Deficient ocean–atmosphere feedbacks constrain seasonal NAO prediction

Erik W. Kolstad[1]

[1]NORCE 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~~ . This is hypothesised here to be partly attributable to deficient representation of ocean–atmosphere feedback mechanisms.

These feedbacks leave an imprint in a lagged relationship between late-autumn ~~North Atlantic sea surface~~ sea-surface temperature (SST) anomalies ~~to~~ and 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 modelperturbation experiments.~~ A subset of these feedbacks and their representation by the seasonal prediction system SEAS5 is studied and compared with the ERA5 reanalysis. It is first confirmed that the model's internal SST–NAO relationship is correlated with its NAO forecast skill. To further examine this result, mediation analysis is applied to identify and quantify feedback pathways.

~~The analysis reveals a physically coherent feedback sequence.~~ Anomalous November SST patterns ~~promote the gradual formation of~~ are found to promote a surface-pressure ~~dipole rotated clockwise relative to the canonical NAO structure. This dipole induces advection anomalies~~ anomaly that induces advection in the western North Atlantic~~, which in turnmodulate~~ . In turn, this modulates surface fluxes in the Subpolar Gyre and lower-tropospheric baroclinicity in the storm-track ~~entry~~ entrance region east of Newfoundland. These changes ~~nudge~~ precondition 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~~ produces weaker mediated effects via both fluxes and baroclinicity than those found in ERA5~~reanalysis. This misrepresentation likely represents a fundamental barrier to improved~~ . Another key result is that the mediated effects in the model are correlated with its 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 By implication, models that are able to reproduce realistic mediation are likely to achieve higher NAO skill.

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. (e.g. Athanasiadis et al., 2017; D
.

Owing to its wide-ranging influence, the NAO is routinely used as a benchmark for mid-latitude seasonal prediction skill. Statistical (empirical) methods have a long history and have achieved useful results potentially useful levels of skill 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). (e.g. Hall et al., 2017; W
. While empirical models offer interpretability and often have at times appeared to outperform dynamical systemsin 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). More recently, empirical models have been enriched with machine-learning and hybrid techniques, yielding higher forecast skill, including for the NAO (e.g. Mu et al., 2023; Sun et al., 2024)
.

In sum, empirical models are susceptible to including non-causal predictors, either because background-state changes render previous relationships invalid or due to spurious correlations. In principle, dynamical coupled prediction systems ought not to be constrained by non-stationarity, since are not constrained by these limitations, as they aim to reproduce the behaviour of the climate system from first principles. Ideally, such systems should Such systems should ideally 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 surface-defined NAO skill reported in some dynamical systems (Scaife et al., 2014). However, ~~subsequent~~ there is a wide range of performance between systems and system upgrades have not ~~improved (or in some cases have even degraded) this level of~~ significantly improved overall 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, Patrizio et al. (2025) showed that skill in decadal NAO forecasts depends on how models represent feedbacks between subpolar SST anomalies and the NAO. 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~~ storm track entrance 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).~~ (Hewitt et al., 2017; Bellucci et al., 2021; Athanasiadis 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, building on both observations or reanalysis data (Czaja and Frankignoul, 2002; Wang et al., 2004; Hall et al., 2017 and models (Rodwell et al., 1999; Watanabe and Kimoto, 2000; Baker et al., 2019; Sun et al., 2024), has shown that characteristic SST patterns can precondition the atmosphere on subseasonal to seasonal timescales~~(Rodwell et al., 1999; Watanabe and Kimoto, 2000 . 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. In particular, the North Atlantic SST tripole (or the similar "horseshoe" pattern) 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 Franki

Building on this perspective, Kolstad and O'Reilly (2024) found A useful distinction emerging from this literature is that observational studies identify these SST–atmosphere linkages directly in the climate record, whereas model-based studies generally recover similar patterns only when the predictable component of variability is isolated. This reflects signal-to-noise issues (Scaife and Smith, 2018; Weisheimer et al., 2024): models can reproduce the relevant mechanisms, but the forced signal is often weaker than in observations, making the pathways more difficult to detect.

Kolstad and O'Reilly (2024) used reanalysis data to demonstrate 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 western part of the Subpolar Gyre region and baroclinicity in the storm-track entry storm track entrance 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 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)(e.g. MacKinnon et a . It is particularly well-suited to climate science applications where feedback loops and indirect mediated 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) (e.g. Granger, 1969; Mosedale et al., 2006; McGraw and Barnes, to more complex methods (e.g., Ebert-Uphoff and Deng, 2012; Hannart et al., 2016; Runge et al., 2019; Docquier et al., 2024) (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 for sensitivity experiments. 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 tointroduces mediation analysis to a broader climate science audience and illustrates its value as a diagnostic tool for ocean–atmosphere feedbacksexamines whether the strength of the November–to–winter SST–NAO linkage in a state-of-the-art seasonal forecast system is related to its skill in predicting the NAO, thereby motivating the subsequent mediation analysis. Second, it attempts to quantify and clarify quantifies and clarifies causal directionality in the relationships linking SSTs , surface between November SSTs and surface heat fluxes, baroclinicity, and the winter NAO. Third, it applies the mediation framework to a state-of-the-art the forecast

system to assess whether ~~systematic~~ biases in these ~~feedbacks~~ relationships can help explain its limited ~~skill in predicting the NAO~~NAO prediction skill.

It is important to emphasise that the mediation pathways examined here do not account for the full influence of November SSTs on the winter NAO, which can be viewed as the combined effect of all possible pathways operating throughout the climate system, both locally and remotely. The present analysis focuses on only a small subset of this much broader interaction network. It zooms in on surface heat fluxes and baroclinicity because there are good physical reasons to expect these mechanisms to participate in SST-induced adjustments of the North Atlantic circulation. In other words, the analysis should be interpreted as isolating two *informed* components of the total SST influence: it has the potential to reveal where these specific pathways reinforce or oppose the SST–NAO relationship, without implying that they represent the full climate system's response.

The following section gives an overview of mediation and partial-correlation analysis~~before describing~~ , before the data and ~~other methods~~ methods are described 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.:

$$X \to Y.$$

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 ~~aa~~ a gridded spatial field representing surface heat fluxes and ~~the Eady growth rate maximum (Hoskins and Valdes, 1990), a measure of~~ a metric for baroclinicity. These mediators are investigated separately through the pathway

$$X \to Z \to Y.$$

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

~~It~~ As mentioned in the Introduction, 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~~ total *effect* τ of the predictor X on the predictand Y:

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

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

~~The third~~ Here, τ represents the correlation between the standardised November SST index and the standardised winter NAO index. The second regression describes $X \to Z \to Y$ by accounting for the ~~mediator~~standardised mediator variables. The effect of X on Y changes to ~~τ_Z,~~ τ', known as the *direct* effect (not through the mediator), and the effect of Z on Y when accounting for X is denoted as β:

$$Y = \tau' X + \beta Z. \tag{2}$$

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

$$Z = \alpha X. \tag{3}$$

An important thing to note is that α encapsulates not just the direct forcing $X \to Z$, but also all the indirect forcing through intermediate variables, crucially including via the pathway $X \to Y \to Z$.

A central concept is the product $\alpha\beta$, known as the *indirect*~~*effect*~~ or *mediated* effect (of X on Y through Z). ~~It~~ The total effect is the sum of the direct and mediated effects: $\tau = \tau' + \alpha\beta$. This also follows from Eqs. (1–3)~~that $\alpha\beta = \tau - \tau_Z$.~~

~~Scaling~~ . Scaling the mediated effect by the total effect yields:

$$\frac{\alpha\beta}{\tau} = 1 - \frac{\tau'}{\tau}. \tag{4}$$

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~~ If $\tau' = 0$ (or is not significantly ~~non-zero~~different from zero), it ~~is plain~~ follows from Eq. ~~4 that the direct and indirect~~ 4 that the total and mediated effects are identical. ~~This means that~~ In this case, the pathway $X \to Y$ is fully accounted for by Z, ~~which again implies~~ indicating that $X \to Z \to Y$ ~~is oneof potentially many correct causal pathways. Partial mediation occurs when $0 < \tau_Z/\tau < 1$.~~ represents a valid causal pathway – though not necessarily the only one. Other mediators may also transmit part of the influence of X on Y.

~~Both perfect and partial mediation can represent~~

2.2 Accounting for reverse pathways

Mediation sometimes represents unambiguous forward-directed pathways in the sense that X changes Z and then Z affects Y. In climate dynamics~~though,~~ , however, feedback mechanisms are common~~. The pathway $X \to Z \to Y$ can be valid even~~

~~when the "reverse" pathway $X \to Y \to Z$ is equally valid. In situations where~~ , and ambiguites may arise because 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 validthan the other. What one can do~~contemporaneously. This means that both the original $X \to Z \to Y$ and the alternative $X \to Y \to Z$ pathways may both be valid. It is, however, ~~is to explore how X affects~~ possible to assess the degree to which Z ~~independently of Y through a partial correlation analysis.~~

2.3 ~~Partial correlation~~

~~The first step to compute the correlation between~~ responds directly to X ~~and Z while accounting for~~ rather than indirectly through Y. This can be done by regressing out Y ~~is to regress out Y from both X and Z~~:

$$Z = \alpha' X + \gamma Y. \tag{5}$$

~~where $\varepsilon_X = X - \gamma_X Y$~~ The new coefficient α' can be compared to the original α in Eq. 3.

It can be argued that it would be prudent to explicitly account for the NAO's autocorrelation to prevent it from confounding the results, even though Kolstad and O'Reilly (2024) showed that the ERA5 NAO autocorrelation was only significant from November to December and not from November to DJF; this was confirmed to be valid for the shorter period examined here for both ERA5 and ~~$\varepsilon_Z = Z - \gamma_Z Y$ are the residuals. The partial correlation is then given by the correlation between these residuals:~~ SEAS5. Labelling the NAO index in November as Y_0, a new regression equation could be defined as:

$$Z = \alpha'' X + \gamma_0 Y_0 + \gamma' Y.$$

~~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 regressed out is also investigated.~~ However, as the coefficient γ_0 was found to be negligible for both mediators, which was expected in light of the missing NAO autocorrelation, Eq. 5 is used in the analysis. The product $\alpha'\beta$ can be interpreted as the SST-forced component of the mediated effect. In other words, this is the influence that would arise through the $X \to Z \to Y$ pathway if the mediator responded only to direct SST forcing, while the NAO retained its full sensitivity to the mediator through β. For simplicity, $\alpha'\beta$ will be referred to as the *SST-forced mediated effect* henceforth. It is not intended as a replacement for the full mediated effect $\alpha\beta$, but as a complementary diagnostic that helps distinguish SST-forced mediation from mediation that is predominantly atmospheric in origin.

To further quantify this balance, it is useful to assess the ratio γ/β, which measures the relative strength of the $Z \to Y$ and $Y \to Z$ paths. If $\gamma/\beta > 1$, the NAO exerts a stronger influence on the mediator than vice versa; values below 1 indicate the opposite. When interpreted together with α' and $\alpha'\beta$, this ratio provides a compact measure for identifying where mediation reflects genuine SST forcing rather than mainly contemporaneous NAO–mediator covariance.

2.3 Suppression

An interesting special case occurs when ~~$\tau_Z/\tau > 1$~~ $\tau'/\tau > 1$, which means that the ~~indirect~~ mediated effect $\alpha\beta$ has the opposite sign to the total effect τ (Eq. 4). In these cases, Z is referred to as a ~~suppressor~~ *suppressor* because the regression coefficient linking X and Y is inflated when Z is accounted for (Muniz and MacKinnon, 2025). In the contect 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 $X \to Y$ 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 $X \to Y$ link is revealed. Put differently, had it not been for the negative feedback through Z, X would have exerted stronger predictive power on Y.

2.4 Sample coefficient notation

Throughout the paper, sample coefficients in Eqs. 1–5 (i.e. coefficients estimated through least-squares fitting) are denoted by carets; for instance, $\hat{\tau}$ is the estimated τ value.

3 Data and methods

3.1 Data

~~This study makes use of reanalysis~~ Reanalysis and seasonal forecast data are used. 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 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. The analysis covers the winters from 1981/82 to 2023/24 (hereafter referred to as 1981–2023).

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~~ but the new ocean model ~~yields~~ nevertheless appears to yield 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).~~ A well-documented feature of SEAS5 relevant for this study is a warm SST bias in the western North Atlantic up to the mid-1990s (Stockdale et al., 2018; Tietsche et al., 2 . Inherited from issues with the ocean reanalysis, this bias allowed SST errors to grow rapidly and produced a local warm

anomaly that affected near-surface temperature and surface heat fluxes. To verify that it did not affect the conclusions of this paper, the core analysis was repeated using only the period 2001–2023; the results were practically unchanged.

SEAS5 reforecasts were used from 1981 to 2016 with 25 ensemble members, and real-time forecasts from 2017 to 2023 with only the first 25 of 51 members used to ensure consistency with the reforecasts. The analysis was based on individual ensemble members (i.e. not ensemble means) unless otherwise specified.

SEAS5 forecasts are ~~produced~~ issued once per month~~; here, only the October initialisations~~. In this study, the November forecasts and reforecasts are used, corresponding by convention to lead times of ~~2–5~~ 1–4 months for November through February. ~~November initialisations were excluded because~~ A potential drawback of using the November initialisations is that 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~~are similar across ensemble members due to oceanic inertia. However, repeating the complete analysis with October initialisations (for which the November SST fields are more diverse) produced qualitatively identical results. I chose to base the analysis on the November runs, as this allows evaluation of the model's skill for the set of forecasts used operationally for predicting winter conditions.

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)~~and the sum of sensible and latent heat flux, hereafter referred to as ~~just "fluxes"~~surface heat flux, or SHF (positive upwards). Baroclinicity is quantified by the Eady growth rate maximum ~~(σ_E henceforth) parameter (e.g., Hoskins and Valdes, 199~~ (e.g. Hoskins and Valdes, 1990), defined for the 700–850 hPa layer as

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

where the unit is day^{-1}, $c = 86400 \times 0.3098$, f is the Coriolis parameter, $\mathbf{v}$ is the wind vector, z is the geopotential 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.

Anomalies were calculated by subtracting the overall mean and dividing by the overall standard deviation, spanning all years and ensemble members.

3.2 Climate indices

Two scalar indices are central to the analysis: the ~~winter (December–February, DJF)~~ DJF NAO index, and an SST-based index representing the November SST anomaly pattern in the extratropical 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°–80°N, 90°W–40°E, using the *eofs* Python package (Dawson, 2016) and applying ~~appropriate latitude-based~~ $\sqrt{\cos\phi}$ latitude 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~~It was a deliberate choice to use the ERA5 loading pattern for both datasets, as the purpose of this study is to assess how SEAS5 represents the real-world NAO pattern.

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~~20°N to 70°N and from 100°W to 20°E (Czaja and Frankignoul, 2002; Kolstad and O'Reilly , 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~~ 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~~by creating 10,000 randomised series through sampling with replacement. To ensure comparability between the two datasets, the bootstrap sample length was set equal to the number of years in the study period for both datasets (i.e. 43). This avoids giving SEAS5 an artificial advantage with respect to ERA5 due to its larger ensemble size (25 members per year). When assessing the significance of a metric (e.g. ~~,~~ a correlation) at a significance level of 5% (used throughout this study), the 2.5$^{\text{th}}$ and 97.5$^{\text{th}}$ percentiles of the correlation coefficient across those ~~1000~~ 10,000 randomised series were computed, and if the interval between these percentiles did not include zero, the correlation was deemed significant.

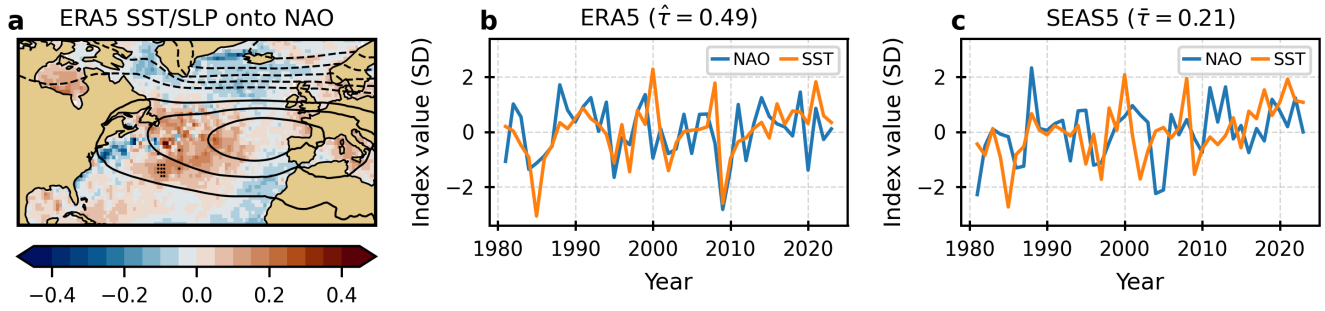

Figure 1. (a) Colours: November SST anomalies in ERA5 projected onto the ERA5 DJF NAO index. ~~Dots~~ The unit is K, and dots mark ~~anomalies~~ coefficients significantly different from zero at the 5% level. Contours: DJF SLP anomalies projected onto the same NAO index (unit: hPa). The contour interval is 1 hPa; solid (dashed) contours indicate positive (negative) ~~anomalies~~coefficients, 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 December month at the start of the winter ~~season~~ (i.e. ~~,~~DJF 1981/82 is labelled 1981). (c) As in (b), but for SEAS5, based on the ensemble mean each year.

4 Results

4.1 ~~Relationship between SSTs and the NAO~~SST–NAO relationship

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~~ display 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, both from ERA5 data. Although only a few of the local SST ~~anomalies~~ coefficients in Figure 1a are significant, the sample correlation between the two indices is relatively high (~~$\hat{\tau} = 0.60$~~$\hat{\tau} = 0.49$, $p \approx 0.001$), 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. It is emphasised that $\hat{\tau}$ does not represent a skill score, as no independent training and evaluation periods were defined.

~~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.~~

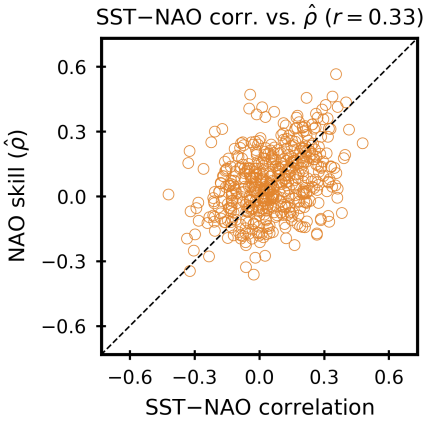

Figure 2. Results for a 10,000-member bootstrap ensemble of SEAS5 series, each of length 43. For each bootstrap sample, the x-axis shows the SEAS5-internal correlation between the November SST index and the DJF NAO index, and the y-axis shows the NAO skill $\hat{\rho}$, defined as the correlation between the DJF NAO indices in SEAS5 and ERA5 (with ERA5 years matched to the bootstrap sample). To enhance readability, only 1000 randomly chosen points are shown.

Turning to the (ensemble mean) 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 ($\hat{\tau} = 0.31$~~it is evident that the SST index covaries with the SST index in ERA5 ($r = 0.91$). However, several differences between the datasets are also apparent. Most important, the SST–NAO correlation is substantially lower ($\bar{\tau} = 0.21$, $p = 0.18$, with the overbar signifying that the ensemble mean was used) than in ERA5~~($\hat{\tau} = 0.60$)~~, demonstrating a discrepancy in the linkages between the observed "NAO-optimal" SST pattern and the winter NAO. Here it is acknowledged that the SST–NAO correlation is weaker than in ERA5 at least partly because both indices were deliberately derived from ERA5-based spatial patterns.

A second point is the ~~weak correspondence~~ non-significant NAO skill: the anomaly correlation coefficient between the NAO ~~indices~~ index in SEAS5 and ERA5~~: the correlation is not statistically significant~~ , denoted henceforth as ρ, is not significant for the ensemble mean ($\bar{\rho} = 0.29$, $p = 0.06$). This low predictive power is consistent with Baker et al. (2024).

When the ensemble members are examined individually, both the SST–NAO correlation and the NAO skill weaken, revealing higher internal noise. This behaviour is consistent with the "signal-to-noise paradox" (e.g. Scaife and Smith, 2018). The sample parameter used in the remainder of the analysis is $\hat{\tau} = 0.06$, which is much lower than $\bar{\tau} = 0.21$. The member-level NAO skill also decreases from $\bar{\rho} = 0.29$ to $\hat{\rho} = 0.07$. Neither $\hat{\tau}$ nor $\hat{\rho}$ is significant.

4.2 Linking the SST–NAO relationship to NAO prediction skill

One of the three main purposes of this paper is to assess whether the SST–NAO relationship in the model is proportional to its NAO skill. Otherwise it would be of limited interest to scrutinise the causal pathways through which the SSTs influence the NAO. To investigate this association, bootstrapping was used to generate an ensemble of 10,~~reflecting the absence of predictive~~

skill in the October-initialised 000 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, series, each with the same length as the number of years in the atmospheric circulation study period (43). The results are not sensitive to this choice of length, and the same bootstrap ensemble is analysed further in Section 4.6.

For each bootstrap sample, two quantities were computed: (1) $\hat{\rho}$, the correlation between the DJF NAO index in SEAS5 would likely have closely resembled that in and the corresponding ERA5 , making it difficult to separate genuine model behaviour from externally constrained predictability. The absence of skill ensures that the values, with ERA5 years selected to match the year of each randomly drawn SEAS5 member; and (2) the SEAS5-internal correlation between the November 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 biasesDJF NAO index.

Figure 2 shows a scatterplot of these parameters for a subset of the samples. The correlation across all the 10,000 samples is positive ($r = 0.33$) and significant at the 5% level. This does not imply that a strong SST–NAO relationship is sufficient or strictly required for high NAO skill, since the NAO is influenced by many processes unrelated to North Atlantic SST variability. Rather, the result provides support for a central premise of this study: the extent to which the model reproduces the observed influence of November SST anomalies on the winter NAO contributes meaningfully to its overall NAO skill.

4.3 ERA5 climatology and SEAS5 bias

Before examining the role of surface fluxes SHF and baroclinicity in mediating the SST–NAO relationship, it is useful to consider the climatological context. In Figure 3, ERA5 climatologies and SEAS5 biases are therefore shown, starting with the mean November SSTs in the North Atlantic in FigFigure 3a. 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 SHF on the warm side of the front (Figure 3b), as well as and they also coincide with strong low-level baroclinicity (Figure 3c). The last panel in the top row, Figure 3d, shows the climatological SLP pattern, consisting of the familiar which is characterised by a dipole between the Icelandic Low and a high west of Gibraltarthe Azores High.

It is striking how poorly The aforementioned warm SST bias in the western North Atlantic (Stockdale et al., 2018; Tietsche et al., 2020) is visible as a tongue-like feature in the east–west direction south of Greenland in Figure 3e. This is also linked to a clearly defined positive DJF SHF bias in Figure 3f. The poor SEAS5 represents representation of the SST gradient along the Gulf Stream seen in Figure 3a . The map in is also of interest. Figure 3e 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 The SHF biases in Figure 3f mirror reflect these SST errors, generally 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 thereAlthough not shown here, these SHF

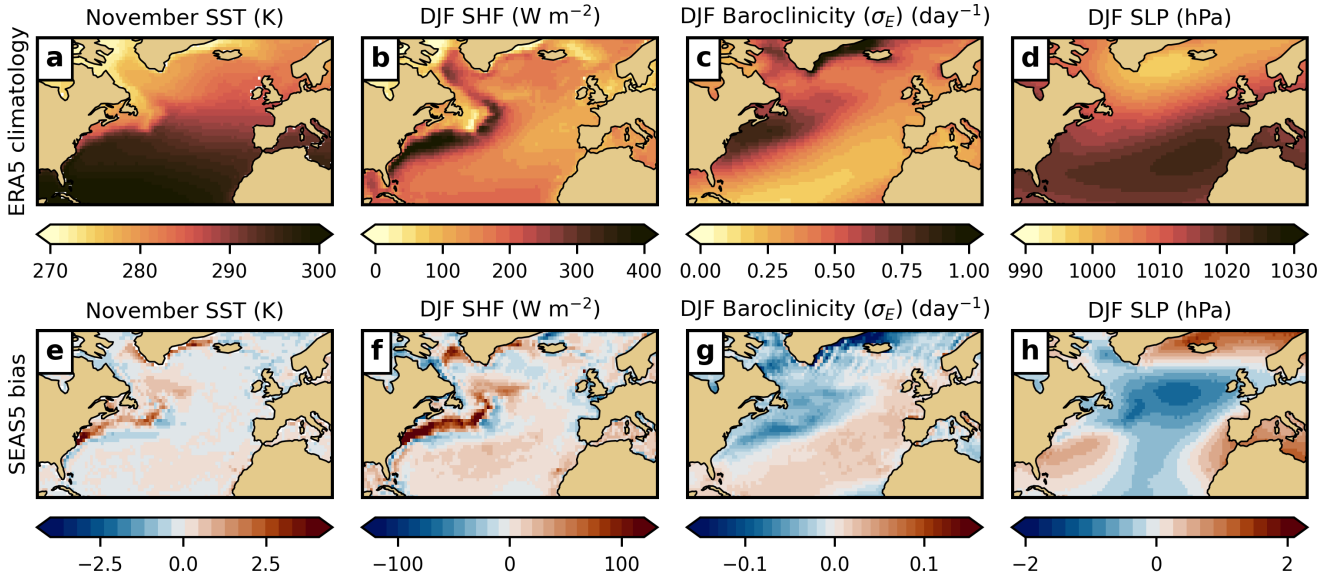

Figure 3. Top row: ERA5 climatologies for (a) November SST (K); (b) DJF surface heat fluxes (W m^{-2}); ~~and~~ (c) DJF Eady growth rate maximum σ_E (day^{-1})~~. Bottom row: SEAS5 biases~~ ; (~~SEAS5 minus ERA5) for~~ (d) ~~November SST (K); (e)~~ DJF ~~surface heat fluxes~~ SLP (~~W m^{-2}~~hPa)~~; and~~ . The bottom row (~~f~~e–h) ~~DJF σ_E~~ show the SEAS5 biases (~~day^{-1}~~SEAS5 minus ERA5) for the same variables as in the top row.

biases project strongly onto the DJF SST bias in the western part of the basin; these are larger in magnitude than the ones for November in Figure 3e, but for the most part they have the same sign, suggesting a growth of the model's SST bias with lead time. Figure 3g ~~further~~ indicates that the underestimated SST gradients along the Gulf Stream are associated with too weak baroclinicity in the ~~storm-track entry~~ storm track entrance region, while σ_E is too high to the south.

~~Taken together, these~~ 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 mean SLP bias pattern in Figure 3h, which shows that SEAS5 ~~fails to reproduce the observed SLP dipoleand exhibits too weak westerly flow between the~~ underestimates the amplitude of the observed NAO-like dipole. Sampled at representative grid points near the two NAO centres of action ~~. These weakened westerlies are~~ (Stykkishólmur, Iceland, and Ponta Delgada, Azores), the mean SLP bias amounts to $+1.0$ hPa and -0.4 hPa, respectively, giving a bias in the north–south difference of $+1.4$ hPa. This confirms that the model's climatological pressure contrast is weaker than observed, implying westerlies that are too weak across the subpolar North Atlantic. The most distinct weakening of the westerlies occurs between the mid-basin negative SLP bias and the positive bias near Iceland. This is commensurate with the negative ~~flux~~ SHF bias in the ~~Gyre~~ region with suppressed westerlies (Figure 3f)~~;~~ , where the fluxes are probably too weak ~~because there is~~ due to less intense cold advection from the west.

4.4 Mediated effects

4.5 ~~Indirect effects, preconditioning, and feedbacks~~

4.4.1 ~~Surface heat fluxes~~

~~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) simulateneous 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.~~

~~Figure 4 a 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~~ Figures 4 and 5 show the sample parameters $\hat{\alpha}$ and $\hat{\beta}$, as well as their product $\hat{\alpha}\hat{\beta}$, for both mediators. The unit is standard devations (SD), as all the variables in the regression equations were standardised prior to estimating the coefficients. It is repeated for emphasis that α, the regression coefficient linking November 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~~ mediator ($X \to Z$) in Eq. 3, captures *all* routes through which SST anomalies influence Z. This includes the indirect effect via the NAO (i.e. the pathway $X \to Y \to Z$~~(where X is the SST index, Y the NAO index, and)~~, as well as other pathways not explicitly considered here. The NAO-independent contribution of SSTs to 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 out from both ($r_{X,Z|Y}$~~, denoted α' in Eq. ~~??)~~ 5, is examined in Section 4.5.

4.4.1 Surface heat fluxes

From Figure 4~~b. The resulting pattern bears some resemblance to the SST anomaly structure used to define the~~ a, it emerges that the November SST index yields positive SHF coefficients in large parts of the Subpolar Gyre in ERA5. These positive $\hat{\alpha}$ values largely coincide with positive $\hat{\beta}$ values (Figure ~~1a~~ 4b). The product $\hat{\alpha}\hat{\beta}$ therefore yields a pronounced mediated effect in the reference region (Figure 4c), consistent with ~~the SST anomalies governing the sign of the heat fluxes . However, in~~ Kolstad and O'Reilly (2024). This implies that heat fluxes in this area play an important role in mediating the effect of November SSTs on the winter NAO.

Limited suppression occurs 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~~mid-basin

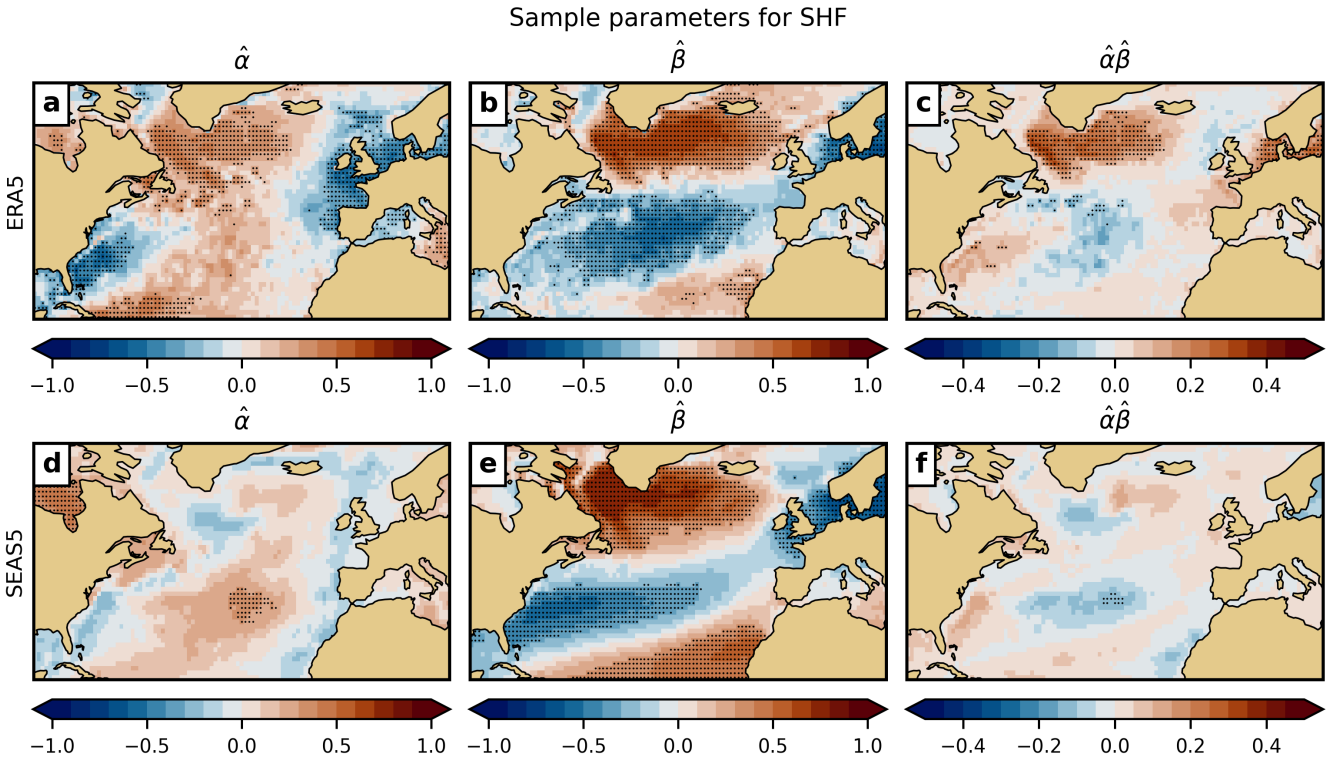

Figure 4. The top row shows sample parameters for ERA5 for the surface heat flux pathway: (a) $\hat{\alpha}$; (b) $\hat{\beta}$; (c) the mediated effect ($\hat{\alpha}\hat{\beta}$). Panels (d)–(f) show the corresponding parameters for SEAS5. Dots indicate where the parameters differ significantly from zero at the 5% level. Unit in all panels: standard deviations (SD).

area, suggesting a negative feedback mechanism. This happens because $\hat{\alpha}$ and $\hat{\beta}$ have opposing signs; in other words, the SST index generates a flux response that counteracts the contemporaneous NAO–SHF relationship.

The white contours in Figure 4b, 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 d reveals that SEAS5 yields barely any significant $\hat{\alpha}$ values. Although there is a fair degree of spatial correspondence with the findings for ERA5, 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. uniformly positive structure inside the reference region seen in the reanalysis is lacking. Instead, $\hat{\alpha}$ is negative values in an area south of Greenland; this partly overlaps with the positive SST and SHF biases identified in Figure 3e,f. The spatial match is not exact, however, and there is no obvious mechanistic link between the bias and the sign reversal. Regardless of its origin, the model produces the wrong sign of the SST–flux relationship in a dynamically important region, in marked contrast to ERA5.

Figure 4c 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 In comparison, the pattern of $\hat{\beta}$ (Figure 4b thus likely helps to set the stage for NAO growth, while Figure 4c 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 e) does resemble the one in ERA5, demonstrating that SEAS5 has a strong and mainly correct contemporaneous SHF–NAO relationship. However, the mediated effect $\hat{\alpha}\hat{\beta}$ in SEAS5 shown in Figure 4a as a genuine two-way air–sea feedback, rather than an artefact of the reverse causal pathway. f diverges from ERA5, with no significant mediation in the Subpolar Gyre region. In light of the strong $\hat{\beta}$ pattern, this suggests that the SHF-related part of this weak SST–NAO correlation is due to the inadequate $\hat{\alpha}$ representation. This is discussed further in Section 4.5.

Turning to SEAS5, the indirect effect of SSTs on the NAO via fluxes is shown in Figure 4d. Its spatial pattern differs substantially from that in Lastly, it is noteworthy that SEAS5 exhibits suppression in the same mid-basin domain as 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 4e). 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 3f, 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

4.4.2 Baroclinicity

The top row of Figure 5 (panels a–c) shows a distinctly positive mediated effect in ERA5 in the western storm track entrance region, in a wide corridor further south, and near Iceland. In all these areas, the sign of $\hat{\alpha}$ and $\hat{\beta}$ is the same, and the spatial structures of these parameters are similar. This likeness could indicate that the mediated effect is mainly due to the effect of the NAO on the baroclinicity (i.e. $Y \to Z$). In that case, $\widehat{\alpha'}$ in Eq. 5 is expected to be near-zero; this is explored in the next section. Even so, feedbacks between the NAO and baroclinicity (i.e., eddy-mean flow feedbacks) still appears to be an important mechanism for maintaining the NAO.

The picture for SEAS5 than ERA5 (Figure 4f vs. 4c) . This muted sensitivity likely reflects the positive mean flux bias in that region (Figure 3f) , itself linked to a positive SST bias (Figure 3e), which may dampen the response of flux anomalies to NAO-related circulation changes.

Beyond the Subpolar Gyre, widespread suppression is evident 5d–f) is similar to ERA5 in the sense that the signs of $\hat{\alpha}$ and $\hat{\beta}$ overlap in two bands 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. . However, the magnitude of $\hat{\beta}$ is distinctly larger than that of $\hat{\alpha}$; clearly the magnitude of the muted mediated effect in panel (f) is dictated by $\hat{\alpha}$. Neither $\hat{\alpha}$ nor $\hat{\alpha}\hat{\beta}$ is significant anywhere.

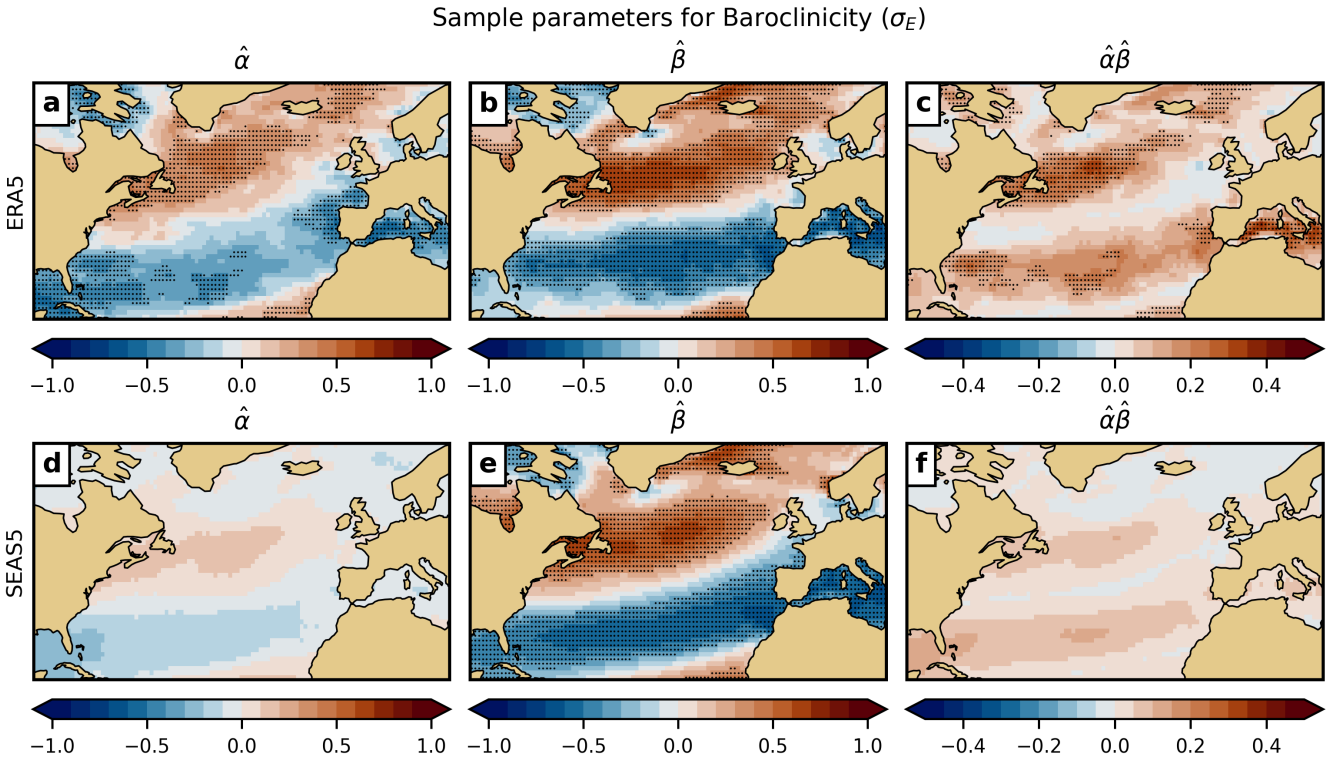

Figure 5. As Figure 4, but for the baroclinicity parameter σ_E.

4.5 Disentangling forcing and feedback

4.5.1 ~~Baroclinicity~~

~~Figure 5a shows the indirect effect of~~ The findings in the previous section raised questions about the directionality of the mediated effects associated with both SHF and baroclinicity. Although the pathway $X \rightarrow Z \rightarrow Y$ is not meant to be interpreted as strictly unidirectional, it is useful to assess the extent to which November SST anomalies ~~on the winter NAOmediated 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).~~ generate responses independently of the NAO. To that end, the leftmost panels in Figure 6 show the sample parameter $\widehat{\alpha'}$ from Eq. 5, which isolates the $X \rightarrow Z$ influence with the NAO regressed out, for SST and SLP. These variables, which are not considered as mediators, are analysed here because they give an indication of changes in the lower boundary (SST) and circulation (SLP).

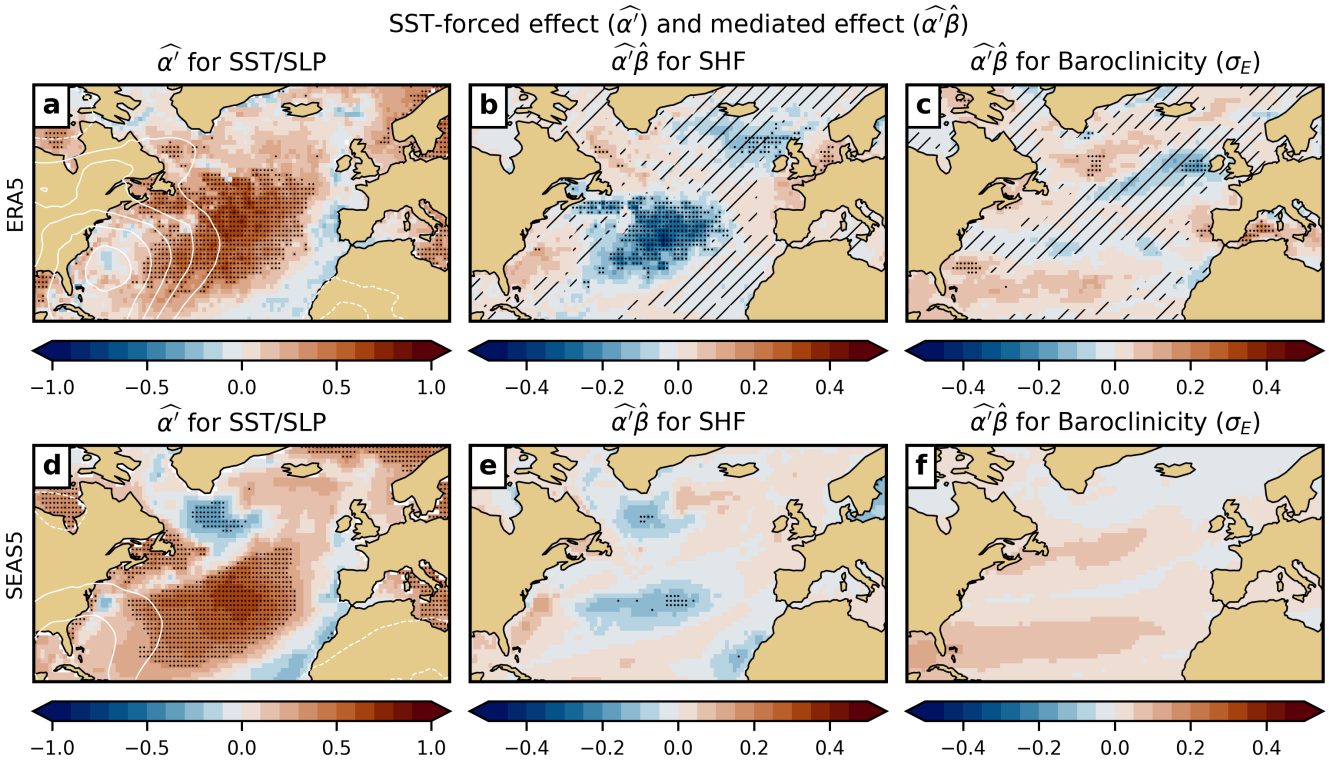

Figure 6. ~~As Figure~~The left column shows the SST-forced effect ($\widehat{\alpha'}$; Eq. ~~4~~5), ~~but~~ on SST (colours) and SLP (contours with interval 0.1 SD; positive solid, negative dashed; zero omitted) in ERA5 (a) and SEAS5 (d). Dots indicates where $\widehat{\alpha'}$ for SST is significantly different from zero at the ~~Eady growth rate maximum~~ 5% level. The remaining panels show the SST-forced mediated effect ($\sigma_E\widehat{\alpha'\beta}$) on ERA5 SHF (b), ERA5 baroclinicity (c), SEAS5 SHF (e), and SEAS5 baroclinicity, all in DJF. Dots denote where $\widehat{\alpha'\beta}$ is significantly different from zero, and hatched areas indicate where $\hat{\gamma}/\hat{\beta}$ (Eqs. 2 and 5) significantly exceeds 1, both at the 5% level.

~~Together, Figures 5b and 5c help resolve potential ambiguity about causality and provide physical context for the indirect effect shown in Figure 5a . Figure 5b demonstrates thatlate-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~~ Starting with ERA5, Figure 6a shows that, once the NAO contribution is removed, November SST anomalies induce an SLP pattern dominated by positive coefficients over the south-western North Atlantic. This pattern implies anomalous northerly advection in positive phases and southerly advection in negative phases. The associated SST response resembles the November antecedent in Figure 1a (as expected from oceanic inertia), but there is more significant mid-basin dominance with positive values.

In SEAS5, 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 5c verifies that once the NAO is established, σ_E co-varies with it in a manner that reinforces the anomaly in~~ $\widehat{\alpha'}$ field in Figure 6d shows an SST structure broadly similar to the ~~storm-track entry region and suppresses it further south, both consistent with a strengthening of the NAO itself~~ERA5 pattern in panel (a), with one notable exception: significant negative values appear south of Greenland. A similar sign discrepancy was already seen for the SHF $\hat{\alpha}$ coefficient (Figure 4d), indicating that the SST–flux response in this region is systematically misrepresented in SEAS5. These negative $\widehat{\alpha'}$ values lie near the well-documented positive SST bias during the early reforecast period (Stockdale et al., 2018; Tietsche et al., 2020), and when the analysis is repeated for 2001–2023, when this bias was much smaller, the negative values largely disappear. This suggests that the sign error may be linked to compensating adjustments associated with the bias, although the spatial correspondence is not exact and the mechanism cannot be established here. However, this issue is not central to the present study – while the bias alters some spatial details in SEAS5, the mediated effect in SEAS5 is not significant in this region in either period, and the skill–mediation covariability discussed in Section 4.6 is unaffected.

~~In SEAS5, however, the contrast to~~ The SST-forced mediated effect $\widehat{\alpha'}\hat{\beta}$ via surface heat fluxes in ERA5 is ~~unmistakable. In the entry region, Figure 5d exhibits only a shadow of the positive indirect effect seen for ERA5. Judging from Figure 5e, it is not clear why this should be the case, as the partial correlation between SSTs and σ_E is significant and positive in~~ shown in Figure 6b. Hatching indicates areas where $\hat{\gamma}/\hat{\beta}$ (Eqs. 2 and 5) is significantly greater than 1, implying that the SHF response is dominated by the $Y \to Z$ pathway rather than by the SST-forced $X \to Z$ pathway. In non-hatched areas, the dominant direction is either ambiguous or primarily reflects $X \to Z$.

Over parts of the Subpolar Gyre~~region. This in itself should strongly precondition the NAO in its northern centre of action, but Figure 5f makes clear that the feedback between the NAO and σ_E is weak and non-significant. This points to a key deficiency in~~ , where β is positive (Figure 4b), $\widehat{\alpha'}\hat{\beta}$ is partly positive or neutral, and some hatching appears. This indicates that the strong total mediation seen in Figure 4c is largely attributable to the $Y \to Z$ pathway; that is, NAO feedbacks on the fluxes dominate in this region. Further south in the North Atlantic, $\widehat{\alpha'}\hat{\beta}$ is significantly negative over a broad area, demonstrating that the SST-forced component of the mediated effect acts to suppress the SST–NAO correlation. A similar but less extensive pattern appeared in the full mediated effect $\hat{\alpha}\hat{\beta}$ in Figure 4c. Accordingly, these results imply that the suppression arises mainly from the NAO-independent SST-forced influence on the fluxes, whereas NAO feedbacks on the fluxes partly counteract this suppression.

In SEAS5 ~~: the feedback that should sustain baroclinicity once the NAO isestablished is too weak, limiting the overall strength of the mediation pathway.~~ (Figure 6e), the spatial structure of $\widehat{\alpha'}\hat{\beta}$ resembles the pattern of the full mediated effect $\hat{\alpha}\hat{\beta}$ in Figure 4c. The absence of hatching proves that the pattern is shaped almost entirely by the SST-forced component rather than by NAO feedbacks. This is consistent with the weak SST–NAO correlation in SEAS5: because November SST anomalies exert little influence on the NAO, there is correspondingly little NAO-driven response in the fluxes.

For baroclinicity, Figure 6c shows that the SST-forced mediated effect is mainly positive in the two regions where the total mediated effect $\hat{\alpha}\hat{\beta}$ is positive and significant (Figure 5c). Although $\widehat{\alpha'}\hat{\beta}$ is only significant in small parts of these areas, the

absence of extensive hatching indicates that $Y \to Z$ is not a clearly dominant pathway; rather, a two-way feedback appears to be operating between baroclinicity and the NAO. SEAS5 similarly produces weak, non-significant but positive $\widehat{\alpha'}\hat{\beta}$ in these two regions. As for SHF, there is no hatching, signifying an SST-driven signal.

In summary, ERA5 exhibits an approximate balance between forcing and feedback for both mediators: SST anomalies exert a demonstrable influence on fluxes and baroclinicity independently of the NAO, while the NAO also feeds back on both variables. SEAS5, in contrast, shows no evidence of NAO–to–mediator feedbacks, consistent with its weak SST–NAO relationship.

These differences motivate the next question: to what extent does the model's essentially one-way representation of these pathways affect its ability to predict the winter NAO? The following section addresses this.

4.6 Relating mediated effects to NAO prediction skill

4.7 ~~Relating feedbacks to skill~~

In Section 4.2, a modest but significant association was shown between the November–DJF SST–NAO correlation and the model's NAO skill $\hat{\rho}$ ($r = 0.33$). This suggests that the mediated effect associated with the SST–NAO linkage may also relate to forecast skill. Although the mediation signal in SEAS5 is weak overall, it is not absent: the positive $\hat{\alpha}\hat{\beta}$ values for baroclinicity (Figure 5f) broadly overlap with those in ERA5 (Figure 5c). For SHF (Figure 5c,f), there is likewise some agreement, apart from the negative values south of Greenland noted earlier.

This section examines whether variations in mediation strength across subsets of SEAS5 realisations are associated with variations in NAO skill. To do so, the 10,000-member bootstrap ensemble introduced in Section 4.2 is revisited. For each bootstrap sample, the SEAS5 mediated effect $\hat{\alpha}\hat{\beta}$ is estimated separately for SHF and σ_E, alongside the NAO skill $\hat{\rho}$ and the model-internal SST–NAO correlation from Section 4.2.

~~This section examines whether the strength of the indirect effect~~ The maps in Figure 7 show where, geographically, the mediated 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~~ co-varies with the NAO skill $\hat{\rho}$ across the SEAS5 ~~. Figures 4a and 5a demonstrated that the indirect effect in~~ bootstrap samples. The most prominent feature is that the strongest positive correlations occur in the regions where ERA5 ~~is substantial within the outlined reference domains. Area-averaged values of~~ exhibits robust positive mediation. For SHF (Figure 7a), this positive covariability appears across the Subpolar Gyre, even south of Greenland, where the overall $\hat{\alpha}\hat{\beta}$ ~~within these boundaries amount to 0.34 for fluxesand 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 4d and 5d that mediation through these feedback pathways is marginal~~ in SEAS5 ~~.~~

~~Figure 2 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~~ is negative (Figure 4f). This indicates that, within SEAS5, bootstrap subsets in which the model produces a mediation pattern more closely resembling ERA5 are also the subsets with higher NAO skill. Conversely, samples that yield negative $\hat{\alpha}\hat{\beta}$ in these regions tend to have lower skill. Thus, even though SEAS5 does not reproduce the magnitude or sign of the ~~same length as the 1981–2023~~

(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.

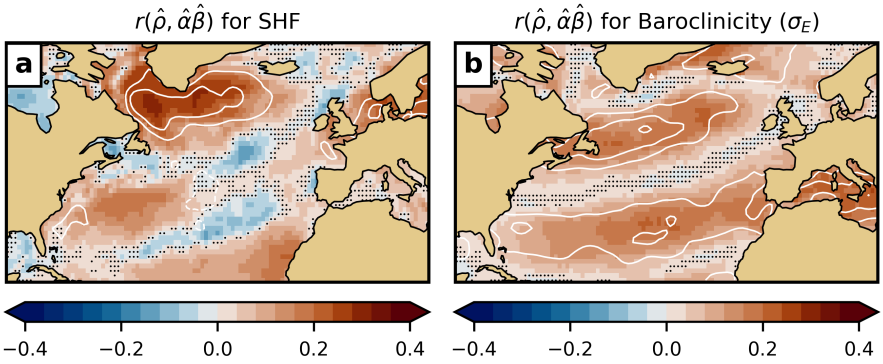

Figure 7. Results for the 10,000 bootstrap samples introduced in Section 4.2. Shading shows the correlation between $\hat{\rho}$ and the mediated effect $\hat{\alpha}\hat{\beta}$, computed from SEAS5 for each bootstrap sample, for (a) surface heat fluxes (SHF) and (b) baroclinicity. Dots indicate correlations that do not significantly differ from zero at the 5% level. White contours reproduce the ERA5 mediated effect $\hat{\alpha}\hat{\beta}$ from Figures 4c and 5c in panels (a) and (b), respectively (contour interval 0.1 SD; positive solid, negative dashed; zero omitted).

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 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. mediated effect perfectly, its internal covariability shows that more realistic mediation pathways are associated with improved NAO prediction skill.

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 2a, the A few regions also display negative correlations between $\hat{\rho}$ and $\hat{\alpha}\hat{\beta}$, but these do not overlap with the key regions where ERA5 point cloud shows a modest diagonal tilt, and exhibits strong SHF-mediated effects in the Subpolar Gyre. For baroclinicity, 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 correspondence between $\hat{\rho}$ and $\hat{\alpha}\hat{\beta}$) is more

uniformly related to the mediated effect in 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~~ (Figure 7b).

Taken together, these patterns reinforce the main conclusion: the clearest and most physically interpretable skill–mediation covariability occurs in the regions where ERA5 displays robust positive mediation. In these areas, 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 between $\hat{\tau}$ 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.~~ achieves higher NAO skill when it incidentally reproduces the observed mediation pathways, underscoring the importance of representing these air–sea feedbacks realistically in seasonal prediction systems.

~~The contrast is even sharper for the baroclinicitypathway (Figure 2b). 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 $r = 0.13$, 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~~

It deserves emphasis that this is not a trivial outcome that would be expected by construction: $\hat{\rho}$ depends only on agreement between the SEAS5 ~~compared to reanalysis. This weakness is particularly striking given that the baroclinicity pathway is the dominant mediator of the SST–NAO link in~~ and ERA5 ~~, explaining a larger fraction of the total effect than the flux pathway. Its near absence~~ NAO indices and contains no explicit information about X, Z, or any feedback pathway. Likewise, $\hat{\alpha}\hat{\beta}$ in SEAS5 ~~therefore points to a fundamental deficiency in~~ depends solely on the model's ~~representation of~~ internal covariance structure among X, Z, and Y. Further, it is independent of ERA5 except for the use of ERA5-based spatial patterns to define the indices. That these two independent quantities co-vary in physically meaningful regions provides strong evidence that the ~~dynamical processes that sustain the storm track and couple it to oceanic anomalies and the NAO~~ pathways identified from ERA5 correspond to mechanisms that matter for NAO predictability in the model.

5 ~~Discussion~~

5 Summary and 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 to~~ The results 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.

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.

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 rebust.~~ Future work could extend this examination to other models, some of which exhibit higher NAO skill than SEAS5 (Baker et al., 2024), or indeed multi-model ensembles~~and incorporate nonlinear mediation~~ , incorporating nonlinear mediation analysis techniques.

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~~ Kolstad and O'Reilly (2024), where processes linking late-autumn SST anomalies and the winter NAO were analysed based on ERA5 data, by investigating these processes in the forecast system SEAS5. It was hypothesised that the SST–NAO relationship has bearing on the NAO prediction skill in that model, and this was confirmed. Although the observed correlation ($r = 0.33$) is modest, it still represents a non-trivial association in light of the many other processes that influence NAO skill, including the stratosphere, tropical SST variability, Arctic sea-ice anomalies, and internal atmospheric dynamics, to mention but a few.

Having established this link between the SST–NAO relationship and NAO skill, the analysis extended Kolstad and O'Reilly (2024) further by revealing a physically coherent sequence of processes underpinning the SST–NAO connection. Independently of the NAO~~itself~~, November SST anomalies induce a ~~surface-pressure dipole pattern resembling a clockwise-rotated NAO. This~~ surface-pressure pattern that preconditions the atmosphere for anomalies in ~~surface fluxes~~ two mediators: surface heat flux (SHF) and baroclinicity in the ~~Subpolar Gyre, which~~ western North Atlantic. These anomalies in turn nudge the NAO~~. Once established, the NAO~~ , which subsequently feeds back on both ~~fluxes and baroclinicity, reinforcing the initial anomalies and sustaining the circulation pattern~~SHF and baroclinicity. A key outcome concerns the directionality of these causal pathways. Across large parts of the North Atlantic, forcing from the NAO onto the mediators was found to dominate. Crucially, however, in the regions with the strongest mediated effects through SHF and baroclinicity, the directionality was ambiguous, consistent with the existence of a two-way feedback mechanism.

It is important to emphasise that these feedbacks do not account for all aspects of NAO variability. ~~Far from it; the~~ The processes identified here represent ~~only~~ one pathway among many, complementing the aforementioned 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~~

shape NAO variability and how its misrepresentation in a prediction system may limit its ability to capture the full ~~spectrum~~ range of NAO behaviour.

~~For a~~ A key finding is that these pathways are ~~considerably weakened or absent~~ substantially weakened 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~~ This is likely linked to the muted SST–NAO relationship in the model relative to ERA5. When this link is weak, 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~~total effect of SSTs on the mediators, labelled α herein, is also necessarily weak. Figure 7, where bootstrap resampling was used to explore the relationship between NAO prediction skill and the mediated SST–NAO effect, illustrates this succinctly. Model samples that exhibit a stronger mediated effect also show higher NAO prediction skill. Conversely, samples that by chance yield higher NAO skill also display a stronger mediated effect. This mutual dependence suggests that if the model were able to reproduce the SST–NAO pathways via SHF and baroclinicity more realistically, it would likely predict the NAO more accurately as well.

~~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~~ However, this counterfactual hypothesis cannot be tested directly because the model does not currently reproduce these pathways. Targeted experiments that enforce more realistic air–sea interactions could help clarify whether strengthening these pathways would indeed improve NAO prediction skill. The study by Roberts et al. (2021) provides a prime example of such experiments. Other improvements, such as enhancing the resolution of the ocean (e.g. Haarsma et al., 2019) or the atmosphere (e.g. Czaja et al., 2019; Wills et al., 2024), or 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).~~ (e.g. Hardiman et al., 2022), have also been shown to produce more precise atmospheric responses to SST forcing.

The results presented here ~~is that they raise important~~ raise interesting questions for the emerging class of ML-based seasonal and subseasonal prediction systems ~~(e.g., Chen et al., 2024; Kent et al., 2025)~~(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 would be able to capture the same preconditioning and feedback structures as the real climate system ~~remains an open question~~is currently unknown.

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 pro-

cess studies and model development/evaluation, advancing our understanding of how ~~ocean–atmosphere coupling shapes~~ both unidirectional and feedback mechanisms shape 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~~ I wish to extend my gratitude to the editor, David Battisti, as well as two anonymous reviewers for improving the manuscript. I also thank the ECMWF for providing access to both ERA5 and SEAS5 data, and Timothy Stockdale at ECMWF for pointing me to the SST bias inherited from the ocean reanalysis. The figures use colour maps developed by Fabio Crameri (Crameri 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.

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.

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, e2024GL108 472, 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.

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.

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.

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.

Crameri, 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.

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, 2002.

Czaja, A., Frankignoul, C., Minobe, S., and Vannière, B.: Simulating the Midlatitude Atmospheric Circulation: What Might We Gain From High-Resolution Modeling of Air-Sea Interactions?, Current Climate Change Reports, 5, 390–406, https://doi.org/10.1007/s40641-019-00148-5, 2019.

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.

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, 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.

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-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 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., Hólm, E., Janisková, M., Keeley, S., Laloyaux, 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.

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.

Johnson, S. J., Stockdale, T. N., Ferranti, L., Balmaseda, M. A., Molteni, F., Magnusson, L., Tietsche, S., Decremer, 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 Ummenhofer, 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 Soloviov, 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 NEMO4-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, e2022GL101 480, 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, 20150 146, 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 Nicolì, 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, e2020GL0911 446, 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.

Stockdale, T., Balmaseda, M. A., Johnson, S., Ferranti, L., Molteni, F., Magnusson, L., Tietsche, S., Vitart, F., Decremer, D., Weisheimer, A., Roberts, C., Balsamo, G., Keeley, S., Mogensen, K., Zuo, H., Mayer, M., and Monge-Sanz, B.: SEAS5 and the Future Evolution of the Long-Range Forecast System, Tech. Rep. 835, ECMWF, https://doi.org/10.21957/z3e92di7y, 2018.

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, 014 024, 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.

Tietsche, S., Balmaseda, M., Zuo, H., Roberts, C., Mayer, M., and Ferranti, L.: The Importance of North Atlantic Ocean Transports for Seasonal Forecasts, Climate Dynamics, 55, 1995–2011, https://doi.org/10.1007/s00382-020-05364-6, 2020.

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.

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.

Wills, R. C. J., Herrington, A. R., Simpson, I. R., and Battisti, D. S.: Resolving Weather Fronts Increases the Large-Scale Circulation Response to Gulf Stream SST Anomalies in Variable-Resolution CESM2 Simulations, Journal of Advances in Modeling Earth Systems, 16, e2023MS004 123, https://doi.org/10.1029/2023MS004123, 2024.

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 Eddying Global Coupled Model, Climate Dynamics, 56, 2895–2913, https://doi.org/10.1007/s00382-020-05621-8, 2021.