



Deficient ocean–atmosphere feedbacks constrain seasonal NAO prediction

Erik W. Kolstad¹

¹NORCE Research, Bjerknes Centre for Climate Research, Bergen, Norway

Correspondence: Erik W. Kolstad (ekol@norceresearch.no)

Abstract.

As the North Atlantic Oscillation (NAO) accounts for a dominant share of wintertime weather variability across the North Atlantic basin, it is a coveted target for seasonal prediction. Yet dynamical forecast systems continue to exhibit limited skill, in part due to deficiencies in representing ocean–atmosphere feedbacks. Here, mediation analysis – a statistical framework from causal inference – is applied to identify and quantify feedback pathways linking late-autumn North Atlantic sea surface temperature (SST) anomalies to the subsequent winter NAO. This approach is attractive because it is straightforward to apply, easy to interpret, and can be used directly on observations-derived data like reanalyses without requiring idealised model perturbation experiments.

The analysis reveals a physically coherent feedback sequence. Anomalous November SST patterns promote the gradual formation of a surface-pressure dipole rotated clockwise relative to the canonical NAO structure. This dipole induces advection anomalies in the western North Atlantic, which in turn modulate surface fluxes in the Subpolar Gyre and lower-tropospheric baroclinicity in the storm-track entry region east of Newfoundland. These changes nudge the NAO, which, once established, feeds back onto the fluxes and baroclinicity, reinforcing the anomaly and sustaining the circulation pattern.

A central finding is that a state-of-the-art seasonal prediction system fails to capture these feedback mechanisms. The baroclinicity pathway, the process through which changes in eddy growth reinforce the circulation anomaly, is particularly deficient, accounting for only 2% of the lagged SST–NAO correlation in SEAS5 compared with 44% in the ERA5 reanalysis. This misrepresentation likely represents a fundamental barrier to improved NAO forecast skill.

More broadly, the results demonstrate the potential of mediation analysis as a diagnostic tool for disentangling coupled feedbacks directly from observations, evaluating their representation in models, and guiding targeted improvements that could enhance seasonal prediction of the NAO.

1 Introduction

It is no wonder that the North Atlantic Oscillation (NAO) has received considerable attention in studies of climate dynamics, given that it accounts for roughly half of the interannual wintertime tropospheric pressure variance over the North Atlantic (Ambaum et al., 2001). Hence, it serves as a good proxy for fluctuations in the strength and latitudinal position of the jet stream (Woollings and Blackburn, 2012) and storm tracks (Rivière and Orlanski, 2007), and by extension for variations in



weather and associated impacts over and around the North Atlantic basin (e.g., Athanasiadis et al., 2017; Degenhardt et al., 2023). I am a witness to the societal relevance of the NAO through my work with firms from multiple sectors – particularly insurance and renewable energy production and trading – to predict it on seasonal, decadal, and multi-decadal timescales.

Owing to its wide-ranging influence, the NAO is routinely used as a benchmark for seasonal prediction skill. Statistical (empirical) methods have a long history and have achieved useful results using predictors such as autumn Arctic sea ice, Eurasian snow cover, tropical and local SSTs, and stratospheric variables (e.g., Hall et al., 2017; Wang et al., 2017). More recently, empirical models have been enriched with machine learning (ML) and hybrid techniques, yielding higher forecast skill including for the NAO (e.g., Mu et al., 2023; Sun et al., 2024). While empirical models offer interpretability and often outperform dynamical systems in specific periods, their reliance on historical relationships makes them vulnerable to non-stationarity and shifts in climate regimes (Hertig et al., 2015; Kolstad and Screen, 2019). Demonstrated skill in one period therefore provides no guarantee of consistent predictability in others (Weisheimer et al., 2017; Baker et al., 2024).

In principle, dynamical coupled prediction systems ought not to be constrained by non-stationarity, since they aim to reproduce the behaviour of the climate system from first principles. Ideally, such systems should be able to integrate any initial condition to a realistic future state. In practice, however, as Suckling and Smith (2013) pointed out, even physics-based models are not independent of the data used in their design. Their apparent ability, evaluated primarily through retroactive forecasts, or *reforecasts* (also known as hindcasts), to reproduce historical variability may therefore overstate true predictive skill. Moreover, in a changing climate, even out-of-sample performance offers no guarantee of future success, given the nonlinear nature of the system's response to external forcing (Stott et al., 2013). This highlights the need to assess not only the overall skill of prediction systems, but also the physical consistency of the processes and feedbacks they represent.

About a decade ago there was a surge of enthusiasm over the high NAO skill reported in some dynamical systems (Scaife et al., 2014). However, subsequent system upgrades have not improved (or in some cases have even degraded) this level of skill (Baker et al., 2024). Several studies have convincingly demonstrated that the performance of dynamical prediction systems depends on how well they represent crucial physical processes. For instance, Haarsma et al. (2019) found that increasing the oceanic resolution of a coupled system strengthened air–sea interactions and enhanced seasonal predictability in the North Atlantic storm-track entry region east of Newfoundland. Similarly, Hardiman et al. (2022) showed that most forecast systems consistently underestimate positive feedback mechanisms between transient eddies and the large-scale flow, leading to weaker eddy forcing of the mean circulation. A related problem concerns the representation of mesoscale oceanic eddies and associated SST gradients in “eddy-rich regions, including the Gulf Stream” (Zhang et al., 2021). Limited horizontal resolution tends to smooth these gradients and weaken the coupling between SST, surface heat fluxes, and low-level atmospheric baroclinicity (Hewitt et al., 2017; Athanasiadis et al., 2022). This, in turn, contributes to the muted air–sea feedbacks (Bellucci et al., 2021; Hardiman et al., 2022).

These deficiencies have been linked to low signal-to-noise ratios (Scaife and Smith, 2018; Weisheimer et al., 2024) in ensemble predictions. Beyond the seasonal timescale, recent work has also demonstrated the importance of ocean–atmosphere feedbacks for decadal NAO predictability. Patrizio et al. (2025) showed that skill in decadal NAO forecasts depends on how models represent feedbacks between subpolar SST anomalies and the NAO.



Together, these findings highlight that two-way ocean–atmosphere coupling is fundamental to NAO variability across timescales, yet it remains misrepresented in current prediction systems. Indeed, for the same family of forecast systems considered here, Roberts et al. (2021) demonstrated that biases in the location and structure of the Gulf Stream substantially degrade subseasonal forecast skill, and that correcting these SST errors online improves the mean state and circulation anomalies across the North Atlantic and downstream into Europe.

A large body of work has shown that characteristic SST patterns can precondition the atmosphere on subseasonal to seasonal timescales (Rodwell et al., 1999; Watanabe and Kimoto, 2000; Czaja and Frankignoul, 2002; Wang et al., 2004; Hall et al., 2017; Baker et al., 2019; Sun et al., 2024). In particular, the North Atlantic SST tripole (or the similar “horseshoe” pattern), with alternating anomalies between the Subpolar Gyre, the mid-latitude Gulf Stream, and the subtropics, has been linked to a feedback loop involving the NAO itself (Peng et al., 2002; Pan, 2005; Mosedale et al., 2006; Cassou et al., 2007; Gastineau and Frankignoul, 2015). Joyce et al. (2019) demonstrated that meridional shifts of the Gulf Stream front and its associated SST gradients tend to lead changes in storm tracks and Greenland blocking by one to three months, suggesting an evolving pathway from autumn ocean conditions to wintertime atmospheric variability.

Building on this perspective, Kolstad and O’Reilly (2024) found that the correlation between November SSTs and the subsequent NAO increases gradually through the winter season, peaking in January and February. They showed that surface heat fluxes in the Subpolar Gyre and baroclinicity in the storm-track entry region in the western North Atlantic act as key mediators in this feedback. The latter result is consistent with the well-established role of diabatic heating and eddy feedbacks in maintaining storm-track baroclinicity (Hardiman et al., 2022).

The statistical framework used by Kolstad and O’Reilly (2024) is known as *mediation analysis* (Nguyen et al., 2021). It is particularly well-suited to climate science applications where feedback loops and indirect pathways are common but difficult to isolate using traditional correlation-based methods. Yet, although it is widely used in the social and medical sciences, mediation analysis has only rarely been applied in climate research. It belongs to the broader family of *causal inference* methods (Pearl et al., 2016), so named because they are designed to identify and quantify causal relationships. Several such approaches have been successfully used in climate science, ranging from easily interpretable approaches like Granger causality (e.g., Granger, 1969; Mosedale et al., 2006; McGraw and Barnes, 2018) to more complex methods (e.g., Ebert-Uphoff and Deng, 2012; Hannart et al., 2016; Runge et al., 2019; Docquier et al., 2024).

A key advantage of causal inference approaches is that they allow pathways to be investigated without manipulating model boundary conditions. Perturbation-based methods, though widely used, can produce unintended consequences. For example, perturbing greenhouse gas concentrations triggers numerous feedbacks on diverse timescales, complicating attribution of the climate system’s response and adjustments (Knutti and Rugenstein, 2015). Even more localised interventions can have undesirable side effects: Lewis et al. (2024) showed that modifying albedo or applying surface heating to force sea-ice loss can generate spurious warming and exaggerate the atmospheric circulation response. As Palmer and Weisheimer (2011) noted, multiple model errors can compensate for one another, making it difficult to diagnose the underlying causes of biases. These considerations further motivate the use of mediation analysis, which relies solely on observed covariances and avoids imposing artificial perturbations.



This study extends the work of Kolstad and O'Reilly (2024) in three ways. First, it aims to introduce mediation analysis to a broader climate science audience and illustrates its value as a diagnostic tool for ocean–atmosphere feedbacks. Second, it attempts to quantify and clarify causal directionality in the relationships linking SSTs, surface fluxes, baroclinicity, and the NAO. Third, it applies the mediation framework to a state-of-the-art forecast system to assess whether systematic biases in these feedbacks can help explain its limited skill in predicting the NAO.

The following section gives an overview of mediation and partial-correlation analysis before describing the data and other methods in Section 3. Section 4 presents the results, and Section 5 discusses their implications for understanding and improving NAO predictability.

2 Mediation analysis

The mediation analysis framework is designed for studying observed causal pathways. Adopting the naming convention of MacKinnon et al. (2000), such a pathway links a predictor variable X to an outcome variable Y , i.e.:

$$X \rightarrow Y.$$

In the analysis to follow, X is an index representing SST anomalies in November and Y is the winter NAO index. Physically, it is obvious that any correlation between these two variables must be mediated by other processes, referred to as *mediators* and denoted Z . Here Z is a gridded spatial field representing surface heat fluxes and the Eady growth rate maximum (Hoskins and Valdes, 1990), a measure of baroclinicity. These mediators are investigated separately through the pathway

$$X \rightarrow Z \rightarrow Y.$$

It is customary to quantify the mediating role of Z and categorizing it as either: a *perfect* mediator if it fully accounts for $X \rightarrow Y$ (Baron and Kenny, 1986); a *partial* mediator if it partly explains $X \rightarrow Y$ (Baron and Kenny, 1986); or a *suppressor* if the correlation between X and Y is strengthened when Z is accounted for (Conger, 1974).

It would be absurd to claim that a pixel value of any one variable uniquely mediates the lagged effect of SSTs on the NAO. In reality, a practically infinite web of interacting processes combine to realise that relationship. Nevertheless, the approach used here is useful for providing a spatial fingerprint of where a single variable exerts the strongest mediating influence. Equally important, the method can be used to identify where a forecast model incorrectly mediates or even suppresses the SST–NAO correlation.

2.1 Regression equations

To test for mediation or suppression, three regression equations are defined (ignoring intercepts and residuals for simplicity). The first describes the *direct effect* τ of the predictor X on the predictand Y :

$$Y = \tau X. \tag{1}$$



125 The effect of X on the mediator Z is labelled here as α in the second equation:

$$Z = \alpha X. \quad (2)$$

The third regression describes $X \rightarrow Z \rightarrow Y$ by accounting for the mediator. The effect of X on Y changes to τ_Z , and the effect of Z on Y when accounting for X is denoted as β :

$$Y = \tau_Z X + \beta Z. \quad (3)$$

130 A central concept is the product $\alpha\beta$, known as the *indirect effect* (of X on Y through Z). It follows from Eqs. (1–3) that $\alpha\beta = \tau - \tau_Z$.

Scaling by the total effect yields:

$$\frac{\alpha\beta}{\tau} = 1 - \frac{\tau_Z}{\tau}. \quad (4)$$

135 According to the standard criteria for mediation laid out by Baron and Kenny (1986), τ , α , and β must all be significantly different from zero. τ and τ_Z must also have the same sign; otherwise the mediation is referred to as *inconsistent*.

If $\tau_Z = 0$ (or in practice is not significantly non-zero), it is plain from Eq. 4 that the direct and indirect effects are identical. This means that the pathway $X \rightarrow Y$ is fully accounted for by Z , which again implies that $X \rightarrow Z \rightarrow Y$ is one of potentially many correct causal pathways. Partial mediation occurs when $0 < \tau_Z/\tau < 1$.

140 Both perfect and partial mediation can represent unambiguous forward-directed pathways in the sense that X changes Z and then Z affects Y . In climate dynamics though, feedback mechanisms are common. The pathway $X \rightarrow Z \rightarrow Y$ can be valid even when the “reverse” pathway $X \rightarrow Y \rightarrow Z$ is equally valid. In situations where Y and Z are evaluated at the same time, as is the case in this study, it can be difficult to establish whether one causal direction is more valid than the other. What one can do, however, is to explore how X affects Z independently of Y through a partial correlation analysis.

2.2 Partial correlation

145 The first step to compute the correlation between X and Z while accounting for Y is to regress out Y from both X and Z :

$$X = \gamma_X Y + \varepsilon_X,$$

$$Z = \gamma_Z Y + \varepsilon_Z,$$

where $\varepsilon_X = X - \gamma_X Y$ and $\varepsilon_Z = Z - \gamma_Z Y$ are the residuals. The partial correlation is then given by the correlation between these residuals:

$$150 \quad r_{X,Z|Y} = \text{corr}(\varepsilon_X, \varepsilon_Z). \quad (5)$$

In the context of this study, $r_{X,Z|Y}$ can be viewed as a measure of preconditioning: it indicates whether November SST anomalies (X) are systematically associated with wintertime flux or baroclinicity anomalies (Z) that are not simply a by-product of the contemporaneous NAO (Y). This helps to disentangle feedback from forcing, and thereby adds diagnostic value when interpreting the spatial structure of air–sea interaction. The partial correlation between SSTs and SLP itself with the NAO
155 regressed out is also investigated.



2.3 Suppression

An interesting special case occurs when $\tau_Z/\tau > 1$, which means that the indirect effect $\alpha\beta$ has the opposite sign to the total effect τ (Eq. 4). In these cases, Z is referred to as a suppressor because the regression coefficient linking X and Y is inflated when Z is accounted for (Muniz and MacKinnon, 2025). In the context of this study, this could mean that X (November SST anomalies) drives changes in Z (e.g., flux anomalies), but the response of Y (the NAO) to those fluxes is of opposite sign to the direct $X \rightarrow Y$ pathway. This can occur because Y itself feeds back onto Z , helping to make β negative. In other words, Z acts as a negative feedback, transmitting a damping influence on Y that partly cancels (suppresses) the predictive signal from X . In the raw correlation, this feedback reduces the apparent strength of X as a predictor of Y , but once Z is controlled for, the hidden strength of the $X \rightarrow Y$ link is revealed. Put differently, had it not been for the negative feedback through Z , X would have exerted stronger predictive power on Y .

3 Data and methods

3.1 Data

This study makes use of reanalysis and seasonal forecast data. The reanalysis reference is ERA5 (Hersbach et al., 2020), produced by the European Centre for Medium-range Weather Prediction (ECMWF), and the forecast system is SEAS5, the fifth generation of the ECMWF's seasonal prediction system (Johnson et al., 2019). The atmospheric component of SEAS5 is the Integrated Forecast System (IFS) atmosphere model. The grid spacing for the ocean model in SEAS5 is 0.25 degrees, which has been shown to yield a decent representation of air–sea interaction along the Gulf Stream front compared to lower-resolution models (Jin and Yu, 2013; Athanasiadis et al., 2022; Patrizio et al., 2023). It seems the resolution will not change in the new SEAS6 system due to be released soon, although the new ocean model yields multiple improvements, including large reductions in SST errors along the Gulf Stream (Keeley et al., 2024, their Figure 2a).

The analysis covers the winters from 1981/82 to 2023/24 (hereafter referred to as 1981–2023). SEAS5 reforecasts were used from 1981 to 2016 with 25 ensemble members, and real-time forecasts from 2017 to 2023 with the first 25 of 51 members used to ensure consistency with the reforecasts.

SEAS5 forecasts are produced once per month; here, only the October initialisations are used, corresponding by convention to lead times of 2–5 months for November through February. November initialisations were excluded because the SST fields at the first lead time were assumed to be too similar to the observed state, which could inflate apparent forecast skill and obscure the role of the model's internal feedbacks. September initialisations, on the other hand, would involve longer lead times and greater forecast drift (Hermanson et al., 2018). The use of October initialisations thus strikes a balance between temporal distance from observed boundary conditions and avoidance of excessive drift, allowing the analysis to focus on the model's intrinsic air–sea coupling behaviour.

In general, ensemble means were used, but when appropriate, derived values like the climate indices described below or the baroclinicity parameter were calculated separately first for each ensemble member.



The variables considered are SST, mean sea level pressure (SLP), latent and sensible surface heat fluxes, and large-scale atmospheric fields used to derive a measure of baroclinicity. Latent and sensible heat fluxes were combined into a single field (positive upward), hereafter referred to as just “fluxes”. Baroclinicity is quantified by the Eady growth rate maximum (σ_E henceforth) parameter (e.g., Hoskins and Valdes, 1990), defined for the 700–850 hPa layer as

$$\sigma_E = 86400 \times 0.3098 f \left| \frac{\partial \mathbf{v}}{\partial z} \right| / N,$$

where the unit is day^{-1} , f is the Coriolis parameter, \mathbf{v} is the wind vector, z is the height, and N is the Brunt–Väisälä frequency, given by

$$N = \sqrt{\frac{g}{\theta} \frac{\partial \theta}{\partial z}},$$

with θ the potential temperature and g the gravitational acceleration.

3.2 Climate indices

Two scalar indices are central to the analysis: the winter (December–February, DJF) NAO index, and an SST-based index representing the November SST anomaly pattern in the North Atlantic most strongly correlated with the following winter’s NAO index.

To construct the NAO index, the first Empirical Orthogonal Function (EOF) of interannual DJF mean ERA5 SLP anomalies was computed over the domain $20^\circ\text{--}80^\circ\text{N}$, $90^\circ\text{W--}40^\circ\text{E}$, using the *eofs* Python package (Dawson, 2016) and applying appropriate latitude-based weighting. For both ERA5 and SEAS5, the corresponding NAO index time series were obtained by projecting their respective gridded SLP anomalies onto the ERA5-based spatial EOF pattern. This ensured that both indices represent the same spatial imprint rather than dataset-specific structures.

The SST index was calculated in a similar way. November SST anomalies from ERA5 were first regressed onto the interannual ERA5 NAO index to obtain a spatial regression pattern. SST anomalies were then projected onto this pattern within a reference domain extending from the Equator to 75°N and from 100°W to 20°E , after which the resulting series was standardised to form the SST index. The selection of this region was done by testing different boundaries and selecting the region that yielded the highest correlation between the SST index and the NAO in the study period. As with the NAO index, the SST index for SEAS5 was computed by projection onto the ERA5-based pattern, not a model-specific optimal pattern, to maintain consistency across the datasets.

Note that no masking for sea ice was applied. In both ERA5 and SEAS5, grid cells covered by sea ice are not missing values but contain subzero SSTs, which remain valid anomalies in this framework. Masking would risk introducing artificial discontinuities in space and time, since the ice edge varies between months and years.

3.3 Statistical significance

Bootstrapping was used to estimate statistical significance, generally by creating 1000 randomised series by drawing N random samples (with replacement) from a series with N data points. When assessing the significance of a metric (e.g., a correlation)

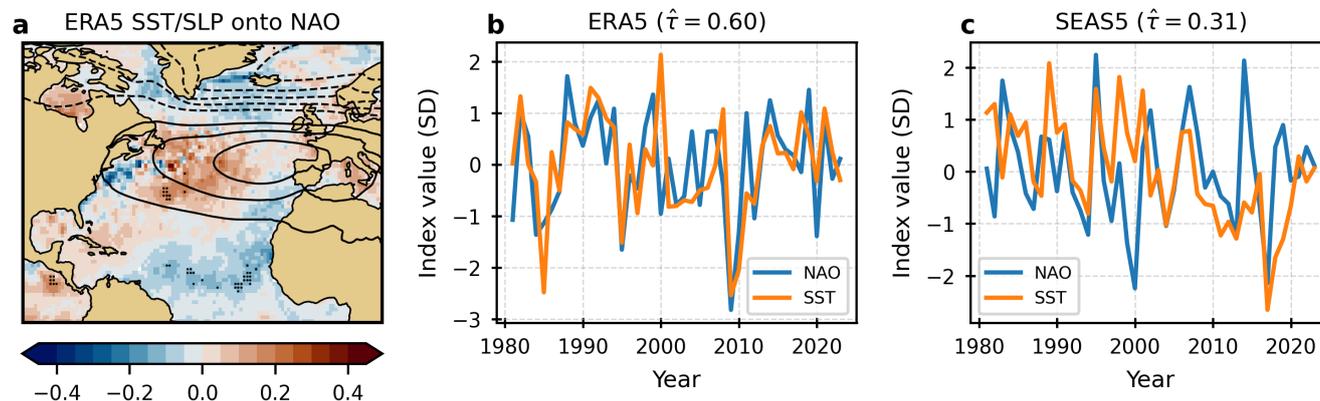


Figure 1. (a) Colours: November SST anomalies in ERA5 projected onto the ERA5 DJF NAO index. Dots mark anomalies significantly different from zero at the 5% level. Contours: DJF SLP anomalies projected onto the same NAO index. The contour interval is 1 hPa; solid (dashed) contours indicate positive (negative) anomalies, and the zero contour is omitted. The map extent corresponds to the region used to define the November SST index. (b) Time series of the November SST index (orange) and the DJF NAO index (blue) in ERA5. Years on the x-axis correspond to the start of the winter season (i.e., DJF 1981/82 is labelled 1981). (c) As in (b), but for SEAS5.

at a significance level of 5% (used throughout this study), the 2.5th and 97.5th percentiles of the correlation coefficient across those 1000 randomised series were computed, and if the interval between these percentiles did not include zero, the correlation was deemed significant.

4 Results

4.1 Relationship between SSTs and the NAO

The ERA5 SST anomaly regression pattern (i.e., November SST anomalies regressed onto the DJF NAO index) is shown with shading in Figure 1a. As expected, it is similar to the pattern in Figure 1f in Kolstad and O'Reilly (2024) – which was also computed based on ERA5 but for a longer period (1940–2022). The contours in Figure 1a displays the regression of DJF SLP anomalies onto the NAO index.

Figure 1b shows the interannual November SST index, obtained by projecting the SST anomalies onto the regression pattern in Figure 1a, together with the winter NAO index. Although only a few of the local SST anomalies in Figure 1a are significant, the correlation between the two indices is relatively high ($\hat{\tau} = 0.60$), underscoring the strong link between late-autumn SSTs and the subsequent winter NAO. (In the following, etimated sample coefficients – values obtained by least-squares fitting – are denoted by carets ; e.g., $\hat{\tau}$ is the sample τ value.) The SST index captures well the two exceptionally negative NAO winters of 2009/10 and 2010/11, as well as the extended positive NAO phase around 1990, though there are also seasons with weak correspondence, such as 2000/01.



235 As noted in Sec. 3b, the reference region used to compute the SST index was selected to maximise its correlation with the NAO index. This required extending the domain to include parts of both the tropical North Atlantic and the eastern tropical North Pacific. When the more limited region used by Kolstad and O'Reilly (2024) and Czaja and Frankignoul (2002) (20°–80°N, 110°W–20°E) was used, the correlation decreased to 0.49, but qualitatively the analysis to follow gave similar results.

Turning to the SEAS5 indices shown in Figure 1c, several differences from ERA5 are apparent. First, the correlation between the SST and NAO indices is much lower in SEAS5 ($\hat{\tau} = 0.31$) than in ERA5 ($\hat{\tau} = 0.60$), demonstrating a discrepancy in the linkages between the observed SST pattern and the winter NAO. A second point is the weak correspondence between the NAO indices in SEAS5 and ERA5: the correlation is not statistically significant, reflecting the absence of predictive skill in the October-initialised SEAS5 forecasts. While this may seem unfavourable from a forecasting perspective, it is in fact advantageous for the present analysis. If the model had exhibited high skill, the atmospheric circulation in SEAS5 would likely have closely resembled that in ERA5, making it difficult to separate genuine model behaviour from externally constrained predictability. The absence of skill ensures that the SST index and the NAO index evolve largely independently in SEAS5 compared to ERA5. This independence allows examination of how the model internally links SST anomalies and the NAO, rather than simply reproducing observed co-variability. It also provides a rationale for using October- rather than November-initialised forecasts, since a later initialisation would likely yield higher skill and thus obscure a fair assessment of the model's feedback biases.

4.2 ERA5 climatology and SEAS5 bias

Before examining the role of surface fluxes and baroclinicity in mediating the SST–NAO relationship, it is useful to consider the climatological context. In Figure 2, ERA5 climatologies and SEAS5 biases are therefore shown, starting with the mean November SSTs in the North Atlantic in Fig 2a. Since the focus here is on the midlatitudes, the displayed region excludes the tropics (cf. Figure 1a). A prominent feature is the strong SST gradient along the boundary between the warm Gulf Stream waters and the much colder waters along the North American coastline. These gradients give rise to intense heat fluxes on the warm side of the front (Figure 2b), as well as strong low-level baroclinicity (Figure 2c). The last panel in the top row, Figure 2d, shows the climatological SLP pattern, consisting of the familiar dipole between the Icelandic Low and a high west of Gibraltar.

It is striking how poorly SEAS5 represents the SST gradient along the Gulf Stream seen in Figure 2a. The map in Figure 2e reveals a pronounced warm bias on the cold side of the front and a weaker cold bias on the warm side, resulting in an overall weakened gradient. Within the Subpolar Gyre, the picture is more mixed, with the strongest warm bias in its southern sector. The flux biases in Figure 2f mirror these SST errors, showing fluxes that are too strong in warm-biased regions and too weak in cold-biased sectors. Notably, the flux bias is negative across much of the Gyre despite modest SST biases there. Figure 2g further indicates that the underestimated SST gradients along the Gulf Stream are associated with too weak baroclinicity in the storm-track entry region, while σ_E is too high to the south.

Taken together, these errors imply a distorted storm track: too few cyclones over the northern North Atlantic and too many further south, which is consistent with the IFS cyclone bias investigations by Jung et al. (2006) and Büeler et al. (2024). This interpretation is supported by the SLP bias pattern in Figure 2h, which shows that SEAS5 fails to reproduce the observed SLP

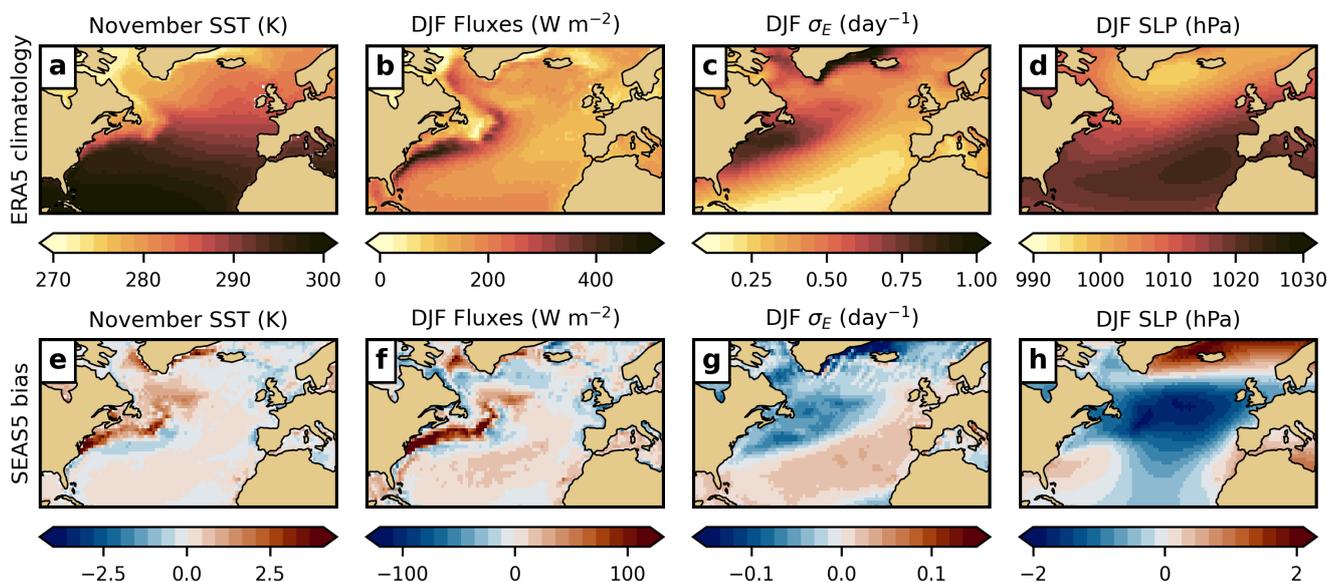


Figure 2. Top row: ERA5 climatologies for (a) November SST (K); (b) DJF surface heat fluxes ($W m^{-2}$); and (c) DJF Eady growth rate maximum σ_E (day^{-1}). Bottom row: SEAS5 biases (SEAS5 minus ERA5) for (d) November SST (K); (e) DJF surface heat fluxes ($W m^{-2}$); and (f) DJF σ_E (day^{-1}).

dipole and exhibits too weak westerly flow between the NAO centres of action. These weakened westerlies are commensurate with the negative flux bias in the Gyre (Figure 2f); the fluxes are probably too weak because there is less intense cold advection from the west.

4.3 Indirect effects, preconditioning, and feedbacks

4.3.1 Surface heat fluxes

Figure 3a displays the indirect effect of November SST anomalies on the winter NAO mediated by surface heat fluxes. Consistent with Kolstad and O'Reilly (2024), the Subpolar Gyre stands out as a hotspot of mediation. Given that the overall SST–NAO correlation is $\hat{\tau} = 0.6$, it is notable that the indirect effect in the western Gyre approaches 0.4, implying that flux variations in this region account for up to two-thirds of the SST–NAO relationship. This does not mean that these fluxes represent the only pathway linking SSTs to the NAO; rather, the correct interpretation is that the SST–NAO link would be weaker without the fluxes and their associated processes. Although not the focus here, positive indirect effects are also found over parts of the North and Baltic Seas, and there is little evidence of suppression anywhere in the domain.

A legitimate concern is that the NAO itself induces flux anomalies in the Subpolar Gyre (Khatri et al., 2022), making the reverse pathway $X \rightarrow Y \rightarrow Z$ (where X is the SST index, Y the NAO index, and Z the fluxes) equally plausible. To explore the influence of November SSTs on fluxes independently of the NAO, the partial correlation between X and Z with Y regressed

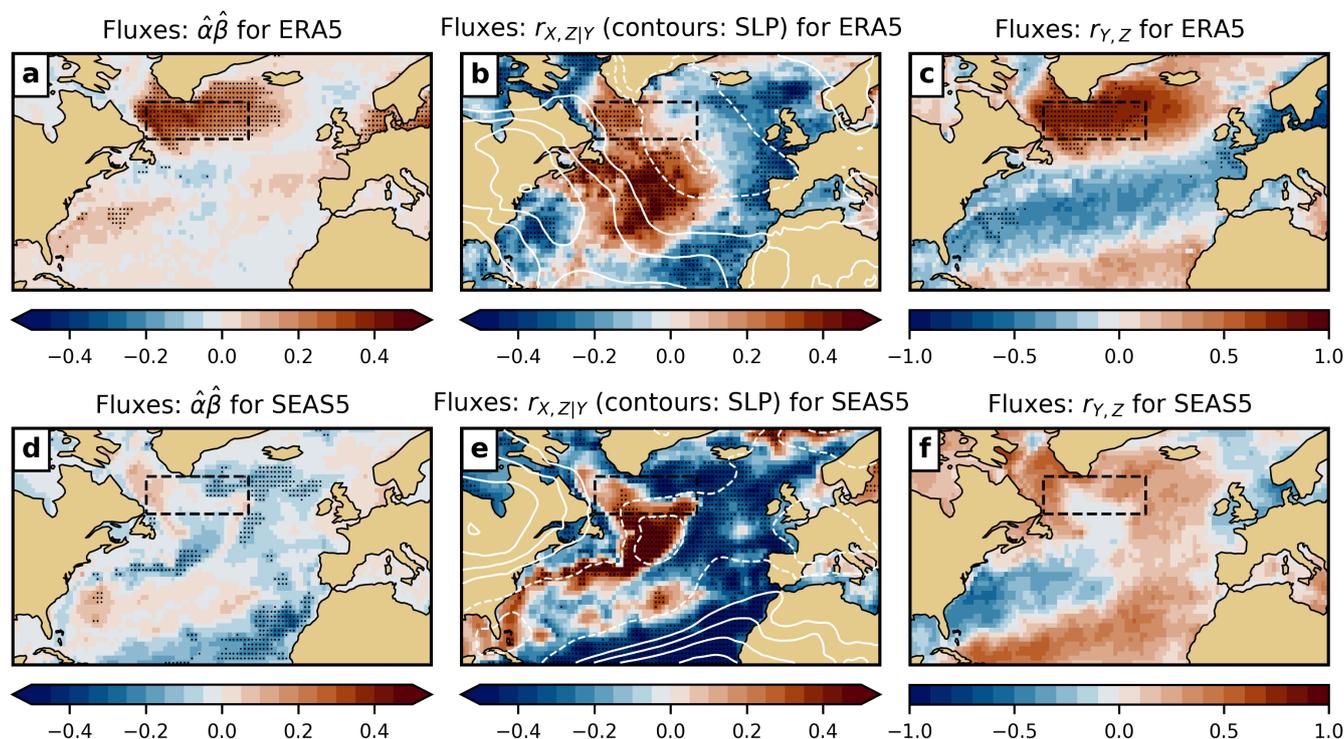


Figure 3. Top row: ERA5 (a) Sample indirect effect ($\hat{\alpha}\hat{\beta}$) via surface heat fluxes; (b) partial correlation ($r_{X,Z|Y}$) for fluxes (shading) and SLP (contours; interval 0.1 SD, positive contours solid, negative contours dashed, zero contour omitted); (c) simultaneous correlation between fluxes and the NAO. Bottom row (d–f): same as (a–c), but for SEAS5. The unit in all the panels is SD.

out from both ($r_{X,Z|Y}$ in Eq. 5) is examined in Figure 3b. The resulting pattern bears some resemblance to the SST anomaly structure used to define the November SST index (Figure 1a), consistent with the SST anomalies governing the sign of the heat fluxes. However, in the western Gyre, $r_{X,Z|Y}$ is significantly positive despite the negative SST anomalies in that area (Figure 1a), suggesting the presence of a process, independent of the NAO, that establishes these flux anomalies.

The white contours in Figure 3b, which depict the partial correlation between November SSTs and SLP (with the NAO removed), reveal a dipole roughly resembling an NAO pattern rotated less than 90° clockwise. This circulation response favours anomalous cold advection from the northwest into the Gyrefor positive realisations of the SST index, thereby enhancing upward heat fluxes and preconditioning the atmosphere for low-pressure development. For negative realisations, the advection and flux tendencies reverse to dampen such a development.

Figure 3c shows the simultaneous correlation between winter NAO and fluxes. Its close resemblance to the indirect-effect pattern in the Subpolar Gyre underscores the feedback nature of this coupling: the preconditioning revealed by Figure 3b thus likely helps to set the stage for NAO growth, while Figure 3c shows how the established NAO strengthens those flux anomalies



in return. This preconditioning–feedback sequence lends physical support to the interpretation of the indirect effect in Figure 3a as a genuine two-way air–sea feedback, rather than an artefact of the reverse causal pathway.

Turning to SEAS5, the indirect effect of SSTs on the NAO via fluxes is shown in Figure 3d. Its spatial pattern differs substantially from that in ERA5. Around the Gyre, significant negative $\hat{\alpha}\hat{\beta}$ values occur, indicating suppression rather than mediation. This appears to arise from two factors. First, a strong negative partial correlation between SSTs and fluxes independent of the NAO is present in the eastern Gyre, coinciding with the area of most negative indirect effect (Figure 3e). This signal does not align with the partial SST–SLP pattern, which shows no consistent advection structure in that region. However, the area of negative $\hat{\alpha}\hat{\beta}$ values overlaps with the negative flux bias seen in Figure 2f, probably linked to too weak westerly flow in SEAS5. Second, the contemporaneous relationship between fluxes and the NAO is much weaker in the southern Gyre in SEAS5 than ERA5 (Figure 3f vs. 3c). This muted sensitivity likely reflects the positive mean flux bias in that region (Figure 2f), itself linked to a positive SST bias (Figure 2e), which may dampen the response of flux anomalies to NAO-related circulation changes.

Beyond the Subpolar Gyre, widespread suppression is evident across the North Atlantic, notably along the Gulf Stream and its extensions, including the southward-flowing Canary Current. The overall picture is therefore one of broadly negative, spatially diffuse flux feedbacks, in sharp contrast to the strong and localised positive mediation seen in ERA5.

310 4.3.2 Baroclinicity

Figure 4a shows the indirect effect of November SST anomalies on the winter NAO mediated by lower-tropospheric baroclinicity, σ_E . Consistent with Kolstad and O'Reilly (2024), the indirect effect is uniformly and significantly positive within the outlined storm track entry region, confirming the important role of baroclinicity in mediating the SST–NAO link. Positive $\hat{\alpha}\hat{\beta}$ values also occur further south, where $\hat{\alpha}$ is negative (not shown), and around Iceland (where $\hat{\alpha}$ is positive).

Together, Figures 4b and 4c help resolve potential ambiguity about causality and provide physical context for the indirect effect shown in Figure 4a. Figure 4b demonstrates that late-autumn SST anomalies can precondition baroclinicity in a way that favours NAO development even after the NAO signal has been regressed out. This preconditioning occurs through anomalous advection in the region between the SLP poles indicated by white contours. When the SST index is positive, anomalous cold advection strengthens horizontal temperature gradients and reduces lower-tropospheric stability, both of which enhance the baroclinicity parameter σ_E . Positive partial correlations for σ_E in the Subpolar Gyre are indicative of conditions conducive to low-pressure development in the northern lobe of the NAO. For negative SST index cases, the advection pattern reverses: anomalous warm advection weakens the temperature gradients and increases stability, reducing σ_E . Figure 4c verifies that once the NAO is established, σ_E co-varies with it in a manner that reinforces the anomaly in the storm-track entry region and suppresses it further south, both consistent with a strengthening of the NAO itself.

In SEAS5, however, the contrast to ERA5 is unmistakable. In the entry region, Figure 4d exhibits only a shadow of the positive indirect effect seen for ERA5. Judging from Figure 4e, it is not clear why this should be the case, as the partial correlation between SSTs and σ_E is significant and positive in the Subpolar Gyre region. This in itself should strongly precondition the NAO in its northern centre of action, but Figure 4f makes clear that the feedback between the NAO and σ_E is weak and

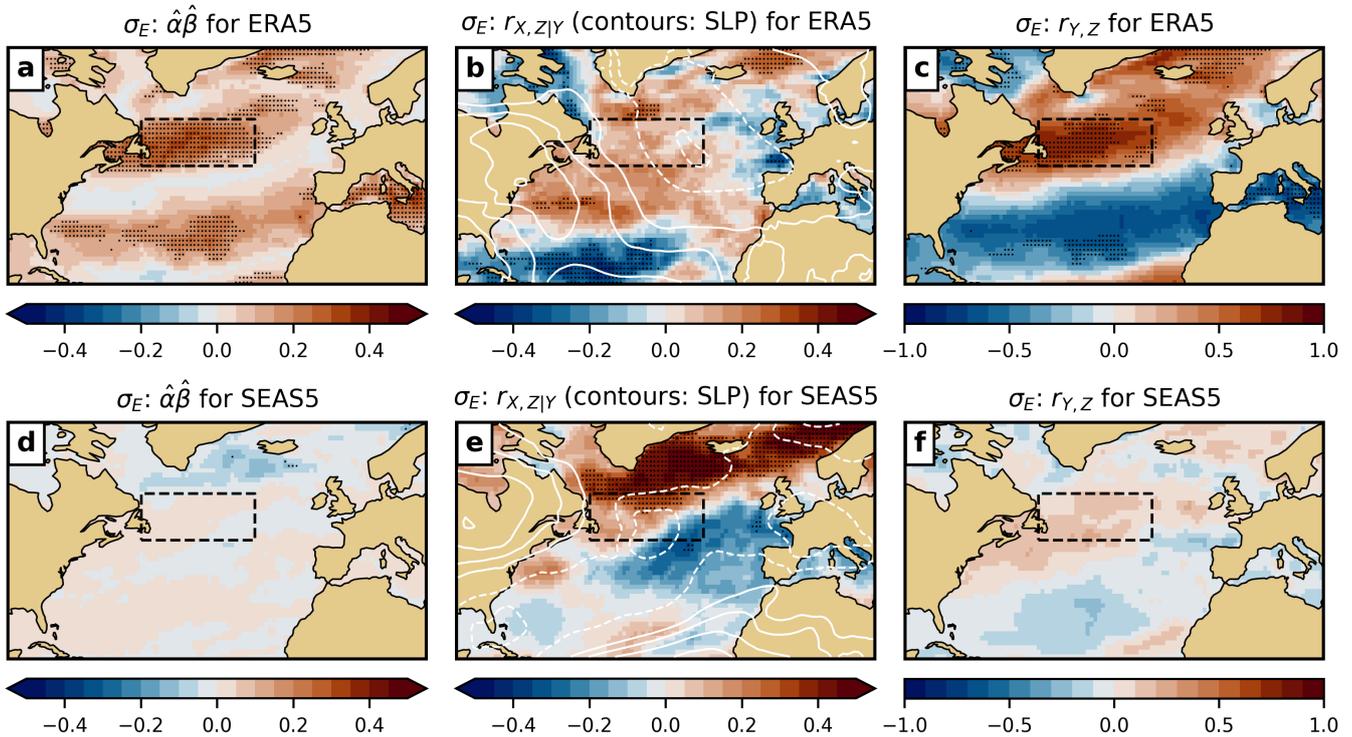


Figure 4. As Figure 3, but for the Eady growth rate maximum (σ_E).

non-significant. This points to a key deficiency in SEAS5: the feedback that should sustain baroclinicity once the NAO is
 330 established is too weak, limiting the overall strength of the mediation pathway.

4.4 Relating feedbacks to skill

This section examines whether the strength of the indirect effect $\hat{\alpha}\hat{\beta}$ scales with $\hat{\tau}$, the strength of the SST–NAO relationship. Section 4.1 revealed that $\hat{\tau}$ is 0.60 in ERA5 and 0.31 in SEAS5. Figures 3a and 4a demonstrated that the indirect effect in ERA5 is substantial within the outlined reference domains. Area-averaged values of $\hat{\alpha}\hat{\beta}$ within these boundaries amount to
 335 0.34 for fluxes and 0.41 for σ_E , corresponding to 57% and 69% of $\hat{\tau}$, respectively. In SEAS5, by contrast, the indirect effect is far weaker: only 10% of $\hat{\tau}$ for fluxes and 8% for σ_E . This confirms the visual impression from Figures 3d and 4d that mediation through these feedback pathways is marginal in SEAS5.

Figure 5 explores the relationship between the total and indirect effects more systematically. Bootstrap resampling was used to generate ensembles of 1000 $(\hat{\tau}, \hat{\alpha}\hat{\beta})$ pairs for both datasets, with each point representing a resampled period of the same
 340 length as the 1981–2023 study window. The dashed ovals enclose the approximate 95% confidence regions, computed as two-dimensional covariance ellipses of the bootstrap samples. The lack of overlap between the two point clouds for both fluxes

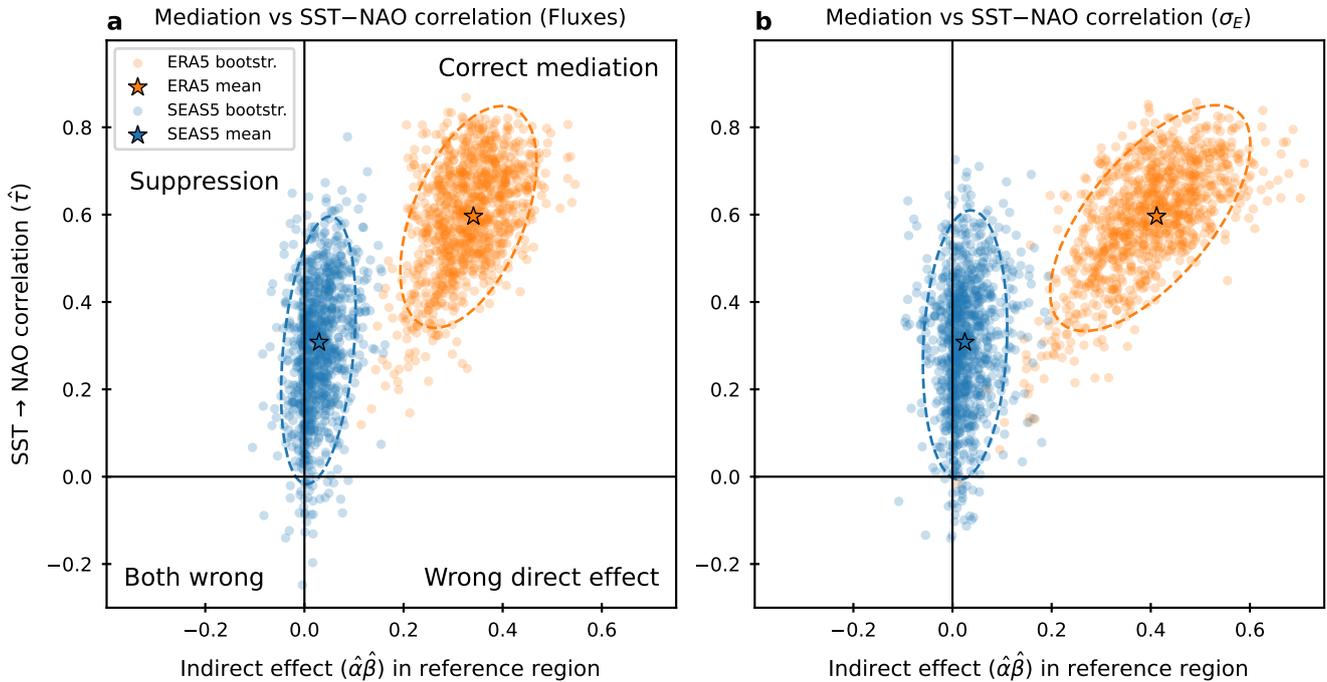


Figure 5. (a) Scatterplot of the indirect effect via fluxes, with $\hat{\alpha}\hat{\beta}$ on the x-axis and the total effect $\hat{\tau}$ on the y-axis. The stars mark the raw values for both datasets (ERA5 in orange and SEAS5 in blue), while the dots represent each of the 1000 bootstrapped $(\hat{\tau}, \hat{\alpha}\hat{\beta})$ pairs. Each bootstrap sample was generated by randomly resampling years with replacement to create synthetic time series of the same length as the original period (1981–2023), from which $\hat{\tau}$ and $\hat{\alpha}\hat{\beta}$ were recalculated. The dashed ovals enclose approximately 95% of these bootstrapped points, providing an estimate of the sampling uncertainty and typical joint variability of $\hat{\tau}$ and $\hat{\alpha}\hat{\beta}$. (b) As (a), but for the Eady growth rate maximum (σ_E). The unit is standard deviations (SD), and the x-axis label “correlation” is appropriate since both indices are standardised.

and baroclinicity highlights fundamental differences between the observational and model-based feedback structures. Before interpreting these results, it is useful to consider what the structure of such point clouds reveals about feedback processes.

If a feedback pathway is physically meaningful and consistently represented, the point cloud is expected to show a positive correlation between $\hat{\tau}$ and $\hat{\alpha}\hat{\beta}$. This implies that stronger total SST–NAO relationships tend to coincide with enhanced indirect effects. Such a cloud will be tilted towards the diagonal. In contrast, if the indirect effect is weak or inconsistently linked to $\hat{\tau}$, the cloud will be oriented along the y-axis, reflecting that some samples may exhibit a strong total effect without any associated change in the indirect effect. Such a pattern suggests a non-existent mediation pathway.

Starting with the flux pathway in Figure 5a, the ERA5 point cloud shows a modest diagonal tilt, and the correlation between $\hat{\tau}$ and $\hat{\alpha}\hat{\beta}$ is $r = 0.49$, indicating that the indirect effect explains roughly a quarter of the variance in the SST–NAO correlation. Nearly all ERA5 samples (orange points) lie in the upper-right quadrant, labelled “correct mediation” because both $\hat{\tau}$ and $\hat{\alpha}\hat{\beta}$ are positive. The SEAS5 point cloud (blue) presents a clear contrast: it is more vertically aligned, and while some resampled periods have relatively high $\hat{\tau}$ values, these are not consistently associated with strong indirect effects. The correlation be-



355 tween $\hat{\tau}$ and $\hat{\alpha}\hat{\beta}$ is lower ($r = 0.31$), though still significant, implying that flux-mediated feedbacks explain about 10% of the variance in $\hat{\tau}$. A non-negligible number of SEAS5 samples fall in the upper-left quadrant, where $\hat{\tau} > 0$ but $\hat{\alpha}\hat{\beta} < 0$, indicating suppression rather than mediation, and a smaller number even appear in the lower quadrants, where the sign of $\hat{\tau}$ is incorrect.

The contrast is even sharper for the baroclinicity pathway (Figure 5b). In ERA5, the correlation between $\hat{\tau}$ and $\hat{\alpha}\hat{\beta}$ is $r = 0.66$, with the indirect effect accounting for 44% of the variance in $\hat{\tau}$. The point cloud is clearly aligned along the diagonal, reflecting a close link between the total effect and mediation strength. In SEAS5, by contrast, the correlation drops to $r = 0.13$, 360 which, although still significant at the 5% level, indicates that only 2% of the variance in $\hat{\tau}$ is associated with mediation through baroclinicity. These results underscore how weakly this feedback is represented in SEAS5 compared to reanalysis. This weakness is particularly striking given that the baroclinicity pathway is the dominant mediator of the SST–NAO link in ERA5, explaining a larger fraction of the total effect than the flux pathway. Its near absence in SEAS5 therefore points to a fundamental deficiency in the model’s representation of the dynamical processes that sustain the storm track and couple it to 365 oceanic anomalies and the NAO.

5 Discussion

In this paper, feedback pathways linking the state of the North Atlantic sea surface in late autumn and the NAO during the following winter have been explored. As in Kolstad and O’Reilly (2024), these pathways were investigated using *mediation analysis*, a branch of statistical causal inference methods that has seen little use in climate dynamics so far. A central aim was 370 to demonstrate that feedbacks previously identified through idealised perturbation experiments in dynamical models can also be diagnosed directly from observational or reanalysis data. One clear advantage of this approach is that it avoids the need to manipulate boundary conditions like SSTs. Such manipulations can elicit compensatory model adjustments that complicate interpretation, particularly when the models themselves suffer from systematic biases. Mediation analysis instead infers causal structure directly from observed covariability, offering a complementary perspective on internal feedback pathways.

375 It must nevertheless be acknowledged that reanalysis products are themselves produced with models – in the case of ERA5, from the same model lineage as SEAS5. Thus, reanalyses are not free from biases, and their depiction of physical relationships may be influenced by model behaviour. Mediation analysis cannot fully resolve such issues, but by contrasting reanalysis-based and model-based feedbacks, it can help to pinpoint where key processes diverge.

380 Additional limitations should be kept in mind. For one, the mediation framework as applied here is linear and does not adequately capture nonlinear feedbacks. Further, SEAS5 is only one dynamical system; different models likely represent feedback differently. The reason only one model is investigated here is that its reforecast period extends back to 1981, while reforecasts are only available from 1993 and onwards for comparable systems – this shorter period would render the mediation analysis less robust. Future work could extend this examination to multi-model ensembles and incorporate nonlinear mediation techniques.

385 Notwithstanding these caveats, this study has extended Kolstad and O’Reilly (2024) by revealing a physically coherent sequence of processes linking late-autumn SST anomalies and the winter NAO. Independent of the NAO itself, November SST anomalies induce a surface-pressure dipole pattern resembling a clockwise-rotated NAO. This preconditions the atmosphere



for anomalies in surface fluxes and baroclinicity in the Subpolar Gyre, which in turn nudge the NAO. Once established, the NAO feeds back on both fluxes and baroclinicity, reinforcing the initial anomalies and sustaining the circulation pattern.

It is important to emphasise that these feedbacks do not account for all aspects of NAO variability. Far from it; the processes identified here represent only one pathway among many, complementing influences from, for example, the stratosphere, tropical SSTs, Arctic sea-ice variability, and internal atmospheric dynamics. Rather than providing a complete explanation, the present results demonstrate how even a single coupled feedback sequence can contribute to shaping NAO variability and how its misrepresentation in a prediction system may limit its ability to capture the full spectrum of NAO behaviour.

For a key finding is that these pathways are considerably weakened or absent in SEAS5. In particular, the baroclinicity feedback – the process by which enhanced or suppressed eddy growth reinforces the circulation anomaly – is nearly non-existent, while the flux-mediated pathway is much weaker than in ERA5. As a result, these processes contribute little to the lagged SST–NAO relationship in SEAS5, which is consistent with the system’s limited skill in predicting the NAO even when initialised with October SST anomalies. This suggests that the misrepresentation of critical feedback mechanisms is a fundamental barrier to improved forecast performance.

These findings have practical implications for model development. Enhancing ocean resolution (Haarsma et al., 2019), reducing Gulf Stream SST biases (Roberts et al., 2021), and improving the representation of eddy–mean flow feedbacks (Hardiman et al., 2022) are all likely to strengthen the feedback pathways highlighted here. Encouragingly, the upcoming SEAS6 system includes a new ocean model that, among other improvements, reduces heat flux biases along the Gulf Stream (Keeley et al., 2024).

The results presented here is that they raise important questions for the emerging class of ML-based seasonal and subseasonal prediction systems (e.g., Chen et al., 2024; Kent et al., 2025). If trained on model-generated data or on reanalyses influenced by model biases, such systems risk inheriting some of the deficiencies documented here. Conversely, ML approaches trained directly on observations might bypass some of these problems — but whether they capture the same preconditioning and feedback structures as the real climate system remains an open question.

Mediation analysis offers a powerful and versatile framework for tackling these research challenges. It can help pinpoint where models fail to represent key causal pathways, assess whether targeted improvements translate into more realistic coupled feedbacks and higher predictive skill, and evaluate whether ML-based forecasts reproduce the same physical linkages observed in nature. In a broader sense, mediation analysis can serve as a bridge between statistical diagnostics and both process studies and model development/evaluation, advancing our understanding of how ocean–atmosphere coupling shapes climate predictability.

Acknowledgements. Funding for this work was provided by the Research Council of Norway through the Climate Futures centre for research-based innovation (grant 309562). The author also wishes to thank the ECMWF for providing access to both ERA5 and SEAS5 data. The figures use colour maps developed by Fabio Cramer (Cramer et al., 2020). Parts of the text were refined with the help of OpenAI’s ChatGPT, which was used to improve clarity and language and to assist with coding tasks, but not to generate scientific content or conclusions.



420 References

- Ambaum, M. H. P., Hoskins, B. J., and Stephenson, D. B.: Arctic Oscillation or North Atlantic Oscillation?, *Journal of Climate*, 2001.
- Athanasiadis, P. J., Bellucci, A., Scaife, A. A., Hermanson, L., Materia, S., Sanna, A., Borrelli, A., MacLachlan, C., and Gualdi, S.: A Multisystem View of Wintertime NAO Seasonal Predictions, *Journal of Climate*, <https://doi.org/10.1175/JCLI-D-16-0153.1>, 2017.
- Athanasiadis, P. J., Ogawa, F., Omrani, N.-E., Keenlyside, N., Schiemann, R., Baker, A. J., Vidale, P. L., Bellucci, A., Ruggieri, P., Haarsma, R., Roberts, M., Roberts, C., Novak, L., and Gualdi, S.: Mitigating Climate Biases in the Midlatitude North Atlantic by Increasing Model Resolution: SST Gradients and Their Relation to Blocking and the Jet, *Journal of Climate*, <https://doi.org/10.1175/JCLI-D-21-0515.1>, 2022.
- Baker, H. S., Woollings, T., Forest, C. E., and Allen, M. R.: The Linear Sensitivity of the North Atlantic Oscillation and Eddy-Driven Jet to SSTs, *Journal of Climate*, 32, 6491–6511, <https://doi.org/10.1175/JCLI-D-19-0038.1>, 2019.
- 430 Baker, L. H., Shaffrey, L. C., Johnson, S. J., and Weisheimer, A.: Understanding the Intermittency of the Wintertime North Atlantic Oscillation and East Atlantic Pattern Seasonal Forecast Skill in the Copernicus C3S Multi-Model Ensemble, *Geophysical Research Letters*, 51, e2024GL108472, <https://doi.org/10.1029/2024GL108472>, 2024.
- Baron, R. M. and Kenny, D. A.: The Moderator–Mediator Variable Distinction in Social Psychological Research: Conceptual, Strategic, and Statistical Considerations., *Journal of personality and social psychology*, 51, 1173–1182, <https://doi.org/10.1037/0022-3514.51.6.1173>, 1986.
- 435 Bellucci, A., Athanasiadis, P. J., Scoccimarro, E., Ruggieri, P., Gualdi, S., Fedele, G., Haarsma, R. J., Garcia-Serrano, J., Castrillo, M., Putrahasan, D., Sanchez-Gomez, E., Moine, M.-P., Roberts, C. D., Roberts, M. J., Seddon, J., and Vidale, P. L.: Air-Sea Interaction over the Gulf Stream in an Ensemble of HighResMIP Present Climate Simulations, *Climate Dynamics*, 56, 2093–2111, <https://doi.org/10.1007/s00382-020-05573-z>, 2021.
- 440 Büeler, D., Sprenger, M., and Wernli, H.: Northern Hemisphere Extratropical Cyclone Biases in ECMWF Subseasonal Forecasts, *Quarterly Journal of the Royal Meteorological Society*, 150, 1096–1123, <https://doi.org/10.1002/qj.4638>, 2024.
- Cassou, C., Deser, C., and Alexander, M. A.: Investigating the Impact of Reemerging Sea Surface Temperature Anomalies on the Winter Atmospheric Circulation over the North Atlantic, *Journal of Climate*, <https://doi.org/10.1175/JCLI4202.1>, 2007.
- Chen, L., Zhong, X., Li, H., Wu, J., Lu, B., Chen, D., Xie, S.-P., Wu, L., Chao, Q., Lin, C., Hu, Z., and Qi, Y.: A Machine Learning Model That Outperforms Conventional Global Subseasonal Forecast Models, *Nature Communications*, 15, 6425, <https://doi.org/10.1038/s41467-024-50714-1>, 2024.
- 445 Conger, A. J.: A Revised Definition for Suppressor Variables: A Guide To Their Identification and Interpretation, *Educational and Psychological Measurement*, 34, 35–46, <https://doi.org/10.1177/001316447403400105>, 1974.
- Cramer, F., Shephard, G. E., and Heron, P. J.: The Misuse of Colour in Science Communication, *Nature Communications*, 11, 5444, <https://doi.org/10.1038/s41467-020-19160-7>, 2020.
- 450 Czaja, A. and Frankignoul, C.: Observed Impact of Atlantic SST Anomalies on the North Atlantic Oscillation, *Journal of Climate*, 15, 606–623, [https://doi.org/10.1175/1520-0442\(2002\)015<0606:OIOASA>2.0.CO;2](https://doi.org/10.1175/1520-0442(2002)015<0606:OIOASA>2.0.CO;2), 2002.
- Dawson, A.: Eofs: A Library for EOF Analysis of Meteorological, Oceanographic, and Climate Data | *Journal of Open Research Software*, *Journal of Open Research Software*, 4, <https://doi.org/10.5334/jors.122>, 2016.
- 455 Degenhardt, L., Leckebusch, G. C., and Scaife, A. A.: Large-Scale Circulation Patterns and Their Influence on European Winter Windstorm Predictions, *Climate Dynamics*, 60, 3597–3611, <https://doi.org/10.1007/s00382-022-06455-2>, 2023.



- Docquier, D., Di Capua, G., Donner, R. V., Pires, C. A. L., Simon, A., and Vannitsem, S.: A Comparison of Two Causal Methods in the Context of Climate Analyses, *Nonlinear Processes in Geophysics*, 31, 115–136, <https://doi.org/10.5194/npg-31-115-2024>, 2024.
- Ebert-Uphoff, I. and Deng, Y.: Causal Discovery for Climate Research Using Graphical Models, *Journal of Climate*,
460 <https://doi.org/10.1175/JCLI-D-11-00387.1>, 2012.
- Gastineau, G. and Frankignoul, C.: Influence of the North Atlantic SST Variability on the Atmospheric Circulation during the Twentieth Century, *Journal of Climate*, <https://doi.org/10.1175/JCLI-D-14-00424.1>, 2015.
- Granger, C. W.: Investigating Causal Relations by Econometric Models and Cross-Spectral Methods, *Econometrica*, pp. 424–438, <https://doi.org/10.2307/1912791>, 1969.
- 465 Haarsma, R. J., García-Serrano, J., Prodhomme, C., Bellprat, O., Davini, P., and Drijfhout, S.: Sensitivity of Winter North Atlantic-European Climate to Resolved Atmosphere and Ocean Dynamics, *Scientific Reports*, 9, 13 358, <https://doi.org/10.1038/s41598-019-49865-9>, 2019.
- Hall, R. J., Scaife, A. A., Hanna, E., Jones, J. M., and Erdélyi, R.: Simple Statistical Probabilistic Forecasts of the Winter NAO, *Weather and Forecasting*, <https://doi.org/10.1175/WAF-D-16-0124.1>, 2017.
- Hannart, A., Pearl, J., Otto, F. E. L., Naveau, P., and Ghil, M.: Causal Counterfactual Theory for the Attribution of Weather and Climate-
470 Related Events, *Bulletin of the American Meteorological Society*, <https://doi.org/10.1175/BAMS-D-14-00034.1>, 2016.
- Hardiman, S. C., Dunstone, N. J., Scaife, A. A., Smith, D. M., Comer, R., Nie, Y., and Ren, H.-L.: Missing Eddy Feedback May Explain Weak Signal-to-Noise Ratios in Climate Predictions, *npj Climate and Atmospheric Science*, 5, 57, <https://doi.org/10.1038/s41612-022-00280-4>, 2022.
- Hermanson, L., Ren, H.-L., Vellinga, M., Dunstone, N. D., Hyder, P., Ineson, S., Scaife, A. A., Smith, D. M., Thompson, V., Tian, B., and
475 Williams, K. D.: Different Types of Drifts in Two Seasonal Forecast Systems and Their Dependence on ENSO, *Climate Dynamics*, 51, 1411–1426, <https://doi.org/10.1007/s00382-017-3962-9>, 2018.
- Hersbach, H., Bell, B., Berrisford, P., Hirahara, S., Horányi, A., Muñoz-Sabater, J., Nicolas, J., Peubey, C., Radu, R., Schepers, D., Simmons, A., Soci, C., Abdalla, S., Abellan, X., Balsamo, G., Bechtold, P., Biavati, G., Bidlot, J., Bonavita, M., De Chiara, G., Dahlgren, P., Dee, D., Diamantakis, M., Dragani, R., Flemming, J., Forbes, R., Fuentes, M., Geer, A., Haimberger, L., Healy, S., Hogan, R. J.,
480 Hólm, E., Janisková, M., Keeley, S., Lalouaux, P., Lopez, P., Lupu, C., Radnoti, G., de Rosnay, P., Rozum, I., Vamborg, F., Villaume, S., and Thépaut, J.-N.: The ERA5 Global Reanalysis, *Quarterly Journal of the Royal Meteorological Society*, 146, 1999–2049, <https://doi.org/10.1002/qj.3803>, 2020.
- Hertig, E., Beck, C., Wanner, H., and Jacobeit, J.: A Review of Non-Stationarities in Climate Variability of the Last Century with Focus on the North Atlantic–European Sector, *Earth-Science Reviews*, 147, 1–17, <https://doi.org/10.1016/j.earscirev.2015.04.009>, 2015.
- 485 Hewitt, H. T., Bell, M. J., Chassignet, E. P., Czaja, A., Ferreira, D., Griffies, S. M., Hyder, P., McClean, J. L., New, A. L., and Roberts, M. J.: Will High-Resolution Global Ocean Models Benefit Coupled Predictions on Short-Range to Climate Timescales?, *Ocean Modelling*, 120, 120–136, <https://doi.org/10.1016/j.ocemod.2017.11.002>, 2017.
- Hoskins, B. J. and Valdes, P. J.: On the Existence of Storm-Tracks, *Journal of Atmospheric Sciences*, 47, 1854–1864, 1990.
- Jin, X. and Yu, L.: Assessing High-Resolution Analysis of Surface Heat Fluxes in the Gulf Stream Region, *Journal of Geophysical Research: Oceans*, 118, 5353–5375, <https://doi.org/10.1002/jgrc.20386>, 2013.
- 490 Johnson, S. J., Stockdale, T. N., Ferranti, L., Balmaseda, M. A., Molteni, F., Magnusson, L., Tietsche, S., Decremmer, D., Weisheimer, A., Balsamo, G., Keeley, S. P. E., Mogensen, K., Zuo, H., and Monge-Sanz, B. M.: SEAS5: The New ECMWF Seasonal Forecast System, *Geoscientific Model Development*, 12, 1087–1117, <https://doi.org/10.5194/gmd-12-1087-2019>, 2019.



- Joyce, T. M., Kwon, Y.-O., Seo, H., and Ummerhofer, C. C.: Meridional Gulf Stream Shifts Can Influence Wintertime Variability in the North Atlantic Storm Track and Greenland Blocking, *Geophysical Research Letters*, 46, 1702–1708, <https://doi.org/10.1029/2018GL081087>, 2019.
- Jung, T., Gulev, S. K., Rudeva, I., and Soloviev, V.: Sensitivity of Extratropical Cyclone Characteristics to Horizontal Resolution in the ECMWF Model, *Quarterly Journal of the Royal Meteorological Society*, 132, 1839–1857, <https://doi.org/10.1256/qj.05.212>, 2006.
- Keeley, S., Mogensen, K., Bidlot, J., Balmaseda, M. A., and Hatfield, S.: Introduction of a New Ocean and Sea-Ice Model Based on NEMO-SI3, *ECMWF Newsletter*, 180, 24–29, 2024.
- Kent, C., Scaife, A. A., Dunstone, N. J., Smith, D., Hardiman, S. C., Dunstan, T., and Watt-Meyer, O.: Skilful Global Seasonal Predictions from a Machine Learning Weather Model Trained on Reanalysis Data, *npj Climate and Atmospheric Science*, 8, 314, <https://doi.org/10.1038/s41612-025-01198-3>, 2025.
- Khatri, H., Williams, R. G., Woollings, T., and Smith, D. M.: Fast and Slow Subpolar Ocean Responses to the North Atlantic Oscillation: Thermal and Dynamical Changes, *Geophysical Research Letters*, 49, e2022GL101480, <https://doi.org/10.1029/2022GL101480>, 2022.
- Knutti, R. and Rugenstein, M. A. A.: Feedbacks, Climate Sensitivity and the Limits of Linear Models, *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 373, 20150146, <https://doi.org/10.1098/rsta.2015.0146>, 2015.
- Kolstad, E. W. and O'Reilly, C. H.: Causal Oceanic Feedbacks onto the Winter NAO, *Climate Dynamics*, 62, 4223–4236, <https://doi.org/10.1007/s00382-024-07128-y>, 2024.
- Kolstad, E. W. and Screen, J. A.: Nonstationary Relationship Between Autumn Arctic Sea Ice and the Winter North Atlantic Oscillation, *Geophysical Research Letters*, 46, 7583–7591, <https://doi.org/10.1029/2019GL083059>, 2019.
- Lewis, N. T., England, M. R., Screen, J. A., Geen, R., Mudhar, R., Seviour, W. J. M., and Thomson, S. I.: Assessing the Spurious Impacts of Ice-Constraining Methods on the Climate Response to Sea Ice Loss Using an Idealized Aquaplanet GCM, *Journal of Climate*, <https://doi.org/10.1175/JCLI-D-24-0153.1>, 2024.
- MacKinnon, D. P., Krull, J. L., and Lockwood, C. M.: Equivalence of the Mediation, Confounding and Suppression Effect, *Prevention Science*, 1, 173–181, <https://doi.org/10.1023/A:1026595011371>, 2000.
- McGraw, M. C. and Barnes, E. A.: Memory Matters: A Case for Granger Causality in Climate Variability Studies, *Journal of Climate*, <https://doi.org/10.1175/JCLI-D-17-0334.1>, 2018.
- Mosedale, T. J., Stephenson, D. B., Collins, M., and Mills, T. C.: Granger Causality of Coupled Climate Processes: Ocean Feedback on the North Atlantic Oscillation, *Journal of Climate*, <https://doi.org/10.1175/JCLI3653.1>, 2006.
- Mu, B., Jiang, X., Yuan, S., Cui, Y., and Qin, B.: NAO Seasonal Forecast Using a Multivariate Air–Sea Coupled Deep Learning Model Combined with Causal Discovery, *Atmosphere*, 14, 792, <https://doi.org/10.3390/atmos14050792>, 2023.
- Muniz, F. B. and MacKinnon, D. P.: Three Approaches to Testing for Statistical Suppression, *Multivariate Behavioral Research*, 60, 817–839, <https://doi.org/10.1080/00273171.2025.2483245>, 2025.
- Nguyen, T. Q., Schmid, I., and Stuart, E. A.: Clarifying Causal Mediation Analysis for the Applied Researcher: Defining Effects Based on What We Want to Learn, *Psychological Methods*, 26, 255–271, <https://doi.org/10.1037/met0000299>, 2021.
- Palmer, T. N. and Weisheimer, A.: Diagnosing the Causes of Bias in Climate Models – Why Is It so Hard?, *Geophysical & Astrophysical Fluid Dynamics*, 105, 351–365, <https://doi.org/10.1080/03091929.2010.547194>, 2011.
- Pan, L.-L.: Observed Positive Feedback between the NAO and the North Atlantic SSTA Tripole, *Geophysical Research Letters*, 32, 1–4, <https://doi.org/10.1029/2005GL022427>, 2005.



- Patrizio, C. R., Athanasiadis, P. J., Frankignoul, C., Iovino, D., Masina, S., Paolini, L. F., and Gualdi, S.: Improved Extratropical North Atlantic Atmosphere–Ocean Variability with Increasing Ocean Model Resolution, *Journal of Climate*, <https://doi.org/10.1175/JCLI-D-23-0230.1>, 2023.
- Patrizio, C. R., Athanasiadis, P. J., Smith, D. M., and Nicoli, D.: Ocean-Atmosphere Feedbacks Key to NAO Decadal Predictability, *npj Climate and Atmospheric Science*, 8, 146, 2025.
- Pearl, J., Glymour, M., and Jewell, N. P.: *Causal Inference in Statistics: A Primer*, John Wiley & Sons, ISBN 978-1-119-18686-1, 2016.
- Peng, S., Robinson, W. A., and Li, S.: North Atlantic SST Forcing of the NAO and Relationships with Intrinsic Hemispheric Variability, *Geophysical Research Letters*, 29, 117–1–117–4, <https://doi.org/10.1029/2001GL014043>, 2002.
- Rivière, G. and Orlanski, I.: Characteristics of the Atlantic Storm-Track Eddy Activity and Its Relation with the North Atlantic Oscillation, *Journal of the Atmospheric Sciences*, <https://doi.org/10.1175/JAS3850.1>, 2007.
- Roberts, C. D., Vitart, F., and Balmaseda, M. A.: Hemispheric Impact of North Atlantic SSTs in Subseasonal Forecasts, *Geophysical Research Letters*, 48, e2020GL0911446, <https://doi.org/10.1029/2020GL091446>, 2021.
- Rodwell, M. J., Rowell, D. P., and Folland, C. K.: Oceanic Forcing of the Wintertime North Atlantic Oscillation and European Climate, *Nature*, 398, 320–323, 1999.
- Runge, J., Bathiany, S., Bollt, E., Camps-Valls, G., Coumou, D., Deyle, E., Glymour, C., Kretschmer, M., Mahecha, M. D., Muñoz-Marí, J., van Nes, E. H., Peters, J., Quax, R., Reichstein, M., Scheffer, M., Schölkopf, B., Spirtes, P., Sugihara, G., Sun, J., Zhang, K., and Zscheischler, J.: Inferring Causation from Time Series in Earth System Sciences, *Nature Communications*, 10, 2553, <https://doi.org/10.1038/s41467-019-10105-3>, 2019.
- Scaife, A. A. and Smith, D.: A Signal-to-Noise Paradox in Climate Science, *npj Climate and Atmospheric Science*, 1, 28, <https://doi.org/10.1038/s41612-018-0038-4>, 2018.
- Scaife, A. A., Arribas, A., Blockley, E., Brookshaw, A., Clark, R. T., Dunstone, N., Eade, R., Fereday, D., Folland, C. K., Gordon, M., Hermanson, L., Knight, J. R., Lea, D. J., MacLachlan, C., Maidens, A., Martin, M., Peterson, A. K., Smith, D., Vellinga, M., Wallace, E., Waters, J., and Williams, A.: Skillful Long-Range Prediction of European and North American Winters, *Geophysical Research Letters*, 41, 2514–2519, <https://doi.org/10.1002/2014GL059637>, 2014.
- Stott, P., Good, P., Jones, G., Gillett, N., and Hawkins, E.: The Upper End of Climate Model Temperature Projections Is Inconsistent with Past Warming, *Environmental Research Letters*, 8, 014024, <https://doi.org/10.1088/1748-9326/8/1/014024>, 2013.
- Suckling, E. B. and Smith, L. A.: An Evaluation of Decadal Probability Forecasts from State-of-the-Art Climate Models, *Journal of Climate*, <https://doi.org/10.1175/JCLI-D-12-00485.1>, 2013.
- Sun, Y., Simpson, I., Wei, H.-L., and Hanna, E.: Probabilistic Seasonal Forecasts of North Atlantic Atmospheric Circulation Using Complex Systems Modelling and Comparison with Dynamical Models, *Meteorological Applications*, 31, e2178, <https://doi.org/10.1002/met.2178>, 2024.
- Wang, L., Ting, M., and Kushner, P. J.: A Robust Empirical Seasonal Prediction of Winter NAO and Surface Climate, *Scientific Reports*, 7, 279, <https://doi.org/10.1038/s41598-017-00353-y>, 2017.
- Wang, W., Anderson, B. T., Kaufmann, R. K., and Myneni, R. B.: The Relation between the North Atlantic Oscillation and SSTs in the North Atlantic Basin, *Journal of Climate*, 17, 4752–4759, <https://doi.org/10.1175/JCLI-3186.1>, 2004.
- Watanabe, M. and Kimoto, M.: Atmosphere-Ocean Thermal Coupling in the North Atlantic: A Positive Feedback, *Quarterly Journal of the Royal Meteorological Society*, 126, 3343–3369, 2000.



- Weisheimer, A., Schaller, N., O'Reilly, C., MacLeod, D. A., and Palmer, T.: Atmospheric Seasonal Forecasts of the Twentieth Century: Multi-Decadal Variability in Predictive Skill of the Winter North Atlantic Oscillation (NAO) and Their Potential Value for Extreme Event Attribution, *Quarterly Journal of the Royal Meteorological Society*, 143, 917–926, <https://doi.org/10.1002/qj.2976>, 2017.
- 570 Weisheimer, A., Baker, L. H., Bröcker, J., Garfinkel, C. I., Hardiman, S. C., Hodson, D. L. R., Palmer, T. N., Robson, J. I., Scaife, A. A., Screen, J. A., Shepherd, T. G., Smith, D. M., and Sutton, R. T.: The Signal-to-Noise Paradox in Climate Forecasts: Revisiting Our Understanding and Identifying Future Priorities, *Bulletin of the American Meteorological Society*, <https://doi.org/10.1175/BAMS-D-24-0019.1>, 2024.
- 575 Woollings, T. and Blackburn, M.: The North Atlantic Jet Stream under Climate Change and Its Relation to the NAO and EA Patterns, *Journal of Climate*, <https://doi.org/10.1175/JCLI-D-11-00087.1>, 2012.
- Zhang, W., Kirtman, B., Siqueira, L., Clement, A., and Xia, J.: Understanding the Signal-to-Noise Paradox in Decadal Climate Predictability from CMIP5 and an Eddy Global Coupled Model, *Climate Dynamics*, 56, 2895–2913, <https://doi.org/10.1007/s00382-020-05621-8>, 2021.