Signal-to-noise errors in early winter Euro-Atlantic predictions linked to weak ENSO teleconnections and pervasive jet biases

Received: 19 January 2024 | Revised: 24 January 2025 | Accepted: 29 January 2025

DOI: 10.1002/qj.4952

Quarterly Journal of the Royal Meteorological Society RMetS

RESEARCH ARTICLE

Signal-to-noise errors in early winter Euro-Atlantic predictions linked to weak ENSO teleconnections and pervasive jet biases

Christopher H. O'Reilly

Department of Meteorology, University of Reading, Reading, UK

Correspondence
Christopher H. O'Reilly, Department of Meteorology, University of Reading, Reading, UK.
Email: c.h.oreilly@reading.ac.uk

Funding information
Royal Society, Grant/Award Number: URF\R1\201230

Abstract

Long-range winter predictions over the Euro-Atlantic sector have demonstrated significant skill but suffer from systematic signal-to-noise errors. Here, I examine sources of early winter seasonal predictability in 16 state-of-the-art seasonal forecasting systems. As in previous studies, these systems demonstrate skill in the hindcasts of the large-scale atmospheric circulation in early winter, associated with the East Atlantic pattern. The predictability is strongly tied to the El Niño–Southern Oscillation (ENSO) teleconnection to the North Atlantic, though the systems' response to ENSO is systematically too weak. The hindcasts of the East Atlantic index exhibit substantial signal-to-noise errors, with the systems' predicted signal generally being smaller than would be expected for the observed level of skill, though there is substantial spread across systems. The signal-to-noise errors are found to be strongly linked to the strength of the ENSO teleconnection in the systems; those with a weaker teleconnection exhibit a larger signal-to-noise problem. The dependency on modelled ENSO teleconnection strength closely follows a simple scaling relationship derived from a toy model. Further analysis reveals that the strength of the ENSO teleconnection in the systems is linked to pervasive climatological biases in the North Atlantic and North Pacific jets. More specifically, systems that better represent the dynamics of the North Atlantic jet—with more frequent poleward jet excursions, less frequent high-latitude blocking, and a more poleward climatological jet position—are better at representing the ENSO teleconnection to the North Atlantic in early winter, with lower associated signal-to-noise errors.

KEYWORDS

atmospheric dynamics, ENSO, large-scale circulation, predictability, seasonal forecasting

1 | INTRODUCTION

The variability in wintertime climate over Europe, as well as parts of North America, is strongly controlled by variability in the large-scale atmospheric circulation over the extratropical North Atlantic. As a result, there is substantial interest in long-range, or "seasonal", forecasts (i.e., lead times of a month or more) of these large-scale circulation anomalies. Historically, long-range forecast skill over the North Atlantic had proven to be elusive (e.g. Johansson, 2007; Smith *et al.*, 2012). However, more recent forecast systems have demonstrated increased levels of skill over the North Atlantic (e.g. Scaife *et al.*, 2014; Dunstone *et al.*, 2016; Baker *et al.*, 2018), opening up new avenues for the application of these long-range forecasts (e.g. Clark *et al.*, 2017; Stringer *et al.*, 2020; Thornton *et al.*, 2019). Previous studies have largely focused on understanding the long-range prediction skill of the North Atlantic Oscillation (NAO) because it is the dominant mode of large-scale circulation variability over the Euro-Atlantic sector (e.g. Hurrell *et al.*, 2003). However, it has recently been shown that early winter (November–January) predictions of the East Atlantic (EA) pattern, the second largest mode of large-scale circulation variability over the Euro-Atlantic sector, are skilful in many state-of-the-art seasonal forecasting systems (Thornton *et al.*, 2023).

The main source of skill in long-range predictions of early winter Euro-Atlantic circulation variability is the El Niño-Southern Oscillation (ENSO) phenomenon in the tropical Pacific Ocean (Thornton *et al.*, 2023). During early winter, ENSO variability is strongly correlated with variability in the EA pattern over the North Atlantic (Ayarzagüena *et al.*, 2018; King *et al.*, 2018), with El Niño years projecting onto a positive phase of the EA pattern, bringing significantly milder and wetter conditions to western Europe, with the opposite conditions typically occurring in La Niña years. The influence of ENSO on the EA pattern in early winter is characterised by the suppression of poleward jet excursions during El Niño years and a zonal extension of the jet (O'Reilly *et al.*, 2024). Recent studies show that whereas the ENSO teleconnection to the North Atlantic in early winter, specifically the link between ENSO and the EA pattern, is robustly reproduced by state-of-the-art seasonal forecasting systems, the teleconnection in the systems is much weaker than that observed in reanalysis datasets (Molteni & Brookshaw, 2023; Thornton *et al.*, 2023). However, the underlying causes for the weak teleconnection, and the associated weak forecast signals, remain unclear.

Weak signals in long-range forecasts of the extratropical large-scale circulation are not unique to the early winter North Atlantic. For many long-range initialised ensemble "hindcasts" (or reforecasts of past seasons), the correlation skill with observations is much larger than should be possible given the signal-to-noise ratio (or perfect model correlation) of the hindcast ensemble. Previous studies have shown that these problems exist for later winter seasonal forecasts (e.g. Scaife *et al.*, 2014; Dunstone *et al.*, 2016; Baker *et al.*, 2018), subseasonal forecasts over the North Pacific (Garfinkel *et al.*, 2022), decadal forecasts of the wintertime North Atlantic (e.g. Smith *et al.*, 2019, 2020; Marcheggiani *et al.*, 2023), and summertime seasonal forecasts over the North Atlantic (e.g. Dunstone, 2018, 2023) and may be related to deficiencies in decadal large-scale circulation variability in free-running climate model simulations (e.g. Bracegirdle *et al.*, 2018; Simpson *et al.*, 2018; O'Reilly *et al.*, 2019a, 2019b, 2021). These signal-to-noise errors have collectively been dubbed the "signal-to-noise problem" (or "signal-to-noise paradox") in the climate science literature (Scaife & Smith, 2018). The signal-to-noise problem is a major challenge within climate science, as these errors significantly limit confidence in regional climate predictions made using model simulations, over a range of time-scales.

A number of theories for the underlying cause, or causes, of the signal-to-noise problem have been proposed. Recent studies have pointed to insufficient atmospheric eddy feedback in models, possibly due to low atmospheric resolution, as being responsible for the weak predicted signal in models (Scaife *et al.*, 2019; Hardiman *et al.*, 2022). Some studies have suggested that the misrepresentation of regime persistence, possibly linked to deficiencies in eddy feedbacks, as a possible explanation of why the signal-to-noise problem emerges (Strommen, 2020; Strommen & Palmer, 2019). Other studies have indicated that models are lacking in their response to specific predictable drivers, such as those associated with midlatitude ocean–atmosphere interactions (Ossó *et al.*, 2020; Zhang *et al.*, 2021) or low-frequency variability in the stratosphere (O'Reilly *et al.*, 2019b; Charlton-Perez *et al.*, 2019). These are not all mutually exclusive and may be of varying importance in the different manifestations of the signal-to-noise problem. Despite there being a number of proposed theories, there remains considerable uncertainty about the origins of the signal-to-noise problem in extratropical circulation variability.

In this study I analyse the predictability of the large-scale circulation over the North Atlantic in a suite of seasonal forecasting systems, aiming to understand the causes of the signal-to-noise errors in the early winter predictions. I find that, for all the systems, the majority of the seasonal forecast skill during this period can be attributed to the ENSO teleconnection, but the ENSO teleconnection is too weak in the systems. The strength of the teleconnection is shown to account for the variation of the

signal-to-noise ratios across the systems, and this scaling can be explained using a toy model of the forecasts. The strength of the ENSO teleconnection is shown to be linked to pervasive biases in the North Atlantic jet—systems whose climatological behaviour is closer to observations are found to have a stronger ENSO teleconnection to the North Atlantic and reduced signal-to-noise issues. These findings provide useful benchmarks for the improvement of operational seasonal forecasting systems and the identification of signal-to-noise errors in other instances.

2 | DATASETS AND METHODS

2.1 | Reanalysis data

We use the state-of-the-art European Centre for Medium-range Weather Forecasts Reanalysis v5 (ERA5) dataset (Hersbach *et al.*, 2020) as the reference dataset in the analysis that follows. ERA5 data are used over the period 1950–2020, comprising 71 winters in total; and a shorter period that is the same as the hindcasts, 1993–2016, is also used in places.

2.2 | Seasonal forecasting systems

In this study I analyse data from hindcasts, which are retrospective forecasts routinely performed by operational forecasting systems. I use hindcasts from 16 different seasonal forecasting systems that are stored in the Copernicus Climate Change Service (C3S) multi-model archive (see Table 1 for details). It should be noted that these systems are from eight different international forecasting centres, so some systems may be expected to be more similar than others, and the 16 systems are not wholly independent. The systems include current and recently operational systems. I have chosen to analyse all the systems in the C3S archive that have hindcasts covering the common period 1993–2016 (i.e., 24 winters) with initialisation dates on or before October 1. Our analysis focuses on the early winter period, November and December, that has been shown to have substantial skill in the hindcasts (Thornton *et al.*, 2023), which is at least in part due to the strong ENSO teleconnection to the North Atlantic during the early winter (e.g. Ayarzagüena *et al.*, 2018). The systems vary in ensemble size, with between 10 and 42 members per hindcast year depending on the system. The C3S hindcast datasets were regridded to a common 2.5° × 2.5° grid for the analysis with the exception of the eddy-driven jet latitude diagnostics, which were performed using U_{850} data regridded to a 1° × 1° grid.

TABLE 1 Seasonal forecast systems from the Copernicus Climate Change Service (C3S) archive analysed in this study.

Model name	Hindcast ensemble size	Centre of origin
CMCC-SPS3	40	Centro Euro-Mediterraneo sui Cambiamenti Climatici (CMCC)
CMCC-SPS3.5	40	CMCC
DWD-GCFS2.0	30	Deutscher Wetterdienst (DWD)
DWD-GCFS2.1	30	DWD
ECCC-CanCM4i	10	Environment and Climate Change Canada (ECCC)
ECCC-GEM-NEMO	10	ECCC
ECCC-GEM5-NEMO	10	EECCC
ECMWF-SEAS5	25	European Centre for Medium-Range Weather Forecasts (ECMWF)
JMA-CPS2	10	Japan Meteorological Agency (JMA)
JMA-CPS3	10	JMA
MF-Sys6	25	Météo-France (MF)
MF-Sys7	25	MF
MF-Sys8	25	MF
NCEP-CFSv2	12	National Centers for Environmental Prediction (NCEP)
UKMO-GloSea5-GC2-LI	42	UK Met Office (UKMO)
UKMO-GloSea6	42	UKMO

Note: Full details for these systems and the datasets are available from the C3S Climate Data Store (https://confluence.ecmwf.int/display/CKB/Description+of+the+C3S+seasonal+multi-system).

2.3 | ENSO index

We use the National Oceanic and Atmospheric Administration's "Oceanic Niño Index" methodology to define ENSO years, using the HadISST dataset (Rayner *et al.*, 2003). The Oceanic Niño Index methodology uses three-month averages of sea-surface temperatures (SSTs) averaged over the Nino-3.4 index region (170° W–120°W, 5° S–5°N). ENSO winters are identified when SST anomalies have a magnitude greater than 0.5 K relative to a moving 30-year averaged climatology (or the last 30 years where this window extends past the last date in the dataset). An additional requirement is that the SST anomaly must remain over the threshold for four consecutive rolling 3-month seasons, one of which must be December–February (DJF). Over the extended ERA5 period a total of 20 El Niño winters and 20 La Niña winters are identified, and over the C3S period a total of seven El Niño winters and eight La Niña winters are identified

(a list of the years is shown in Supporting Information Table S1). For the interannual correlations, the 3-month DJF winter Nino-3.4 SST index is used, calculated as already detailed from the HadISST dataset.

2.4 | EA index

The EA index is defined here as the second empirical orthogonal function (EOF) of the early winter (November–December) area-weighted mean-sea-level pressure (SLP) anomalies over the Euro-Atlantic sector (90° W–40° E, 20°–70° N). The ERA5 data over 1950–2020 are used to calculate the reference patterns and indices. The reference EOF patterns are shown in Supporting Information Figure S1. The C3S indices are calculated by projecting the SLP anomalies from each system onto the pattern of the EA from the ERA5 dataset and then renormalised. This is repeated for all of the C3S systems to generate the hindcast EA indices. The conclusions are not qualitatively sensitive to the choice of EA index definition, and signal-to-noise errors are also found if using the systems' own EA pattern (the equivalent to Figure 3 is shown for the systems' own EOF2 in Supporting Information Figure S3).

2.5 | Blocking event diagnostic

To assess the behaviour of atmospheric blocking I apply a two-dimensional large-scale wave-breaking index, which has been commonly used to identify blocking events in the literature (e.g. Woollings *et al.*, 2008). Here, I follow the methodology outlined in Masato *et al.* (2013). The blocking index uses daily averaged Z_{500} fields and identifies meridional reversals of the climatological Equator-to-Pole gradient, calculated over regions spanning 15° to the north and south of each point in the northern midlatitudes. Events must also extend at least 15° in longitude and are required to persist for at least 5 days to be identified as blocking events.

2.6 | North Atlantic eddy-driven jet diagnostic

In the following analysis I analyse the behaviour of the daily North Atlantic eddy-driven jet, its variability, and response to ENSO. To identify the latitude of the eddy-driven jet over the North Atlantic I broadly follow the method of Woollings *et al.* (2010). The daily zonal wind in the lower troposphere (at 850 hPa) is zonally averaged between 0° and 60° W, retaining values from 15° N to 75°N. The daily zonal mean zonal wind is then low-pass filtered using a 10-day Lanczos filter to identify changes in the jet on time-scales longer than those of individual synoptic systems. The North Atlantic eddy-driven jet latitude is identified as the latitude of the maximum wind speed for each day. These daily jet latitudes are used to compute probability distributions of the jet latitude using a kernel density estimate, with standard bandwidth $h = 1.06\sigma n^{-1/5}$, where σ is the standard deviation and n is the sample size (Silverman, 1981). In the probability density functions presented herein, I use the same h calculated from ERA5 to smooth the probability density functions from the C3S simulations, which provides a fairer comparison between the reanalysis and model data.

2.7 | Ratio of predictable components

To quantify the signal-to-noise in the hindcasts I compute the "ratio of predictable components" (RPC), which has previously been used in various studies evaluating forecast skill (e.g. Eade *et al.*, 2014; Scaife & Smith, 2018). The RPC is the ratio of the correlation skill between the ensemble mean hindcast and the observations (r_{mo}) and the correlation skill of the hindcast ensemble mean predicting a single ensemble member (r_{mm}):

$$\mathrm{RPC} = \frac{r_{\mathrm{mo}}}{r_{\mathrm{mm}}}. \quad (1)$$

To calculate r_{mm}, which can be referred to as a perfect model correlation, I remove one ensemble member from each season at random to create an individual realisation. The ensemble mean is calculated from the remaining ensemble members and correlated with the individual realisation. This is repeated 10,000 times and the resulting r^2 values are averaged; the square-root of this average gives the perfect model correlation, r_{mm}.

2.8 | Confidence intervals

To calculate confidence intervals I used a Monte Carlo bootstrap resampling (with replacement). The resampling was repeated at random 10,000 times to provide estimates of the confidence intervals shown.

3 | RESULTS

3.1 | Overview of early winter hindcast skill and ENSO teleconnection in the C3S systems

We begin by examining the ensemble mean hindcast correlation skill of the SLP in the C3S systems initialised

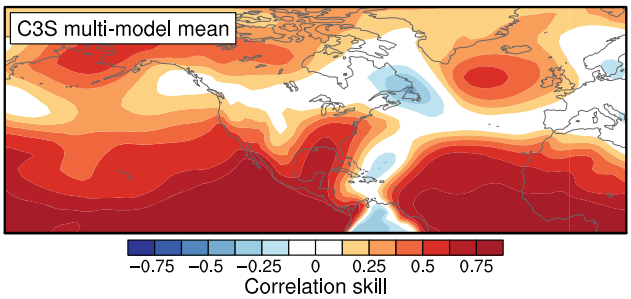

FIGURE 1 Multisystem ensemble mean hindcast correlation skill for early winter (November–December) sea-level pressure in the Copernicus Climate Change Service (C3S) systems, over the hindcast period 1993–2016. The equivalent plots for each of the 16 C3S systems are shown in Supporting Information Figure S2. [Colour figure can be viewed at wileyonlinelibrary.com]

on October 1, which is shown for the multisystem mean in Figure 1 (maps for each individual system are shown in Supporting Information Figure S2). As is typical for seasonal forecast systems, there is substantial correlation skill in the Tropics and over much of the North Pacific. The multisystem mean and most of the individual systems exhibit a local maximum in SLP correlation skill located over the extratropical North Atlantic, consistent with the results shown by Thornton *et al.* (2023) for the ensemble mean of a smaller subset of these C3S systems, albeit for the sightly different November–January season (here, I focus on the November–December early winter season, as this period has the strongest and most consistent ENSO teleconnection; O'Reilly *et al.* (2024)). The local maximum in SLP correlation skill in the eastern North Atlantic in the C3S systems projects onto the EA pattern (i.e., Supporting Information Figure S1), a region strongly associated with the early winter ENSO teleconnection.

To examine the representation of the early winter ENSO teleconnections in the C3S systems, I now examine the SLP difference between El Niño and La Niña years; these are shown in Figure 2 for each C3S system and ERA5 (correlations between the Nino-3.4 index and the SLP anomaly in each ensemble member are also shown, in contours). The C3S systems all show ENSO differences with negative SLP (and negative correlations) over the eastern North Atlantic, though there is substantial variation in the magnitude of the ENSO influence. In all systems, however, the ENSO teleconnection to the North Atlantic appears weaker than that seen in ERA5. This is most clear for the C3S reference period (i.e., 1993–2016), though the systems' teleconnections are also substantially weaker for the extended ERA5 period, which might be considered a statistically more robust measure of the observed teleconnection.

The weak early winter ENSO teleconnection to the North Atlantic is not only evident in the SLP anomaly. Figure 3 shows the C3S average teleconnection in terms of zonal wind anomalies, alongside the equivalent teleconnection estimated from ERA5. For the upper tropospheric winds (U_{200}) it is clear that the influence of ENSO on the North Atlantic jet anomalies is weaker in the C3S systems than in reanalysis. In terms of upper level winds, it seems the disparities are most obvious in the North Atlantic, with the North Pacific teleconnection being of similar strength in the C3S systems and the reanalysis. The ENSO impact on early winter blocking events is also shown in Figure 3. A recent observational study showed that the ENSO influence on the North Atlantic jet is established through changes in the frequency of poleward jet excursions and associated Iberian wave-breaking events (O'Reilly *et al.*, 2024), evident in the change in ERA5 blocking frequency shown here. In contrast, there are only very modest changes in the frequency of Iberian wave-breaking events associated with ENSO in the C3S systems. Together, these provide a consistent picture of the dynamical response to ENSO over the North Atlantic being weaker in the C3S systems than in observations.

3.2 | Signal-to-noise of the EA index hindcasts and link to the ENSO teleconnection strength

To compare the hindcast skill across the C3S systems more quantitatively it is useful to analyse the skill of the EA index (see Section 2.4). The EA index is a useful measure as it captures the main areas of skill over the North Atlantic during early winter and also dominates the ENSO teleconnection to the North Atlantic during this period. The ensemble mean hindcast correlation skill (i.e., r_{mo}) of the early winter EA index in the C3S systems is shown in Figure 4a. Skill varies across the systems, but the vast majority of the systems exhibit skill levels above $r = 0.3$, with only three of the systems exhibiting correlation skills with $P > 0.05$ (based on a t-test). Also shown in Figure 4a is the perfect model correlation (i.e., r_{mm}) for each of the C3S systems. For all but two of the systems, the perfect model correlation is lower than the hindcast correlation skill, and in some cases it is much lower. The signal-to-noise of the hindcast EA indices (in terms of RPC; see Section 2.7) is shown in Figure 4b. The C3S systems nearly all have RPC > 1, demonstrating that predictions of the early winter EA index are generally underconfident—consistent with the findings of Thornton *et al.* (2023).

To examine how ENSO influences the early winter EA index I computed the correlation between the early winter EA index and Nino-3.4 index across ensemble members for each C3S system; these are shown in Figure 5a, along with the equivalent correlation in ERA5. The correlation

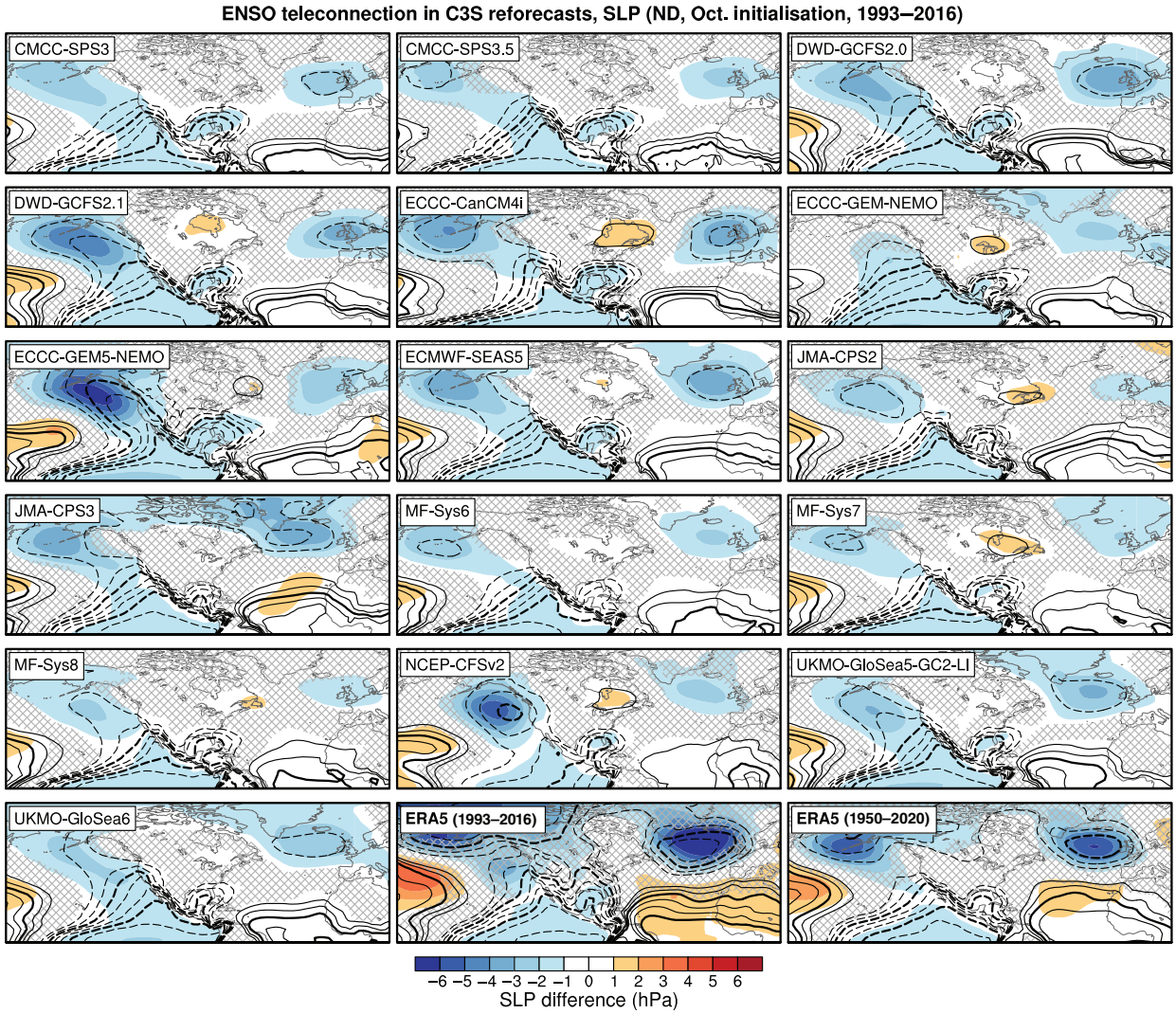

FIGURE 2 Early winter (November–December, ND) El Niño–Southern Oscillation (ENSO) teleconnection calculated using sea-level pressure (SLP) anomaly in each of the Copernicus Climate Change Service (C3S) systems over the period 1993–2016. Shading shows the composite difference between El Niño and La Niña years (defined using the Oceanic Niño Index; see Section 2.3), and the contours show the correlation between the SLP anomaly in each ensemble member and the observed Nino-3.4 sea-surface temperature index. Contours start from 0.2 with an interval of 0.1, and are emboldened at 0.4 and 0.7; negative contours are indicated by dashed lines. Also shown are the equivalent plots for the European Centre for Medium-range Weather Forecasts Reanalysis v5 (ERA5) data over an extended period (1950–2020). Hatching shows regions where the 5–95% confidence interval of the composite difference (calculated using a bootstrap resampling) intersects zero. [Colour figure can be viewed at wileyonlinelibrary.com]

values vary between $r = 0.1$ and 0.4 for the C3S systems, but these are all less than in ERA5. For the C3S period (1993–2016) the correlation in ERA5 is 0.57, though this short period is subject to substantial sampling uncertainty. However, even when estimated over a longer period (1950–2020) the correlation between the EA index and Nino-3.4 is 0.44, higher than any individual C3S system. The weak influence of ENSO on the EA index in early winter is consistent with the weak teleconnection patterns shown in Figures 2 and 3.

We now compare the signal-to-noise in the predictions of the EA index, in terms of RPC, with the strength of the ENSO teleconnection to the EA index in early winter, shown in Figure 5b. Previous studies have highlighted that the RPC is a more useful measure of the signal-to-noise ratio in model predictions that exhibit significant levels of skill (Hardiman *et al.*, 2022); following this convention, I plot the systems that have ensemble mean hindcast skill with $P < 0.05$ (see Figure 4), though the conclusions drawn from the analysis are not sensitive to this specific criterion. From the distribution of the points in Figure 5b it is clear that, across the C3S systems, those with weaker ENSO teleconnections generally have larger signal-to-noise errors, with a linear

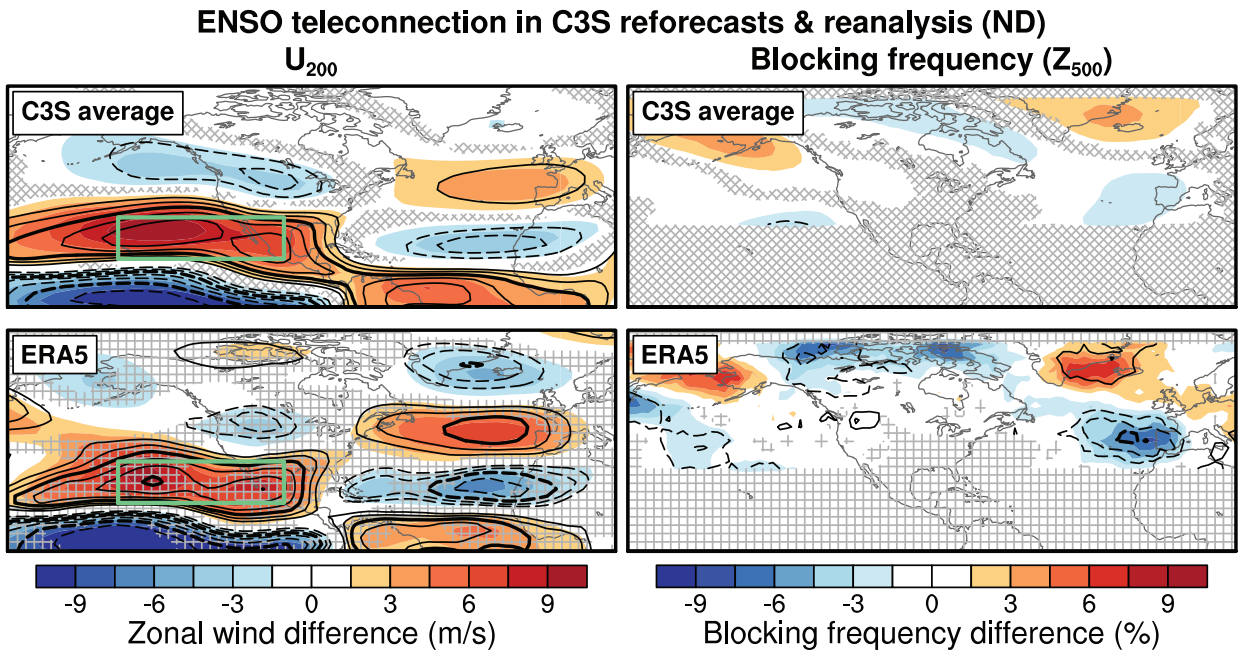

FIGURE 3 Early winter (November–December, ND) El Niño–Southern Oscillation (ENSO) teleconnection averaged across the 16 different Copernicus Climate Change Service (C3S) systems over the period 1993–2016, calculated for U_{200} and blocking frequency (see Section 2.5). The teleconnection metrics were calculated for each system individually, then averaged. Shading shows the composite difference between El Niño and La Niña years (defined using the Oceanic Niño Index; see Section 2.3) and the contours show the correlation between each variable anomaly in each ensemble member and the observed Nino-3.4 sea-surface temperature index. Contours start from 0.2 with an interval of 0.1, and are emboldened at 0.4 and 0.7; negative contours are indicated by dashed lines. Also shown are the equivalent plots for the European Centre for Medium-range Weather Forecasts Reanalysis v5 (ERA5) data over an extended period (1950–2020). Hatching for the reanalysis plots shows regions where the value is outside the 10–90% range of the C3S systems. Hatching for the C3S average plots shows regions where less than 75% of the systems have composite differences with the same sign. [Colour figure can be viewed at wileyonlinelibrary.com]

correlation of $r = -0.76$. The teleconnection strength (β_{mem}) is only weakly related to the EA hindcast skill ($r = 0.17$) but is strongly tied to the EA perfect model skill ($r = 0.88$), indicating that the teleconnection strength is impacting the RPC by reducing model signal. These results suggest that the weak ENSO teleconnection across the systems is responsible for causing the early winter signal-to-noise problem over the North Atlantic.

To provide some further insight into the relationship between the RPC and the ENSO teleconnection strength, I consider a toy model of the hindcasts, which I outline here. We will first model the EA index in the observations as being linearly dependent on ENSO:

$$\text{EA}^*_{\text{obs}} = \beta_{\text{obs}} N^*_{34} + \epsilon_{\text{obs}}, \tag{2}$$

where EA^*_{obs} is the normalised EA index, N^*_{34} is the normalised observed Nino-3.4 index, β_{obs} is a dimensionless regression coefficient, and ϵ_{obs} is a random residual term with a mean of zero. Similarly, we can model the (normalised) forecast ensemble mean EA index, EA^*_{em}, and the (normalised) forecast ensemble member EA indices, EA^*_{mem}, as

$$\begin{aligned} \text{EA}^*_{\text{em}} &= \beta_{\text{em}} N^*_{34} + \epsilon_{\text{em}}, \\ \text{EA}^*_{\text{mem}} &= \beta_{\text{mem}} N^*_{34} + \epsilon_{\text{mem}}. \end{aligned} \tag{3}$$

Note here that the normalised "observed" Nino-3.4 index, N^*_{34}, is included in the linear models as the seasonal forecasts of the Nino-3.4 index are very skilful over the lead times considered here, and this simplifies the expressions that follow (though does not materially affect the resulting scaling). Using the expressions for different normalised EA indices, we can now evaluate expectations of the correlations used to calculate the RPC:

$$\begin{aligned} r_{\text{mo}} &= \text{corr}(\text{EA}^*_{\text{obs}}, \text{EA}^*_{\text{em}}) \\ &= \text{corr}(\beta_{\text{obs}} N^*_{34} + \epsilon_{\text{obs}}, \beta_{\text{em}} N^*_{34} + \epsilon_{\text{em}}) \\ &\approx \beta_{\text{obs}} \beta_{\text{em}} \\ r_{\text{mm}} &= \text{corr}(\text{EA}^*_{\text{mem}}, \text{EA}^*_{\text{em}}) \\ &= \text{corr}(\beta_{\text{mem}} N^*_{34} + \epsilon_{\text{mem}}, \beta_{\text{em}} N^*_{34} + \epsilon_{\text{em}}) \\ &\approx \beta_{\text{mem}} \beta_{\text{em}}. \end{aligned} \tag{4}$$

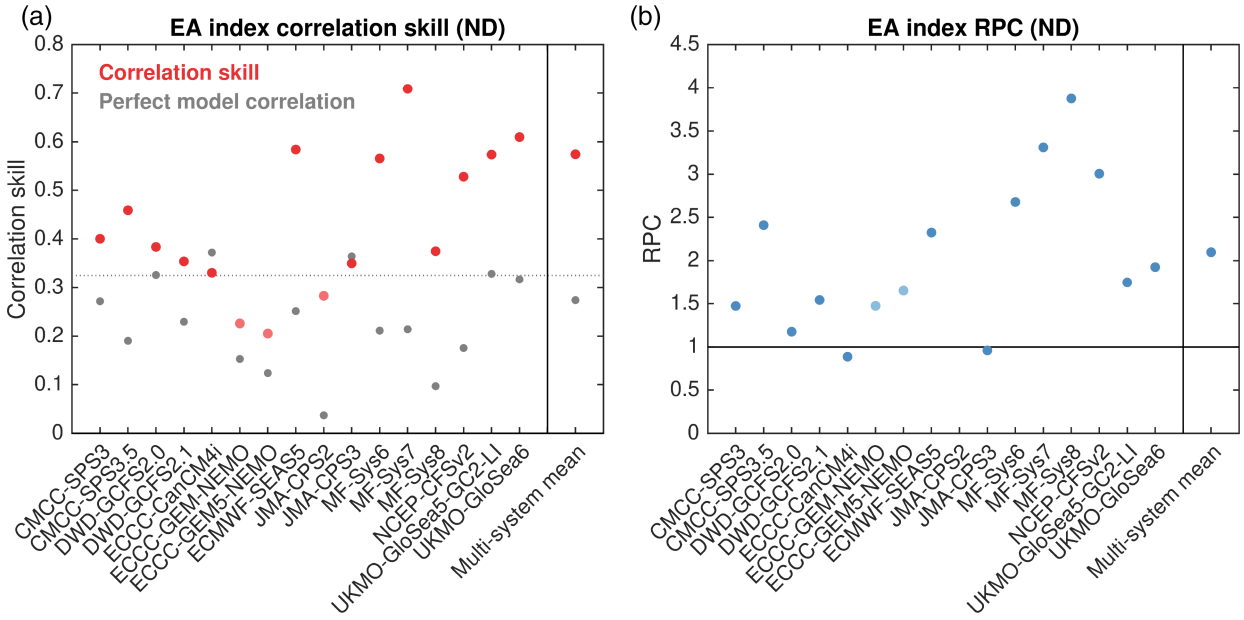

FIGURE 4 (a) Ensemble mean hindcast correlation skill (red circles) and the perfect model correlation skill (grey circles) for the East Atlantic (EA) index over the period 1993–2016 for each of the Copernicus Climate Change Service (C3S) systems (the dotted line indicates the correlation skill corresponding to $P = 0.05$ based on a t test). (b) Ratio of predictable components (RPC) for the EA index hindcasts for each of the C3S systems (the solid line indicates where RPC = 1, which would indicate a reliable forecast by this measure and systems with RPC > 1 being underconfident). Models that have hindcast correlation skills with $P < 0.05$ are indicated by lighter shaded circles in both panels. Values for a simple multisystem mean, produced by combining ensemble members from all systems, are also shown in both panels. [Colour figure can be viewed at wileyonlinelibrary.com]

Here, we have made the assumption that the covariance between the residual terms is zero and that the residual terms average to zero[1]. This results in a simple expected scaling of the RPC:

$$\mathrm{RPC} = \frac{r_{\mathrm{mo}}}{r_{\mathrm{mm}}} \approx \frac{\beta_{\mathrm{obs}}}{\beta_{\mathrm{mem}}}. \tag{5}$$

Therefore, if this is an appropriate model, we should expect the RPC of the EA index hindcasts to be dependent on the ratio of the observed ENSO teleconnection strength β_{obs} and the ensemble member ENSO teleconnection strength β_{mem}.

There are various assumptions that go into this toy model. An important assumption is that ENSO alone is responsible for the forecast skill in the EA index and that it does so in linear way. To test this assumption, I calculated the skill of a simple linear model fit to each hindcast system separately and compared this with the actual hindcast skill; the fraction of the actual hindcast skill (i.e., r^2_{mo}) that can be accounted for by the linear model is plotted in Supporting Information Figure S4. In all the C3S systems, the majority of the skill can be recovered in the simple linear ENSO model, indicating that the linear model is a reasonable approach. It is important to note that the systems that exhibit the highest hindcast correlation skill are those that cannot be fully explained by this linear model, which indicates some of the hindcast skill arises from other, more complex, sources. A related assumption is that the residual terms have zero covariance and zero mean (e.g., other sources of skill would result in a positive correlation between the residual terms). In reality, owing to finite ensemble sizes and short hindcast periods, these terms will not be exactly uncorrelated and any deviations from zero will deteriorate the fit of the scaling expression.

The predicted scaling of RPC for the early winter EA indices is plotted with the actual RPC values in Figure 5b. The green curve shows the RPC scaling for the observed ENSO teleconnection over the C3S period, along with shading that shows the sampling uncertainty. The RPC scaling from the toy model broadly captures the relationship between the actual RPC values calculated from the hindcasts and the actual ENSO teleconnection strength in the systems (note that, because the linear models are normalised, β_{mem} is equal to the correlation between the EA index and Nino-3.4 over ensemble members). In particular, the scaling highlights the expected nonlinearity of the RPC with respect to teleconnection strength, with consistent behaviour seen in the actual RPC values. The nonlinear scaling of the RPC in terms of teleconnection strength also highlights a potential difficulty in using RPC to discriminate between systems, as the

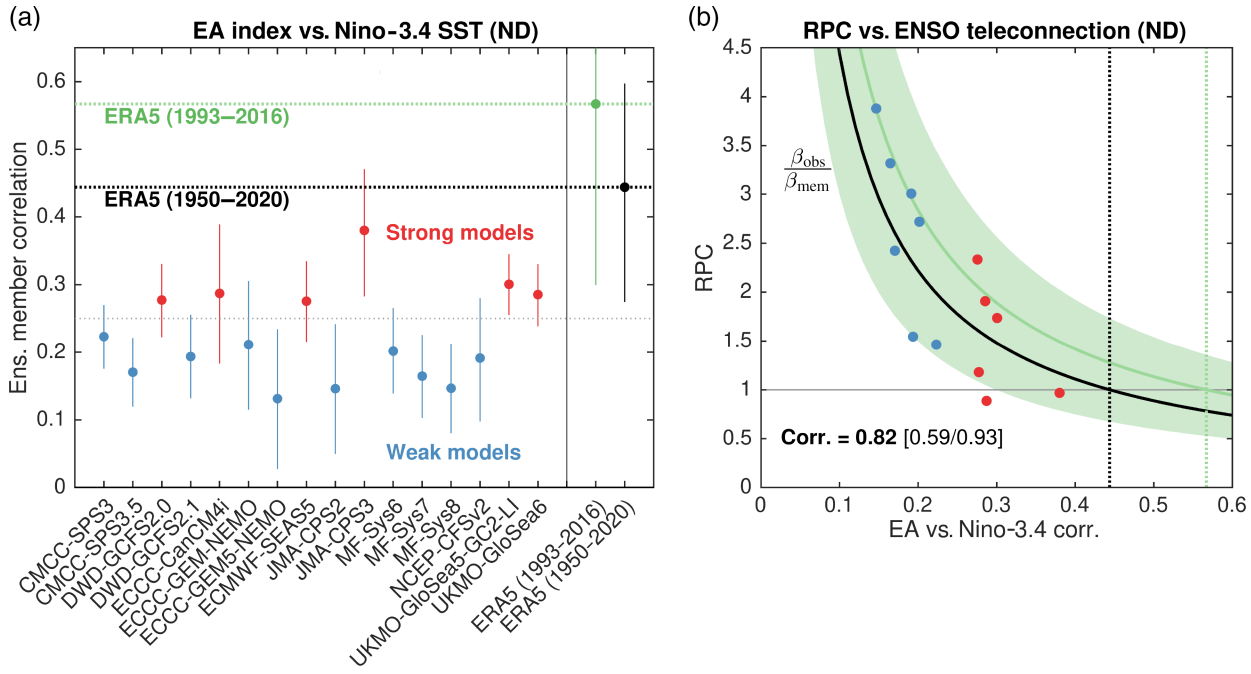

FIGURE 5 (a) Ensemble member correlation between the early winter (November–December, ND) East Atlantic (EA) index and the Nino-3.4 sea-surface temperature (SST) index for each of the Copernicus Climate Change Service (C3S) systems (1993–2016). Also shown in thick dashed lines are the equivalent correlation for the early winter EA index calculated from reanalysis data for an extended period (1950–2020) and a shorter period that matches the C3S systems (1993–2016). The systems are separated into two subsets based on the strength of this correlation, with systems greater than $r = 0.25$ corresponding to the "strong" subset (in red) and with systems less than $r = 0.25$ corresponding to the "weak" subset (in blue). The lines show the 5–95% confidence interval of the correlation coefficients calculated using a bootstrap resampling. (b) Relationship between the ratio of predictable components (RPC) and the EA index versus Nino-3.4 correlation—also equal to β_{mem} in the linear El Niño–Southern Oscillation (ENSO) mode (see text). Curves of the scaling for a linear ENSO model, β_{obs}/β_{mem}, are also shown for values of β_{obs} from the European Centre for Medium-range Weather Forecasts Reanalysis v5 (ERA5) data over an extended period (1950–2020, in black) and a shorter period that matches the C3S systems (1993–2016, in green). The shading shows a 5–95% confidence interval for the green β_{obs}/β_{mem} curve, estimated using a Monte Carlo resampling (random bootstrapping with replacement over years in the sample, repeated 10,000 times). [Colour figure can be viewed at wileyonlinelibrary.com]

expected RPC becomes more similar for systems as their ENSO teleconnections approach the observed strength. These difficulties are of course exacerbated by the sampling uncertainties due to the short 24-year hindcast period.

In this section I have shown that the C3S systems have robust but varying hindcast skill for the early winter EA index. However, the signals in the hindcasts, as measured by the perfect model correlation r_{mm}, are generally too weak in the systems, resulting in substantial signal-to-noise errors (i.e., Figure 4). The ENSO teleconnection to the EA index is too weak in all the hindcasts but shows substantial variability across the C3S systems. Further analysis shows that systems with a weaker teleconnection generally exhibit larger RPC values and, therefore, clearer signal-to-noise errors. Finally, I demonstrated that a toy model of the ENSO teleconnection to the early winter EA index can broadly explain the magnitude and RPC scaling across the systems, depending only on the modelled teleconnection strength.

3.3 | Exploring causes of the weak ENSO teleconnection in the C3S systems

In the previous sections we have seen that the weak ENSO teleconnection is largely responsible for the signal-to-noise errors observed in the early winter hindcasts. I now turn attention to the causes of this weak ENSO teleconnection to the North Atlantic in the early winter.

To begin, it is useful to revisit the mechanisms through which ENSO influences the North Atlantic circulation during the early winter. In a recent article, O'Reilly *et al.* (2024) showed that the ENSO teleconnection to the North Atlantic is largely through the modification of the poleward jet excursions, which project onto the EA pattern. The response of the North Atlantic jet is sensitive, on both subseasonal and seasonal time-scales, to the jet and storm track anomalies over the eastern North Pacific. A schematic view of this is shown in the simple causal chain diagram in Figure 6a—for example, following Kretschmer *et al.* (2021). Such a model is likely an oversimplification;

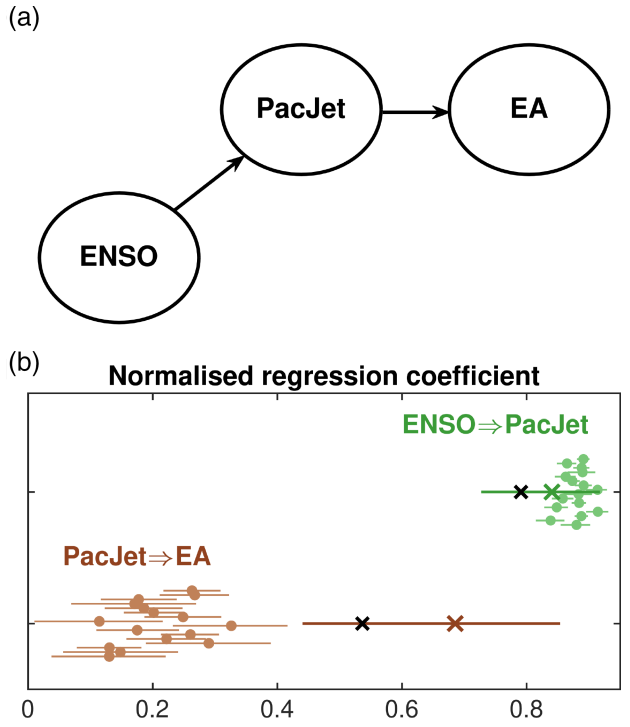

FIGURE 6 (a) Schematic of the causal chain—for example, following Kretschmer *et al.* (2021)—linking El Niño–Southern Oscillation (ENSO) variability to the East Atlantic (EA) index variability over the North Atlantic—based on O'Reilly *et al.* (2024). Here, "ENSO" refers specifically to the normalised Nino 3.4 SST index, "PacJet" refers to an index for the Pacific jet, defined as the normalised U_{200} anomaly averaged over the eastern North Pacific (shown by box in Figure 3), and "EA" refers to the normalised EA index. (b) The circles show the linear regression coefficients between the indices in the causal chain for each Copernicus Climate Change Service (C3S) system, and the crosses indicate the coefficients calculated from the European Centre for Medium-range Weather Forecasts Reanalysis v5 (ERA5) dataset (green for the C3S period 1993–2016 and black for the extended period 1950–2020). The horizontal lines show the 5–95% confidence interval of the correlation coefficients calculated using a bootstrap resampling. A jitter has been added to the *y*-axis to aid visualisation of the individual points. [Colour figure can be viewed at wileyonlinelibrary.com]

however, explicitly stating the causal chain in this way allows us to interrogate each step in the teleconnection in the systems as well as reanalysis and to identify any key differences. I define normalised indices for ENSO, the Pacific jet, and the EA pattern and use linear regression between these indices to calculate the strength of these connections in each C3S system and in ERA5; see Figure 6b.

The C3S systems all exhibit strong relationships between ENSO and the Pacific jet (green points), which are all very similar to the value calculated from ERA5 data. However, the link between the Pacific jet and the early winter EA index is much more variable across the C3S systems (brown points), though all systems are substantially weaker than the link between the Pacific jet and EA index calculated from ERA5 data. This simple analysis suggests that the biggest differences in the total ENSO teleconnection pathway stem from the deficiencies in the response of the North Atlantic circulation to upstream circulation anomalies over the North Pacific. This conclusion is supported by the average C3S U_{200} teleconnection maps, shown in Figure 3, which show similar anomalies to ERA5 over the North Pacific (and within the C3S model range) but much weaker responses over the North Atlantic. Analysis of the ENSO SST and precipitation anomalies in the Tropics reveals relatively modest differences across the C3S models (Supporting Information Figure S5). Taken together, these results point to an extratropical deficiency in the teleconnection pathway and are broadly consistent with the experimental results of Knight *et al.* (2022), which implicated the extratropical dynamics as a cause of the signal-to-noise errors in seasonal hindcasts of the (DJF) NAO in the UKMO-GloSea5-GC2-LI system. The results from the causal chain analysis (Figure 6) indicate that differences in the ENSO teleconnection originate from differences in the behaviour of the North Atlantic jet across the C3S systems.

To explore the possible causes for the differences in ENSO teleconnection strength over the North Atlantic, I define two C3S system subsets based on the strength of the correlation between the EA index and the Nino-3.4 index across all ensemble members (shown in Figure 5a). The threshold was set at $\beta_{\text{mem}} = 0.25$, since from visual inspection (of Figure 5a) this provided the clearest separation of the systems; this threshold results in six systems in the "strong" model subset and 10 systems in the "weak" model subset. To examine the differences in model behaviour I first examine the climatologies of the zonal wind in the C3S systems; the differences between the strong and weak subsets and the average C3S system bias with respect to ERA5 are shown in Figure 7.

There are some clear differences between the strong and weak model subsets over the North Atlantic, with the stronger models having stronger zonal winds over the northern part of the basin (Figure 7a). Over the North Atlantic, the midlatitude jet is generally too far south in the C3S systems, with significantly weaker winds over the northern part of the basin (Figure 7b). The strong subset of models, therefore, have reduced biases over a northern band of the North Atlantic basin (i.e., between the southern tip of Greenland and Scotland). To compare the North Atlantic jet in the C3S systems and observations further, the early winter ENSO teleconnection strength (i.e., β_{mem})

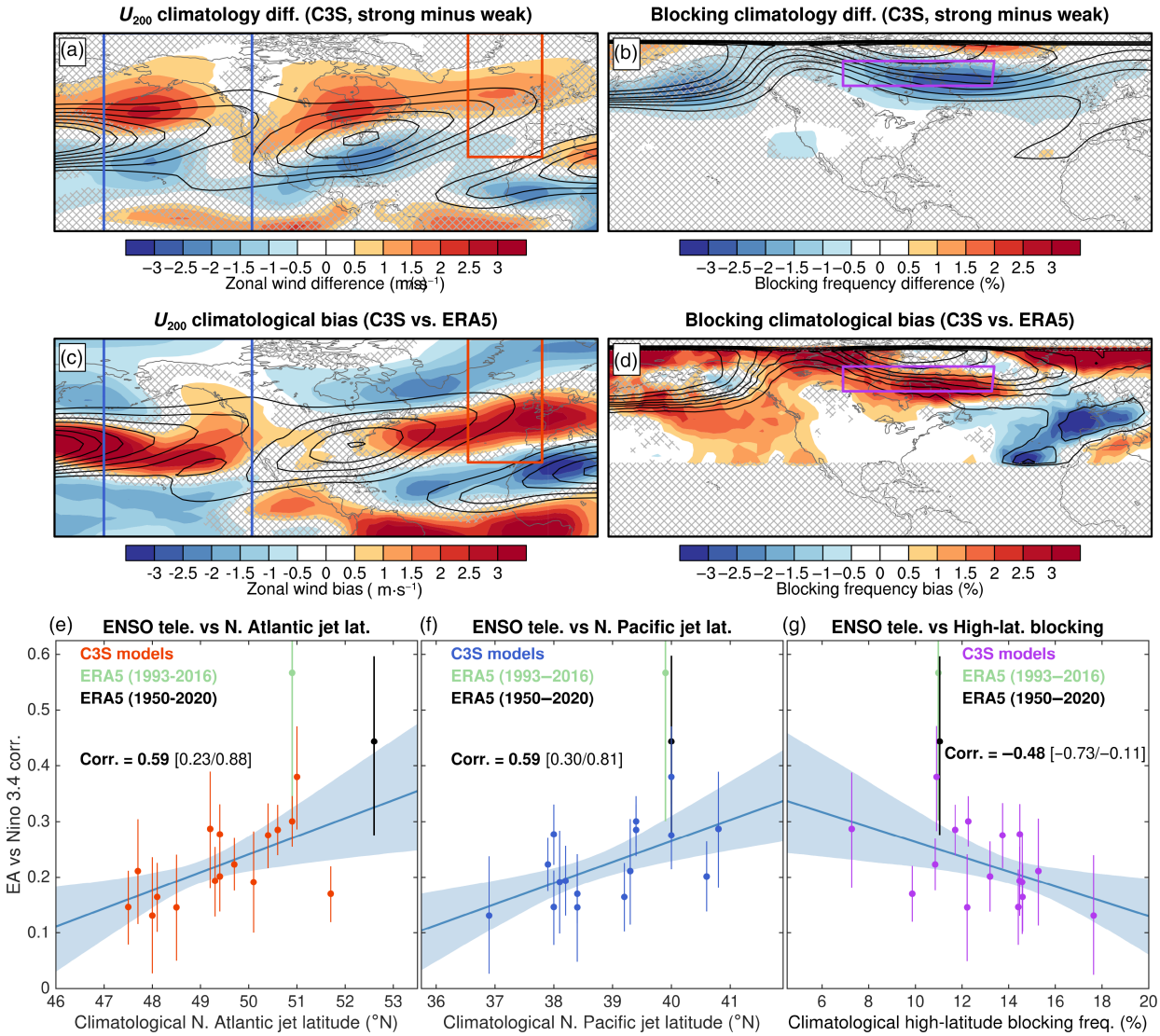

FIGURE 7 (a) The composite U_{200} difference between the climatologies of the strong and weak subsets of Copernicus Climate Change Service (C3S) systems (as defined in the text and Figure 5a). The C3S average climatology is shown in black contours every 5 m · s^{-1} from 20 m · s^{-1}. (c) The C3S average climatological U_{200} bias with respect to European Centre for Medium-range Weather Forecasts Reanalysis v5 (ERA5; defined C3S minus ERA5). The ERA5 climatology (1950–2020) is shown in black contours every 5 m · s^{-1} from 20 m · s^{-1}s. (b, d) As in (a) and (c) but for blocking frequency (see Section 2.5); climatology is shown in black contours at 5%, 7.5%, 10%, 12.5%, 15%, 20%, and 25%. Hatching shows where the 5–95% confidence interval of the difference/bias crosses zero; the confidence intervals are estimated using a Monte Carlo resampling (random bootstrapping with replacement, repeated 10,000 times). (e) Scatter plot of the El Niño–Southern Oscillation (ENSO) teleconnection strength (i.e., Figure 5a) against climatological North Atlantic jet latitude—defined as the latitude of the maximum U_{200} zonally averaged over the jet extension region outlined in the red box in (a) and (c). (f) As in (e) but for the climatological North Pacific jet latitude calculated over the region outlined in the blue box in (a) and (c). (g) As in (e) but for the climatological high-latitude blocking frequency, averaged over the Canada–Greenland region–shown by the purple boxes in (b) and (d). The blue lines in (e), (f), and (g) show the line of best fit across the C3S systems, and the shading shows the 5–95% confidence interval of the fit, calculated using a bootstrap resampling (the associated correlation across the systems is also shown). The vertical lines in (e), (f), and (g) show the 5–95% confidence interval of the correlation coefficients calculated using a bootstrap resampling. [Colour figure can be viewed at wileyonlinelibrary.com]

is plotted against the climatological North Atlantic jet latitude in Figure 7e (where jet latitude is defined as the latitude of the maximum of the zonal mean U_{200}, averaged over the jet extension region shown in red in Figure 7). It is clear in Figure 7e that the latitude of the climatological North Atlantic jet over early winter tends to be located further south in the C3S systems and, moreover, the strength of the ENSO teleconnection to the North Atlantic seems to be clearly related to the climatological jet latitude across the C3S systems ($r = 0.59$, Figure 7e).

There are also substantial differences between the strong and weak model subsets in the upper level jet upstream over the North Pacific (Figure 7a), indicating a potential influence on the ENSO teleconnection to the North Atlantic. In observational data, equatorward shifts in the North Pacific jet and storm track tend to reduce the frequency of the poleward jet shifts downstream over the North Atlantic, which are crucial for shaping the early winter ENSO teleconnection (O'Reilly *et al.*, 2024). The North Pacific jet in the C3S systems shows a strong and equatorward bias, with the jet extending further downstream than in reanalysis (Figure 7c), and the climatological jet latitude seems to be clearly related to the ENSO teleconnection to the North Atlantic across the C3S systems ($r = 0.59$, Figure 7f). The climatological North Pacific and North Atlantic jet latitudes, interestingly, seem to be only weakly related across the C3S systems ($r = 0.28$), and they both exhibit an independent influence over the early winter teleconnection strength. Together, the analysis of the climatological jets here suggests that a better model representation of climatological North Atlantic and North Pacific jet behaviour improves the fidelity of the early winter ENSO teleconnection.

To further examine the differences in North Atlantic jet behaviour across the C3S systems, we now analyse distributions of the daily North Atlantic eddy-driven jet latitude (see Section 2.6), shown for the C3S systems and ERA5 in Figure 7a. The C3S systems clearly underestimate the frequency of the poleward jet excursions, around 55–60° N, and generally overestimate the jet frequency further south. This is consistent with the climatological mean jets shown in Figure 8. The strong model subset exhibits higher frequencies of poleward jet excursions on average, and their behaviour is closer to that seen in the reanalysis, compared with the weak model subset. The weak model subset tends to more strongly overestimate the southern jet frequency, around 35–40° N, compared with the strong model subset.

A consistent picture that emerges from this jet latitude analysis is that systems with a better representation of the poleward jet events generally have stronger early winter ENSO teleconnections. To demonstrate this more clearly I have plotted the difference between the jet latitude between El Niño and La Niña years (in Figure 8b). The ENSO teleconnection in observations is strongly connected to changes in the occurrence of poleward jet excursions, which occur more frequently in La Niña years (O'Reilly *et al.*, 2024). The C3S systems that struggle to simulate these events often enough in the climatology (i.e., Figure 8a) tend to be those systems that show a smaller jet latitude frequency difference in response to ENSO and, therefore, show a weaker ENSO teleconnection.

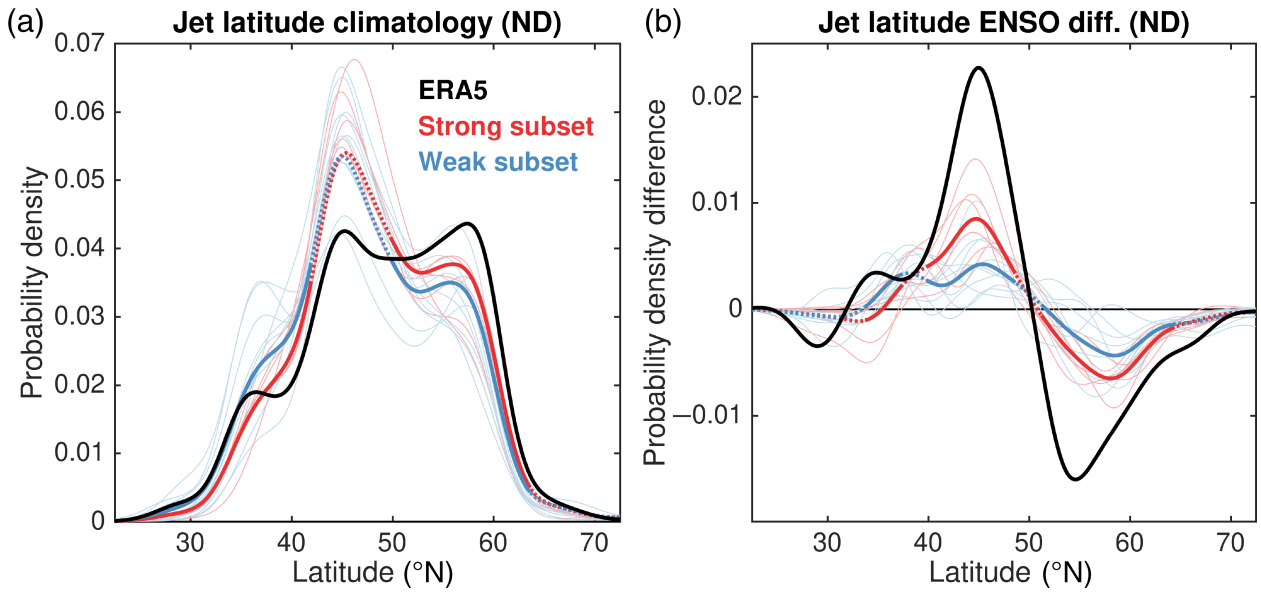

FIGURE 8 (a) Climatological eddy-driven jet latitude probability density functions (pdfs; see Section 2.6) shown for each individual Copernicus Climate Change Service (C3S) system in light-coloured lines and for European Centre for Medium-range Weather Forecasts Reanalysis v5 (ERA5) in black. The averages of these pdfs for the strong and weak subsets of models are shown in the thick red and blue lines respectively. (b) As in (a) but for the difference in eddy-driven jet latitude pdfs between El Niño and La Niña years. The dotted thick red/blue lines shows where the 5–95% confidence interval of the difference between the strong and weak subsets crosses zero; the confidence intervals are estimated using a Monte Carlo resampling (random bootstrapping with replacement, repeated 10,000 times). [Colour figure can be viewed at wileyonlinelibrary.com]

The early winter ENSO teleconnection is strongly linked to changes in large-scale blocking frequency near the Iberian Peninsula in observations (i.e., Figure 3b; O'Reilly *et al.* (2024)), so we will now examine the climatological behaviour of large-scale blocking in the C3S systems. The C3S systems generally have too little blocking over the Iberian Peninsula (Figure 7d); however, this problem seems to plague all the systems and is only weakly linked to the teleconnection strength (Figure 7b). The clearest difference between the strong and weak models is found further upstream, over the Canada–Greenland region, where significantly more high-latitude blocking occurs in the weak subset. On average, the C3S systems exhibit too much blocking over this region (Figure 7d), with the stronger subset of models demonstrating better agreement with observations compared with the weaker models. High-latitude blocking events over this region are typically associated with southward shifts in the jet (e.g. Woollings *et al.*, 2010), so the higher blocking frequency in weaker subsets of models is broadly consistent with the equatorward North Atlantic jet bias in these systems (i.e., Figure 7a,e). To explore this further, Figure 7g shows the climatological high-latitude blocking over this Canada/Greenland region plotted against the ENSO teleconnection strength. Although this relationship does not fully account for the discrepancies between the C3S systems and observations, the systems with reduced biases in high-latitude blocking generally have stronger teleconnections ($r = -0.48$). Climatological blocking frequency and climatological North Atlantic jet latitude (i.e., Figure 7e) are anticorrelated, $r = -0.55$, indicating that models with reduced upstream high-latitude blocking biases tend to be those models with a more realistic North Atlantic jet and exhibit an improved ENSO teleconnection in the forecast systems.

The results in this section demonstrate that, across the C3S systems, the strength of the ENSO teleconnection is linked to the climatological behaviour of the North Atlantic jet, as well as the upstream North Pacific jet. Models that have a stronger ENSO teleconnection tend to have stronger jets over the northern part of the North Atlantic basin and more northerly climatological jet positions, associated with an increased frequency of poleward jet excursions. There is also a link to the North Pacific, where the equatorward jet biases common to the C3S systems are linked to weaker teleconnection strength over the North Atlantic. The systems that have a weaker ENSO teleconnection tend to exhibit stronger biases in upstream high-latitude blocking, over the Canada–Greenland region. These differences in dynamical behaviour across the C3S systems motivate the need for further work, including specific targeted experiments, to interrogate the relationships between pervasive errors in jet behaviour and the strength of the ENSO teleconnection to the North Atlantic. In the previous section, we found that the weak ENSO teleconnection appears to be an important source of signal-to-noise errors in early winter hindcasts. The results here suggest that systematic biases in the climatological behaviour of the North Pacific and North Atlantic jets are contributing to the weak teleconnection and associated signal-to-noise errors in the C3S systems.

4 | DISCUSSION

In this study I have focused on early winter Euro-Atlantic predictability in an ensemble of state-of-the-art seasonal forecasting systems. The majority of the systems analysed show skill in the hindcasts of the extratropical large-scale atmospheric circulation in early winter, which mostly projects onto the EA pattern, consistent with the findings of Thornton *et al.* (2023). The predictability is strongly tied to the ENSO teleconnection to the North Atlantic, which is generally skilfully captured but the teleconnection is typically too weak in the C3S systems (Thornton *et al.*, 2023). The hindcasts of the EA index generally exhibit a substantial signal-to-noise error, with the signals being lower than would be expected for the demonstrated level of hindcast skill (i.e., RPC > 1), though there is a variation in this error across systems. I demonstrate that the signal-to-noise error is strongly linked to the strength of the ENSO teleconnection in the systems, with systems that exhibit a weaker teleconnection displaying a larger signal-to-noise problem. The dependency on ENSO teleconnection strength closely follows a simple scaling relationship derived from a toy model. Further analysis reveals that the strength of the ENSO teleconnection in the systems is linked to climatological biases in the behaviour of the North Atlantic jet. Models that better represent the dynamics of the jet over the northern part of the basin, with more frequent poleward jet excursions and less frequent high-latitude blocking events, are typically better at representing the strength of the ENSO teleconnection to the North Atlantic in early winter, with lower associated signal-to-noise errors.

Our analysis has highlighted the weak ENSO teleconnection, as well as associated biases in the behaviour of the North Atlantic jet, as being strongly linked to many of the signal-to-noise errors seen in the early winter hindcasts. It is worthwhile to compare how our findings fit with previous theories on the origins of signal-to-noise errors. One prominent theory, proposed in Scaife *et al.* (2019) and further investigated in Hardiman *et al.* (2022), is that deficiencies in the eddy feedbacks are responsible for the weak predictable signal in models. Hardiman *et al.* (2022)

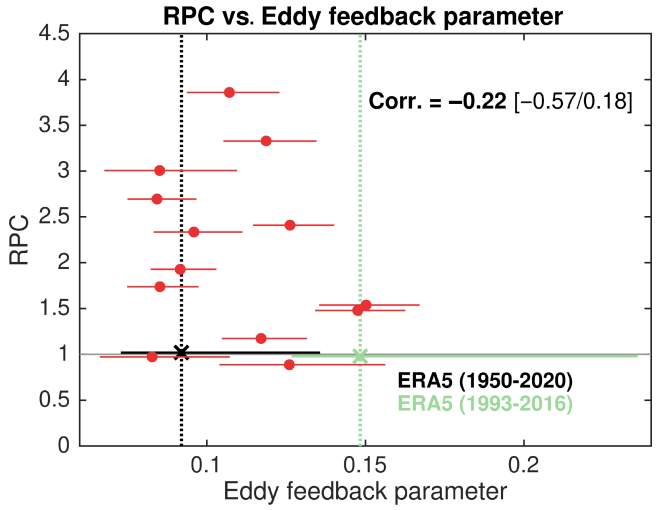

FIGURE 9 Ratio of predictable components (RCP) for the early winter (November–December) East Atlantic index hindcasts from the Copernicus Climate Change Service (C3S) systems plotted against the eddy feedback parameter defined in Hardiman *et al.* (2019) (and following (Smith *et al.*, 2022)). Specifically, the eddy feedback parameter is calculated as the area-weighted average of $\mathrm{corr}(\overline{u}, \nabla \cdot \boldsymbol{F}_{\mathrm{H}})^2$ between 20° N and 72° N (calculated at 500 hPa), where $\boldsymbol{F}_{\mathrm{H}}$ is the horizontal quasi-geostrophic Eliassen–Palm flux, $\overline{u}$ is the zonal mean zonal wind, and the correlation is calculated on seasonally averaged data. Also shown are the values of the eddy feedback parameter from the European Centre for Medium-range Weather Forecasts Reanalysis v5 (ERA5) dataset, calculated over both the C3S period and an extended period. The horizontal lines show the 5–95% confidence intervals for the eddy feedback parameters, calculated using a bootstrap resampling. The systems are the same as those plotted in Figure 5b. The full latitudinal variation of the correlation term, $\mathrm{corr}(\overline{u}, \nabla \cdot \boldsymbol{F}_{\mathrm{H}})$, also shows very good agreement across the systems and ERA5 (Supporting Information Figure S6). [Colour figure can be viewed at wileyonlinelibrary.com]

showed that, for winter (DJF) seasonal hindcasts of the Arctic Oscillation and NAO indices, the RPC is correlated with an "eddy feedback parameter", which is a particular measure of the feedback of the horizontal Eliassen–Palm flux on the zonal mean jet (following Smith *et al.* (2022); see caption of Figure 9 for specific definition). Moreover, the strength of the winter (DJF) ENSO–Arctic Oscillation teleconnection was found to be better in systems that exhibited a more realistic eddy feedback parameter. To examine how well the eddy feedback parameter of Hardiman *et al.* (2022) explains the signal-to-noise errors in the early winter EA index hindcasts, I computed the eddy feedback parameter and compared it with the RPC (shown in Figure 9). Overall, the eddy feedback parameter for the C3S systems analysed in this study is similar to the ERA5 value over the extended 1950–2020 period or slightly lower than over the equivalent C3S period—consistent with the high sampling uncertainty in this parameter highlighted by Saffin *et al.* (2024). Most notably, Figure 9 shows the variation in the eddy feedback parameter is not strongly related to the RPC. The correlation between the eddy feedback parameter and RPC is $r = -0.22$, which is small in magnitude compared with the correlation between the ENSO teleconnection strength and the theoretical scaling of the RPC, $r = 0.82$ (i.e., Figure 5b).

Although we have seen that the specific eddy feedback parameter analysed in Hardiman *et al.* (2022) to be not strongly linked with the signal-to-noise error over the North Atlantic in these early winter hindcasts, this does not mean that deficiencies in eddy feedbacks are not playing a role. Transient eddy feedbacks are crucial in shaping blocking events and shifts in the North Atlantic jet, and these are important in determining the early winter ENSO teleconnection in observations (O'Reilly *et al.*, 2024). In the preceding analysis, it was found that climatological biases in the eddy-driven jet latitude distributions underlie the weak ENSO teleconnections, which does broadly represent a deficiency in eddy feedbacks. However, it may be more useful to consider the signal-to-noise errors to stem from the systematic biases in the representation of the North Atlantic jet—this will not necessarily directly relate to a zonally averaged measure, as shown here for the eddy feedback parameter. The biases in the jet latitude distributions of the C3S systems here also exhibit some consistencies with the regime hypothesis of Strommen and Palmer (2019); here, the frequency of the poleward jet events, or regimes, are systematically underestimated by the systems and display a muted change in frequency in response to predictable ENSO forcing (e.g., Figure 8b).

Beyond this study, the approach applied here could provide a useful framework for exploring the origins of signal-to-noise errors in other seasons, regions, and over different time-scales—broadly similar to the approach in Hardiman *et al.* (2022). Specifically, the general process of identifying important predictable drivers of large-scale circulation anomalies and exploring how biases in the model behaviour are undermining the predictable model signals. Here, the dominance of the ENSO signal in early winter motivates the use of a relatively simple conceptual model of the processes driving the weak model signals, but other instances are likely more challenging. For example, using a similar approach to study the predictable signals of the later winter NAO would be more complex because there are multiple important drivers (e.g. Dunstone *et al.*, 2016; Folland *et al.*, 2012) and hindcast skill levels are typically not as high (e.g. Hardiman *et al.*, 2022). Nonetheless, applying this process and identifying biases in the model behaviour that are linked to the weak predictable signals provides a practical approach towards developing further

understanding of other signal-to-noise errors in coupled climate models.

ACKNOWLEDGEMENTS

COR was supported by a Royal Society University Research Fellowship (URF\R1\201230).

DATA AVAILABILITY STATEMENT

The data used in this article are all open access datasets available on public servers.

ENDNOTE

[1]The full Pearson correlation coefficient between the two normalised indices, $r = \mathrm{corr}(\beta_a N^*_{34} + \epsilon_a, \beta_b N^*_{34} + \epsilon_b)$, simply reduces to the covariance $\mathrm{cov}(\beta_a N^*_{34} + \epsilon_a, \beta_b N^*_{34} + \epsilon_b)$. Expanding this fully gives the correlation coefficient as $r = \mathrm{cov}(\beta_a N^*_{34}, \beta_b N^*_{34}) + \mathrm{cov}(\beta_a N^*_{34}, \epsilon_b) + \mathrm{cov}(\epsilon_a, \beta_b N^*_{34}) + \mathrm{cov}(\epsilon_a, \epsilon_b)$. The expected value of the second, third, and fourth terms on the right-hand side of this expression are all zero since they include covariance with a random residual with zero mean, ϵ. Therefore, the expectation of the correlation coefficient reduces to $r \approx \mathrm{cov}(\beta_a N^*_{34}, \beta_b N^*_{34}) = \beta_a \beta_b$.

ORCID

Christopher H. O'Reilly https://orcid.org/0000-0002-8630-1650

REFERENCES

Ayarzagüena, B., Ineson, S., Dunstone, N.J., Baldwin, M.P. & Scaife, A.A. (2018) Intraseasonal effects of El Niño–southern oscillation on north atlantic climate. *Journal of Climate*, 31(21), 8861–8873. Available from: https://doi.org/10.1175/JCLI-D-18-0097.1

Baker, L.H., Shaffrey, L.C., Sutton, R.T., Weisheimer, A. & Scaife, A.A. (2018) An Intercomparison of skill and overconfidence/Underconfidence of the wintertime North Atlantic oscillation in multimodel seasonal forecasts. *Geophysical Research Letters*, 45(15), 7808–7817. Available from: https://doi.org/10.1029/2018GL078838

Bracegirdle, T.J., Lu, H., Eade, R. & Woollings, T. (2018) Do CMIP5 models reproduce observed low-frequency North Atlantic jet variability? *Geophysical Research Letters*, 45(14), 7204–7212. Available from: https://doi.org/10.1029/2018GL078965

Charlton-Perez, A.J., Bröcker, J., Stockdale, T.N. & Johnson, S. (2019) When and where do ECMWF seasonal forecast systems exhibit anomalously low signal-to-noise ratio? *Quarterly Journal of the Royal Meteorological Society*, 145(725), 3466–3478. Available from: https://doi.org/10.1002/qj.3631

Clark, R.T., Bett, P.E., Thornton, H.E. & Scaife, A.A. (2017) Skilful seasonal predictions for the European energy industry. *Environmental Research Letters*, 12(2), 024002. Available from: https://doi.org/10.1088/1748-9326/aa57ab

Dunstone, N. (2018) Skilful seasonal predictions of summer European rainfall. *Geophysical Research Letters*, 45(7), 3246–3254. Available from: https://doi.org/10.1002/2017GL076337

Dunstone, N. (2023) Skilful predictions of the summer North Atlantic oscillation. *Communications Earth & Environment*, 4(1), 1–11. Available from: https://doi.org/10.1038/s43247-023-01063-2

Dunstone, N., Smith, D., Scaife, A., Hermanson, L., Eade, R., Robinson, N. et al. (2016) Skilful predictions of the winter North Atlantic oscillation one year ahead. *Nature Geoscience*, 9(11), 809–814. Available from: https://doi.org/10.1038/ngeo2824

Eade, R., Smith, D., Scaife, A., Wallace, E., Dunstone, N., Hermanson, L. et al. (2014) Do seasonal-to-decadal climate predictions underestimate the predictability of the real world? *Geophysical Research Letters*, 41(15), 5620–5628. Available from: https://doi.org/10.1002/2014GL061146

Folland, C.K., Scaife, A.A., Lindesay, J. & Stephenson, D.B. (2012) How potentially predictable is northern European winter climate a season ahead? *International Journal of Climatology*, 32(6), 801–818. Available from: https://doi.org/10.1002/joc.2314

Garfinkel, C.I., Chen, W., Li, Y., Schwartz, C., Yadav, P. & Domeisen, D. (2022) The winter North Pacific teleconnection in response to ENSO and the MJO in operational subseasonal forecasting models is too weak. *Journal of Climate*, 35(24), 8013–8030. Available from: https://doi.org/10.1175/JCLI-D-22-0179.1

Hardiman, S.C., Dunstone, N.J., Scaife, A.A., Smith, D.M., Comer, R., Nie, Y. et al. (2022) Missing eddy feedback may explain weak signal-to-noise ratios in climate predictions. *npj Climate and Atmospheric Science*, 5(1), 1–8. Available from: https://doi.org/10.1038/s41612-022-00280-4

Hardiman, S.C., Dunstone, N.J., Scaife, A.A., Smith, D.M., Ineson, S., Lim, J. et al. (2019) The impact of strong El Niño and La Niña events on the North Atlantic. *Geophysical Research Letters*, 46(5), 2874–2883. Available from: https://doi.org/10.1029/2018GL081776

Hersbach, H., Bell, B., Berrisford, P., Hirahara, S., Horányi, A., Muñoz-Sabater, J. et al. (2020) The ERA5 global reanalysis. *Quarterly Journal of the Royal Meteorological Society*, 146(730), 1999–2049. Available from: https://doi.org/10.1002/qj.3803

Hurrell, J.W., Kushnir, Y., Ottersen, G. & Visbeck, M. (2003) An overview of the North Atlantic oscillation. The North Atlantic oscillation: climatic significance and environmental impact. *American Geophysical Union*, 134, 1–35. Available from: https://doi.org/10.1029/134GM01

Johansson, A. (2007) Prediction skill of the NAO and PNA from daily to seasonal time scales. *Journal of Climate*, 20(10), 1957–1975. Available from: https://doi.org/10.1175/JCLI4072.1

King, M.P., Herceg-Bulić, I., Bladé, I., García-Serrano, J., Keenlyside, N., Kucharski, F. et al. (2018) Importance of late fall ENSO teleconnection in the euro-Atlantic sector. *Bulletin of the American Meteorological Society*, 99(7), 1337–1343. Available from: https://doi.org/10.1175/BAMS-D-17-0020.1

Knight, J.R., Scaife, A.A. & Maidens, A. (2022) An extratropical contribution to the signal-to-noise paradox in seasonal climate prediction. *Geophysical Research Letters*, 49(23), e2022GL100471. Available from: https://doi.org/10.1029/2022GL100471

Kretschmer, M., Adams, S.V., Arribas, A., Prudden, R., Robinson, N., Saggioro, E. et al. (2021) Quantifying causal pathways of teleconnections. *Bulletin of the American Meteorological Society*, 102(12),

E2247–E2263. Available from: https://doi.org/10.1175/BAMS-D-20-0117.1

Marcheggiani, A., Robson, J., Monerie, P.-A., Bracegirdle, T.J. & Smith, D. (2023) Decadal predictability of the North Atlantic Eddy-driven jet in winter. *Geophysical Research Letters*, 50(8), e2022GL102071. Available from: https://doi.org/10.1029/2022GL102071

Masato, G., Hoskins, B.J. & Woollings, T. (2013) Wave-breaking characteristics of northern hemisphere winter blocking: a two-dimensional approach. *Journal of Climate*, 26(13), 4535–4549. Available from: https://doi.org/10.1175/JCLI-D-12-00240.1

Molteni, F. & Brookshaw, A. (2023) Early- and late-winter ENSO teleconnections to the euro-Atlantic region in state-of-the-art seasonal forecasting systems. *Climate Dynamics*, 61, 2673–2692. Available from: https://doi.org/10.1007/s00382-023-06698-7

O'Reilly, C.H., Befort, D.J., Weisheimer, A., Woollings, T., Ballinger, A. & Hegerl, G. (2021) Projections of northern hemisphere extratropical climate underestimate internal variability and associated uncertainty. *Communications Earth & Environment*, 2(1), 1–9. Available from: https://doi.org/10.1038/s43247-021-00268-7

O'Reilly, C.H., Drouard, M., Ayarzagüena, B., Ambaum, M.H.P. & Methven, J. (2024) The role of storm-track dynamics in the intraseasonal variability of the winter ENSO teleconnection to the North Atlantic. *Quarterly Journal of the Royal Meteorological Society*, 150(761), 2069–2086. Available from: https://doi.org/10.1002/qj.4691

O'Reilly, C.H., Weisheimer, A., Woollings, T., Gray, L.J. & MacLeod, D. (2019) The importance of stratospheric initial conditions for winter North Atlantic oscillation predictability and implications for the signal-to-noise paradox. *Quarterly Journal of the Royal Meteorological Society*, 145(718), 131–146. Available from: https://doi.org/10.1002/qj.3413

O'Reilly, C.H., Zanna, L. & Woollings, T. (2019) Assessing external and internal sources of Atlantic multidecadal variability using models, proxy data, and early instrumental indices. *Journal of Climate*, 32(22), 7727–7745. Available from: https://doi.org/10.1175/JCLI-D-19-0177.1

Ossó, A., Sutton, R., Shaffrey, L. & Dong, B. (2020) Development, amplification, and decay of Atlantic/European summer weather patterns linked to spring North Atlantic Sea surface temperatures. *Journal of Climate*, 33(14), 5939–5951. Available from: https://doi.org/10.1175/JCLI-D-19-0613.1

Rayner, N.A., Parker, D.E., Horton, E.B., Folland, C.K., Alexander, L.V., Rowell, D.P. et al. (2003) Global analyses of sea surface temperature, sea ice, and night marine air temperature since the late nineteenth century. *Journal of Geophysical Research: Atmospheres*, 108(D14), 4407. Available from: https://doi.org/10.1029/2002JD002670

Saffin, L., McKenna, C.M., Bonnet, R. & Maycock, A.C. (2024) Large uncertainties when diagnosing the Eddy feedback parameter and its role in the signal-to-noise paradox. *Geophysical Research Letters*, 51(11), e2024GL108861. Available from: https://doi.org/10.1029/2024GL108861

Scaife, A.A., Arribas, A., Blockley, E., Brookshaw, A., Clark, R.T., Dunstone, N. et al. (2014) Skillful long-range prediction of European and north American winters. *Geophysical Research Letters*, 41(7), 2514–2519. Available from: https://doi.org/10.1002/2014GL059637

Scaife, A.A., Camp, J., Comer, R., Davis, P., Dunstone, N., Gordon, M. et al. (2019) Does increased atmospheric resolution improve seasonal climate predictions? *Atmospheric Science Letters*, 20(8), e922. Available from: https://doi.org/10.1002/asl.922

Scaife, A.A. & Smith, D. (2018) A signal-to-noise paradox in climate science. *npj Climate and Atmospheric Science*, 1(1), 1–8. Available from: https://doi.org/10.1038/s41612-018-0038-4

Silverman, B.W. (1981) Using kernel density estimates to investigate multimodality. *Journal of the Royal Statistical Society: Series B*, 43(1), 97–99. Available from: https://doi.org/10.1111/j.2517-6161.1981.tb01155.x

Simpson, I.R., Deser, C., McKinnon, K.A. & Barnes, E.A. (2018) Modeled and observed multidecadal variability in the North Atlantic jet stream and its connection to sea surface temperatures. *Journal of Climate*, 31(20), 8313–8338. Available from: https://doi.org/10.1175/JCLI-D-18-0168.1

Smith, D.M., Eade, R., Andrews, M.B., Ayres, H., Clark, A., Chripko, S. et al. (2022) Robust but weak winter atmospheric circulation response to future Arctic sea ice loss. *Nature Communications*, 13(1), 727. Available from: https://doi.org/10.1038/s41467-022-28283-y

Smith, D.M., Eade, R., Scaife, A.A., Caron, L.P., Danabasoglu, G., DelSole, T.M. et al. (2019) Robust skill of decadal climate predictions. *npj Climate and Atmospheric Science*, 2(1), 1–10. Available from: https://doi.org/10.1038/s41612-019-0071-y

Smith, D.M., Scaife, A.A., Eade, R., Athanasiadis, P., Bellucci, A., Bethke, I. et al. (2020) North Atlantic climate far more predictable than models imply. *Nature*, 583(7818), 796–800. Available from: https://doi.org/10.1038/s41586-020-2525-0

Smith, D.M., Scaife, A.A. & Kirtman, B.P. (2012) What is the current state of scientific knowledge with regard to seasonal and decadal forecasting? *Environmental Research Letters*, 7(1), 015602. Available from: https://doi.org/10.1088/1748-9326/7/1/015602

Stringer, N., Knight, J. & Thornton, H. (2020) Improving meteorological seasonal forecasts for hydrological modeling in European winter. *Journal of Applied Meteorology and Climatology*, 59(2), 317–332. Available from: https://doi.org/10.1175/JAMC-D-19-0094.1

Strommen, K. (2020) Jet latitude regimes and the predictability of the North Atlantic oscillation. *Quarterly Journal of the Royal Meteorological Society*, 146(730), 2368–2391. Available from: https://doi.org/10.1002/qj.3796

Strommen, K. & Palmer, T.N. (2019) Signal and noise in regime systems: a hypothesis on the predictability of the North Atlantic oscillation. *Quarterly Journal of the Royal Meteorological Society*, 145(718), 147–163. Available from: https://doi.org/10.1002/qj.3414

Thornton, H.E., Scaife, A.A., Hoskins, B.J., Brayshaw, D.J., Smith, D.M., Dunstone, N. et al. (2019) Skilful seasonal prediction of winter gas demand. *Environmental Research Letters*, 14(2), 024009. Available from: https://doi.org/10.1088/1748-9326/aaf338

Thornton, H.E., Smith, D.M., Scaife, A.A. & Dunstone, N.J. (2023) Seasonal predictability of the East Atlantic pattern in late autumn and early winter. *Geophysical Research Letters*, 50(1), e2022GL100712. Available from: https://doi.org/10.1029/2022GL100712

Woollings, T., Hannachi, A. & Hoskins, B. (2010) Variability of the North Atlantic eddy-driven jet stream. *Quarterly Journal of the Royal Meteorological Society*, 136(649), 856–868. Available from: https://doi.org/10.1002/qj.625

Woollings, T., Hoskins, B., Blackburn, M. & Berrisford, P. (2008) A new Rossby wave–breaking interpretation of the North Atlantic oscillation. *Journal of the Atmospheric Sciences*, 65(2), 609–626. Available from: https://doi.org/10.1175/2007JAS2347.1

Zhang, W., Kirtman, B., Siqueira, L., Clement, A. & Xia, J. (2021) Understanding the signal-to-noise paradox in decadal climate predictability from CMIP5 and an eddying global coupled model. *Climate Dynamics*, 56(9), 2895–2913. Available from: https://doi.org/10.1007/s00382-020-05621-8

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: O'Reilly, C.H. (2025) Signal-to-noise errors in early winter Euro-Atlantic predictions linked to weak ENSO teleconnections and pervasive jet biases. *Quarterly Journal of the Royal Meteorological Society*, e4952. Available from: https://doi.org/10.1002/qj.4952