal in dynamical forecasting systems (Patterson et al. 2022; Dunstone et al. 2023). Additional circulation patterns include the East Atlantic pattern (EA), East Atlantic – Western Russia pattern (EAWR) and Scandinavian pattern (SCAND), which are also known to impact European surface climate (Barnston and Livezey 1987; Bueh and Nakamura 2007; Woollings et al. 2010; Lim 2015; Wang and Tan 2020). Recent studies suggest that some of these lower-order modes may be predictable at certain times of year, for example in summer (Lledó et al. 2020) and specifically EA in late autumn – early winter (Thornton et al. 2023), but the literature has thus far centred predominantly on NAO as the leading predictable mode of variability in winter.

Here, we quantify the skill of a large ensemble global ocean-atmosphere coupled dynamical seasonal forecasting system, namely the UK Met Office Global Seasonal Forecasting System (GloSea), in making predictions of the key end-user relevant NWS variable, sea surface temperature (SST), using retrospective forecasts (hereafter hindcasts) across 1993–2016. We extend existing NWS seasonal forecasting work (Tinker et al. 2018; Tinker and Hermanson 2021) to consider both winter (DJF) and summer (JJA) seasons, when the dynamics of the marine environment are markedly different, and focus on the impact of atmospheric variability and the inherent predictability offered by the dynamics of the shelf sea waters themselves.

In this paper we directly assess the seasonal predictability of NWS SST. To do so, we first characterise the influence of atmospheric circulation patterns on NWS SST (Sect. 3.1), and their predictability (Sect. 3.2), before presenting the current skill levels of GloSea and persistence forecasts of NWS SST (Sect. 3.3) and tests of upper limits to NWS SST skill with the use of idealised atmospheric fields (Sect. 3.4 and

3.5). In Sect. 4, we discuss the implications of our findings and consider the potential for improved predictability on the NWS. We describe our data and methodology in Sect. 2.

2 Data and methodology

2.1 Persistence hindcasts

Persistence analysis is employed here for two purposes. First, it serves as a skill benchmark which the more sophisticated GloSea system should aim to beat. Second, persistence will be a scientific tool for quantifying the inherent memory of waters (thermal inertia) as a potential source of predictability on the NWS.

Here, persistence hindcasts are empirical models built using reanalysis data of the physical marine environment of the NWS. The data is sourced from version 5 of the regional physical reanalysis of the NWS produced by the UK Met Office and distributed by the Copernicus Marine Environment Monitoring Service (hereafter CMEMS-v5; Renshaw et al. 2021). The reanalysis is based on the Forecasting Assimilation Model 7 km Atlantic Margin model (FOAM AMM7) using the NEMO version 3.6 ocean model (Madec and the NEMO team 2016). The model is run for the domain 20°W – 13°E, 40°N – 65°N at ~7 km horizontal resolution, across the period 1993 – near-present. The original model simulation has 51 terrain-following vertical levels and data is distributed on 24 vertical geopotential levels. CMEMS-v5 has been extensively validated (Renshaw et al. 2021) and forms a ‘best estimate’ of the observational truth on the shelf. At ~7 km horizontal resolution, CMEMS-v5 includes dynamic tides and resolves key shelf sea processes such as seasonal stratification.

The process for building a timeseries of persistence hindcasts at 2–4 months lead time from CMEMS-v5 data per target season is as follows: SST anomaly data are averaged over a 15-day period centred on the first day of the month preceding the target season, per hindcast year. For example, a winter (DJF) predictor field is formulated from the 15-day average centred on the November 1st immediately preceding that winter. The hindcast prediction is then taken as persistence of this fixed field through the winter. The 15-day averaging period is selected to reflect the span of start dates and information available for a typical GloSea hindcast set-up (as discussed in Sect. 2.2). The hindcast data are constrained to winter and summer of 1993–2016 to match the length of the GloSea hindcast period.

2.2 GloSea large ensemble hindcasts

We use operational hindcast data from versions 5 and 6 of the UK Met Office Global Seasonal Forecasting System (GloSea; MacLachlan et al. 2015). Both versions use the HadGEM3 ocean-atmosphere-land-sea-ice coupled climate model. GloSea6 uses HadGEM3-GC3 (Williams et al. 2018) whereas GloSea5 uses HadGEM3-GC2 (Williams et al. 2015) but both versions show similar skill. The model is initialised with observational analyses of the atmosphere, land, ocean and sea ice. The global ocean component of the model (NEMO; Megann et al. 2014) is run on the ORCA025 grid, i.e. a 0.25° horizontal resolution tri-polar grid with 75 vertical levels, 18 (24) of which are in the top 50 m (100 m) of the water column. The model's atmospheric component is run at N216 resolution (0.8° in latitude and 0.5° in longitude) with 85 vertical levels. Operationally, GloSea produces real-time forecasts as well as hindcasts for the period 1993–2016, which we use here for skill assessment. GloSea employs a lagged ensemble technique; the hindcasts are generated on four calendar dates per month: the 1st, 9th, 17th and 25th and are integrated forward for the next six simulated months. For each start date, seven individual members (differing only by stochastic perturbation of atmospheric initial conditions; i.e. Bowler et al. 2009) are initialised. All initialisations use only conditions which would have been available at the time.

To obtain an ensemble of hindcast simulations at 2–4 months lead time per target season, we combine the 21 members from start dates centred on the 1st of the month preceding the target season. For example, a winter (DJF) hindcast ensemble combines members initialised on October 25th, November 1st, November 9th (3 start dates \times 7 members) per hindcast year. Multiple hindcasts have been run and we are able to combine all available data from five hindcasts to form a larger ensemble of 105 members (5 \times 21-member ensembles) per year. We take this step to provide a hindcast ensemble for skill assessment which best represents the real-time forecasting set up, which is soon due to increase from a 42- to 100-member ensemble. Using such a large ensemble has the notable advantage of being able to better extract predictable signals in the extra-tropics, which are thought to be underrepresented in models (Eade et al. 2014; Scaife and Smith 2018).

The linear trend across the hindcast period is preserved in all analyses (except when defining atmospheric circulation patterns, see Sect. 2.3) to match real-time forecasting procedures, which are capable of capturing skill originating from long-term trends.

2.3 Atmospheric circulation identification

Atmospheric circulations over the North Atlantic – European domain likely impact NWS SST variability due to the fast response of shallow seas to atmospheric forcing (Becker and Pauly 1996; Dippner 1997; Sharples et al. 2006). Here, we investigate dominant circulation patterns in winter and summer as identified by the three leading modes of atmospheric variability. As the circulations will have different centres of action which change across seasons, a fixed point (or box) difference definition for each mode is not suitable for our investigation. We employ an empirical orthogonal function (EOF; Hannachi et al. 2007; Wilks 2011) analysis of ERA5 (Hersbach et al. 2020) geopotential height anomalies at 500 hPa (hereafter Z500), temporally detrended and weighted by the square root of the cosine of the latitude (North et al. 1982) across the domain 90°W–40°E and 20°–80°N (as in Hall and Hanna 2018), to define the three leading atmospheric circulation patterns per season. ERA5 Z500 fields are projected onto the EOFs to generate observation-based principal component timeseries per circulation pattern. EOFs are calculated using the xeof Python package (Rieger and Levang 2024) across the period 1979–2021. Z500 EOFs are denoted by EOF_{Z500} subscript notation.

The loading pattern of each ERA5-derived atmospheric circulation is presented in Fig. 2. In both seasons, EOF1_{Z500} closely resembles the North Atlantic Oscillation (NAO) and EOF2_{Z500} resembles the East Atlantic pattern (EA). EOF3_{Z500} most closely resembles the East Atlantic – Western Russia pattern in the winter (EAWR; though given the domain constraints it may be missing the eastern low pressure node) and the Scandinavian pattern (SCAND) in the summer. GloSea full ensemble Z500 fields from the same domain are projected onto the ERA5-derived loading patterns to get GloSea principal component predictions for the three atmospheric circulation patterns per season and for each individual member.

2.4 Hindcast skill measures

Hindcast skill is assessed by validation against available reanalyses (hereafter observation-based products). We focus predominantly on the Pearson Anomaly Correlation Coefficient skill measure (ACC; Jolliffe and Stephenson 2011), i.e.

$$ACC = \frac{\sum_{i=1}^N (h_i - \bar{h})(v_i - \bar{v})}{\sqrt{\sum_{i=1}^N (h_i - \bar{h})^2 (v_i - \bar{v})^2}} \quad (1)$$

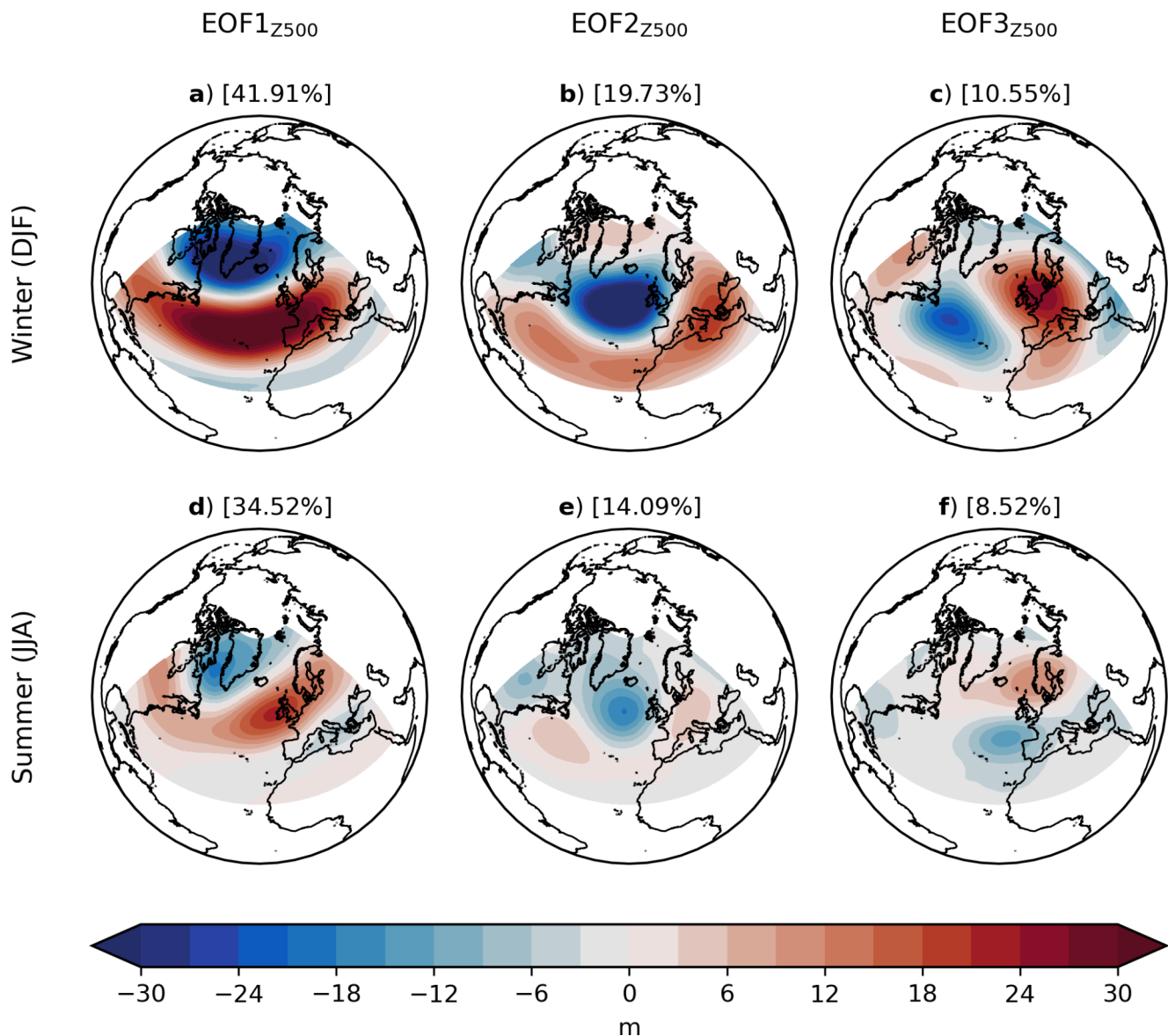


Fig. 2 Leading modes of North Atlantic – European atmospheric variability. The three leading modes of atmospheric variability (EOF_{Z500} 1–3; panel columns) derived from ERA5 Z500 anomalies, across the 1979–2021 reanalysis period, in winter (DJF; row a – c) and summer

(JJA; d – f). Percentage of total variance explained by each mode is marked in square brackets. Loading patterns are expressed as linear regressions on the standardised principal components

where N is the number of years in the hindcast period, h_i and v_i are the hindcast and verification fields, respectively, and overbars denote timeseries means. For validation of winter and summer NWS SST skill, v is taken as CMEMS-v5 SST seasonal mean fields. Note that for GloSea validation, NWS SST data on the ORCA025 tri-polar grid are re-gridded to CMEMS-v5 specification by nearest-neighbour interpolation. For atmospheric circulation validation, v is taken as ERA5-derived winter and summer principal components. ACC is a useful measure of the extent to which the simulated phase of variability matches observations, where a score of 1 implies perfect association between simulation

and observation-based estimate. Significance (p value) is calculated using a positive one-tailed t test as we are testing for positive correlation only.

We assess the significance of the difference between GloSea and persistence hindcast ACC skill by bootstrapping. That is, we randomly select 105 members (with replacement) per year, calculate the ACC score of the ensemble mean and build a distribution of ACC scores by repeating this process 1000 times. GloSea ensemble mean ACC is deemed significantly different from persistence where persistence lies outside the 97.5th and 2.5th percentiles of the randomly selected GloSea ensembles distribution.

A limitation of the ACC statistic is its insensitivity to errors in magnitude, which can mask poor skill associated with the strength of signals arising from unrealistic variability in the model ensemble. In some cases, particularly in the northern extra-tropics in winter, it is possible for the ACC skill of the signal extracted from the mean of a large ensemble of model simulations to be high but for the ensemble mean variance to be too small (Eade et al. 2014; Scaife et al. 2014; Scaife and Smith 2018; Smith et al. 2020). In such cases, there is a relatively high proportion of noise in the ensemble and therefore unrealistic disagreement between individual members. This suggests that each member cannot truly be interpreted as a plausible alternative realisation of the real world (Scaife and Smith 2018). This behaviour is described as the “signal-to-noise paradox” (Scaife and Smith 2018) as it results in the ensemble mean being better able to predict the real world than its own ensemble members. Here we use a measure of the ratio of predictable components (RPC) to quantify the error in the signal-to-noise ratios in the hindcasts. Using the definition by Scaife and Smith (2018), which is itself an iteration of the definition by Eade et al. (2014): RPC is calculated as the ratio of the predictable component of observed variability (PC_{obs}) and the predictable component of model variability (PC_{model}). PC_{obs} is estimated by ACC (Eq. 1) between model ensemble mean and the observations (ACC_{mo}), whereas PC_{model} is taken as the average ACC between each model member and the remaining ensemble mean (ACC_{mm}). Therefore,

$$RPC = \frac{PC_{obs}}{PC_{model}} = \frac{ACC_{mo}}{ACC_{mm}} \quad (2)$$

A perfect forecast system, with infinite samples and ensemble members, exhibits an RPC of 1, indicating that the predictable fraction of variance in the observations matches the predictable fraction of the model itself. When the model ensemble mean is more skilful in predicting the observed signal than individual members the RPC is greater than 1. With an RPC greater than 1, the signal-to-noise ratio is erroneously low and the magnitude of variability in the ensemble mean will be suppressed.

2.5 Ensemble sub-selection

In Sect. 3.4, we employ an observation-matching approach to sub-select the full 105-member GloSea ensemble, based on comparisons of the modelled to observation-based state of the three atmospheric circulations investigated (EOF_{Z500} 1–3, as defined in Sect. 2.3). That is, for each year and for each atmospheric mode, the 20 members which have predicted principal components closest (least absolute difference) to ERA5-derived principal components are selected.

The NWS SST fields from this 20-member sub-sample are averaged to create a new sub-selected ensemble mean. This technique relies on having observation-based estimates available against which to compare the predicted principal components, and as such it does not have relevance for real-time forecasting application. Instead this procedure represents a set of “atmospheric mode matched” experiments which enable us to ask how GloSea would perform if it could accurately predict the evolution of each atmospheric circulation pattern, to provide an idealised upper limit associated with atmospheric variability forecasts. Of course, the true limit of predictability will be lower than this.

The statistical significance of the sub-selected ensemble mean results is assessed by bootstrapping. That is, 20 members are chosen at random (with replacement) per year, and this process is repeated 1000 times to build a distribution of sub-selected ensemble mean ACC scores. The mode-matched sub-ensemble mean is deemed significantly different from random chance where it falls outside the 97.5th and 2.5th percentiles of the random-selection distribution.

3 Results

3.1 Influence of atmospheric circulations on NWS SST

We first investigate the impact of atmospheric circulations on NWS SST in observation-based and modelled (GloSea) fields. In both observation-based and GloSea versions, EOF_{Z500} 1–3 show strong correlation against NWS SST (Fig. 3). In winter, EOF1_{Z500} (NAO) is positively correlated with large areas of English Channel and North Sea SST, and the spatial impact patterns match closely between observation-based (Fig. 3a[i]) and GloSea fields (Fig. 3b[i]). Winter EOF2_{Z500} (EA) is strongly positively correlated with SST around the Southern Bight (see Fig. 1a for NWS region locations) in both observation-based (Fig. 3a[ii]) and GloSea versions (Fig. 3b[ii]). Winter EOF3_{Z500} (EAWR) demonstrates strong positive correlation in the Celtic Sea, outer shelf and northern North Sea regions, again in both observation-based (Fig. 3a[iii]) and GloSea fields (Fig. 3b[iii]). The broad similarity between observation-based and modelled atmospheric circulation impact on NWS SST in the winter indicates that GloSea tends to correctly simulate the spatial impact of winter atmospheric circulation variability on NWS SST. However, the correlations tend to be weaker in GloSea compared to observation-based fields and there are some cases where GloSea displays the wrong sign of correlation between winter atmospheric circulations and NWS SST. For example, the model- and observation-based correlations are of opposite sign for EOF1_{Z500} (NAO) in the

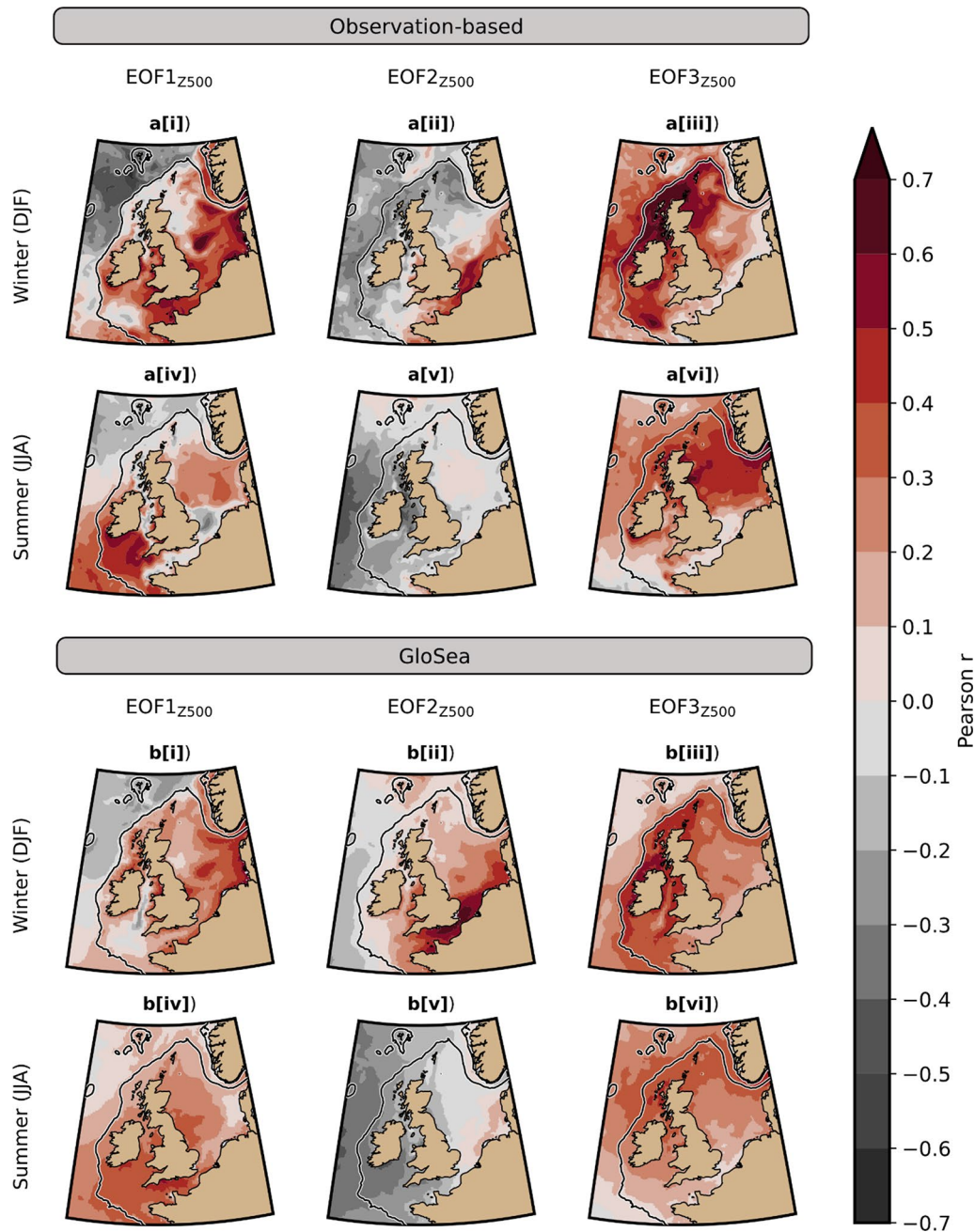


Fig. 3 Observation-based and modelled influence of atmospheric variability on NWS SST. Panels **a[i] – a[vi]**: correlation (Pearson r) between ERA5-derived EOF_{Z500} 1–3 principal components (panel columns) against CMEMS-v5 SST, in winter (DJF; row i – iii) and summer (JJA; iv – vi), across 1993–2021 (common period between datasets; $n=28$). Panels **b[i] – b[vi]**: correlation (Pearson r) between

GloSea predicted EOF_{Z500} 1–3 principal components (panel columns) against GloSea predicted SST, in winter (DJF; row i – iii) and summer (JJA; iv – vi). Note, in panels **b[i] – b[vi]** (GloSea) correlations are calculated for all individual ensemble members and hindcast years ($n=2520$) to avoid artificially smoothing variability within the ensemble. Black line marks the 200 m isobath

Irish Sea (Fig. 3a[i] vs. Figure 3b[i]) and EOF2_{Z500} (EA) in the north and west portions of the NWS (Fig. 3a[ii] vs. Figure 3b[ii]).

According to summer observation-based estimates, EOF1_{Z500} (NAO) correlates strongly with SST in the Celtic Sea (Fig. 3a[iv]), but EOF2_{Z500} (EA) demonstrates little to

no substantial on-shelf correlation (Fig. 3a[v]). Summer EOF3_{Z500} (SCAND) is positively correlated with central and northern North Sea and northern outer shelf region SST (Fig. 3a[vi]). Despite being only third in terms of contribution to total atmospheric variance, EOF3_{Z500} (SCAND) has a substantial influence on NWS summer SST. The

spatial patterns of modelled (GloSea) atmospheric circulation impact on NWS SST are approximately correct in the summer, by comparison to observation-based estimates, but do exhibit some notable inconsistencies (Fig. 3 summer observation-based vs. GloSea panels), suggesting that GloSea may produce errors in the spatial impacts of summer atmospheric circulations on NWS SST in some cases. For example, EOF1_{Z500} (NAO) is negatively correlated with SST in the English Channel, southern North Sea and Irish Sea regions in observation-based estimates but the correlations are positive in GloSea (Fig. 3a[iv] vs. Figure 3b[iv]).

3.2 Prediction skill of atmospheric circulation patterns

We now consider how skilfully GloSea predicts each atmospheric circulation pattern per season, which will likely impact the performance of GloSea NWS SST given the clear influence of atmospheric variability on the NWS. We do this to begin to attribute the sources of NWS SST prediction skill, and subsequently understand potential limits to predictability. In the winter, only predictions of EOF1_{Z500} (NAO) are skilful and significant at the 95% confidence level in GloSea (ACC=0.57; $p=0.002$; Fig. 4a). In addition, the signal-to-noise ratio is anomalously low and hence the RPC is above

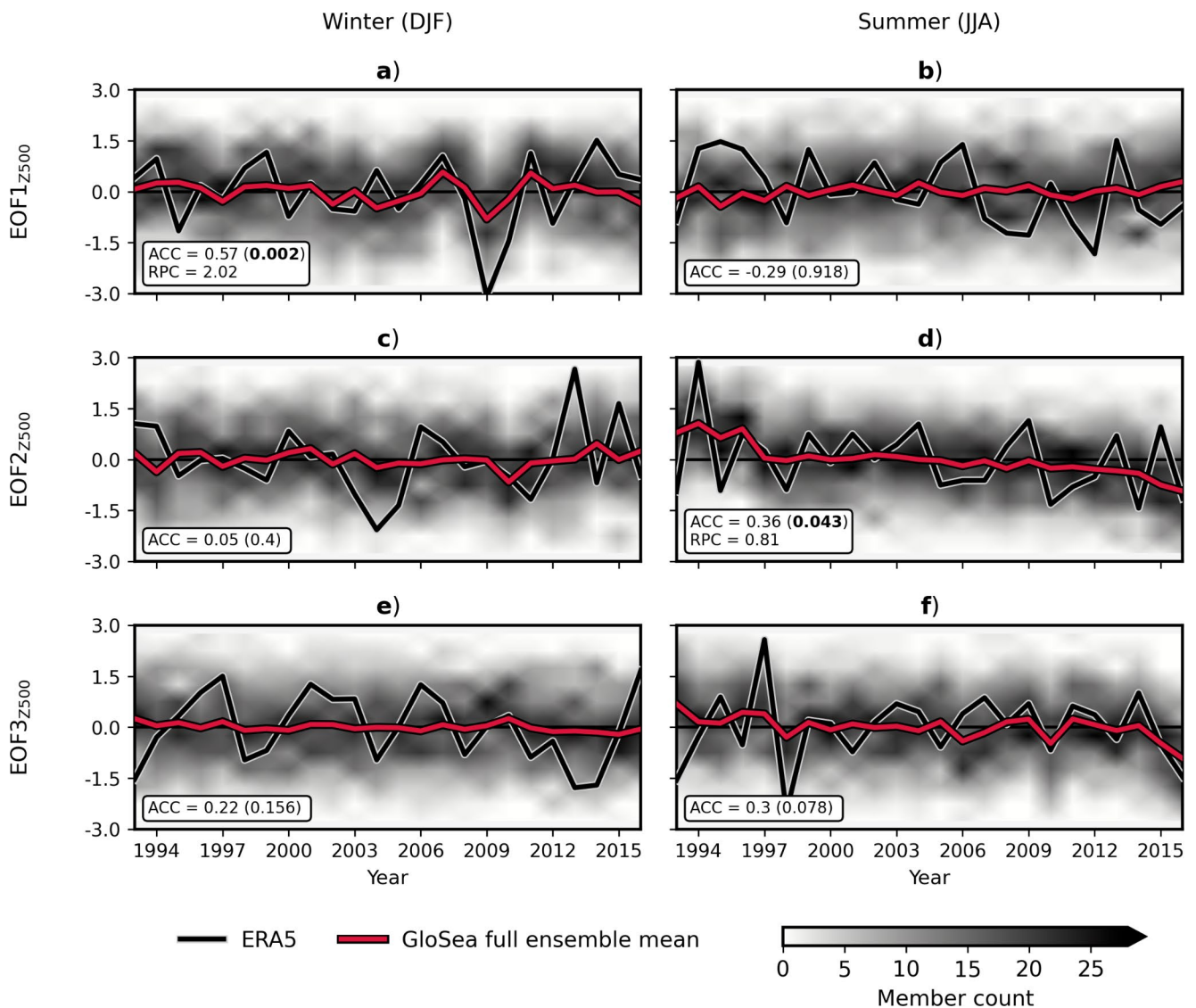


Fig. 4 Timeseries of predicted and observation-based atmospheric circulation patterns. GloSea full ensemble mean principal components (red lines) for EOF1_{Z500} (a, b), EOF2_{Z500} (c, d) and EOF3_{Z500} (e, f) in winter (DJF; panel column a, c, e) and summer (JJA; column b, d, f), at 2–4 months lead time across 1993–2016 hindcast period. ERA5 principal components are shown by the black line. All principal com-

ponent values are normalised by their respective timeseries standard deviations. Background shading represents the density (count) of GloSea individual member principal components per year. ACC scores (with p value in brackets) are marked in the white boxes. Where significant at the 95% confidence level, p values are emboldened and RPC values are also displayed

1 for winter EOF1_{Z500} (NAO). Winter EOF2_{Z500} (EA) and EOF3_{Z500} (EAWR) demonstrate no significant predictability (Fig. 4c and e). In the summer, signals are generally weaker and only EOF2_{Z500} (EA) is significantly predictable at the 95% confidence level (ACC=0.36, $p=0.043$; Fig. 4d). Predictions of summer EOF3_{Z500} (SCAND) are significant at 90% confidence (ACC=0.3, $p=0.078$; Fig. 4f), whilst summer EOF1_{Z500} (NAO) skill is negative and insignificant (ACC = -0.29; $p=0.918$; Fig. 4b). Summer EOF2_{Z500} (EA) significance is degraded when hindcasts are temporally detrended (not shown), suggesting that a portion of the skill is attributable to low-frequency decadal signals or the trend. The opposite is true for summer EOF3_{Z500} (SCAND); predictions are significant at the 95% confidence level when detrended.

3.3 GloSea and persistence NWS SST skill

We now assess the skill of GloSea in predicting NWS SST in winter and summer. Both GloSea and persistence demonstrate high skill across much of the NWS in winter (ACC > 0.6; Fig. 5a and b) and show similar spatial patterns, suggesting that persistence contributes to GloSea prediction skill. High persistence in winter is likely attributable to the thermal inertia of the waters, which are generally fully mixed in winter and thus suppress the influence of atmospheric variability on SST. Persistence and GloSea winter skill is low in the English Channel and southern North Sea regions (ACC < 0.4), which are some of the shallowest portions of the NWS. Lower thermal inertia associated with shallow bathymetry is likely to tie SST closer to atmospheric

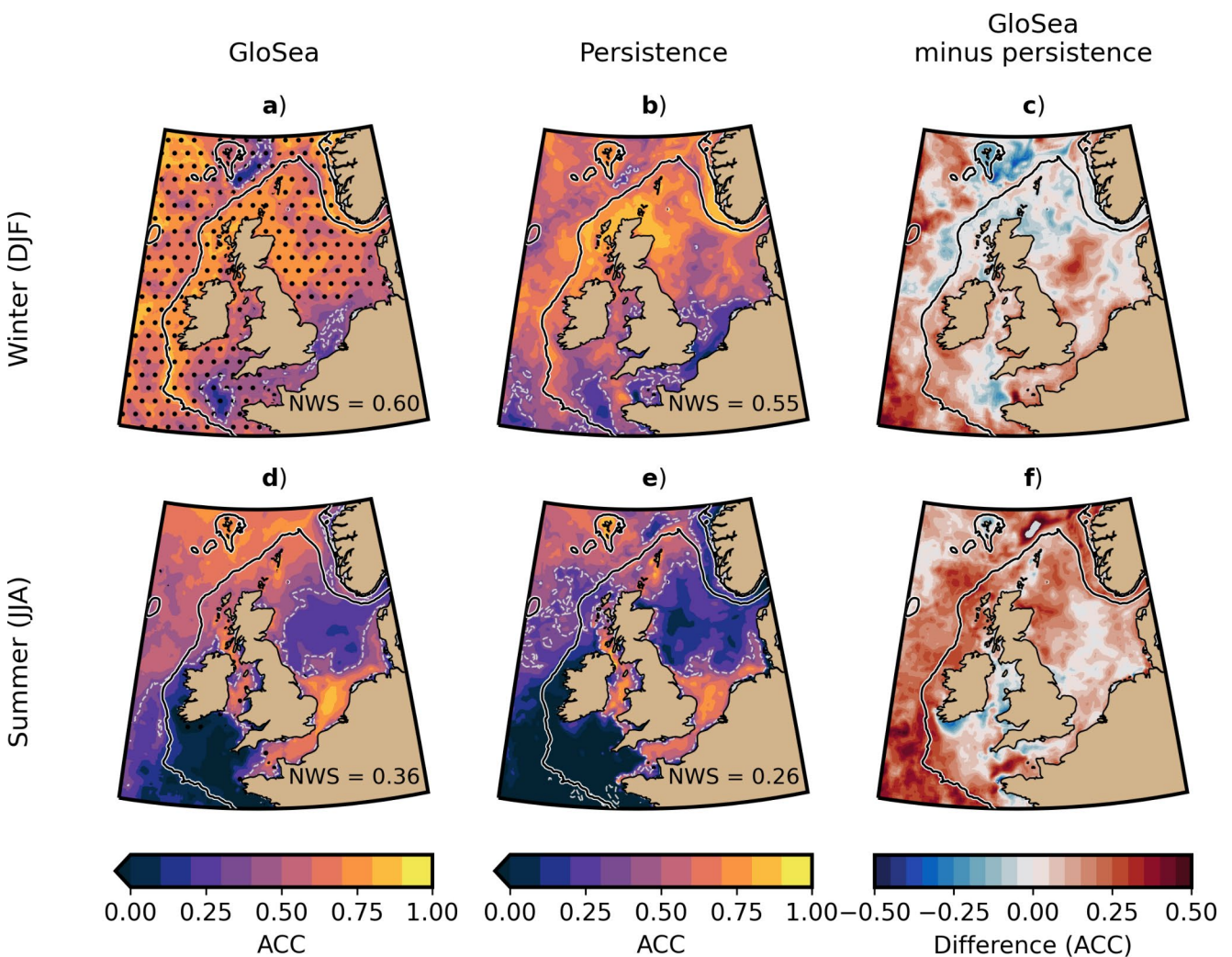


Fig. 5 Dynamical (GloSea) and persistence NWS SST prediction skill. GloSea full ensemble mean ($n=105$) SST ACC (panel column a, d), persistence ACC (column b, e), and GloSea minus persistence (column c, f), for winter (DJF; top panel row) and summer (JJA; bottom row), at 2–4 months lead time across the hindcast period 1993–2016. The $p=0.05$ statistical significance level is marked by grey dashed

contours, where the side of the contour with higher ACC scores marks $p < 0.05$. p values are calculated using a one-tailed test. Mean ACC for the NWS domain is marked in the bottom right. Black line marks the 200 m isobath. Stippling marks locations where GloSea is significantly different from persistence, as assessed by bootstrapping method described in Sect. 2.4

variability in these parts and therefore lead to accelerated decoupling from initial conditions. GloSea shows higher skill than persistence overall during winter (Fig. 5c). The NWS-wide improvements are modest in GloSea (area mean ACC=0.60 for GloSea vs. ACC=0.55 for persistence) but spatially coherent and significantly different from persistence across large areas, including the North Sea. The majority of improvement comes in regions which are shown to be impacted by winter EOF1_{Z500} (NAO), in both observation-based (Fig. 3a[i]) and modelled fields (Fig. 3b[i]), noting that EOF1_{Z500} (NAO) is skilfully predicted in GloSea (Fig. 4a). That is, the skilful simulation of winter EOF1_{Z500} (NAO) in GloSea likely leads to improved NWS SST predictions over persistence, which lacks information on the evolution of the atmosphere over the predicted period. SST in the Southern Bight region, where EOF2_{Z500} (EA) is most influential (Fig. 3a[ii] and 3b[ii]), remains poorly simulated in GloSea, likely because EOF2_{Z500} (EA) is not skilfully predicted (Fig. 4c). The areas of impact associated with winter EOF3_{Z500} (EAWR), namely the Celtic Sea, outer shelf and northern North Sea regions (Fig. 3c), show statistically significant but relatively small changes in GloSea as persistence is already high in these areas (Fig. 5b).

Summer NWS SST skill in both GloSea and persistence hindcasts is low in large areas (ACC<0.3; Fig. 5d and e), particularly in the Celtic Sea and North Sea regions where ACC is almost entirely statistically insignificant in both systems. Moreover, there is no statistically significant difference between GloSea and persistence across most of the NWS. Persistence in the Celtic Sea and North Sea regions is low in summer likely due to seasonal stratification in the warmer summer months (shown by mixing fronts in Fig. 1b and d, and van Leeuwen et al. 2015; Pingree and Griffiths 1978) which results in a thin mixed layer (reduced thermal inertia), thus amplifying the impact of atmospheric variability on SST. Though there is moderate improvement in GloSea skill relative to persistence on-shelf (within the 200 m isobath where this study is focussed) large areas remain statistically insignificant (Fig. 5f). Summer EOF1_{Z500} (NAO) is shown to impact the Celtic Sea in observation-based (Fig. 3a[iv]) and modelled fields (Fig. 3b[iv]) but is not skilfully predicted in GloSea (Fig. 4b), meaning there is no associated meaningful improvement in SST skill in GloSea (Fig. 5d). Summer EOF3_{Z500} (SCAND) is shown to be both influential on North Sea SST (Fig. 3) and somewhat skilfully predicted by GloSea (Fig. 4f) yet there is little meaningful improvement in GloSea SST skill in these parts relative to persistence. Potential explanations may include one or a combination of (i) the skill of summer EOF3_{Z500} (SCAND) being still too low (ACC=0.3, Fig. 4f) to adequately constrain SST evolution in the particularly sensitive stratifying regions, (ii) the weak and spatially erroneous

response of SST to summer EOF3_{Z500} (SCAND) in GloSea (Fig. 3a[vi] vs. Figure 3b[vi]), potentially due to (iii) deficiencies in the GloSea global ocean component associated with locally erroneous representation of key NWS physics such as seasonal stratification (discussed in greater detail in Sect. 3.5). Summer EOF2_{Z500} (EA) is skilfully predicted by GloSea (Fig. 4d) but is shown to have weaker impact on on-shelf SST (Fig. 3a[v] and 3b[v]) meaning GloSea NWS SST prediction skill is unaffected.

Exceptions to low summer skill include the English Channel, southern North Sea, Irish Sea regions and around the Fair Isle current (see Fig. 1a for NWS region locations) where, unlike winter for the most part, skill is high in both GloSea and persistence systems (ACC>0.6; Fig. 5d and e). These regions tend to remain fully mixed throughout the year due to strong tidal mixing (see mixing fronts in Fig. 1b and d, as well as van Leeuwen et al. 2015), which typically favours higher thermal inertia. However, the high summer skill seemingly contradicts the reasoning for low skill in these regions in winter when the waters are also fully mixed but nonetheless shallow and therefore likely to have low thermal inertia.

We propose that the high persistence in waters with low thermal inertia, seen in the shallow English Channel and southern North Sea regions in summer (Fig. 5e), occurs because these regions are only weakly influenced by the dominant summer atmospheric circulations according to observation-based estimates (Fig. 3a summer panels). This state contrasts with the situation in the winter (Fig. 3a winter vs. summer panels). However, this cannot explain the English Channel, southern North Sea and Irish Sea skill in GloSea, which falsely simulates summer EOF1_{Z500} (NAO) as being impactful on SST in these regions (Fig. 3a[iv] vs. Figure 3b[iv]). The ORCA025 ocean grid used by GloSea is relatively coarse and parameterises tidal mixing (i.e. Simmons et al. 2004) rather than including dynamic tides of varying amplitude. The parameterisation of tides is expected to be an underestimation for the NWS, which can lead to artificial stratification in the English Channel, southern North Sea and Irish Sea regions (Tinker et al. 2018), therefore artificially amplifying air-sea exchange. Intuitively, this would suggest that the true SST persistence in the English Channel, southern North Sea and Irish Sea regions should be degraded in GloSea due to the artificial influence of EOF1_{Z500} (NAO), yet summer SST skill is maintained in these regions in GloSea (Fig. 5d). We explain this as a cancellation of errors between the GloSea EOF1_{Z500} (NAO) prediction and SST response. That is, the GloSea summer SST response to EOF1_{Z500} (NAO) has the incorrect sign in these regions (Fig. 3a[iv] vs. Figure 3b[iv]), in parallel to summer EOF1_{Z500} (NAO) prediction skill displaying the wrong sign (ACC = -0.29; Fig. 4b).

GloSea summer SST skill in the outer shelf region is high (Fig. 5d) and sees improvements when compared with persistence (Fig. 5f), albeit the difference is not statistically significant. We note that the same region displays lower substantially skill when hindcast data are temporally detrended (not shown). Skill in the outer shelf region may be associated with low-frequency decadal signals in the North Atlantic. The advection of these signals onto the NWS should be captured by the GloSea hindcasts.

3.4 Observation-matched GloSea atmospheric circulation patterns

We have shown that GloSea can be skilful in predicting NWS SST, particularly in winter. However, large amounts of variance remain unexplained at current levels of skill so we now ask what the prediction skill of NWS SST would be with improved forecasts of atmospheric circulation patterns. The “atmospheric mode matching” procedure gives an upper estimate of the skill that could be achieved if we had improved atmospheric forecasts with respect to each individual circulation pattern. To do so, we sub-select members from the full GloSea ensemble which best simulate observation-based estimates of each atmospheric circulation pattern’s principal components per hindcast year (see Sect. 2.5). In both winter and summer, ACC skill for the “atmospheric mode matched” sub-selected ensemble mean for each circulation pattern (blue lines in Fig. 6) is close to 1 and is significantly improved over the GloSea full ensemble mean (red lines). This idealised method achieves near perfect skill for each mode in terms of capturing the amplitude, sign and phase of variability ($ACC > 0.9$ across all seasons and circulation patterns) and it corrects any anomalous signal-to-noise ratios ($RPC \approx 1$). Note, however, we find that there is little overlap between members in sub-ensembles selected for each atmospheric circulation pattern (on average across the hindcast period, 5.62% and 3.33% of members overlap in winter and summer, respectively), indicating that individual members tend not to skilfully predict multiple circulation patterns simultaneously.

3.5 Sub-selected ensemble GloSea NWS SST skill

Using only the members whose predictions of each atmospheric circulation pattern track observation-based estimates ($n = 20$), a new sub-selected NWS SST ensemble mean field is generated per mode from the members’ corresponding ocean fields. In the winter and for all three atmospheric circulation patterns, the sub-selected ensemble mean shows improvement in SST skill across large areas of the NWS, with improvements on the order of $ACC > 0.25$ relative to the GloSea full ensemble mean (Fig. 7). Concerning winter

EOF1_{Z500} (NAO) and EOF2_{Z500} (EA), when matched to observation-based estimates, the areas which perform worst in the full ensemble mean, namely the English Channel and southern North Sea (Fig. 5a), show the largest improvements (Fig. 7d and e). This is consistent with the influence of EOF1_{Z500} (NAO) and EOF2_{Z500} (EA) in these regions in winter (Fig. 3). Regarding winter EOF3_{Z500} (EAWR), there is widespread improvement in skill across most of the NWS (Fig. 7f). Skill in the small area in the southeast Celtic Sea/west of France is boosted but remains low (Fig. 7c).

In the summer, the sub-selected ensembles matched to observation-based estimates of EOF1_{Z500} (NAO) show considerable improvement in skill (Fig. 8a and d) in areas of the NWS, namely the Celtic Sea and North Sea, where the GloSea full ensemble mean is both very low in skill (Fig. 5d) and where summer EOF1_{Z500} (NAO) is shown to be impactful in the summer (Fig. 3). Therefore, summer EOF1_{Z500} (NAO) forecasts offer the greatest opportunity for improvement of NWS summer SST predictions. However, large areas in the Celtic Sea and parts of the eastern North Sea remain low in skill and statistically insignificant. This is likely due to one or a combination of: poor model representation of NWS summer stratification (i.e. Tinker et al. 2018) in the seasonally stratifying regions of the Celtic Sea and North Sea, deficiencies in simulating the impacts of meteorological events such as storms/rainfall/freshwater inputs on the water column (i.e. Jardine et al. 2023), or errors in exchange between the shelf and open ocean in the GloSea global ocean model component, which is not best suited to the fine-scale processes of the shelf-break (Graham et al. 2018). Also of note is the fact that improved summer EOF1_{Z500} (NAO) predictions lead to degraded SST performance in regions such as the English Channel (Fig. 8d). This is consistent with the interpretation of results presented in Sect. 3.3 where full ensemble GloSea SST skill is suggested to be high in such regions due to a cancellation of errors between summer EOF1_{Z500} (NAO) predictions and SST response. That is, improved EOF1_{Z500} (NAO) predictions will degrade SST predictions in areas where the spatial impacts of atmospheric circulations are erroneous, likely due to deficiencies in the ocean model component such as in the case of artificial stratification in the English Channel. The sub-selected ensembles matched to observation-based estimates of summer EOF3_{Z500} (SCAND) also show meaningful improvement in skill across parts of the NWS (Fig. 8c and f) but again, ocean model deficiencies may limit improvement in summer. In the case of matching to observation-based estimates of summer EOF2_{Z500} (EA), there is little improvement and skill is often significantly degraded relative to the full ensemble mean (Fig. 8b and e). This is explained by the weak impact by EOF2_{Z500} on NWS summer SST (Fig. 3) meaning that the sub-selection procedure

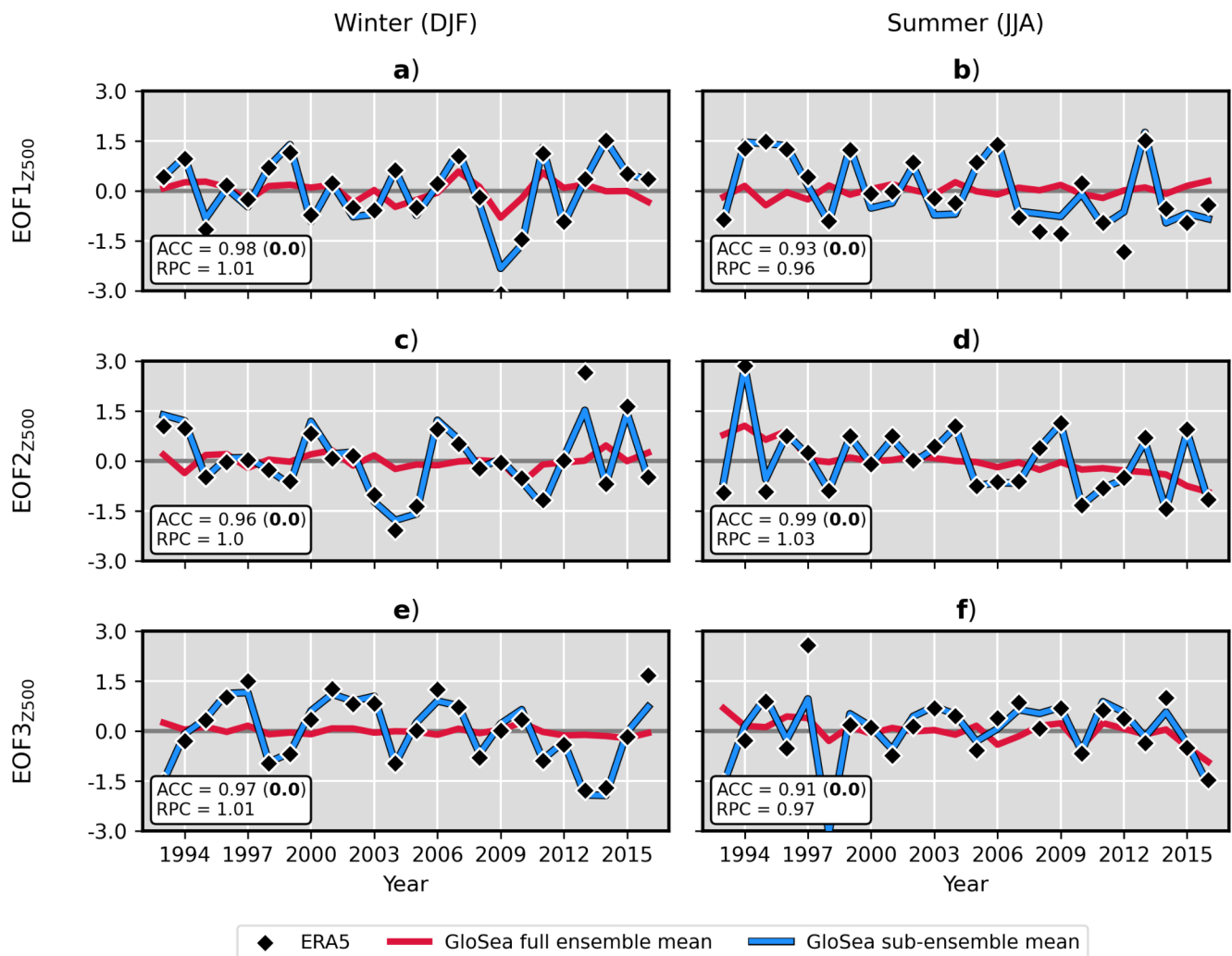


Fig. 6 GloSea “atmospheric mode matched” tests. Sub-selected ensemble mean principal component predictions (blue lines) with respect to EOF1_{Z500} (a, b), EOF2_{Z500} (c, d) and EOF3_{Z500} (e, f) in winter (DJF; panel column a, c, e) and summer (JJA; column b, d, f), at 2–4 months lead time across 1993–2016 hindcast period. Red line shows the original full ensemble mean (as in Fig. 4). ERA5 principal components are

shown by the black scatter markers. All principal component values are normalised by their respective timeseries standard deviations. ACC (with p value in brackets, calculated using a one-tailed test) and RPC scores marked in the white boxes are for GloSea “atmospheric mode matched” sub-selected ensembles

serves only to reduce the ensemble size without the benefit of better representing impactful atmospheric circulations.

4 Summary and discussion

We have quantified the current levels of skill in predicting boreal winter and summer European North-West shelf seas sea surface temperature using a dynamical forecasting system and have understood potential sources of skill improvement. Approximate locations where we find high and low current skill, as well as locations where there may be improvements in skill with improved atmospheric forecasts, are summarised in the schematic in Fig. 9. Whilst it is unrealistic to expect GloSea to ever produce near perfect

predictions of atmospheric circulation patterns, the “atmospheric mode matched” analysis presented in this study is a useful way of identifying locations where more NWS SST skill may be attainable with some future increases in atmospheric skill, as well what skill remains unexplained in GloSea under improved atmospheric circulation prediction. In this study, we have evaluated NWS SST predictability from the perspective of atmospheric circulation drivers. Alternative approaches could involve methods to explicitly link predictable modes of the variable of interest (i.e. NWS SST) to potential sources of predictability (e.g. DelSole and Chang 2003; Fan et al. 2020; Zhang et al. 2023; Chen et al. 2024).

GloSea winter NWS SST prediction skill is generally high and beats persistence, particularly in shallow regions

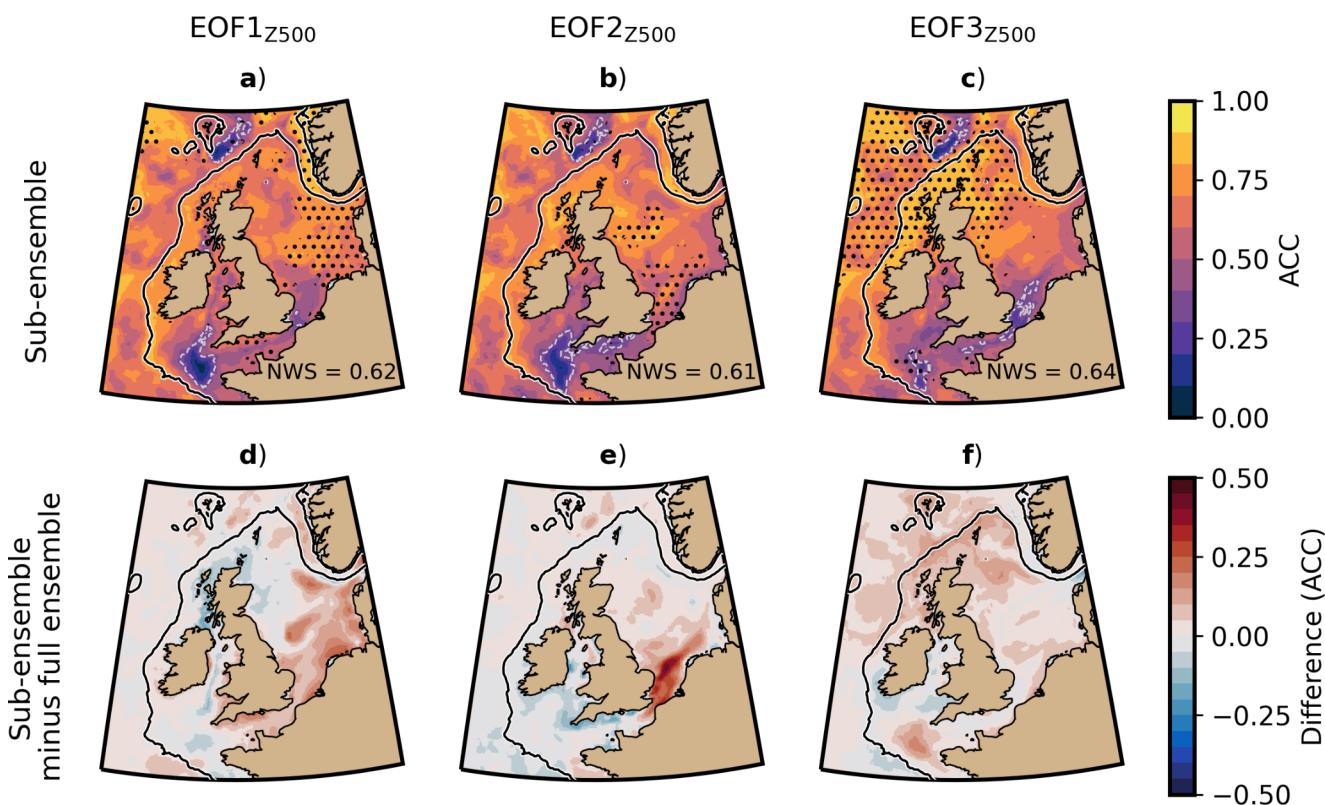


Fig. 7 GloSea NWS winter SST prediction skill if improved atmospheric predictions were possible. Top panel row: Winter (DJF) only EOF1_{Z500} (panel column a, d), EOF2_{Z500} (column b, e) and EOF3_{Z500} (column c, f) GloSea sub-selected ensemble mean SST ACC (n members=20), at 2–4 months lead time across the hindcast period 1993–2016. Bottom panel row: sub-selected ensemble mean (i.e. top row) minus full ensemble mean (i.e. Figure 5a). The $p=0.05$ statistical

significance level is marked by grey dashed contours, where the side of the contour with higher ACC scores marks $p < 0.05$. p values are calculated using a one-tailed test. Mean ACC for the NWS domain is marked in the bottom right for the top row. Black line marks the 200 m isobath. Stippling shows locations where the sub-selected ensemble mean is significantly different from random member selection, as assessed by bootstrapping method described in Sect. 2.5

where EOF1_{Z500} (NAO) exerts a strong control. This builds on the well-documented understanding of GloSea winter NAO prediction skill (e.g. Scaife et al. 2014) and echoes the NWS winter predictability results of Tinker and Hermanson (2021). In the summer, GloSea NWS SST skill is generally lower when the stratified surface mixed layer is particularly sensitive to atmospheric variability. For both seasons, prospects for future seasonal forecast skill in dynamical forecasting systems are identified through GloSea “atmospheric mode matched” simulations. Potential increases in NWS SST skill stem primarily from improved EOF1_{Z500} (NAO) predictions in both winter and summer, in line with well understood surface climate impacts of winter and summer NAO (Hurrell 1995; Folland et al. 2009; Hurrell and Deser 2009; Bladé et al. 2012). Winter NAO has previously been shown to impact SST across large areas of the NWS (Becker and Pauly 1996; Dippner 1997; Tinker et al. 2018) meaning its role in producing skilful simulations of NWS SST is expected. However, we also demonstrate for the first time the key contribution of summer NAO to NWS SST variability and its potential for improving summer SST simulations

if predicted well. EOF2_{Z500} (EA) in winter and EOF3_{Z500} (SCAND) in summer also contribute to NWS SST predictability. Winter EOF2_{Z500} (EA) is linked to precipitation over North-West continental Europe and Britain (Casanueva et al. 2014; Hall and Hanna 2018; West et al. 2021), indicating a potential mechanism for influencing the NWS via variability in coastal runoff. With increased freshwater runoff we might expect increased stratification in the otherwise well-mixed Southern Bight area (Simpson et al. 1993; van Leeuwen et al. 2015) with consequences for SST variability. Despite being the lowest order EOF in summer, EOF3_{Z500} (SCAND) exerts strong control over summer NWS SST, reflective of its broad links to European surface climate (Bueh and Nakamura 2007; Wang and Tan 2020). If predicted well, it offers potential for skill improvement across the NWS in summer.

Even in “atmospheric mode matched” tests, only ~50% of the variance in NWS SST on seasonal timescales can be explained. Whilst unpredictable noise associated with atmospheric variability (e.g. short-term weather) contributes an inherent limit to total NWS SST predictability, the impact of

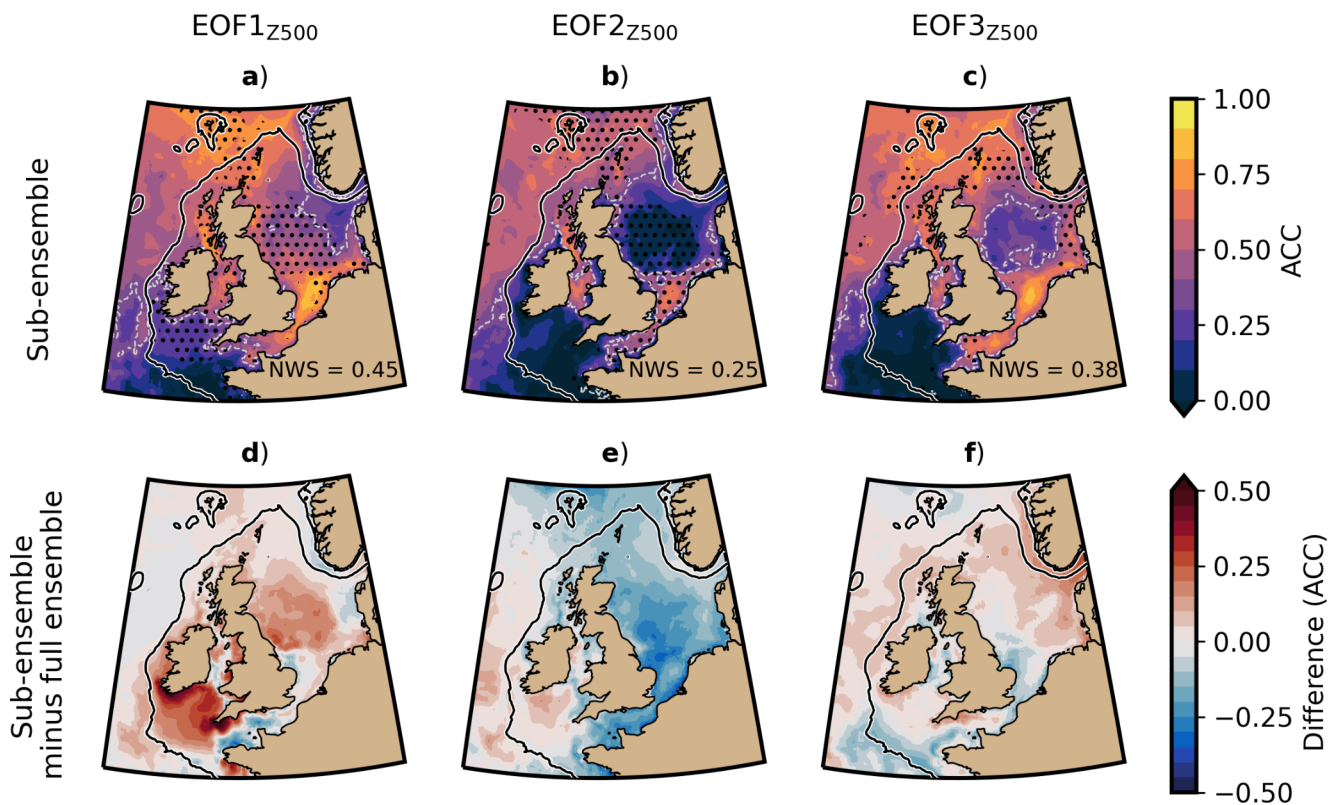


Fig. 8 GloSea NWS summer SST prediction skill if improved atmospheric predictions were possible. Top panel row: Summer (JJA) only EOF1_{Z500} (panel column a, d), EOF2_{Z500} (column b, e) and EOF3_{Z500} (column c, f) GloSea sub-selected ensemble mean SST ACC (n members=20), at 2–4 months lead time across the hindcast period 1993–2016. Bottom panel row: sub-selected ensemble mean (i.e. top row) minus full ensemble mean (i.e. Figure 5d). The $p=0.05$ statistical

significance level is marked by grey dashed contours, where the side of the contour with higher ACC scores marks $p < 0.05$. p values are calculated using a one-tailed test. Mean ACC for the NWS domain is marked in the bottom right for the top row. Black line marks the 200 m isobath. Stippling shows locations where the sub-selected ensemble mean is significantly different from random member selection, as assessed by bootstrapping method described in Sect. 2.5

fine-scale ocean processes on the NWS remains a potentially significant source of skill and has not been explored in detail in this study. The GloSea ocean model component, which is relatively coarse resolution for this regional application and lacks dynamic tides, may fail to properly resolve NWS processes, such as stratification, shelf-break transport and shelf-edge exchange (Holt et al. 2017). Therefore, alongside continuing development aimed at improving atmospheric variability simulation in dynamical forecast systems, there is significant scope to derive further NWS SST skill through ocean component development with greater attention given to resolving shelf seas processes. Moreover, in certain cases, improvements to the ocean component of GloSea may be required to unlock skill offered by any improvements to atmospheric forecasts. For example, we have shown that errors in the spatial impact of summer EOF1_{Z500} (NAO) on SST in the English Channel within GloSea, potentially due to artificial stratification, results in degraded SST skill in this region when predictions of summer EOF1_{Z500} (NAO) are improved. The prospect of dynamically downscaling the ocean component of a GloSea large ensemble is appealing

(i.e. Tinker and Hermanson 2021) but would likely be prohibitively expensive. Future work could instead aim to force a simple shelf seas model (e.g. “S2P3”; Halloran et al. 2021) to explore the benefits of improved shelf dynamics at lower expense, or to improve representation of tides in GloSea either by using a shelf-enabled global NEMO ocean component (in development) or by improving their parameterisation (e.g. Tinker et al. 2022). Whilst we have focussed here on seasonal SST predictability, it will be important to consider the predictability of other variables, averaging periods and lead times. For example, predictability of sub-surface temperature fields could be critical for ecosystem applications (e.g. Smyth et al. 2010; Marsh et al. 2015), and exploring prospects for early-warning systems of NWS extreme events, including marine heatwaves (e.g. Berthou et al. 2024), could be important for climate adaptation.

We have demonstrated that considerable NWS SST seasonal forecasting skill already exists through both the persistence of anomalies in the initial conditions and the predictability of atmospheric circulation. Further skill can likely be derived from improvements to both atmospheric

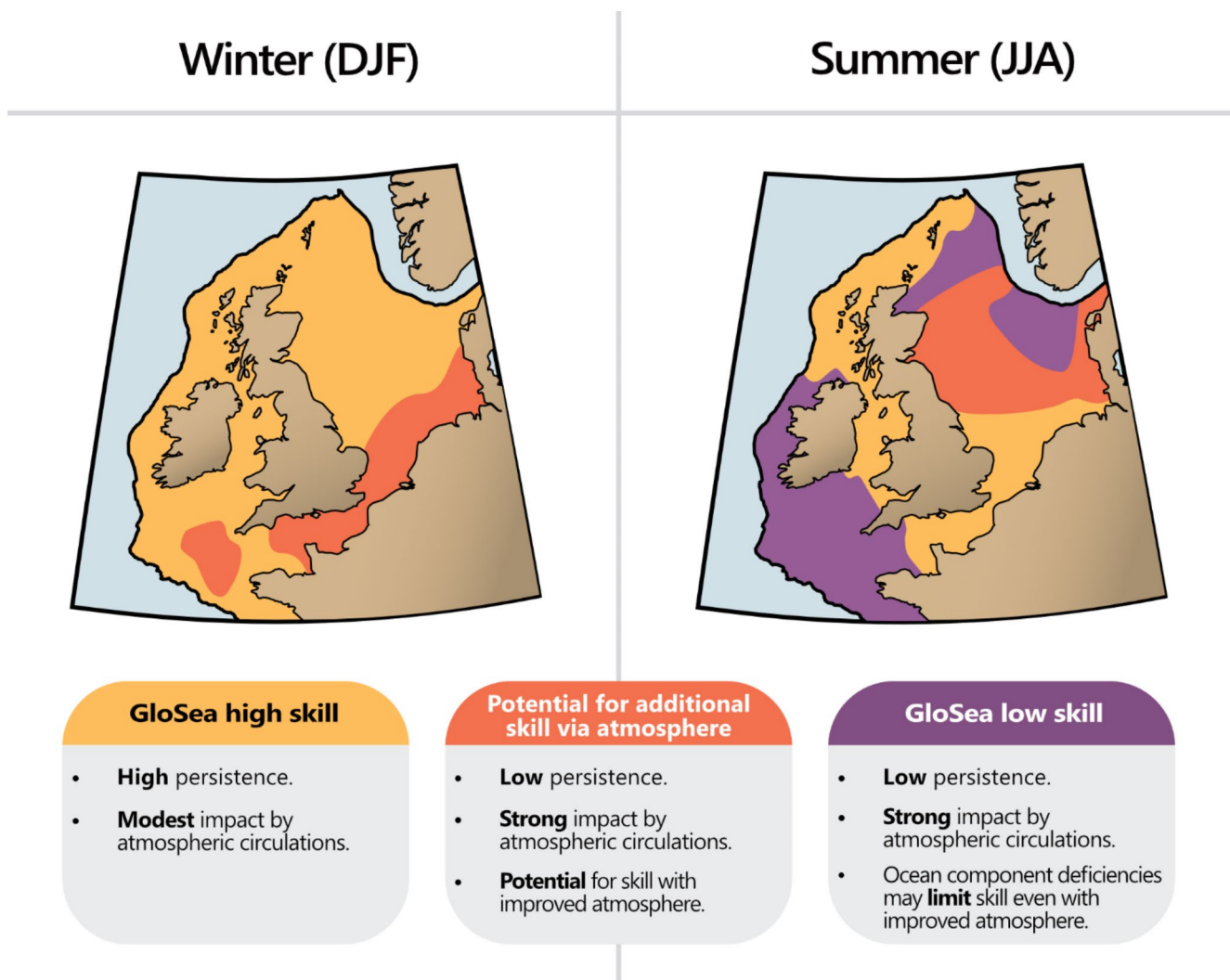


Fig. 9 Current skill and potential skill around the NWS. Schematic to summarise the *approximate* locations of high skill (yellow), low skill (purple) and potential for skill indicated by “atmospheric mode matched” tests (orange) in GloSea predictions of NWS SST for winter

(DJF) and summer (JJA) seasons. Annotations provide more detail on the sources and limits to skill in each category, based on the analysis presented on GloSea and persistence predictability in Sect. 3.3 and the results of “atmospheric mode matched” experiments in Sect. 3.5

and oceanic components of dynamical forecasting systems. However, we encourage the exploitation of the existing skill for managing the demands on the NWS in response to human pressures.

Author contributions All authors contributed to the study conception and design. Material preparation, data collection and analysis were performed by Jamie Atkins. The first draft of the manuscript was written by Jamie Atkins and all authors commented on previous versions of the manuscript. All authors read and approved the final manuscript.

Funding This work was supported by a GW4+ Doctoral Training Partnership studentship from the Natural Environmental Research Council (NE/S007504/1) and the UK Met Office. AAS and JT were supported by the Met Office Hadley Centre Climate Programme (HCCP) funded by the UK Department for Science, Innovation and Technology (DSIT), the UK Public Weather Service and the UK Department for Environment, Food and Rural Affairs (DEFRA). JAG was supported by the Cefas Seedcorn studentship programme.

Data availability GloSea5 and GloSea6 SST and Z500 hindcast data used in this study (on native grids) were accessed directly from the UK Met Office but are also freely available (interpolated to 1° × 1° grid) from the Copernicus Climate Change Service (C3S; <https://doi.org/10.24381/cds.68dd14c3>). ERA5 data is also freely available from C3S (<https://doi.org/10.24381/cds.6860a573>). CMEMS-v5 data is freely available from the Copernicus Marine Environment Monitoring Service (<https://doi.org/10.48670/moi-00059>).

Declarations

Competing interests The authors have no relevant financial or non-financial interests to disclose.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this

article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Athanasiadis PJ, Bellucci A, Scaife AA et al (2017) A multisystem view of Wintertime NAO seasonal predictions. *J Clim* 30:1461–1475. <https://doi.org/10.1175/JCLI-D-16-0153.1>
- Baker LH, Shaffrey LC, Sutton RT et al (2018) An intercomparison of skill and Overconfidence/Underconfidence of the Wintertime North Atlantic Oscillation in Multimodel Seasonal forecasts. *Geophys Res Lett* 45:7808–7817. <https://doi.org/10.1029/2018GL078838>
- Barnston AG, Livezey RE (1987) Classification, seasonality and persistence of low-frequency Atmospheric circulation patterns. *Mon Weather Rev* 115:1083–1126. [https://doi.org/10.1175/1520-0493\(1987\)115%3C1083:CSAPOL%3E2.0.CO;2](https://doi.org/10.1175/1520-0493(1987)115%3C1083:CSAPOL%3E2.0.CO;2)
- Becker GA, Pauly M (1996) Sea surface temperature changes in the North Sea and their causes. *ICES J Mar Sci* 53:887–898. <https://doi.org/10.1006/jmsc.1996.0111>
- Bell VA, Davies HN, Kay AL et al (2017) A national-scale seasonal hydrological forecast system: development and evaluation over Britain. *Hydrol Earth Syst Sci* 21:4681–4691. <https://doi.org/10.5194/hess-21-4681-2017>
- Berthou S, Renshaw R, Smyth T et al (2024) Exceptional atmospheric conditions in June 2023 generated a northwest European marine heatwave which contributed to breaking land temperature records. *Commun Earth Environ* 5:1–11. <https://doi.org/10.1038/s43247-024-01413-8>
- Bladé I, Liebmann B, Fortuny D, van Oldenborgh GJ (2012) Observed and simulated impacts of the summer NAO in Europe: implications for projected drying in the Mediterranean region. *Clim Dyn* 39:709–727. <https://doi.org/10.1007/s00382-011-1195-x>
- Bowler NE, Arribas A, Beare SE et al (2009) The local ETKF and SKEB: upgrades to the MOGREPS short-range ensemble prediction system. *Q J R Meteorol Soc* 135:767–776. <https://doi.org/10.1002/qj.394>
- Bueh C, Nakamura H (2007) Scandinavian pattern and its climatic impact. *Q J R Meteorol Soc* 133:2117–2131. <https://doi.org/10.1002/qj.173>
- Casanueva A, Rodríguez-Puebla C, Frías MD, González-Reviriego N (2014) Variability of extreme precipitation over Europe and its relationships with teleconnection patterns. *Hydrol Earth Syst Sci* 18:709–725. <https://doi.org/10.5194/hess-18-709-2014>
- Ceglar A, Toreti A, Prodhomme C et al (2018) Land-surface initialisation improves seasonal climate prediction skill for maize yield forecast. *Sci Rep* 8:1322. <https://doi.org/10.1038/s41598-018-19586-6>
- Chen M, Kumar A, L'Heureux M et al (2024) Why do DJF 2023/24 Upper-Level 200-hPa Geopotential Height forecasts look different from the expected El Niño Response? *Geophys Res Lett* 51:e2024GL108946. <https://doi.org/10.1029/2024GL108946>
- Clark RT, Bett PE, Thornton HE, Scaife AA (2017) Skilful seasonal predictions for the European energy industry. *Environ Res Lett*. <https://doi.org/10.1088/1748-9326/AA57AB.12>
- DelSole T, Chang P (2003) Predictable component analysis, canonical correlation analysis, and autoregressive models. *J Atmos Sci* 60:409–416. [https://doi.org/10.1175/1520-0469\(2003\)060%3C0409:PCACCA%3E2.0.CO;2](https://doi.org/10.1175/1520-0469(2003)060%3C0409:PCACCA%3E2.0.CO;2)
- Dippner JW (1997) SST anomalies in the north sea in relation to the North Atlantic Oscillation and the influence on the theoretical spawning time of fish. *Dtsch Hydrographische Z* 49:267–275. <https://doi.org/10.1007/BF02764038>
- Dunstone N, Smith DM, Hardiman SC et al (2023) Skilful predictions of the Summer North Atlantic Oscillation. *Commun Earth Environ* 4:1–11. <https://doi.org/10.1038/s43247-023-01063-2>
- Eade R, Smith D, Scaife A et al (2014) Do seasonal-to-decadal climate predictions underestimate the predictability of the real world? *Geophys Res Lett* 41:5620–5628. <https://doi.org/10.1002/2014GL061146>
- Ely CR, Brayshaw DJ, Methven J et al (2013) Implications of the North Atlantic Oscillation for a UK–Norway renewable power system. *Energy Policy* 62:1420–1427. <https://doi.org/10.1016/j.enpol.2013.06.037>
- Fan H, Wang L, Zhang Y et al (2020) Predictable Patterns of Wintertime Surface Air Temperature in Northern Hemisphere and Their Predictability Sources in the SEAS5. <https://doi.org/10.1175/JCLI-D-20-0542.1>
- Folland CK, Knight J, Linderholm HW et al (2009) The Summer North Atlantic Oscillation: past, Present, and Future. *J Clim* 22:1082–1103. <https://doi.org/10.1175/2008JCLI2459.1>
- Graham JA, Rosser JP, O'Dea E, Hewitt HT (2018) Resolving Shelf Break Exchange around the European Northwest Shelf. *Geophys Res Lett* 45:12,386–12,395. <https://doi.org/10.1029/2018GL079399>
- Hall RJ, Hanna E (2018) North Atlantic circulation indices: links with summer and winter UK temperature and precipitation and implications for seasonal forecasting. *Int J Climatol* 38:e660–e677. <https://doi.org/10.1002/joc.5398>
- Halloran PR, McWhorter JK, Arellano Nava B et al (2021) S2P3-R v2.0: computationally efficient modelling of shelf seas on regional to global scales. *Geosci Model Dev* 14:6177–6195. <https://doi.org/10.5194/gmd-14-6177-2021>
- Hannachi A, Jolliffe IT, Stephenson DB (2007) Empirical orthogonal functions and related techniques in atmospheric science: a review. *Int J Climatol* 27:1119–1152. <https://doi.org/10.1002/joc.1499>
- Hersbach H, Bell B, Berrisford P et al (2020) The ERA5 global reanalysis. *Q J R Meteorol Soc* 146:1999–2049. <https://doi.org/10.1002/qj.3803>
- Hobday AJ, Spillman CM, Paige Eveson J, Hartog JR (2016) Seasonal forecasting for decision support in marine fisheries and aquaculture. *Fish Oceanogr* 25:45–56. <https://doi.org/10.1111/fog.12083>
- Holt J, Wakelin S, Lowe J, Tinker J (2010) The potential impacts of climate change on the hydrography of the northwest European continental shelf. *Prog Oceanogr* 86:361–379. <https://doi.org/10.1016/j.pocean.2010.05.003>
- Holt J, Hyder P, Ashworth M et al (2017) Prospects for improving the representation of coastal and shelf seas in global ocean models. *Geosci Model Dev* 10:499–523. <https://doi.org/10.5194/gmd-10-499-2017>
- Hurrell JW (1995) Decadal trends in the North Atlantic Oscillation: Regional temperatures and Precipitation. *Science* 269:676–679. <https://doi.org/10.1126/science.269.5224.676>
- Hurrell JW, Deser C (2009) North Atlantic climate variability: the role of the North Atlantic Oscillation. *J Mar Syst* 78:28–41. <https://doi.org/10.1016/j.jmarsys.2008.11.026>
- Huthnance JM (1991) Physical oceanography of the North Sea. *Ocean Shorel Manage* 16:199–231. [https://doi.org/10.1016/0951-8312\(91\)90005-M](https://doi.org/10.1016/0951-8312(91)90005-M)
- Huthnance JM, Holt JT, Wakelin SL (2009) Deep ocean exchange with west-european shelf seas. *Ocean Sci* 5:621–634. <https://doi.org/10.5194/os-5-621-2009>

- Jardine JE, Palmer M, Mahaffey C et al (2023) Rain triggers seasonal stratification in a temperate shelf sea. *Nat Commun* 14:3182. <https://doi.org/10.1038/s41467-023-38599-y>
- Jolliffe I, Stephenson (2011) *Forecast Verification: a practitioner's guide in Atmospheric Science*, 2nd edn. Wiley, New York
- Kim H-M, Webster PJ, Curry JA (2012) Seasonal prediction skill of ECMWF System 4 and NCEP CFSv2 retrospective forecast for the Northern Hemisphere Winter. *Clim Dyn* 39:2957–2973. <https://doi.org/10.1007/s00382-012-1364-6>
- Kröger S, Parker R, Cripps G, Williamson P (2018) Shelf seas: the Engine of Productivity: Underpinning Science for Policy and Management. CEFAS Lowestoft
- Lim Y-K (2015) The East Atlantic/West Russia (EA/WR) teleconnection in the North Atlantic: climate impact and relation to Rossby wave propagation. *Clim Dyn* 44:3211–3222. <https://doi.org/10.1007/s00382-014-2381-4>
- Lledó L, Cionni I, Torralba V et al (2020) Seasonal prediction of Euro-Atlantic teleconnections from multiple systems. *Environ Res Lett* 15:074009. <https://doi.org/10.1088/1748-9326/ab87d2>
- MacLachlan C, Arribas A, Peterson KA et al (2015) Global seasonal forecast system version 5 (GloSea5): a high-resolution seasonal forecast system. *Q J R Meteorol Soc* 141:1072–1084. <https://doi.org/10.1002/qj.2396>
- Madec G, the NEMO team (2016) Note du Pôle de Modélisation, Institut Pierre-Simon Laplace (IPSL), France, No 27 ISSN No 1288–1619
- Marsh R, Hickman AE, Sharples J (2015) S2P3-R (v1.0): a framework for efficient regional modelling of physical and biological structures and processes in shelf seas. *Geosci Model Dev* 8:3163–3178. <https://doi.org/10.5194/gmd-8-3163-2015>
- Megann A, Storkey D, Aksenov Y et al (2014) GO5.0: the joint NERC–Met Office NEMO global ocean model for use in coupled and forced applications. *Geosci Model Dev* 7:1069–1092. <https://doi.org/10.5194/gmd-7-1069-2014>
- Mills KE, Pershing AJ, Hernández CM (2017) Forecasting the Seasonal timing of Maine's lobster Fishery. *Front Mar Sci* 4. <https://doi.org/10.3389/fmars.2017.00337>
- North GR, Bell TL, Cahalan RF, Moeng FJ (1982) Sampling errors in the estimation of empirical orthogonal functions. *Mon Weather Rev* 110:699–706. [https://doi.org/10.1175/1520-0493\(1982\)110%3C0699:SEITEO%3E2.0.CO;2](https://doi.org/10.1175/1520-0493(1982)110%3C0699:SEITEO%3E2.0.CO;2)
- Otto L, Zimmerman JTF, Furnes GK et al (1990) Review of the physical oceanography of the North Sea. *Neth J Sea Res* 26:161–238. [https://doi.org/10.1016/0077-7579\(90\)90091-T](https://doi.org/10.1016/0077-7579(90)90091-T)
- Palin EJ, Scaife AA, Wallace E et al (2016) Skillful Seasonal forecasts of Winter disruption to the U.K. Transport System. *J Appl Meteorol Climatol* 55:325–344. <https://doi.org/10.1175/JAMC-D-15-0102.1>
- Pätsch J, Burchard H, Dieterich C et al (2017) An evaluation of the North Sea circulation in global and regional models relevant for ecosystem simulations. *Ocean Model* 116:70–95. <https://doi.org/10.1016/j.ocemod.2017.06.005>
- Patterson M, Weisheimer A, Befort DJ, O'Reilly CH (2022) The strong role of external forcing in seasonal forecasts of European summer temperature. *Environ Res Lett* 17:104033. <https://doi.org/10.1088/1748-9326/ac9243>
- Pauly D, Christensen V, Guénette S et al (2002) Towards sustainability in world fisheries. *Nature* 418:689–695. <https://doi.org/10.1038/nature01017>
- Payne MR, Hobday AJ, MacKenzie BR et al (2017) Lessons from the first generation of marine ecological forecast products. *Front Mar Sci* 4:289. <https://doi.org/10.3389/fmars.2017.00289/BIBTEX>
- Payne MR, Hobday AJ, MacKenzie BR, Tommasi D (2019) Editorial: Seasonal-to-decadal prediction of marine ecosystems: opportunities, approaches, and applications. *Front Mar Sci* 6:100. <https://doi.org/10.3389/fmars.2019.00100/BIBTEX>
- Pingree RD, Griffiths DK (1978) Tidal fronts on the shelf seas around the British Isles. *J Geophys Res: Oceans* 83:4615–4622. <https://doi.org/10.1029/JC083iC09p04615>
- Renshaw R, Wakelin S, Golbeck I, O'Dea E (2021) Quality Information Document - North West European Shelf Production Centre NWSHELF_MULTIYEAR_PHY_004_009. Copernicus Marine Environment Monitoring Service
- Riddle EE, Butler AH, Furtado JC et al (2013) CFSv2 ensemble prediction of the wintertime Arctic Oscillation. *Clim Dyn* 41:1099–1116. <https://doi.org/10.1007/s00382-013-1850-5>
- Rieger N, Levang SJ (2024) Xeof: Comprehensive EOF analysis in Python with xarray. *JOSS* 9:0606. <https://doi.org/10.21105/joss.06060>
- Scaife AA, Smith D (2018) A signal-to-noise paradox in climate science. *npj Clim Atmos Sci* 1:1–8. <https://doi.org/10.1038/s41612-018-0038-4>
- Scaife AA, Arribas A, Blockley E et al (2014) Skillful long-range prediction of European and north American winters. *Geophys Res Lett* 41:2514–2519. <https://doi.org/10.1002/2014GL059637>
- Scott T, McCarroll RJ, Masselink G et al (2021) Role of Atmospheric indices in describing Inshore Directional Wave Climate in the United Kingdom and Ireland. *Earth's Future* 9. <https://doi.org/10.1029/2020EF001625>. e2020EF001625
- Sharples J, Ross ON, Scott BE et al (2006) Inter-annual variability in the timing of stratification and the spring bloom in the North-Western North Sea. *Cont Shelf Res* 26:733–751. <https://doi.org/10.1016/j.csr.2006.01.011>
- Simmons HL, Hallberg RW, Arbic BK (2004) Internal wave generation in a global baroclinic tide model. *Deep Sea Res Part II* 51:3043–3068. <https://doi.org/10.1016/j.dsr2.2004.09.015>
- Simpson J, Bos W, Schirmer F et al (1993) Periodic stratification in the rhine ROFI in the north-sea. *Oceanol Acta* 16:23–32
- Smith DM, Scaife AA, Eade R et al (2020) North Atlantic climate far more predictable than models imply. *Nature* 583:796–800. <https://doi.org/10.1038/s41586-020-2525-0>
- Smyth TJ, Fishwick JR, AL-Moosawi L et al (2010) A broad spatio-temporal view of the western English Channel observatory. *J Plankton Res* 32:585–601. <https://doi.org/10.1093/plankt/fbp128>
- Stock CA, Pegion K, Vecchi GA et al (2015) Seasonal sea surface temperature anomaly prediction for coastal ecosystems. *Prog Oceanogr* 137:219–236. <https://doi.org/10.1016/j.pocean.2015.06.007>
- Thornton HE, Smith DM, Scaife AA, Dunstone NJ (2023) Seasonal predictability of the East Atlantic Pattern in Late Autumn and Early Winter. *Geophys Res Lett* 50. <https://doi.org/10.1029/2022GL100712>. e2022GL100712
- Tinker J, Hermanson L (2021) Towards Winter Seasonal predictability of the North West European Shelf seas. *Front Mar Sci* 8:979. <https://doi.org/10.3389/fmars.2021.698997/BIBTEX>
- Tinker J, Lowe J, Pardaens A et al (2016) Uncertainty in climate projections for the 21st century northwest European shelf seas. *Prog Oceanogr* 148:56–73. <https://doi.org/10.1016/j.pocean.2016.09.003>
- Tinker J, Krijnen J, Wood R et al (2018) What are the prospects for seasonal prediction of the marine environment of the North-West European Shelf? *Ocean Sci* 14:887–909. <https://doi.org/10.5194/OS-14-887-2018>
- Tinker J, Polton JA, Robins PE et al (2022) The influence of tides on the North West European shelf winter residual circulation. *Front Mar Sci* 9. <https://doi.org/10.3389/fmars.2022.847138>
- Tonani M, Bruciaferri D, Pequignet C et al (2019) Quality Information Document - North West European Shelf Production Centre NORTHWESTSHELF_ANALYSIS_FORECAST_PHYS_004_013. Copernicus Marine Environment Monitoring Service

- Turco M, Ceglar A, Prodhomme C et al (2017) Summer drought predictability over Europe: empirical versus dynamical forecasts. *Environ Res Lett* 12:084006. <https://doi.org/10.1088/1748-9326/aa7859>
- van Leeuwen S, Tett P, Mills D, van der Molen J (2015) Stratified and nonstratified areas in the North Sea: long-term variability and biological and policy implications. *J Geophys Research: Oceans* 120:4670–4686. <https://doi.org/10.1002/2014JC010485>
- Wang M, Tan B (2020) Two types of the scandinavian pattern: their formation mechanisms and climate impacts. *J Clim* 33:2645–2661. <https://doi.org/10.1175/JCLI-D-19-0447.1>
- West H, Quinn N, Horswell M (2021) Monthly rainfall signatures of the North Atlantic Oscillation and East Atlantic Pattern in Great Britain. *Atmosphere* 12:1533. <https://doi.org/10.3390/atmos12111533>
- Wilks DS (2011) *Statistical methods in the atmospheric sciences*, 3rd edn. Academic Press
- Williams KD, Harris CM, Bodas-Salcedo A et al (2015) The Met Office Global coupled model 2.0 (GC2) configuration. *Geosci Model Dev* 8:1509–1524. <https://doi.org/10.5194/gmd-8-1509-2015>
- Williams KD, Copsey D, Blockley EW et al (2018) The Met Office Global coupled Model 3.0 and 3.1 (GC3.0 and GC3.1) configurations. *J Adv Model Earth Syst* 10:357–380. <https://doi.org/10.1002/2017MS001115>
- Woollings T, Hannachi A, Hoskins B (2010) Variability of the North Atlantic eddy-driven jet stream. *Q J R Meteorol Soc* 136:856–868. <https://doi.org/10.1002/qj.625>
- Zhang L, Delworth TL, Yang X, Zeng F (2023) Skillful multiyear to decadal predictions of sea level in the North Atlantic Ocean and U.S. East Coast. *Commun Earth Environ* 4:1–10. <https://doi.org/10.1038/s43247-023-01093-w>

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