UAB
Universitat Autònoma
de Barcelona

Faculty of Science

Atmospheric Resolution and the Signal-to-Noise Problem: Investigating Eddy Feedback and Teleconnection Mechanisms in North Atlantic Winter Predictions

Harry Wolimba Hall

Supervised by Dr. Pablo Ortega and PhD candidate Ángel Sánchez Lorente
from the Barcelona Supercomputing Centre

Thesis submitted as a requirement for the degree of
MSc in Modelling for Science & Engineering

Date: January 30, 2026

Abstract

The North Atlantic Oscillation (NAO) is a key driver of climate variability across Europe and eastern North America, yet its prediction at seasonal timescales remains challenging. Recent studies have identified a "signal-to-noise problem" in climate predictions where models predict the observed NAO better than their own ensemble members - particularly in the Atlantic sector. This research investigates whether high-resolution atmospheric only models, capable of resolving the effects of mesoscale eddies, provide improved skill in predicting the NAO compared to standard resolution models. We use outputs from the European Eddy-Rich Earth System Models (EERIE) project's experiments at TCo399 (∼27 km) and TCo1279 (∼9 km) resolutions to systematically assess differences in the representation of key teleconnection strength and eddy feedback parameters. Findings demonstrate that mid-tropospheric eddy feedback remains systematically deficient in both configurations, whilst upper-level processes show complex, level-dependent resolution sensitivity. We hypothesise that even higher resolution simulations may be needed to realistically resolve small-scale eddies and other processes, like convection, that enable the currently missing positive feedback onto larger-scale anomalous flows. Our methodology combines statistical analysis of model performance metrics with a mechanistic investigation of the eddy feedback parameter across resolution configurations. This research has significant implications for seasonal forecasting systems, particularly for addressing whether computational resources should prioritise increased model resolution or larger ensemble sizes.

Acknowledgements

Firstly, I would like to thank Pablo and Paco for responding to my interest in this area; I would not have had the experience with the Barcelona Supercomputing Centre otherwise.

I would like to thank Pablo again for his patient supervision of the process, his time every week and sharing his deep knowledge of the area.

Thanks Ángel for all the help throughout the process, from getting me set up at the start with the infrastructure and initial diagnostics to challenging questions that resulted in a more thorough work.

I'd also like to acknowledge Matthias Aengenheyster from the European Centre for Medium-range Weather Forecasting (ECMWF) for help sharing the simulation data.

Finally, thanks to everyone else at the BSC that has made it such an enjoyable and fulfilling process.

Contents

List of Figures

List of Tables

List of Abbreviations

AMIP	Atmospheric Model Intercomparison Project
ATOVS	Advanced TIROS Operational Vertical Sounder
BSC	Barcelona Supercomputing Centre
C3S	Copernicus Climate Change Service
CDS	Copernicus Climate Data Store
CERES EBAF	Clouds and the Earth's Radiant Energy System Energy Balanced and Filled
CMIP	Coupled Model Intercomparison Project
CMIP6	Coupled Model Intercomparison Project Phase 6
DestinE	Destination Earth
EA	East Atlantic
ECMWF	European Centre for Medium-range Weather Forecasting
EDA	Ensemble of Data Assimilations
EERIE	European Eddy-Rich Earth System Models
EFP	Eddy Feedback Parameter
EKE	Eddy Kinetic Energy
ENSO	El Niño Southern Oscillation
EOF	Empirical Orthogonal Function
ERA5	ECMWF Reanalysis Version 5
ESA CCI SST V3	European Space Agency Climate Change Initiative Sea Surface Temperature Version 3
ESMs	Earth System Models
EUMETSAT	European Organisation for the Exploitation of Meteorological Satellites
GPCP	Global Precipitation Climatology Project

GRIB2	General Regularly-Distributed Information in Binary Form Version 2
IMERG	Integrated Multi-satellitE Retrievals for Global Precipitation Measurement
IFS	Integrated Forecasting System
MIR	Meteorological Interpolation and Regridding
MSLP	Mean Sea Level Pressure
NAO	North Atlantic Oscillation
NetCDF	Network Common Data Form
NextGEMS	Next Generation Earth Modelling Systems
OSI-SAF	Ocean and Sea Ice Satellite Application Facility
OSTIA	Operational Sea Surface Temperature and Ice Analysis
PNA	Pacific North America
RPC	Ratio of Predictable Components
RPCA	Rotated Principal Component Analysis
SLP	Sea Level Pressure
SST	Sea Surface Temperature
TOA	Top Of Atmosphere
WP	West Pacific

Chapter 1

Introduction & Literature Review

This chapter provides an overview of the research and literature that addresses the predictability of the NAO, the signal-to-noise problem in Earth System Models, the representation of teleconnections, and the relevance of eddy feedback mechanisms. Through this overview, an argument for continued research into this area is sustained. The established research justifies the hypothesis that additional horizontal resolution has the potential to provide mechanistic improvements that address the signal-to-noise problem.

1.1 North Atlantic Oscillation

The NAO is the dominant mode of extratropical atmospheric variability in the Northern Hemisphere (Hurrell 1995), as it accounts for the largest fraction of variance in Sea Level Pressure (SLP) and geopotential height fields during the boreal winter (Patrizio et al. 2025). Physically, it is characterised by a redistribution of atmospheric air masses between the subtropical high-pressure region centred near the Azores, and the subpolar low-pressure region that extends across Iceland and the Nordic Seas (Hurrell et al. 2003). This large-scale meridional dipole in SLP anomalies dictates the speed and orientation of the prevailing westerly flow across the North Atlantic, as the circulation from each pressure centre accelerates the flow most when the difference is greatest (Figure 1.1). It thereby serves as the primary conductor of heat, moisture, and momentum transport from the Atlantic Ocean onto the Eurasian continent, impacting its climate (Hurrell 1995).

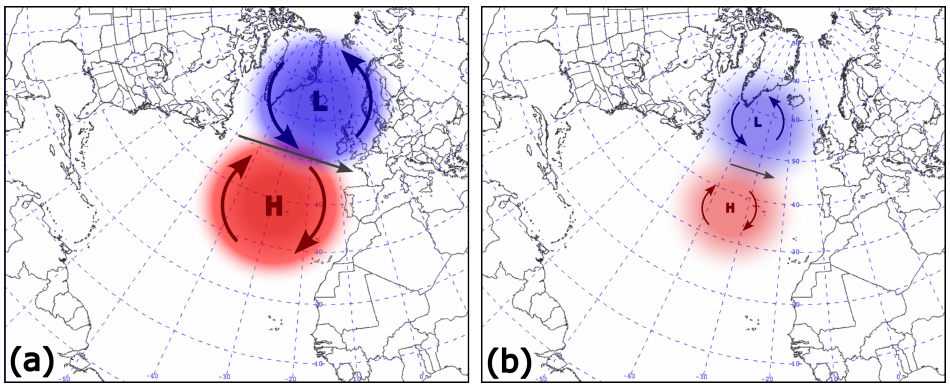

Figure 1.1: A simple visualisation of the NAO. The positive phase is when the difference between the pressure centres is greatest (a) and the negative phase least (b).

The scientific characterisation of the NAO has evolved significantly since its early identification. Foundational work by Wallace et al. (1981) identified the NAO as one of the fundamental "teleconnections" - standing oscillations in the planetary wave field - alongside the Pacific North America (PNA), West Pacific (WP), and East Atlantic (EA) patterns. While Wallace et al. (1981) utilised station-based cross-correlation matrices, subsequent work from Barnston et al. (1987) employing Rotated Principal Component Analysis (RPCA) confirmed the NAO's robustness and distinctness from other regional modes, for instance the East Atlantic pattern, previously considered a southward-shifted NAO (Mellado-Cano et al. 2018), which have different dynamical drivers and impacts (Barnston et al. 1987).Thompson et al. (1998) and Thompson et al. (2000) argued that the NAO is simply the regional manifestation of a hemispheric-scale, zonally symmetric mode of variability. However, the debate remains active; DelSole et al. (2009) and others have questioned whether the annular structure is a physical entity or a statistical artefact of EOF analysis, noting that the Atlantic and Pacific sectors often decorrelate on shorter timescales.

There are two primary methods for defining the NAO index, with different uses. The station-based index is calculated as the normalised SLP difference between stations such as Lisbon (Portugal) and Reykjavik (Iceland). This benefits from longer observational records - back to the mid-19th century - allowing for multidecadal trend analysis (Hurrell 1995). However, this index is fixed in space and can be sensitive to the migration of pressure centres of action (Hurrell et al. 2003). The PC-based index, derived from the leading Empirical Orthogonal Function (EOF) of the SLP field, captures the spatial variability of the pattern more dynamically but is constrained by the availability of gridded reanalysis data which cover a much shorter temporal span (Hurrell et al. 2003; Hannachi 1997).

The NAO is not static nor confined to the surface; it is a highly dynamic phenomenon driven by internal atmospheric processes. Feldstein (2003) analysed the daily evolution of the NAO, showing that it exhibits a life cycle of growth and decay on the order of two weeks. This life cycle is fuelled by high-frequency transient eddy fluxes — the "noise" of daily weather systems — which act to reinforce the low-frequency anomaly (Feldstein 2003; Limpasuvan et al. 2000). The vertical structure of the NAO is barotropic, meaning the anomalies maintain the same sign with height but increase in amplitude into the stratosphere (Thompson et al. 1998). This coupling allows for significant stratosphere-troposphere interaction. Limpasuvan et al. (2000) demonstrated that the annular modes are maintained by a positive feedback between the zonal mean flow and eddy momentum fluxes, a mechanism that helps explain the persistence of the NAO beyond the timescale of individual synoptic storms.

The societal relevance of the NAO is profound, as a main driver of the winter climate of the Euro-Atlantic sector. In its positive phase, the enhanced pressure gradient drives strong westerlies that steer mild, moist oceanic air over Northern Europe and Russia, resulting in warm and wet winters (Hurrell 1995). Conversely, the negative phase blocks this flow, allowing cold Arctic air to plunge southward, leading to severe winters in Central and Northern Europe and Eastern North America, with the opposite impacts occurring over the Mediterranean (Hurrell et al. 2003).

The strong impacts that the NAO exerts on the European climate in turn have significant environmental and societal consequences. Some examples are provided below:

- Trigo et al. (2004) documented that the negative NAO phase is crucial for replenishing

water resources in the Iberian Peninsula, with the majority of major river flow anomalies and filling of reservoirs linked to negative NAO winters.

- Bojariu et al. (2002) extended this to extreme events, showing that the NAO modulates the frequency of heavy precipitation and flood risks across Southern Europe.
- Thompson et al. (2001) and Hurrell et al. (2003) reviewed impacts ranging from crop yields and energy demand to the phenology of plant flowering.
- Hurrell et al. (2009) detailed how NAO-driven wind stress and heat flux anomalies alter sea surface temperatures (Sea Surface Temperature (SST)s) and mixed layer depths, driving variability in North Atlantic fish stocks and plankton distributions.
- Osborn et al. (2006) placed these modern impacts in a millennial context, using proxy data to show that the late 20th-century trend toward the positive NAO contributed significantly to the observed warming of the Northern Hemisphere landmasses, a trend that may be unprecedented in the last 1200 years.

The scale of these impacts underscores the importance of accurately predicting the NAO, as it affects critical sectors including water resources, extreme weather events, agriculture, and marine ecosystems, which all benefit from informed long-term planning (Bruno Soares et al. 2018).

Yet, despite its clearly demonstrated importance, the NAO has proved difficult to predict on seasonal timescales due to the chaotic dynamics of the extratropical atmosphere (Patrizio et al. 2025). Until relatively recently, the NAO was considered fundamentally unpredictable beyond timescales of atmospheric internal dynamics (10-14 days; Hurrell et al. 2003; O'Reilly et al. 2017). The dominant paradigm viewed NAO variations as arising primarily from stochastic processes internal to the atmosphere (Feldstein 2003; Hurrell et al. 2009).

However, recent advances in seasonal forecasting systems have demonstrated increased levels of skill over the North Atlantic, challenging the previously accepted inherent unpredictability (Scaife et al. 2014; Dunstone et al. 2016). This progress stems from improved understanding of the NAO's connections to slower-varying components of the climate system, particularly ocean circulation patterns and sea surface temperature anomalies, as well as stratospheric influences (Patrizio et al. 2025; Scaife et al. 2018).

Nevertheless, current prediction systems face significant challenges, most notably the "unrealistically weak signal-to-noise ratio of the NAO in prediction systems" (Patrizio et al. 2025; Scaife et al. 2018). This signal-to-noise problem, where models show higher correlation with observations than with their own ensemble members, suggests fundamental issues in how atmospheric variability is represented in current Earth System Models (ESMs). Understanding and addressing these model deficiencies requires examining how ESMs represent the key processes that govern NAO variability, including the representation of atmospheric dynamics, air-sea interactions, and the parametrisation of sub-grid scale processes.

1.2 Earth System Models

Understanding and addressing the limitations in NAO prediction skill outlined above requires examining how Earth System Models (ESMs) represent the key processes that govern NAO

variability. ESMs are the primary tools for climate prediction on seasonal-to-decadal timescales (Gelbrecht et al. 2023), built on coupled atmosphere-ocean-land-ice components (Eyring et al. 2016). However, as will be shown, systematic biases and representational deficiencies — particularly in atmospheric dynamics — compromise their ability to accurately simulate NAO behaviour and its predictability.

Horizontal and vertical resolution fundamentally affect model dynamics rather than providing mere cosmetic improvements (Scaife et al. 2019; Roberts et al. 2018b). The HighResMIP protocol (Roberts et al. 2018b) enables systematic assessment across modelling centres, with configurations spanning from standard Coupled Model Intercomparison Project (CMIP) resolutions (1°, ∼120km) to eddy-rich scales (<0.25°, ∼27km). Increasing oceanic resolution to resolve mesoscale processes significantly reduces the persistent cold SST bias in the central North Atlantic (Athanasiadis et al. 2022). This is not merely a local improvement; the cold SST bias affects low-level baroclinicity and diabatic heating patterns, which cascade into errors in atmospheric blocking frequency and jet stream variability (Athanasiadis et al. 2022). When oceanic resolution increases to eddy-permitting scales (∼0.25°), models show substantial improvements in Gulf Stream separation, SST frontal gradients, and air-sea interactions—all crucial for realistic NAO simulation (1.2) (Athanasiadis et al. 2022; Roberts et al. 2018b).

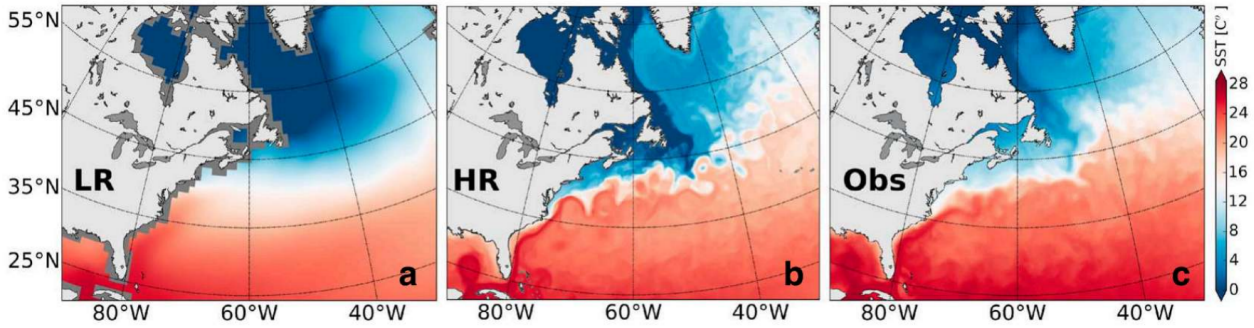

Figure 1.2: A comparison of sea surface temperature (°C) from low resolution (LR; i.e. 100km) model output on the left, high resolution (HR; i.e. <10 km) in the centre, and satellite observations (OBS) on the right. From (Siqueira et al. 2016).

Atmospheric resolution similarly impacts critical processes. Scaife et al. (2019) demonstrated that while doubling resolution from ∼0.8° to ∼0.3° showed limited NAO skill improvements, further increases to even higher resolution (0.14°) revealed significant improvements in eddy momentum flux convergence, approaching observed values. This eddy feedback — the positive interaction between transient eddies and low-frequency flow anomalies — is fundamental to NAO maintenance and persistence. The insufficient representation of this feedback in models has been linked to the signal-to-noise problem (Scaife et al. 2019; O'Reilly et al. 2025). Crucially, Scaife et al. (2019) found that even at doubled resolution, signal-to-noise ratios showed no improvement, suggesting resolution increases alone cannot resolve fundamental representational deficiencies without addressing underlying issues in eddy-mean flow interactions.

Despite these advances, persistent biases remain pervasive. Many models suffer from excessive "zonalisation" of flow, failing to capture correct blocking frequency and duration over Greenland and Europe (Athanasiadis et al. 2022; Davini et al. 2012). These blocking

deficiencies are dynamically linked to errors in eddy-driven jet representation, with models exhibiting reduced poleward tilt of the climatological jet (Athanasiadis et al. 2022). O'Reilly et al. (2025) found prevailing jet latitude biases in seasonal forecast systems, where modelled jets are often too zonal or shifted equatorward, leading to substantially weakened teleconnections between tropical drivers such as El Niño Southern Oscillation (ENSO) and the NAO. Stratospheric representation presents additional challenges. Palmeiro et al. (2023) identified significant biases in stratospheric temperature profiles that distort polar vortex strength and its downward coupling to the troposphere. Since stratospheric variability — particularly sudden stratospheric warmings — represents a key source of NAO predictability on seasonal timescales (Scaife et al. 2016), these biases directly compromise forecast skill.

A critical methodological consideration involves the distinction between fully coupled and atmosphere-only model configurations. Atmosphere-only simulations, where SST and sea ice are prescribed from observations (AMIP-type experiments), provide an important tool for isolating atmospheric model biases from coupled ocean-atmosphere interaction errors (Flato et al. 2013; Weisheimer et al. 2020). Weisheimer et al. (2020) found that both atmosphere-only (ASF-20C) and coupled (CSF-20C) seasonal hindcasts (using a model to "predict" the past) show similar temporal variations in NAO skill, including the mid-century skill drop. The finding that atmosphere-only simulations exhibit similar skill patterns suggests that changes in predictability are linked to intrinsic atmospheric dynamics rather than merely to ocean-atmosphere coupling errors or observational data quality (Weisheimer et al. 2020; O'Reilly et al. 2017).

Atmosphere-only configurations offer several advantages for mechanistic studies: they eliminate ocean model biases and drift, provide a benchmark for maximum potential atmospheric predictability, and enable controlled experiments to test specific hypotheses about atmospheric response to SST forcing (Athanasiadis et al. 2022; Sigmond et al. 2013). Several studies have used atmosphere-only experiments with perturbed SSTs to demonstrate causal links between North Atlantic SST biases and atmospheric circulation errors (Athanasiadis et al. 2022; Gastineau et al. 2013). DeWeaver et al. (2000) established foundational understanding of NAO dynamics using atmosphere-only approaches to diagnose zonal-eddy feedbacks and momentum budgets, demonstrating the value of isolating atmospheric processes. Kumar (2009) provided comprehensive treatment of skill estimation and sampling uncertainty in AMIP-style predictions, highlighting both the advantages and limitations of this approach.

The challenges outlined above — insufficient eddy feedback, persistent mean state biases, inadequate stratosphere-troposphere coupling, and weak teleconnections — all point to fundamental deficiencies in how current ESMs represent extratropical atmospheric variability. While increased resolution and improved parametrisation address specific biases, the pervasive signal-to-noise problem suggests more fundamental representational issues require investigation. Understanding the physical basis of this problem represents a central challenge for improving seasonal-to-decadal NAO prediction and Euro-Atlantic climate forecasts more broadly.

1.3 Signal-to-Noise Problem

The "signal-to-noise problem" represents one of the most significant unresolved challenges in modern climate prediction. First fully articulated by Scaife et al. (2014) and Eade et al. (2014), the problem describes a counter-intuitive situation where climate models are better at

predicting the real world than they are at predicting their own ensemble members. In a perfect forecasting system, the correlation between the ensemble mean and observations (r_{mo}, the skill) should equal the correlation between the ensemble mean and individual members (r_{mm}, the model's internal predictability). However, in North Atlantic predictions, r_{mo} consistently and substantially exceeds r_{mm}, often by large margins (Figure 1.3) (Scaife et al. 2018).

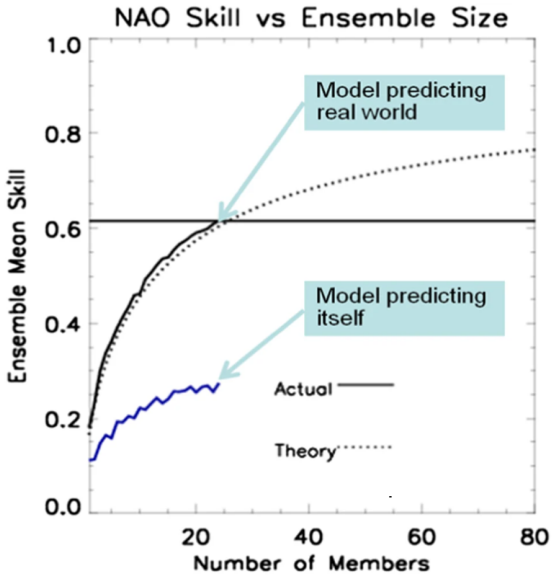

Figure 1.3: Predictability of the North Atlantic Oscillation in the real world (r_{mo}, black) is higher than the predictability in the model (r_{mm}, blue) (Scaife et al. 2018)

This discrepancy is quantified through the Ratio of Predictable Components (RPC), defined as the ratio of observed-to-modeled predictable variance (Eade et al. 2014). An ideal forecast system should thus exhibit RPC = 1, indicating that the model's internal signal-to-noise ratio matches reality. Instead, operational seasonal forecasts of the winter NAO show $RPC < 2$ (Scaife et al. 2014; Eade et al. 2014), implying that the signal in models is systematically less than half the strength it should be compared to the unpredictable "noise". Critically, the problem stems from insufficient signal strength rather than excessive noise (Scaife et al. 2018; O'Reilly et al. 2017). In particular, Scaife et al. (2018) demonstrated that total ensemble standard deviation closely matches observed NAO variability, meaning that there is not excessive noise and therefore that the signal is too weak. This has profound practical consequences: massive ensembles of 40 - 100+ members are required to average out the noise and extract the weak predictable signal (Eade et al. 2014; Smith et al. 2020).

Baker et al. (2018) conducted a multi-model intercomparison showing that signal-to-noise issues are pervasive across operational seasonal forecast systems, not artefacts of a single modelling centre. The problem extends beyond seasonal timescales: Smith et al. (2020) demonstrated that decadal predictions of the NAO face similar challenges, with a 169–member multi-model ensemble showing high anomaly correlation coefficients (ACC) but low mean squared skill scores (MSSS), requiring very large ensembles to extract predictable signals. Additionally, with large ensembles the spread becomes poorly related to forecast error, unlike in medium-range weather prediction where spread provides useful uncertainty estimates (O'Reilly et al. 2017). This fundamentally limits the utility of probabilistic seasonal and decadal forecasts for decision-making.

Several physical mechanisms - and their misrepresentation - have been identified as contributing to the weak signals. Hardiman et al. (2022) presents the most compelling recent evidence that insufficient atmospheric eddy feedback represents a primary cause. By computing the magnitude of feedback between transient eddies and large-scale flow anomalies across multiple forecast systems, they demonstrated that models underestimate eddy feedback by factors of 2-3 compared to reanalyses. A more accurate representation of the eddy feedback is linked to stronger predictable signals and teleconnection patterns, and its misrepresentation may explain weak responses to both tropical (ENSO) and extratropical (stratospheric) predictability sources. O'Reilly et al. (2025) further demonstrated that ENSO-NAO teleconnections are substantially weaker in seasonal forecast models than observed, and linked weak atmospheric responses to deficient eddy feedback and misrepresented regime persistence.

On decadal timescales, Patrizio et al. (2025) identified an additional mechanism: underrepresented ocean-atmosphere feedbacks. They showed that NAO decadal predictability depends on positive feedback between subpolar North Atlantic SSTs and the NAO, which varies in strength across decadal prediction systems but appears too weak even in skillful systems compared to observational estimates. Models exhibiting stronger positive SST-NAO feedback show higher NAO skill and less severe signal-to-noise problems, suggesting that insufficient air-sea coupling contributes to weak predictable signals on longer timescales.

Stratospheric processes provide another source of predictability, yet they fail to be well represented in current models. O'Reilly et al. (2017) demonstrated that errors in the Quasi-Biennial Oscillation (QBO) partially explains the signal-to-noise problem in seasonal NAO hindcasts. They performed two ensembles of ERA-40/Interim initialized forecast, one including and another removing the QBO influence, that had RPC values of 1.51 and 1.12, respectively. While this suggests the QBO contributes to the problem, the fact that RPC remains above unity even without QBO influence indicates that insufficient stratosphere-troposphere coupling represents only part of the issue, with deficient tropospheric processes also playing a crucial role.

The signal-to-noise problem exhibits pronounced temporal variability, challenging assumptions based on short evaluation periods. Weisheimer et al. (2020) analyzed 110-year seasonal hindcasts, revealing that RPC varies substantially across the 20th century. Recent decades and the early 1900s show RPC > 1, while the mid-century period (1940s-1970s) exhibits RPC < 1—coinciding with the period of lowest NAO skill. This dynamic nature suggests that the strength of predictability sources and their representation in models varies on multi-decadal timescales, possibly linked to changes in ENSO teleconnection strength or Atlantic multi-decadal variability. Siegert et al. (2016) addressed this through a rigorous Bayesian framework, quantifying uncertainty in signal-to-noise estimates and confirming with >99% certainty that GloSea5 exhibits anomalously low signal-to-noise ratios ($SNR_{model} = 0.21$ vs. $SNR_{obs} = 1.73$).

The origins of the signal-to-noise problem have also been investigated through targeted experiments. Knight et al. (2022) tested whether the problem arises from errors in simulating tropical dynamics versus extratropical atmospheric processes. Using relaxation experiments that imposed observed tropical atmospheric state, they found the signal-to-noise problem persists despite eliminating tropical errors. This implicates deficiencies in extratropical atmospheric processes—particularly the atmospheric response to predictability sources — as

the primary cause.

Attempts to resolve the problem through model improvements have met with limited success. Scaife et al. (2019) demonstrated that doubling atmospheric resolution from ∼0.8° to ∼0.3° showed no improvement in signal-to-noise ratios despite improvements in some aspects of eddy feedback, and even very high resolution (0.14°) did not fully resolve the issue. This suggests that resolution increases alone are insufficient, and more fundamental issues in how models represent eddy-mean flow interactions require addressing. Weisheimer et al. (2014) showed that stochastic physics parameterizations, designed to improve ensemble reliability by representing model uncertainty, can improve forecast calibration. However, in regions where RPC > 1 (indicating models are already "underconfident"), techniques aimed at increasing ensemble spread may exacerbate rather than solve the underlying problem of insufficient signal strength.

The signal-to-noise problem remains a major unresolved challenge in climate science, affecting confidence in regional climate predictions and projections across timescales (Scaife et al. 2018; O'Reilly et al. 2025). For NAO prediction specifically, the weak signals limit the practical utility of forecasts despite demonstrated correlation skill. Understanding the physical basis of insufficient atmospheric responses—particularly deficient eddy feedback, weak teleconnections, and misrepresented regime behaviour—represents a central motivation for detailed investigation of the atmospheric dynamics underlying NAO variability.

1.4 Teleconnections

Teleconnections — statistically significant correlations between atmospheric circulation patterns at widely separated geographical locations (Wallace et al. 1981; Barnston et al. 1987) — provide crucial pathways through which NAO predictability can arise from remote forcing (Hardiman et al. 2019). Understanding teleconnection dynamics is essential for exploiting predictability sources such as tropical sea surface temperatures, stratospheric variability, and oceanic conditions. However, as will be shown, systematic model deficiencies in representing teleconnection strength and structure contribute directly to the signal-to-noise problem and limit NAO prediction skill.

The most extensively studied extratropical teleconnection involves the El Niño-Southern Oscillation (ENSO). Classical theory explains tropical-extratropical teleconnections through Rossby wave trains generated by tropical convective heating anomalies and propagating along great circle paths to higher latitudes (Hoskins et al. 1981). While ENSO's influence on the PNA sector is robust and well-established (Hurrell et al. 2009), its impact on the North Atlantic is weaker, less spatially coherent, and more challenging to demonstrate (King et al. 2018). The PNA pattern represents the extratropical arm of ENSO, whereas the NAO does not solely owe its existence to ocean-atmosphere coupling, as demonstrated by atmospheric model experiments without SST or sea ice variability (Hurrell et al. 2009).

Despite demonstrated skill in capturing ENSO variability itself, seasonal forecast models substantially under represent the ENSO-NAO teleconnection strength. O'Reilly et al. (2025) analysed sixteen Copernicus Copernicus Climate Change Service (C3S) forecasting systems, finding that while all exhibit strong correlations between ENSO and Pacific jet anomalies comparable to reanalysis, the subsequent connection between Pacific jet anomalies and North Atlantic circulation is substantially weaker than observed. This systematic deficiency points

to errors in the extratropical North Atlantic atmospheric response to upstream circulation anomalies rather than deficiencies in tropical forcing or tropical-Pacific atmospheric response. Knight et al. (2022) reached similar conclusions through relaxation experiments that imposed observed tropical atmospheric states, finding the signal-to-noise problem persists, thus implicating extratropical processes.

The role of transient eddies in amplifying and maintaining teleconnection patterns provides crucial context. Kang et al. (2014) demonstrated that models with stronger transient eddy activity produce substantially larger extratropical circulation responses to ENSO despite similar tropical forcing. While transient eddies contribute to both predictable signals (through amplification of forced responses) and unpredictable noise (through generation of internal variability), models with stronger transient activity more closely match observed teleconnection amplitudes. Feldstein (2003) established that NAO lifecycle—growth, maintenance, and decay—is fundamentally mediated by synoptic eddy vorticity fluxes, with NAO anomalies arising from the propagation and merger of preexisting circulation features over 8-12 day timescales. This mechanistic understanding links the weak teleconnections in models directly to the insufficient eddy feedback identified in previous sections (Hardiman et al. 2022; Scaife et al. 2019).

Stratospheric variability represents another critical source of NAO predictability on seasonal timescales. The strength of the stratospheric polar vortex exhibits a two-way interaction with tropospheric circulation: tropospheric planetary waves propagate upward rapidly (days), while stratospheric anomalies exert slower downward influence (days to weeks) affecting the Arctic Oscillation (AO) and NAO (Baldwin et al. 2001; Andrews et al. 2019). Sudden stratospheric warmings (SSWs) — rapid breakdowns of the winter polar vortex — are typically followed by negative NAO phases, while strong polar vortex (SPV) events precede positive NAO phases, with surface impacts persisting for weeks to months (Scaife et al. 2016).

The importance of stratospheric pathways for NAO seasonal prediction is demonstrated compellingly by Scaife et al. (2016), who analysed the Met Office GloSea5 seasonal forecast system. While the system achieves impressive NAO correlation skill ($\sim$0.6), this skill vanishes entirely when ensemble members containing SSW events are excluded from the hindcasts. This striking result indicates that stratospheric variability is not merely a supplementary predictability source but rather an essential component of current seasonal NAO forecast skill. Sigmond et al. (2013) further demonstrated that dynamical forecasts initialised during SSW onset show significantly enhanced skill for surface temperature and precipitation months ahead compared to unconditional forecasts, with the improvement arising specifically from atmospheric initialisation rather than SST predictability.

The stratospheric influence pathway involves modulation of upward-propagating planetary wave activity. The Quasi-Biennial Oscillation (QBO) in tropical stratospheric winds modulates this pathway: during QBO easterly phases, more planetary wave energy is confined to Northern Hemisphere extratropics, favouring SSWs and a weakened polar vortex, while QBO westerly phases allow greater wave penetration into equatorial regions, favouring a strengthened vortex (Andrews et al. 2019). However, stratospheric temperature profile biases in models distort polar vortex strength and the timing and magnitude of downward coupling to the troposphere (Palmeiro et al. 2023), compromising this predictability source. Well-resolved stratospheres (high-top models) are important but not sufficient—accurate representation of coupling mechanisms and vortex-NAO relationships matters equally (King et al. 2018; Scaife

et al. 2016).

Ocean-atmosphere feedbacks provide predictability on longer timescales. The NAO forces a characteristic SST tripole pattern through surface heat flux anomalies and Ekman transport: warm anomalies in the central North Atlantic, cold anomalies in the subpolar and subtropical regions (Gastineau et al. 2013; Wanner et al. 2001). Reciprocally, SST anomalies—particularly the North Atlantic Horseshoe (NAH) pattern appearing in late summer—influence subsequent winter NAO with multi-year lags, acting as positive feedback (Gastineau et al. 2013). Patrizio et al. (2025) demonstrated that NAO decadal predictability in retrospective forecasts is fundamentally linked to the strength of this two-way ocean-atmosphere interaction.

The fast component of ocean-atmosphere feedback operates on monthly timescales: subpolar North Atlantic SST anomalies affect meridional SST gradients, which modulate low-level baroclinicity and diabatic heating, thereby influencing atmospheric jet strength and position, which projects onto the NAO (Patrizio et al. 2025). The slow component involves oceanic circulation responses: positive NAO phases bring cold air masses from continental Canada into the Labrador Sea fostering deep ocean mixing, which accelerates the Atlantic Meridional Overturning Circulation (AMOC) with multi-year lags, producing delayed subpolar SST anomalies that oppose the initial forcing, acting as a negative feedback (Patrizio et al. 2025; Gastineau et al. 2013). Decadal prediction systems exhibiting significant NAO skill show stronger representations of these feedbacks, though even skilful systems under represent the positive SST-NAO feedback compared to observational estimates (Patrizio et al. 2025). This under representation likely reflects the same atmospheric deficiencies affecting ENSO teleconnections—insufficient atmospheric responses to SST gradient forcing.

The synthesis of teleconnection research reveals that multiple pathways contribute to NAO predictability: ENSO provides modest influence with strong intraseasonal dependence, stratospheric variability offers essential but conditional predictability on seasonal timescales, and ocean feedbacks enable predictability extending to decadal timescales. However, systematic model deficiencies in representing teleconnection strength—particularly weak extratropical atmospheric responses to upstream forcing—directly contribute to the signal-to-noise problem. It is possible that a common thread linking weak teleconnections - is insufficient transient eddy feedback in models (Hardiman et al. 2022; O'Reilly et al. 2025), suggesting that addressing teleconnection biases requires fundamental improvements in how models represent atmospheric dynamics—particularly eddy-mean flow interactions, jet variability, and regime behavior—rather than merely refining tropical, stratospheric, or oceanic forcing sources.

1.5 Eddy Feedback

The interaction between transient synoptic-scale eddies and the large-scale mean flow — termed eddy feedback — represents one of the most fundamental dynamical processes in the extratropical atmosphere. Positive eddy feedback occurs when eddies generated by baroclinic instabilities in the mean flow subsequently act to reinforce and amplify flow anomalies through momentum and vorticity transports (Lorenz et al. 2003; Barnes et al. 2010). This feedback is central to the North Atlantic Oscillation: the growth, maintenance, and persistence of NAO anomalies are all fundamentally mediated by eddy-mean flow interactions (Feldstein 2003; DeWeaver et al. 2000). Recent research has identified insufficient eddy feedback in climate models as the primary physical mechanism explaining the signal-to-noise problem

discussed in previous sections, providing a unifying explanation for weak teleconnections, deficient stratospheric responses, and underestimated predictable signals.

The physical mechanisms underlying positive eddy feedback are well established. Lorenz et al. (2003) demonstrated that high-frequency transient eddies (2-8 day synoptic systems) interact systematically with low-frequency mean flow: eddies accelerate regions of strong westerlies and decelerate regions of weak westerlies or easterlies, thereby amplifying existing anomalies rather than damping them toward climatology. This positive feedback arises from the systematic organisation of eddy momentum and vorticity fluxes relative to the mean flow configuration. For the NAO specifically, Feldstein (2003) established that the entire 8 - 12 day life cycle - growth, subsequent maintenance, and eventual decay — is fundamentally mediated by synoptic eddy vorticity fluxes. NAO anomalies do not arise spontaneously but rather emerge from the propagation and merger of preexisting circulation anomalies, with high-frequency transient eddies providing the forcing that drives these processes. The maintenance phase similarly depends on continued eddy forcing that counteracts dissipative processes.

Scaife et al. (2019) provided the first clear argument that the deficient eddy feedback in models could be directly linked to the prediction challenges of the NAO, however the EFP was first defined by Smith et al. (2022) and Hardiman et al. (2022), who provided breakthrough evidence directly linking eddy feedback deficiencies to signal-to-noise errors across seasonal forecast systems. Smith et al. (2022) first defined the EFP quantifying the relationship between atmospheric wave driving and zonal-mean wind: specifically, the squared correlation between horizontal Eliassen-Palm flux divergence and zonal-mean zonal wind, area-weighted over Northern Hemisphere extratropical latitudes at 200 - 600 hPa. Recent work however, from Saffin et al. (2024) found that the EFP is sensitive to the sample period it is computed for and proposed an alternative measurement - the barotropic energy generation rate - which has less sensitivities.

Hardiman et al. (2022) analysed seventeen systems from multiple modelling centres and found that they all exhibited deficient eddy feedback compared to ERA5 reanalysis. However, the magnitude of this deficiency varied substantially across systems. Critically, systems with stronger (more realistic) eddy feedback showed significantly higher model-model skill (r_{mm}), higher model-observation skill (r_{mo}), and ratios of predictable components closer to unity. The correlation between EFP and the Arctic Oscillation RPC was -0.57 (negative because higher EFP reduces RPC toward the ideal value of 1), and the linear regression line crosses RPC = 1 very close to the observed EFP value from reanalysis. This suggests that correcting eddy feedback to realistic levels would substantially alleviate or eliminate the signal-to-noise problem.

The physical chain linking eddy feedback to predictable signals may operate through teleconnection strength. Hardiman et al. (2022) demonstrated a strong positive correlation ($r = 0.69$) between EFP and ENSO teleconnection strength to the Arctic Oscillation across forecast systems. Systems with improved eddy feedback capture stronger and more realistic extratropical responses to tropical forcing. This stronger teleconnection, in turn, produces larger predictable ensemble-mean signals (correlation between teleconnection strength and signal variance of 0.89), directly addressing the weak signal problem. By estimating potential skill gains if eddy feedback were brought to realistic levels, Hardiman et al. (2022) projected that correlation skill could approximately double in some extratropical regions, representing a

profound improvement in seasonal prediction capability.

Barnes et al. (2010) demonstrated that eddy feedback is crucial for NAO persistence on intraseasonal timescales. During positive NAO phases, synoptic eddies reinforce anomalous westerlies through vorticity flux feedbacks; during negative phases, eddies reinforce the weakened or easterly flow. Without these feedbacks, NAO persistence would be dramatically reduced. Gerber et al. (2007) showed through idealised modelling that positive eddy-zonal flow feedback extends the persistence timescale of zonal wind anomalies from approximately 10 days (in the absence of feedback) to 20-30 days (with realistic feedback). Models with weak eddy feedback therefore exhibit less persistent, more "red noise" behaviour in atmospheric indices, consistent with observed model deficiencies in capturing low-frequency variability.

The self-maintenance of midlatitude jets through eddy feedback represents a fundamental dynamical principle. Robinson (1991) and Robinson (2006) developed theoretical frameworks showing that jets maintain themselves through a positive feedback loop: the jet generates baroclinic eddies through instability, these eddies transport vorticity and momentum in ways that accelerate the jet, and this acceleration promotes further eddy generation. Woollings et al. (2018) documented an inverse relationship between jet speed and meridional variability that holds across observations, comprehensive climate models, and highly idealised models (dry dynamical cores, barotropic vorticity equation models). Faster jets are more zonally oriented with less meridional meandering, while slower jets exhibit greater latitude variability. This relationship emerges as a fundamental consequence of barotropic wave dynamics—specifically wave refraction and breaking—rather than depending on baroclinic growth processes, and it directly links jet behaviour to eddy-mean flow interactions.

Both stationary and transient eddies contribute to NAO dynamics through their interactions with the mean flow. DeWeaver et al. (2000) used a linear primitive equation model to demonstrate that zonal-mean flow changes provide the dominant forcing to stationary waves, which then feed back to maintain the flow anomalies. This represents mutual adjustment between zonal-mean and eddy components, with interactions between climatological eddies and low-frequency anomalies playing crucial roles. The relative importance of stationary versus transient eddy contributions varies with the spatial structure and timescale of NAO variability, but both components participate in the feedback processes that amplify and maintain anomalies.

The relationship between model resolution and eddy feedback has been investigated. Scaife et al. (2019) tested seasonal hindcasts at standard ($\sim$0.8° or $\sim$90km), higher ($\sim$0.3° or $\sim$35km), and very high (0.14° or $\sim$15km) atmospheric resolutions. While eddy momentum flux convergence over the North Atlantic improved modestly with increased resolution, it remained too weak even at the higher resolution, showing only small improvements across the tested range. Eddy vorticity forcing regressed onto the NAO index showed convergence toward observational values only at the very high (0.14°) resolution approaching mesoscale-resolving scales. Most critically, the signal-to-noise ratio of NAO predictions remained essentially unchanged across the resolution range despite these incremental eddy feedback improvements. This indicates that resolution increases alone are insufficient to address eddy feedback deficiencies, suggesting that fundamental representational issues—potentially involving convection parametrisations, boundary layer schemes, or numerical diffusion—also play important roles.

Regional and structural aspects of eddy feedback may be more relevant than zonal-mean

measures for some applications. O'Reilly et al. (2025) noted that the specific zonally-averaged EFP defined by Hardiman et al. (2022) showed only weak correlation (r = -0.22) with early winter North Atlantic RPC, despite the strong relationship found for winter-mean Arctic Oscillation. This suggests that for regional phenomena like the early winter East Atlantic response to ENSO, structural aspects of the eddy feedback — such as its spatial distribution, its relationship to blocking frequencies, or its interaction with specific jet configurations — may matter more than hemispheric-average measures. Climatological biases in jet latitude distributions and the frequency of jet regime occurrences may better capture the eddy feedback deficiencies affecting regional teleconnections. Athanasiadis et al. (2022) demonstrated that Gulf Stream Extension SST gradient biases are linked to eddy-driven jet position biases and high-latitude blocking deficiencies through a positive coupled feedback, illustrating how mean state biases, eddy dynamics, and teleconnection representations are intimately connected.

The amplification of predictable signals by eddy feedback extends to multiple forcing sources. Kang et al. (2014) demonstrated that seasonal prediction models with stronger climatological transient activity produce substantially larger extratropical circulation responses to ENSO despite similar tropical forcing and precipitation anomalies. While enhanced transient activity amplifies both the predictable signal (forced response) and unpredictable noise (internal variability), the net effect is beneficial: models with strong transient activity more closely match observed teleconnection amplitudes. Willison et al. (2015) showed through high-resolution (20 km) regional simulations that enhanced eddy kinetic energy responses to warming emerge when mesoscale latent heating processes within cyclones are better resolved, suggesting that eddy feedback involves complex thermodynamic as well as momentum transport processes.

Insufficient eddy feedback in current models represents the primary physical cause of weak predictable signals and the signal-to-noise problem that limits seasonal-to-decadal prediction skill (Hardiman et al. 2022). This deficiency may provide a unifying explanation connecting previously disparate findings: weak ENSO-NAO teleconnections result from insufficient eddy amplification of tropical forcing; deficient stratospheric-tropospheric coupling reflects inadequate eddy responses to polar vortex variations; underestimated SST-NAO feedbacks arise from weak atmospheric responses that depend on eddy dynamics. Until now, increased resolution hasn't reached a threshold that addresses the problem (Scaife et al. 2019; Patrizio et al. 2025), however with ever increasing computational resources, resolutions that explicitly resolve mesoscale eddy processes are now available and have the potential to provide a significant mechanistic improvement.

1.6 Research Objective

The literature considered above establishes a clear challenge: climate models consistently underestimate the predictable signal of the NAO relative to internal noise, despite demonstrating skill in forecasting observed variations (Scaife et al. 2014; Eade et al. 2014). This signal-to-noise problem compromises the practical utility of seasonal-to-decadal predictions for the Euro-Atlantic sector and limits confidence in climate projections of regional circulation changes.

Recent research has suggested insufficient atmospheric eddy feedback (Hardiman et al. 2022; Smith et al. 2022) as the primary physical mechanism underlying weak predictable signals. Models underestimate the positive feedback between transient synoptic eddies and large-scale

flow anomalies that maintains and amplifies NAO variability. This deficiency may provide a unifying explanation for multiple observed model failures: weak ENSO-NAO teleconnections (O'Reilly et al. 2025), deficient stratosphere-troposphere coupling (O'Reilly et al. 2017), and underestimated persistence of atmospheric regimes. While increased model resolution shows promise—with very high resolution (0.14°) approaching observed eddy feedback values (Scaife et al. 2019)—no systematic investigation has examined whether resolution-driven improvements in eddy dynamics translate to enhanced NAO predictability and reduced signal-to-noise ratios.

This thesis addresses this gap through the following research question:

> **Does increased atmospheric resolution improve the representation of eddy feedback and associated teleconnection mechanisms, and therefore help address the signal-to-noise problem in NAO predictions?**

Specifically, this research investigates:

1. How eddy feedback strength varies between 27 km and 9 km horizontal atmospheric resolution
2. Whether higher resolution (9 km) produces more realistic teleconnection patterns
3. The relationship between resolution-dependent eddy dynamics, teleconnection patterns and the NAO

This research aims to contribute a mechanistic understanding of how atmospheric resolution affects NAO predictability through improved representation of fundamental dynamical processes. By establishing whether resolution-driven improvements in eddy feedback can address the signal-to-noise problem, this work informs development strategies for the next generation of seasonal prediction systems and climate models, with direct implications for the practical utility of Euro-Atlantic climate forecasts.

Chapter 2

Methodology

This chapter explains the data sources and diagnostic methods that were used in the analysis. The aim was to provide results that are broadly relevant and reproducible - focusing on the mechanistic differences rather than attempting to quantify them. This is reflected in the methodological considerations.

2.1 Data Sources

2.1.1 AMIP-EERIE Simulations

This study utilises atmosphere-only simulations from the EERIE project, conducted using the ECMWF Integrated Forecasting System (IFS). These simulations follow the AMIP protocol, where the atmospheric model is forced with prescribed observed daily sea surface temperature and sea ice concentration fields rather than being coupled to an interactive ocean model. The AMIP configuration enables isolation of atmospheric dynamical processes without complications from ocean model biases or coupled atmosphere-ocean drift, making it particularly suitable for investigating eddy feedback mechanisms and teleconnection dynamics (Flato et al. 2013; Weisheimer et al. 2020).

The simulations employ IFS Cycle 48r1 with modifications developed during the Next Generation Earth Modelling Systems (NextGEMS) and Destination Earth (DestinE) projects (Rackow et al. 2024). These modifications include activation of tracer mass-fixers to improve global energy and moisture conservation, and parameter tuning to improve the Top Of Atmosphere (TOA) radiation balance. The atmospheric component uses a two-time-level semi-implicit, semi-Lagrangian dynamical core with computations alternated between spectral and cubic octahedral reduced Gaussian (TCo) grid-point representations (Simmons et al. 1989; Diamantakis et al. 2022). All configurations employ 137 vertical levels extending from the surface to 0.01 hPa and use 32-bit single precision for computational efficiency (Lang et al. 2021). Both configurations use identical physics parametrisations without resolution-dependent retuning, enabling isolation of pure resolution effects.

Two horizontal resolutions are analysed in this study:

TCo399 Configuration (∼27 km): The lower-resolution configuration uses a TCo399 spectral truncation with approximately 27 km horizontal grid spacing. The model employs a 900-second (15-minute) time step and consists of a 10-member ensemble with perturbed

atmospheric initial conditions for 1 January 1980. However, daily outputs of 3D variables were only saved for 5 members, to save storage. Initial condition perturbations follow operational ECMWF ensemble forecast procedures, combining singular vectors and the ERA5 Ensemble of Data Assimilations (EDA) (Lock et al. 2019). The simulation period spans 1980 - 2023, providing 44 years of model output for analysis, or 43 winters.

TCo1279 Configuration (∼9 km): The higher-resolution configuration uses TCo1279 spectral truncation with approximately 9 km horizontal grid spacing. Due to computational constraints, this configuration employs a shorter 450-second (7.5-minute) time step to maintain numerical stability and consists of a single ensemble member. The simulation covers the same 1980 - 2023 period as the lower-resolution ensemble. Only one simulation was performed at this resolution, due to its much higher computational cost. Both configurations are scientifically identical apart from horizontal resolution and time step, enabling direct assessment of resolution impacts on atmospheric dynamics. The models do not use stochastic perturbations to parameterised physics, ensuring that differences arise solely from deterministic processes at different scales.

Boundary Conditions

Ocean and sea-ice boundary conditions are prescribed using high-resolution satellite-based observations. Sea surface temperatures are specified from the European Space Agency Climate Change Initiative Sea Surface Temperature Version 3 (ESA CCI SST V3), which provides daily-mean values representative of 20 cm depth with global coverage from 1980 to present (Embury et al. 2024). ESA CCI SST V3 was selected for its superior temporal consistency and accuracy compared to previous versions, particularly during the 1980s and early 1990s. Sea ice concentrations are specified using data from the European Organisation for the Exploitation of Meteorological Satellites (EUMETSAT) Ocean and Sea Ice Satellite Application Facility (OSI-SAF), specifically the OSI-450 Climate Data Record (1980–2021) and OSI-430-b Intermediate Climate Data Record (2022–2023) products distributed alongside ESA CCI SST V3.

Original SST and sea ice data on a $0.05^\circ \times 0.05^\circ$ latitude-longitude grid were conservatively interpolated to the required TCo grids using ECMWF's Meteorological Interpolation and Regridding (MIR) software (Maciel et al. 2017). It should be noted that while distributed on a fine grid, the effective feature resolution of ESA CCI SST V3 varies temporally, being generally lower during the earlier period (pre-1985) due to limited high-resolution satellite coverage (Embury et al. 2024). This temporal evolution of observational resolution is an important consideration when analysing atmospheric responses to mesoscale SST features.

External Forcings

External radiative forcings generally follow Coupled Model Intercomparison Project Phase 6 (CMIP6) protocols (Eyring et al. 2016). For the historical period (1980–2014), greenhouse gas concentrations (CO_2, CH_4, N_2O, CFC-11, CFC-12), volcanic aerosols, ozone, and total solar irradiance follow CMIP6 historical specifications (Roberts et al. 2018a). Tropospheric aerosols are specified as time-evolving climatologies generated within the CONFESS project (Stockdale et al. 2022; Stockdale et al. 2024) using IFS simulations with prognostic atmospheric

chemistry driven by CMIP6 emissions data. Each aerosol climatology represents a 9-year running mean sampled at 3-year intervals from 1975 to 2015.

For the post-2014 period where CMIP6 historical data are unavailable, forcings are specified as follows:

- Greenhouse gases and ozone follow the SSP3-7.0 scenario
- Volcanic aerosol optical depths are held constant using a representative year without major eruptions
- Solar forcing uses the CMIP6 reference estimate of most likely solar activity from 2015 to 2300
- Tropospheric aerosol climatologies remain constant at 2015 levels
- Land surface properties are specified as seasonally-varying climatologies without inter-annual evolution

Model Output

All output variables follow General Regularly-Distributed Information in Binary Form Version 2 (GRIB2) standards as encoded by ecCodes 2.36.1, with descriptions available in the ECMWF parameter database. Data were archived on both the native TCo grids and on a regular $0.25^\circ \times 0.25^\circ$ latitude-longitude grid (conservative interpolation via MIR). Output frequencies include 6-hourly instantaneous values (or 6-hourly accumulations for flux variables), daily means, and monthly means. Daily and monthly means are computed online from hourly native-resolution data using MultIO (Sarmany et al. 2024), ensuring that flux averages are exact rather than approximated from lower-frequency samples. For this study, we re-interpolated all the data to a $1^\circ \times 1^\circ$ grid for reducing storage and enabling cheaper analysis computations.

Variables analysed in this study include: sea level pressure, geopotential height on pressure levels (particularly 500 hPa), zonal and meridional wind components (surface and upper levels), temperature (2-metre and pressure levels), and various diagnostic fields for eddy feedback analysis. All data was kindly shared by Matthias Aengenheyster from ECMWF, who produced the simulations within the context of the EERIE project.

2.1.2 ERA5 Reanalysis

The ERA5 reanalysis from the ECMWF serves as the primary observational reference for model validation (Hersbach et al. 2020). ERA5 represents ECMWF's fifth-generation atmospheric reanalysis, providing hourly estimates of atmospheric, land, and ocean-wave variables from 1940 to near-present (with typical 2–3 month delay for quality control).

ERA5 employsIFS Cycle 41r2 (released 2016) with 4-dimensional variational (4D-Var) data assimilation. The native atmospheric model resolution is approximately 31 km (reduced Gaussian grid N320, which is different from the EERIE native grids), with 137 vertical levels identical to the AMIP simulations. The 4D-Var assimilation uses 12-hour windows with analysis updates every 12 hours (00:00 and 12:00 UTC). An EDA with 10 members provides uncertainty estimates, though the deterministic analysis is used as the primary reference in this study.

This study uses ERA5 data on the standard 1° × 1° regular latitude-longitude grid, re-gridded from the native reduced Gaussian grid. The analysis period focuses on 1980–2023 to match the AMIP simulation period, though ERA5 extends back to 1940 and reaches up to present. Data quality is highest from 1979 onward, coinciding with the beginning of the satellite era and substantially increased observation coverage (Hersbach et al. 2020).
Variables: Key variables used in this study include:

- *Pressure-level fields* (37 standard levels from 1000 hPa to 1 hPa): geopotential height, temperature, zonal and meridional wind components and vertical velocity
- *Single-level fields*: mean sea level pressure, 2-metre temperature and dewpoint, 10-metre winds and precipitation (accumulated)

ERA5 uses the Operational Sea Surface Temperature and Ice Analysis (OSTIA) for its lower boundary condition, which differs from the ESA CCI SST V3 employed in the AMIP simulations. OSTIA provides daily, global SST analysis at 0.05° resolution through near-real-time data assimilation of satellite and *in situ* observations. While both products are observationally based and highly correlated, minor differences exist in processing methodology, quality control, and the specific satellite retrievals assimilated. ESA CCI SST V3 represents a reprocessed climate data record optimised for temporal stability, whereas OSTIA prioritises operational timeliness. These differences are minor for large-scale patterns but may introduce small discrepancies in SST gradient structure, particularly in regions with sparse observations.

ERA5's primary strength lies in its observational constraint through continuous data assimilation, ensuring that the atmospheric state remains consistent with observed weather patterns. This makes ERA5 an appropriate reference for validating model climatologies, synoptic variability, and teleconnection patterns. The reanalysis benefits from ECMWF's advanced data assimilation system, comprehensive observation usage (satellite radiances, radiosondes, aircraft reports, surface stations), and sophisticated bias correction schemes (Hersbach et al. 2020).

However, several limitations of ERA5 are relevant to this study:

1. **Model dependency**: Variables not directly observed (e.g., precipitation, surface fluxes, diabatic heating) depend substantially on model physics and may inherit model biases. Precipitation in particular shows known biases in intensity and frequency that are not fully constrained by observations (Hersbach et al. 2020).

2. **Resolution constraints**: The 31 km native resolution cannot fully resolve mesoscale atmospheric features below approximately 60–100 km (2–3 times the grid spacing). Synoptic-scale cyclones and fronts are well represented, but small convective systems and fine-scale orographic effects remain partially parameterised.

3. **Temporal inhomogeneities**: While ERA5 uses a single model version throughout, changes in the observing system — particularly major satellite transitions in 1979 (microwave sounding units), 1998 with the introduction of Advanced TIROS Operational Vertical Sounder (ATOVS), and ongoing improvements in satellite retrievals — can introduce subtle discontinuities. The 1979 transition is particularly significant, with substantially improved quality in the satellite era.

4. **Stratospheric biases**: ERA5 exhibits a known cold bias in the upper stratosphere, qualitatively similar to that found in high-resolution IFS free-running simulations. This bias reflects limitations in both model physics and observational constraints at high altitudes (Hersbach et al. 2020).

5. **Regional uncertainty**: Uncertainty is elevated in data-sparse regions, particularly the Southern Ocean, polar regions, and the pre-satellite era. While model physics provides first-guess fields in these regions, the observational constraint is substantially weaker than in well-observed areas like the Northern Hemisphere mid-latitudes.

To account for these limitations, model validation employs a multi-dataset approach. While ERA5 serves as the primary reference for large-scale circulation patterns (sea level pressure, geopotential height, wind fields), supplementary observational datasets are used where ERA5 is less reliable: Clouds and the Earth's Radiant Energy System Energy Balanced and Filled (CERES EBAF) for top-of-atmosphere radiation fluxes (Loeb et al. 2018), and consideration of multiple precipitation products (Integrated Multi-satellitE Retrievals for Global Precipitation Measurement (IMERG), Global Precipitation Climatology Project (GPCP)) for assessing precipitation biases, following the approach of Roberts et al. (2018a) and the EERIE project protocol.

The comparison between ERA5 (IFS Cycle 41r2, constrained by observations) and AMIP simulations (IFS Cycle 48r1, free-running with prescribed SST) isolates several factors:

(i) model physics improvements between cycles; (ii) effects of observational constraint versus free atmospheric evolution; (iii) SST product differences (OSTIA vs. ESA CCI SST V3).

For large-scale circulation features like the NAO, these differences are generally small compared to the signal being analyzed, justifying ERA5's use as a validation reference (Hersbach et al. 2020; Roberts et al. 2018a).

ERA5 data was downloaded through the Copernicus Climate Data Store (CDS), with pre-processed subsets stored on the Barcelona Supercomputing Centre (BSC) computing facilities for efficient analysis.

2.2 Diagnostic Methods

2.2.1 Computational Environment and Software

All of the following analyses were performed using Python 3.11 with standard scientific computing libraries on BSC's MareNostrum 5 supercomputer and local workstations. Multidimensional climate data arrays were handled using `xarray` (Hoyer et al. 2017), which provides labelled N-dimensional array operations optimised for Network Common Data Form (NetCDF) and GRIB2 formats. Numerical operations employed `NumPy` (Harris et al. 2020), whilst statistical analyses utilised `SciPy` (Virtanen et al. 2020), including the `scipy.stats` module for correlation analysis, t-tests, and linear regression, and the `scipy.signal` module for digital filter design. Time series manipulation was performed with `pandas` (McKinney et al. 2010). Visualisation utilised `Matplotlib` (Hunter 2007) for standard plots and `Cartopy` (Met Office 2015) for map projections, particularly polar stereographic projections for Northern Hemisphere diagnostics.

2.2.2 NAO Index Calculation

The NAO index was calculated using the station-based method, which computes the normalised SLP difference between two fixed geographical locations representing the centres of action of the NAO dipole. This approach was selected for its simplicity, reproducibility, and historical precedence, following the methodology established by Hurrell (1995) and widely adopted in climate research.

The station-based method offers several advantages over other alternative approaches such as EOF-based indices. Most importantly, this methodology is simple, efficient and easily reproducible across different datasets and model configurations, making it ideal for systematic comparison of multiple simulations. Additionally, unlike EOF-based approaches that can be sensitive to domain selection and may conflate different modes of variability, the station-based index provides a stable, consistent metric focused specifically on the North Atlantic sector (Hannachi 1997).

The NAO index was calculated identically for both ERA5 reanalysis and all AMIP-EERIE model simulations, ensuring direct comparability across datasets. The two representative stations that were selected are typical as they have the longest observational record: Lisbon, Portugal (38.7°N, 9.1°W) representing the subtropical high-pressure centre, and Reykjavik, Iceland (64.1°N, 21.9°W) representing the subpolar low-pressure centre. These locations correspond to the traditional centres of action identified in early NAO studies and capture the fundamental meridional pressure gradient that characterises the oscillation.

Monthly SLP data were processed to extract DJF (December January February) seasonal averages for each winter from 1980/81 to 2023/24. The DJF seasonal averaging carefully handled the year boundary, where December of year $n-1$ was combined with January and February of year n to create the winter season for year n. For gridded datasets (ERA5 and model simulations), pressure values at the nearest grid points to the specified coordinates were extracted.

The raw pressure difference was calculated as $\Delta P = P_{\text{Lisbon}} - P_{\text{Reykjavik}}$, then standardised using the long-term mean and standard deviation to produce the final NAO index:

$$\text{NAO}_{\text{index}} = \frac{\Delta P - \overline{\Delta P}}{\sigma_{\Delta P}}$$

This standardisation ensures that positive values correspond to the positive NAO phase (stronger than normal westerlies) and negative values correspond to the negative phase (weaker westerlies), with the index having zero mean and unit variance over the analysis period. The identical methodology applied across all datasets enables robust assessment of how well different model resolutions capture observed NAO variability and its relationship to large-scale circulation patterns.

2.2.3 Teleconnection Indices & Composites

Teleconnection analysis employed both regression and composite methodologies to analyse the relationship between ENSO and North Atlantic atmospheric circulation. The ENSO index was derived from SST anomalies in the Niño 3.4 region (5°S-5°N, 170°W-120°W) using ERA5 sea surface temperature data. Monthly SST values were converted to Celsius and processed into DJF seasonal averages following the same temporal construction as the NAO index.

The Niño 3.4 index was calculated as the area-weighted average of SST anomalies relative to the long-term climatological mean, then standardised to unit variance. El Niño years were defined as winters with Niño 3.4 index $>$ 1.0, La Niña years as index $<$ -1.0, and neutral conditions for intermediate values, consistent with standard meteorological practices (King et al. 2018).

Linear regression was performed between the standardised Niño 3.4 index and atmospheric fields (SLP, geopotential height) at each grid point. The regression coefficient represents the change in the atmospheric variable (in pascals per standard deviation for SLP, metres per standard deviation for geopotential height) associated with a one standard deviation change in the ENSO index. Statistical significance was assessed using Pearson correlation coefficients with p-values calculated assuming normal distribution and independence of annual values.

Composite maps were constructed by averaging atmospheric fields separately for El Niño and La Niña years using the standard definition threshold of 0.5, with the teleconnection pattern represented as the difference between these composites. Statistical significance of composite differences was not shown on the result plots - as the composite analysis highlights the differences between conditions with much smaller sample sizes, making point-by-point significance testing less statistically robust and often less relevant to interpreting the overall spatial pattern.

Both methodologies provide complementary qualitative perspectives on ENSO-North Atlantic teleconnections: regression analysis captures the linear relationship strength across the full ENSO spectrum, whilst composite analysis emphasises the atmospheric response to extreme ENSO events that may exhibit nonlinear characteristics (O'Reilly et al. 2025).

2.2.4 Eddy Kinetic Energy

Synoptic-scale EKE was calculated to quantify the intensity of transient atmospheric disturbances and assess their representation across different model resolutions. EKE represents the kinetic energy associated with time-varying wind fluctuations and provides a direct measure of synoptic storm activity, which is fundamental to eddy feedback processes and North Atlantic variability (Limpasuvan et al. 2000).

The calculation employed a band-pass filtering approach to isolate synoptic-scale motions from daily wind data. Daily zonal (u) and meridional (v) wind components at 500 hPa were processed using a Butterworth band-pass filter with cutoff periods of 2-10 days, following established methodology for identifying synoptic-scale eddies (Blackmon et al. 1977). The filter was applied using fourth-order Butterworth coefficients with frequencies $f_{\text{low}} = 1/10$ day^{-1} and $f_{\text{high}} = 1/2.001$ day^{-1}.

To minimise edge effects inherent in digital filtering, the time series were extended by mirror-padding the data at both ends before filtering. The filtered wind components were then processed using zero-phase filtering (filtfilt) to prevent phase distortion. EKE was calculated as:

$$\text{EKE} = \frac{1}{2}\left(u'^2 + v'^2\right)$$

where u' and v' represent the band-pass filtered (synoptic-scale) components of the zonal and meridional winds, respectively.

DJF seasonal averages were computed for each winter from 1980-2023. Regional analyses focused on the North Atlantic storm track region (40°N-70°N, 60°W-0°W) where EKE values are typically highest and most relevant for NAO dynamics. Area-weighted spatial averages were computed using cosine of latitude as the weighting factor to account for meridional grid convergence.

2.2.5 Eddy Feedback Parameter

The EFP quantifies the degree to which transient eddies reinforce large-scale atmospheric anomalies, following the methodology established by Hardiman et al. (2022). This parameter represents a critical diagnostic for understanding how effectively the models are able to represent the eddy feedback mechanisms, where deficiencies are suspected to result in poor teleconnection representation and weak predictable signals.

The EFP was calculated as the squared correlation coefficient between DJF-mean zonal wind and horizontal Eliassen-Palm (E-P) flux divergence at each latitude, following Edmon Jr et al. (1980). The calculation proceeded in several steps:

1. **Zonal mean and eddy decomposition:** Daily wind fields were decomposed into zonal mean and eddy components: $u = \bar{u} + u'$, $v = \bar{v} + v'$, where overbar denotes the zonal average and prime denotes deviation from the zonal mean.

2. **Eddy momentum flux calculation:** The meridional eddy momentum flux was computed as $\overline{u'v'}$, representing the poleward transport of zonal momentum by eddies.

3. **E-P flux divergence:** The horizontal component of E-P flux divergence was calculated as:
$$\nabla \cdot \mathbf{F}_H = -\frac{1}{a\cos^2\phi}\frac{\partial}{\partial\phi}\left(\overline{u'v'}\cos^2\phi\right)$$
where a is Earth's radius, ϕ is latitude, and the E-P flux divergence represents the acceleration of the zonal-mean flow due to eddy momentum transport.

4. **Seasonal averaging and correlation:** Both the zonal-mean zonal wind $\bar{u}$ and E-P flux divergence were averaged over each DJF season. The EFP was then computed as the squared correlation coefficient between these quantities across all winters:
$$\text{EFP} = r^2(\bar{u}_{\text{DJF}}, \nabla \cdot \mathbf{F}_{H,\text{DJF}})$$

The final EFP was calculated as an area-weighted average over Northern Hemisphere extratropical latitudes (25°N-72°N), consistent with Hardiman et al. (2022). This parameter ranges from 0 to 1, and in the literature is historically underrepresented by the models. Therefore, values closer to those calculated from the reanalysis representing more realistic dynamical coupling essential for maintaining NAO anomalies and enabling strong teleconnection responses.

Chapter 3

Results & Discussion

This chapter presents the systematic evaluation of the AMIP-EERIE TCo1279 and TCo399 simulations against ERA5 reanalysis, focusing on the representation of the NAO and its underlying dynamical processes. The analysis begins with an assessment of the model climatology and mean state to identify systematic biases that may affect NAO dynamics, a second part evaluates the representation of the NAO and a third provides a comparison of resolution-dependent processes including eddy activity, eddy feedback and teleconnection mechanisms.

3.1 Climatology & Mean State

Accurate representation of the climatological mean state is fundamental to realistically simulate the NAO variability (Stronmen et al. 2020). Systematic biases in mean sea level pressure, particularly in the centres of action over the Azores and Iceland, can distort the background state upon which NAO anomalies develop. Similarly, errors in the North Atlantic storm track position and intensity affect the eddy forcing that maintains NAO anomalies (Feldstein 2003). This subsection evaluates the fidelity of the AMIP-EERIE simulations in capturing the observed wintertime climatology, with emphasis on features directly relevant to NAO dynamics.

3.1.1 T-Tests

Statistical significance testing provides essential context for interpreting model-observation differences, distinguishing genuine systematic biases from sampling variability. The t-test is particularly valuable in this context because it accounts for both the magnitude of differences and the internal variability of the model ensemble. For atmosphere-only simulations prescribed with observed SST, differences between model and reanalysis should be small in regions where atmospheric variability is strongly constrained by boundary forcing. Conversely, larger statistically significant differences indicate deficiencies in atmospheric dynamics or physics that persist despite correct boundary conditions.

We focus the t-test analysis on MSLP for two principal reasons. First, MSLP directly defines the NAO through the pressure gradient between the Azores High and Icelandic Low, making it the most relevant variable for assessing model fidelity in simulating this mode of variability. Second, MSLP integrates the effects of the full atmospheric column, providing

a robust diagnostic less sensitive to vertical interpolation errors or level-specific biases than upper-air fields. Whilst examination of upper-level fields (e.g., 500 hPa geopotential height) would provide additional insights into the vertical structure of biases, the surface pressure field captures the fundamental representation of the NAO's centres of action.

Table 3.1 presents t-test statistics comparing DJF-mean MSLP between ERA5 and the AMIP-EERIE simulations at both resolutions. Tests were performed separately for the Azores High region (35°N–40°N, 15°W–10°W) and the Icelandic Low region (63°N–67°N, 25°W–15°W), representing the two centres of action of the NAO. For the TCo399 ensemble, statistics are shown both for the ensemble mean and for individual members, whilst the TCo1279 configuration provides a single realisation.

Table 3.1: T-test statistics comparing DJF-mean MSLP between ERA5 and AMIP-EERIE simulations for the NAO centres of action; statistically significant results ($p<0.05$) in bold.

		TCo1279	Ensemble TCo399	Members									
			Mean	1	2	3	4	5	6	7	8	9	10
Azores	t value	-0.24	0.74	-0.09	0.02	1.03	0.68	1.31	-0.06	0.50	1.03	2.19	-0.80
High	p value	0.81	0.46	0.93	0.98	0.31	0.50	0.19	0.95	0.62	0.31	**0.03**	0.42
Icelandic	t value	-0.10	-0.14	0.80	0.66	-1.00	-0.69	0.02	0.46	-0.61	-0.34	-2.09	1.50
Low	p value	0.92	0.89	0.42	0.51	0.32	0.49	0.98	0.65	0.54	0.73	**0.04**	0.14

The results reveal generally good agreement between the models and ERA5, with most t-values falling well within the range expected from sampling variability (typically $|t| < 2$ for 43 winters). Neither the ensemble mean nor the majority of individual members show statistically significant biases in either region at the conventional $p < 0.05$ threshold. The TCo1279 high-resolution simulation similarly shows no significant bias in either centre of action. Member 9 of the TCo399 ensemble exhibits statistically significant differences in both regions ($t = 2.19$, $p = 0.03$ for Azores; $t = -2.09$, $p = 0.04$ for Iceland), though these appear to represent sampling extremes within the ensemble spread rather than systematic model deficiencies, as other members and the ensemble mean show no such bias.

The absence of systematic MSLP biases in the NAO centres of action suggests that the AMIP-EERIE simulations capture the large-scale pressure distribution reasonably well when forced with observed SSTs. This provides confidence that differences in NAO characteristics, eddy feedback, or teleconnections identified in subsequent sections reflect dynamical processes rather than merely compensating for mean state errors.

3.1.2 Global Temperature and Precipitation Biases

Systematic biases in temperature and precipitation fields provide important context for evaluating model performance beyond the NAO-specific diagnostics. Whilst atmosphere-only simulations with prescribed SST eliminate ocean model biases and reduce atmosphere-ocean coupling errors, systematic errors in atmospheric physics—particularly in convection, cloud processes, and radiation—can still produce temperature and precipitation biases (Flato et al. 2013). These biases are particularly relevant for NAO studies if they affect diabatic heating gradients, baroclinicity, and the energetics of synoptic-scale eddies that maintain NAO anomalies (Willison et al. 2015).

Figure 3.1 presents DJF-mean 2-metre temperature biases relative to ERA5 for the TCo399 ensemble mean (a) and TCo1279 (b) configurations. The absolute difference between the biases $|(a)| - |(b)|$ is shown in (c). Both resolutions exhibit remarkably small temperature biases globally, with the vast majority of grid points showing differences of less than 1°C from reanalysis. This close agreement demonstrates that the atmospheric model physics, when forced with observed SST and sea ice, reproduces observed near-surface temperatures with high fidelity.

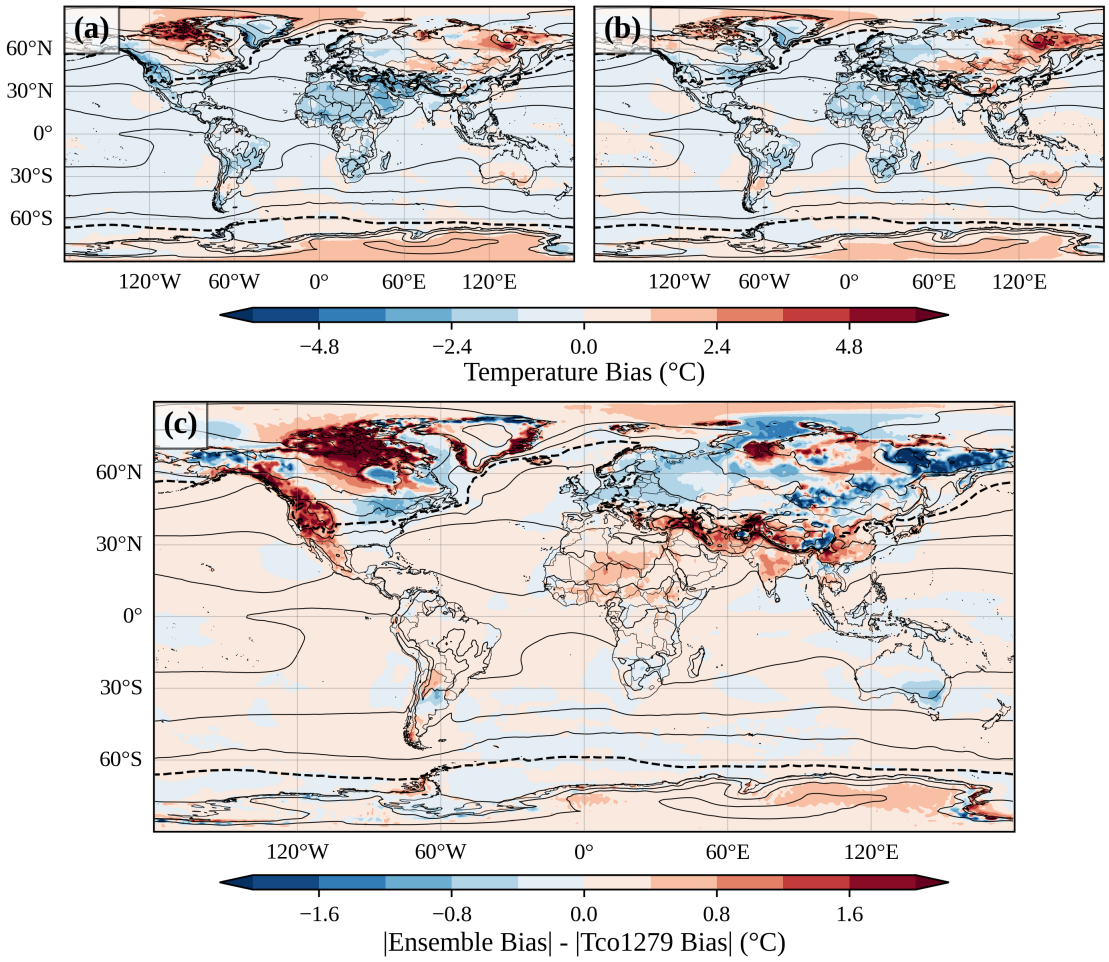

Figure 3.1: Mean 2-metre temperature bias (°C) during DJF 1980–2023 for TCo399 ensemble mean **(a)** and TCo1279 **(b)** relative to ERA5. Positive values indicate the model is warmer than reanalysis. **(c)** Shows $|(a)| - |(b)|$. The greater bias is shown by blue if TCo1279 and red for TCo399 Ensemble. Contours show the ERA5 Temperature climatology at ±(5,15,25)°C and the thick dashed line shows 0°C.

The spatial patterns of temperature bias are nearly identical between the two resolutions. The most prominent systematic bias is a warm bias of 3–6°C over Central Asia and Siberia, likely reflecting challenges in representing the stable boundary layer and surface energy balance over snow-covered continental interiors. A modest cold bias of 1–3°C appears over parts of northern North America. Critically for this study, the North Atlantic and European sectors—the regions most directly affected by NAO variability—show minimal temperature biases, typically less than 0.5°C. This suggests that the models accurately capture the thermal structure relevant to Atlantic storm track dynamics and NAO-related temperature anomalies.

The similarity of biases between TCo399 and TCo1279 indicates that increasing horizontal resolution from ∼28 km to ∼9 km does not substantially alter the representation of near-surface temperature climatology. This is consistent with temperature being primarily controlled by large-scale circulation and boundary conditions (prescribed SST, incoming radiation) rather than by mesoscale processes that higher resolution would better resolve.

Figure 3.2 (c) shows the difference in precipitation biases between the two model configurations: (ERA5 – TCo399 Ensemble) – (ERA5 – TCo1279). Blue shading indicates regions where TCo1279 has a greater bias (further from ERA5), whilst red shading indicates regions where the TCo399 ensemble mean has a greater bias. The magnitude of these differences is modest, with most regions showing differences of less than 0.6 mm day^{-1}—small relative to climatological precipitation rates and to the uncertainty in ERA5 precipitation itself (Hersbach et al. 2020).

Over the North Atlantic and European sectors, the precipitation bias differences between resolutions are minimal, typically less than 0.3 mm day^{-1}. Some tropical regions show larger differences (up to 1.8 mm day^{-1}), suggesting that increased resolution may modestly improve the representation of convective precipitation organisation and mesoscale features in these regions however, globally this is offset by the TCo1279 relative deficiency in South America, although this could be due to the fact that only one member is available, for which the evaluation of the mean bias is more uncertain. In general, the small magnitude of differences over the extratropics indicates that both resolutions capture the large-scale precipitation climatology comparably well.

The overall assessment from temperature and precipitation diagnostics is encouraging: both model configurations reproduce observed wintertime climatology with minimal systematic biases, particularly over the North Atlantic sector. The prescribed SST boundary condition strongly constrains the lower-tropospheric thermal structure, whilst the atmospheric physics—including convection, cloud, and boundary layer schemes—appears adequate for representing large-scale precipitation patterns. Importantly, the near-absence of substantial biases over the North Atlantic means that any differences in NAO characteristics, eddy feedback, or teleconnections between the models and observations or between resolutions cannot be attributed to compensating for mean state errors. This provides confidence that subsequent dynamical diagnostics reflect genuine differences in atmospheric processes rather than artefacts of biased climatology.

3.1.3 Zonal Wind

The position and intensity of the North Atlantic eddy-driven jet stream are fundamental to NAO dynamics. The jet serves as a waveguide for synoptic-scale disturbances, determines the pathways of Atlantic storm tracks, and exhibits variability that projects strongly onto the NAO pattern (Woollings et al. 2018).

Systematic biases in jet latitude have been identified as a pervasive deficiency in climate models, with many exhibiting equatorward-shifted or excessively zonal jets that compromise teleconnection strength and blocking frequency (Stronmen et al. 2020; O'Reilly et al. 2025). Accurate representation of the jet's climatological position is therefore essential for realistic simulation of eddy feedback processes, NAO variability and their associated climate impacts, in particular over Europe.

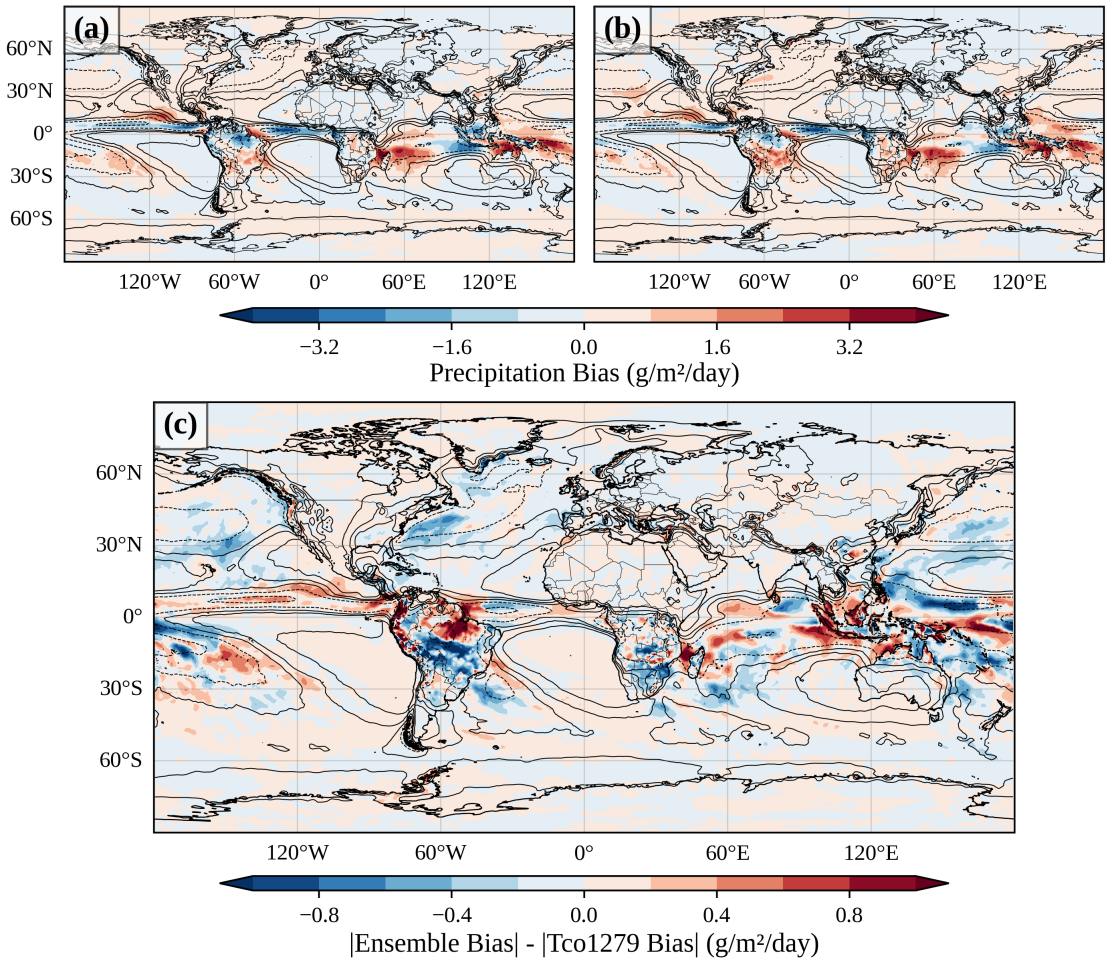

Figure 3.2: Mean precipitation biases (mm/day) during DJF 1980–2023 for TCo399 ensemble mean **(a)** and TCo1279 **(b)** relative to ERA5. **(c)** Shows $|(a)| - |(b)|$. The greater bias is shown by blue if TCo1279 and red for TCo399 Ensemble. Black contours indicate observed ERA5 precipitation climatology (solid lines: 1-3 mm/day; dashed lines: 5 mm/day increments).

Figure 3.3 presents DJF-mean zonal wind biases at 850 hPa over the North Atlantic sector for both model configurations. The 850 hPa level is selected as it captures the lower-tropospheric jet structure whilst avoiding direct surface friction effects. Both configurations exhibit a characteristic dipole bias pattern: easterly (negative, blue) biases of 0.3–0.6 m s^{-1} over the subtropical North Atlantic (30°N–45°N) and westerly (positive, red) biases of similar magnitude over the subpolar region (50°N–65°N). This dipole structure is the signature of a northward-shifted jet compared to ERA5, rather than a simple uniform speed bias (Stronmen et al. 2020).

The TCo1279 high-resolution configuration exhibits a more pronounced northward shift than the TCo399 ensemble mean. The positive biases north of 50°N reach 0.9 m/s or greater in TCo1279, compared to 0.3–0.6 m/s in the ensemble mean, whilst the negative biases to the south are similarly enhanced. This suggests that increasing horizontal resolution in this modelling system does not reduce the jet latitude bias; rather, it modestly amplifies it. This finding contrasts with some studies showing improved jet representation at higher resolution (Scaife et al. 2019), but is consistent with research demonstrating that jet biases can be

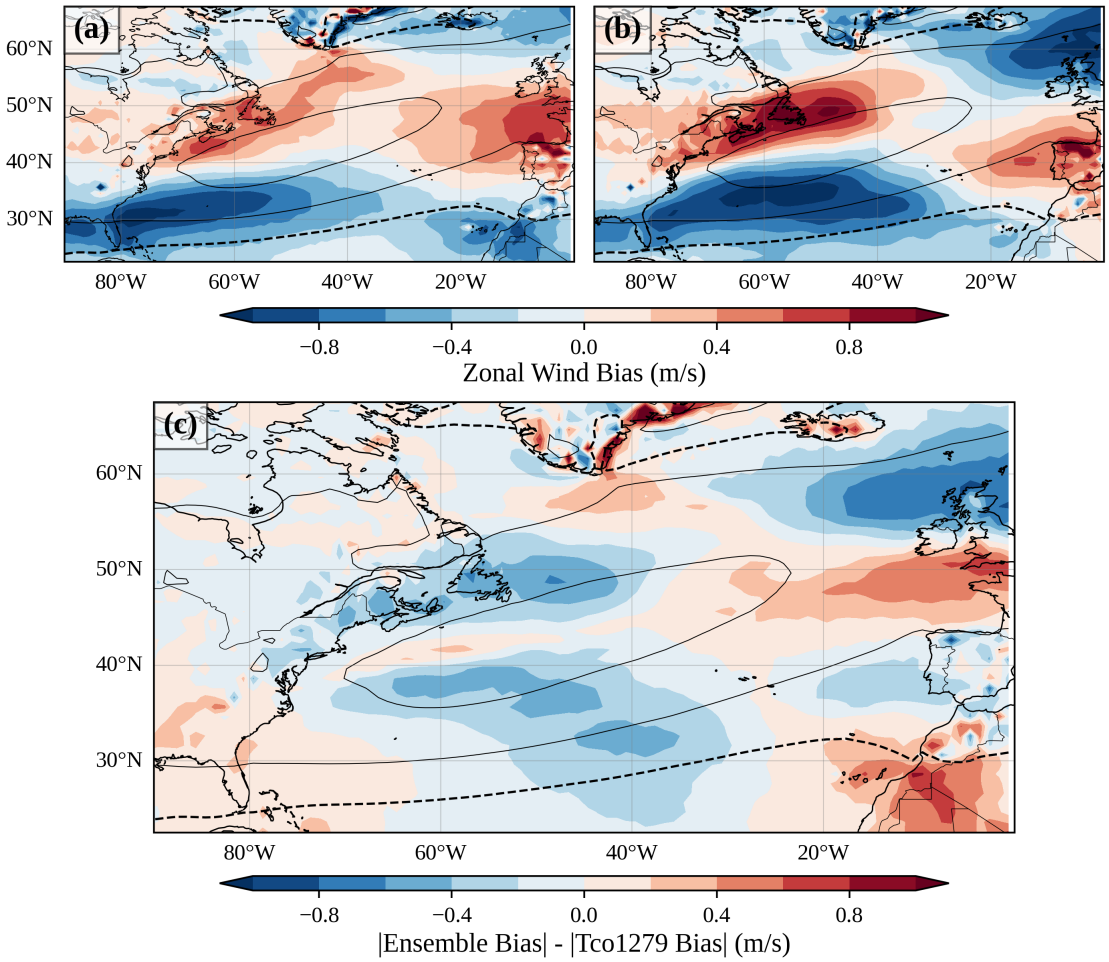

Figure 3.3: Mean zonal wind bias (m s^{-1}) at 850 hPa during DJF 1980–2023 over the North Atlantic sector for TCo399 ensemble mean **(a)** and TCo1279 **(b)**, both relative to ERA5. **(c)** Shows $|(a)| - |(b)|$. The greater bias is shown by blue if TCo1279 and red for TCo399 Ensemble. The contours show the observed climatology with 5 m/s increments. 0 m/s is shown by the dotted line.

insensitive to resolution if they stem from parametrisation deficiencies in convection, gravity wave drag, or other sub-grid processes (Stronmen et al. 2020).

Critically, the magnitude of the speed biases themselves is small: across the majority of the North Atlantic, differences remain below 1 m/s, representing less than 10% of typical climatological jet speeds (10–15 m/s at this level). Such modest speed biases are not inherently problematic for NAO simulation. What matters more for NAO dynamics is the jet's meridional position and its temporal variability (Woollings et al. 2018). The NAO reflects shifts in jet latitude: the positive phase corresponds to a northward-shifted, intensified jet, whilst the negative phase corresponds to a southward-shifted, weakened jet. A model with a modest northward bias in the climatological jet position can still accurately represent NAO variability if it correctly simulates the amplitude and timescale of meridional shifts about that biased mean state.

The observed northward jet bias has several potential implications for NAO-related processes. First, it may affect the representation of atmospheric blocking over Greenland and the Nordic Seas, as blocking frequency is sensitive to the background jet latitude (Davini et al.

2012). A more poleward jet provides less favourable conditions for blocking establishment in some configurations. Second, the bias may influence eddy propagation characteristics and eddy-mean flow feedbacks, as eddies preferentially grow in regions of enhanced baroclinicity associated with the jet (Robinson 2006). Third, teleconnection patterns that depend on Rossby wave propagation along the jet—such as the ENSO-NAO connection—may be affected by altered jet position (O'Reilly et al. 2025).

However, several factors mitigate concerns about these biases for the present analysis. The atmosphere-only configuration eliminates coupled ocean-atmosphere feedback biases that can amplify jet position errors in fully coupled models (Stronmen et al. 2020). The prescribed observed SST forces the low-level baroclinicity to remain close to reality, constraining jet position errors compared to free-running coupled simulations. Most importantly, subsequent sections will demonstrate that both configurations reproduce observed NAO spatial patterns, variability, and persistence characteristics with good accuracy despite the modest jet position bias, indicating that the models capture the fundamental dynamics governing NAO behaviour.

The resolution dependence observed here—where TCo1279 shows a slightly enhanced northward bias compared to TCo399—suggests that increased horizontal resolution alone does not automatically improve jet representation in this modelling system. This finding reinforces conclusions from Scaife et al. (2019) that resolution increases must be accompanied by appropriate re-tuning of parametrisation schemes to realise benefits for circulation biases.

3.1.4 Vertical Structure of Zonal Wind Biases

The vertical structure of zonal wind biases provides crucial insights into the three-dimensional circulation errors and their potential impacts on stratosphere-troposphere coupling, which represents a key source of NAO predictability on seasonal timescales (Scaife et al. 2016; Baldwin et al. 2001). Stratospheric variability—particularly the strength of the polar vortex and the occurrence of sudden stratospheric warmings—exerts downward influence on tropospheric circulation that persists for weeks to months, with strong vortex states favouring positive NAO phases and weak vortex states favouring negative phases (Sigmond et al. 2013).

Figure 3.4 presents latitude-height cross-sections of DJF-mean zonal wind biases over the Atlantic sector (80°W–0°W) for both model configurations. The vertical extent spans from the surface (1000 hPa) through the troposphere to the top of the stratosphere (1 hPa), with pressure plotted on a logarithmic scale to emphasise the upper levels. The black contour lines indicate the zero-bias boundary, separating regions where the model wind is too strong (westerly bias, red) from regions where it is too weak (easterly bias, blue).

In the troposphere (below 100 hPa), both configurations exhibit the dipole bias structure identified in the horizontal analysis: easterly biases of 4.5–9.0 m s^{-1} centred around 35°N–40°N, and westerly biases of similar magnitude centred around 55°N–60°N. This pattern extends coherently from the surface through the mid-troposphere (500 hPa), confirming the barotropic nature of the northward jet displacement. The TCo399 ensemble mean shows a more vertically coherent structure with the maximum biases concentrated between 400–700 hPa, whilst TCo1279 exhibits somewhat more vertical variability in the bias structure.

Crucially, substantial biases emerge in the stratosphere (above 100 hPa) for both configurations. The TCo399 ensemble mean simulation shows particularly pronounced stratospheric negative biases exceeding 6.0 m/s poleward of 50°N throughout the stratosphere, extending

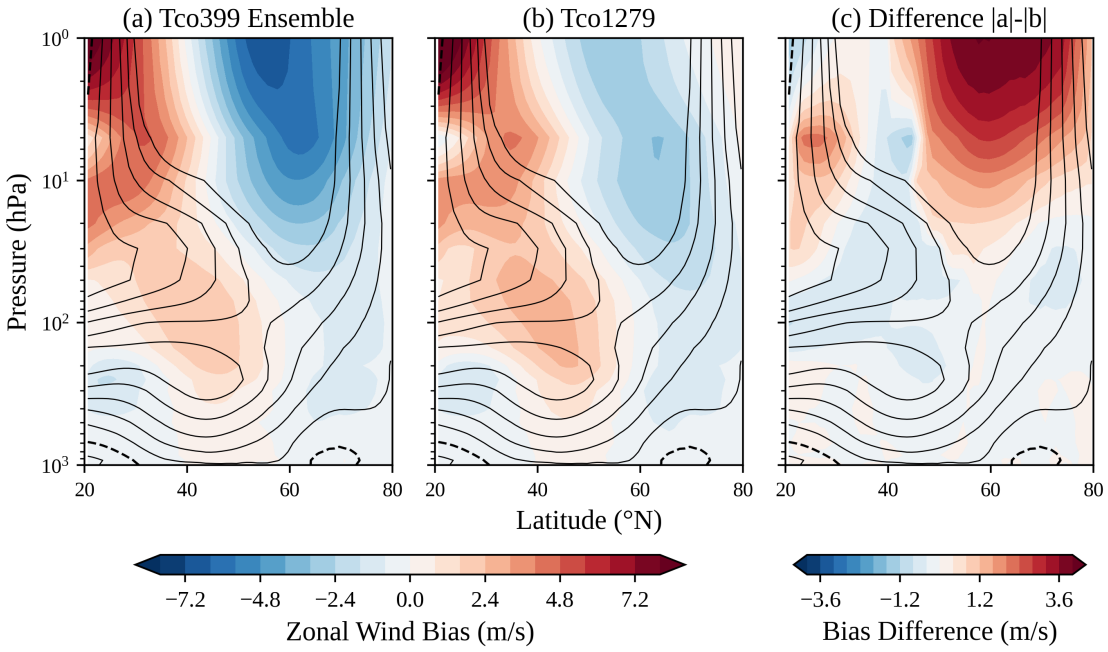

Figure 3.4: Latitude-height cross-section of DJF - mean zonal wind bias (m s^{-1}) over the Atlantic sector (80°W–0°W) for TCo399 ensemble mean **(a)** and TCo1279 **(b)**, both relative to ERA5 for 1980–2023. The bias between the two ($|(a)| - |(b)|$) is shown in **(c)**. Pressure is plotted on a logarithmic scale. The contours show the observed climatology with 5 m/s increments, 0 m/s is shown by the dotted line.

to the uppermost levels shown (1 hPa). The TCo1279 exhibits more modest negative biases of 2–3 m/s in the same region. These stratospheric biases indicate that the models simulate a systematically weaker polar vortex than observed in ERA5 during boreal winter, a problem that is more prominent in the TCo399 runs.

A weaker-than-observed polar vortex has several important implications for NAO simulation and predictability. First, it suggests that the models may underestimate the frequency or intensity of sudden stratospheric warmings—the rapid breakdowns of the polar vortex that provide substantial predictability for negative NAO phases (Scaife et al. 2016). Scaife et al. (2016) demonstrated that removing winters containing sudden stratospheric warmings from seasonal hindcasts eliminates virtually all NAO prediction skill, highlighting the critical importance of correctly representing stratospheric variability. Second, a persistently weak vortex bias may favour negative NAO phases through downward coupling, potentially skewing the distribution of NAO states. Third, biases in stratospheric winds affect the propagation of planetary waves from the troposphere, as the ability of waves to propagate vertically depends critically on the background zonal wind structure (Andrews et al. 2019).

However, it is important to contextualise these stratospheric biases within the broader understanding of reanalysis uncertainty. Hersbach et al. (2020) noted that ERA5 exhibits a known cold bias in the upper stratosphere, which is coupled to wind biases through thermal wind balance. Additionally, stratospheric representation in reanalysis is constrained primarily by satellite radiance observations rather than direct wind measurements, introducing model-dependent elements even in the "observational" reference. The progression from TCo399 to TCo1279 shows reduced stratospheric biases, indicating that increased horizontal resolution does produce improved stratospheric representation without corresponding adjustments to

upper-level parametrisations such as gravity wave drag and stratospheric ozone chemistry.

The presence of systematic stratospheric biases raises questions about potential impacts on NAO predictability through the stratospheric pathway. The coherent vertical structure of the tropospheric jet bias—extending from surface through the upper troposphere—indicates a fundamental dynamical rather than surface-layer issue. This barotropic structure is consistent with the bias arising from systematic errors in eddy-mean flow interactions or planetary wave propagation rather than from boundary layer processes. The examination of eddy feedback parameters in subsequent sections will investigate whether the northward-shifted, stronger-than-observed configuration affects the positive feedback between transient eddies and the mean flow that is essential for NAO growth and maintenance (Lorenz et al. 2003).

3.2 NAO Representation

Having established that both model configurations reproduce the observed climatological mean state with minimal systematic biases over the North Atlantic sector, we now examine their representation of NAO variability. The NAO index was calculated following the station-based methodology described in Chapter 2, using MSLP values from grid points closest to Reykjavik (64.1°N, 21.9°W) and Lisbon (38.7°N, 9.1°W). The analysis encompasses all 43 available DJF winters from 1980/81 through 2022/23.

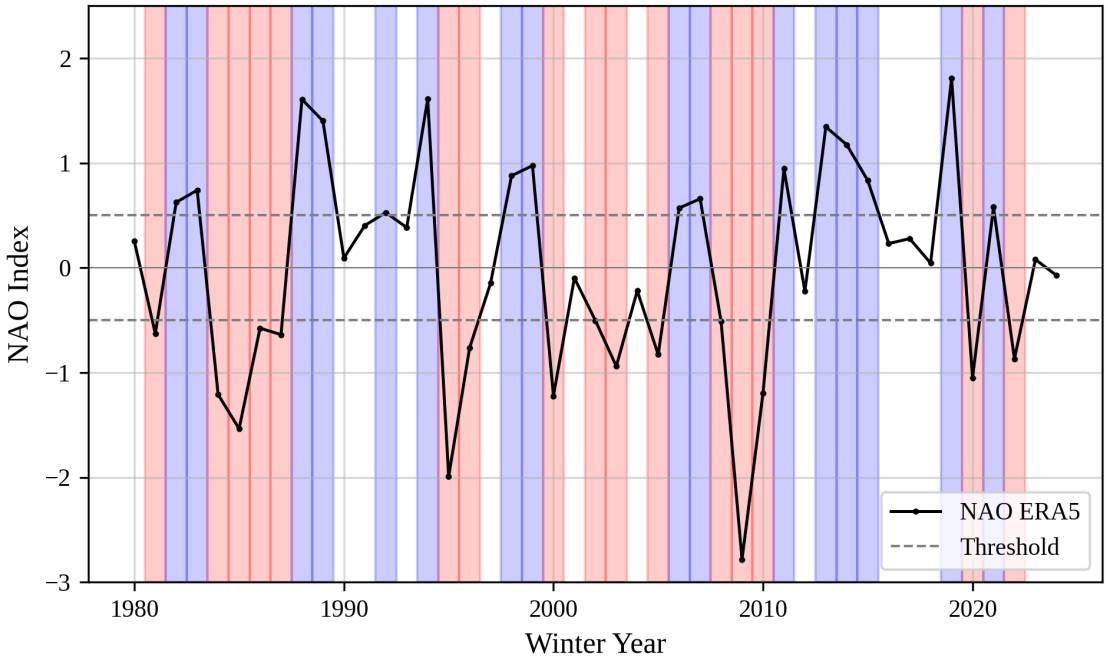

Figure 3.5: NAO calculated from ERA5 for DJF winters 1980–2023. Red and blue dotted lines indicate the ±0.5 threshold, with shading for years that exceed it.

3.2.1 Observed NAO Variability

Figure 3.5 presents the ERA5 NAO time series, revealing the characteristic high-frequency variability that defines this mode of atmospheric circulation. The observed NAO exhibits substantial interannual variability ranging from strongly negative phases (index < -2.0) to

strongly positive phases (index > +1.5) over the 43-year period. Notably, the period shows a predominance of positive NAO winters during the late 1980s and early 1990s, consistent with the well-documented trend towards the positive phase during this epoch (Hurrell 1995; Osborn et al. 2006). The most extreme negative NAO winter occurred in 2009/10 (index -2.8), a year noted for severe cold outbreaks across Northern Europe. Conversely, the strongest positive phase occurred in 2019/20 (index +1.9), associated with exceptionally mild and wet conditions over Northern Europe.

The temporal distribution shows that approximately 30% of winters (13 of 43) exceed the +0.5 threshold, whilst 26% (11 of 43) fall below -0.5, with the remaining 44% classified as neutral. This asymmetry towards positive phases during the analysis period reflects documented multi-decadal variability in NAO behaviour (Hurrell et al. 2003). Persistence characteristics are evident in the time series, with positive phases often sustained across multiple consecutive winters (e.g., 1988–1995) and negative phases similarly clustered. However, rapid year-to-year transitions also occur frequently, demonstrating the fundamentally chaotic nature of extratropical atmospheric variability that limits predictability beyond seasonal timescales (Feldstein 2003).

3.2.2 Simulated NAO Variability

Figure 3.6 displays the TCo399 ensemble NAO indices, with individual members shown as thin lines, the ensemble mean as the thick red line, and the $\pm 1\sigma$ ensemble spread shaded in grey. The ensemble exhibits several noteworthy characteristics. First, individual members show substantial year-to-year variability comparable to observations, with extreme values occasionally exceeding ±2.0 standard deviations. This demonstrates that the model physics generate realistic temporal fluctuations in the NAO dipole structure. Second, the ensemble spread is considerable, with individual members frequently disagreeing on the sign of the NAO phase for any given winter. This large spread reflects the dominance of internal atmospheric variability over external forcing in determining NAO behaviour on seasonal timescales.

Critically, the ensemble mean NAO (thick red line) exhibits substantially reduced amplitude compared to both the individual members and the ERA5 observations. This reduction arises from the averaging process: when individual members disagree on the NAO phase, their contributions partially cancel in the ensemble mean. The ensemble mean rarely exceeds ±0.5 standard deviations, despite individual members frequently reaching ±1.5 or beyond. In atmosphere-only simulations this behaviour implies that NAO variability is dominated by unpredictable internal dynamics rather than by responses to prescribed boundary forcing (Weisheimer et al. 2020).

Figure 3.7 directly compares the NAO indices from ERA5 (black), TCo1279 (blue), and the TCo399 ensemble mean (red). It is clear that there is no relationship between any of the series, however it demonstrates that the TCo1279 produces a realistic temporal variability.

Table 3.2 quantifies the temporal correlation between modelled and observed NAO indices. The TCo1279 high-resolution simulation achieves a correlation of r = 0.11 (p = 0.46) with ERA5, indicating essentially no linear relationship between the simulated and observed year-to-year NAO evolution. The TCo399 ensemble mean similarly shows r = 0.11 (p = 0.47), whilst individual ensemble members exhibit correlations ranging from r = -0.23 to r = 0.30, with only member r06 approaching marginal statistical significance (r = 0.30, p = 0.04),

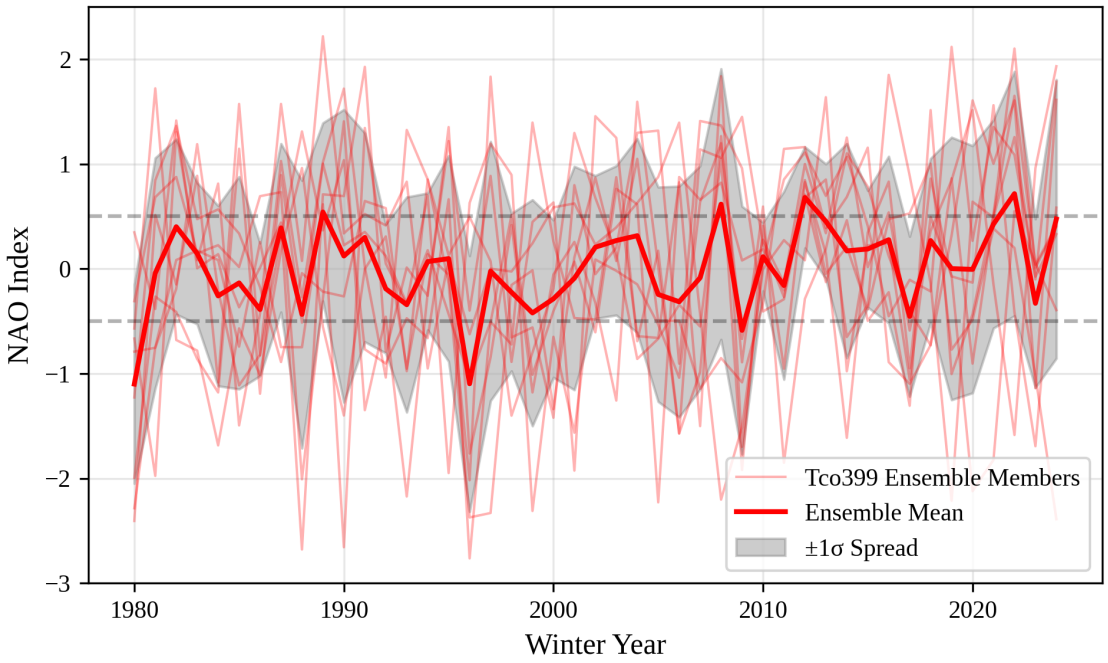

Figure 3.6: NAO calculated from the TCo399 ensemble for DJF winters 1980–2023. Thin red lines show individual ensemble members, thick red line shows the ensemble mean, and grey shading indicates $\pm 1\sigma$ ensemble spread. Dotted lines indicate the ± 0.5 threshold.

although this is most probably by chance, given the results of the other members.

Table 3.2: Correlation between modelled and observed NAO indices for DJF 1980–2023. Bold p-values indicate statistical significance at $p < 0.05$.

	TCo1279	**TCo399**										
		Mean	r01	r02	r03	r04	r05	r06	r07	r08	r09	r10
ERA5 correlation	0.11	0.11	-0.09	0.02	0.19	0.25	-0.23	0.30	-0.22	0.17	0.02	0.02
p-value	0.46	0.47	0.56	0.89	0.23	0.10	0.13	**0.04**	0.14	0.27	0.89	0.87

The critical question is whether - despite not representing the observed NAO sequences - the model at least generates NAO variability with realistic statistical properties. The fact that correlations hover near zero (rather than being systematically negative or positive) confirms that the models produce NAO-like variability that is neither systematically in-phase nor out-of-phase with observations, but rather represents plausible alternative realisations of the climate system forced by the same boundary conditions. This suggests that the models correctly represent the balance between forcing (baroclinic instability, eddy momentum transport) and damping (friction, diabatic heating) that determines NAO amplitude (Lorenz et al. 2003).

Comparing the TCo1279 single-member simulation with the TCo399 ensemble members reveals no systematic improvement in correlation skill with increased resolution. The high-resolution simulation ($r = 0.11$) performs equivalently to the ensemble mean ($r = 0.11$) and falls within the range of correlations exhibited by individual TCo399 members (-0.23 to +0.30). This equivalence suggests that horizontal resolution increases from ∼27 km to ∼9 km do not substantially enhance the atmospheric response to prescribed SST forcing in terms of NAO phase selection. This finding aligns with previous research demonstrating that

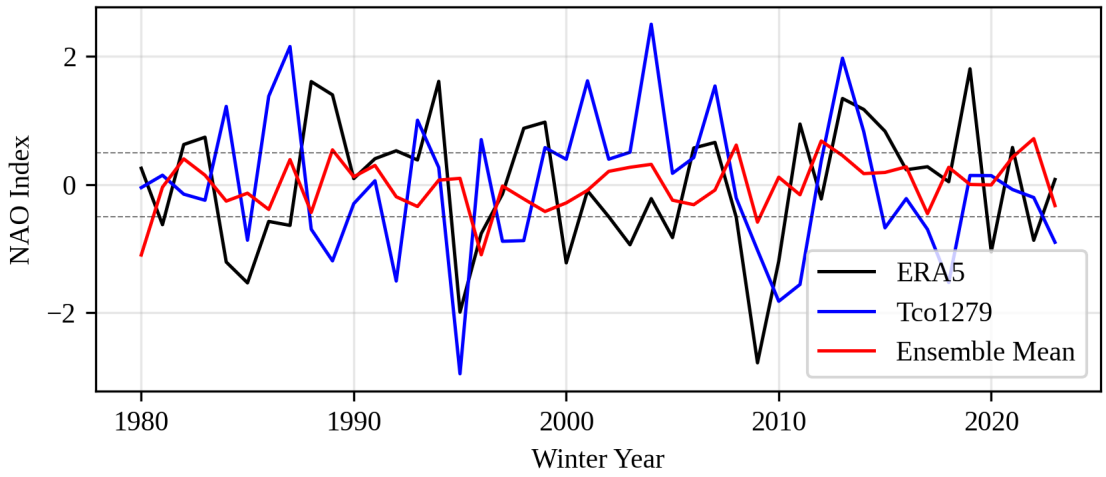

Figure 3.7: Direct comparison of NAO indices: ERA5 (red), TCo1279 (blue), and TCo399 ensemble mean (green) for DJF winters 1980–2023. Dotted lines indicate the ±0.5 threshold.

NAO predictability on seasonal timescales depends more critically on the initialisation of the atmosphere, the representation of eddy feedback mechanisms, stratospheric coupling, and teleconnection pathways than on resolution alone (Scaife et al. 2019; O'Reilly et al. 2025).

However, as mentioned earlier, correlation with observations represents only one aspect of model fidelity. Subsequent sections will examine whether the models correctly represent the spatial structure of NAO patterns, the eddy feedback mechanisms that maintain NAO anomalies, and the teleconnection pathways through which remote forcing influences North Atlantic circulation.

3.3 Resolution Comparison

3.3.1 Eddy Kinetic Energy

Synoptic-scale EKE provides a direct measure of transient atmospheric disturbances and storm track activity, serving as a fundamental diagnostic for assessing how well models represent the energetics of weather systems. High EKE values indicate regions of intense cyclonic activity where synoptic eddies extract energy from the mean flow through baroclinic instability. For NAO dynamics, the North Atlantic storm track—characterised by a prominent EKE maximum extending from the eastern seaboard of North America towards northwestern Europe—represents the primary source of high-frequency forcing that maintains low-frequency NAO anomalies. We examine EKE distributions at two atmospheric levels: 500 hPa, which captures mid-tropospheric storm track structure where baroclinic development is most active (Lau 1988; Gastineau et al. 2013), and 200 hPa, which represents the upper-tropospheric jet exit region where eddies transfer momentum to the large-scale flow (DeWeaver et al. 2000; Hoskins et al. 1983).

500 hPa Eddy Kinetic Energy

Figure 3.8 presents the North Atlantic distribution of DJF-mean EKE biases at 500 hPa for both model configurations relative to ERA5. The observed EKE climatology exhibits a

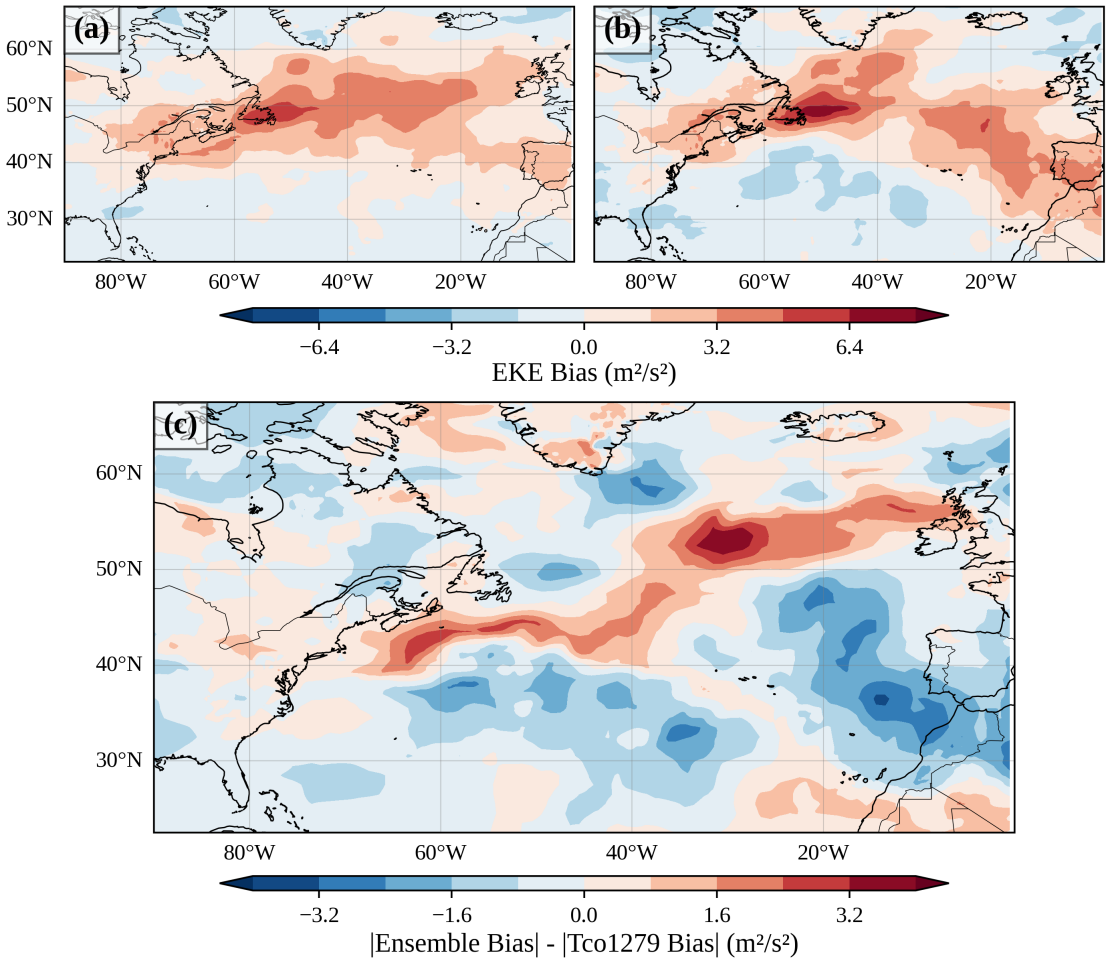

Figure 3.8: Mean EKE bias (m^2s^{-2}) at 500 hPa during DJF 1980–2023 over the North Atlantic sector for TCo399 ensemble mean **(a)** and TCo1279 **(b)** relative to ERA5. Positive values (red) indicate excessive eddy activity; negative values (blue) indicate deficient activity. **(c)** Shows the difference in bias magnitude, with blue indicating larger TCo1279 bias and red indicating larger TCo399 bias.

well-defined storm track maximum extending from Cape Hatteras towards Iceland and the Norwegian Sea, with peak values typically exceeding 200 m^2s^{-2}.

The TCo399 ensemble mean (panel a) exhibits a coherent band of excessive EKE extending from the western Atlantic near 40°N eastward across the basin towards the British Isles. Peak positive biases of 5–6 m^2s^{-2} occur along the main storm track axis around 45–50°N, particularly concentrated over the eastern North Atlantic. Notably, the bias pattern shows spatial heterogeneity: whilst the core storm track shows excessive activity, regions to the south (around 30–40°N) and in parts of the subpolar North Atlantic display negative biases (blue shading), indicating deficient eddy activity in these locations.

The TCo1279 configuration (panel b) displays a qualitatively different bias structure. The positive bias along the main storm track is greater in magnitude and reduced spatial extent compared to the ensemble mean. However, TCo1279 develops a more complex pattern with stronger negative biases appearing in the subtropical Atlantic (30–40°N) and positive biases emerging over western Europe and the Iberian Peninsula. This suggests that increased resolution not only affects the magnitude of EKE but also modifies its spatial distribution,

potentially shifting the storm track position or altering the preferred regions of cyclogenesis.

The difference plot (panel c) reveals a spatially complex resolution dependence. Varied shading over the central North Atlantic storm track core around 40 - 55°N, 40 - 10°W, suggests that TCo1279 exhibits no clear advantage over the ensemble mean in this critical region—representing approximately $2 - 3m^2s^{-2}$ variation. However, red shading appears over the eastern Atlantic and European sector around 50 - 55°N, indicating that the ensemble mean configuration develops larger biases in these northern downstream regions. The western subtropical Atlantic (south of 40°N) shows predominantly blue shading, suggesting TCo1279 poorly realises the reduced eddy activity in this region where observed EKE values are inherently lower.

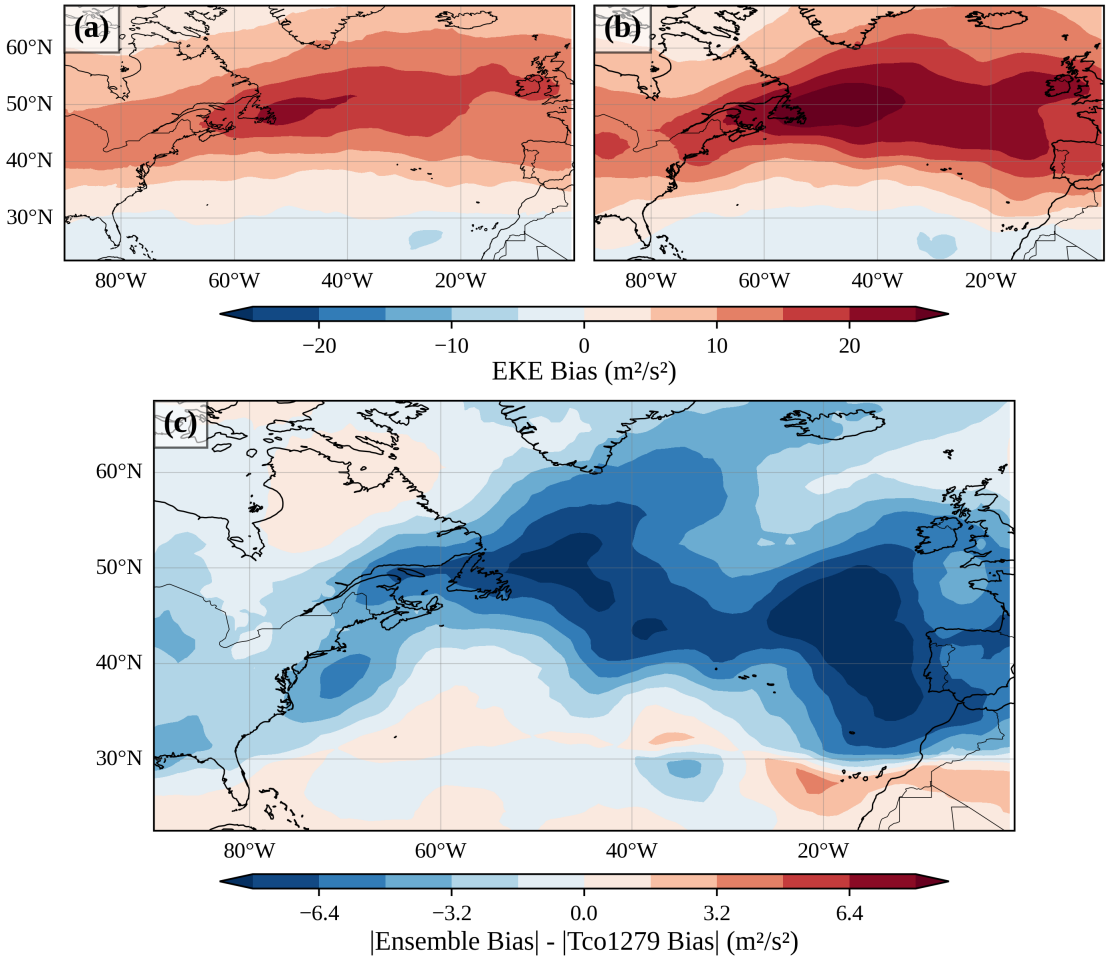

Figure 3.9: Mean EKE bias (m^2s^{-2}) at 200 hPa during DJF 1980–2023 over the North Atlantic sector for TCo399 ensemble mean **(a)** and TCo1279 **(b)** relative to ERA5. **(c)** Shows the difference in bias magnitude, with blue indicating smaller TCo1279 bias and red indicating smaller TCo399 bias.

The mixed nature of the resolution impact—improvement in some regions, degradation in others—suggests that increased resolution modifies the spatial organisation of storm track activity rather than producing uniform improvements. The central Atlantic improvements may reflect better representation of baroclinic growth processes or cyclogenesis mechanisms, whilst the eastern Atlantic changes could indicate altered downstream development or modified eddy propagation characteristics at higher resolution.

200 hPa Eddy Kinetic Energy

Figure 3.9 presents the corresponding analysis at 200 hPa, revealing systematically larger bias magnitudes than observed at 500 hPa. Both model configurations exhibit enhanced positive biases in the upper troposphere, with the TCo1279 showing excessive EKE of 10–25 m^2s^{-2} across broad regions of the North Atlantic storm track—approximately 2–3 times larger than the 500 hPa biases.

The TCo399 ensemble mean (panel a) shows a broad region of excessive upper-level EKE extending across the entire North Atlantic basin between 40–60°N, with peak positive biases up to 20 m^2s^{-2} over the central and eastern sectors. TCo1279 (panel b) exhibits even further exageration in these upper-level biases. The region of excessive EKE expands spatially and increases in magnitude, with peak biases now greater than 25 m^2s^{-2}. The difference plot (panel c) shows extensive blue shading across the North Atlantic, confirming that the high-resolution configuration produces greater biases throughout most of the domain. The magnitude of the relative excessive EKE in the TCo1279 ranges from 3–6 m^2s^{-2} across large areas, representing approximately 30–40% bias. Unlike the 500 hPa level where the resolution impact was spatially mixed, the 200 hPa appear more systematic and widespread.

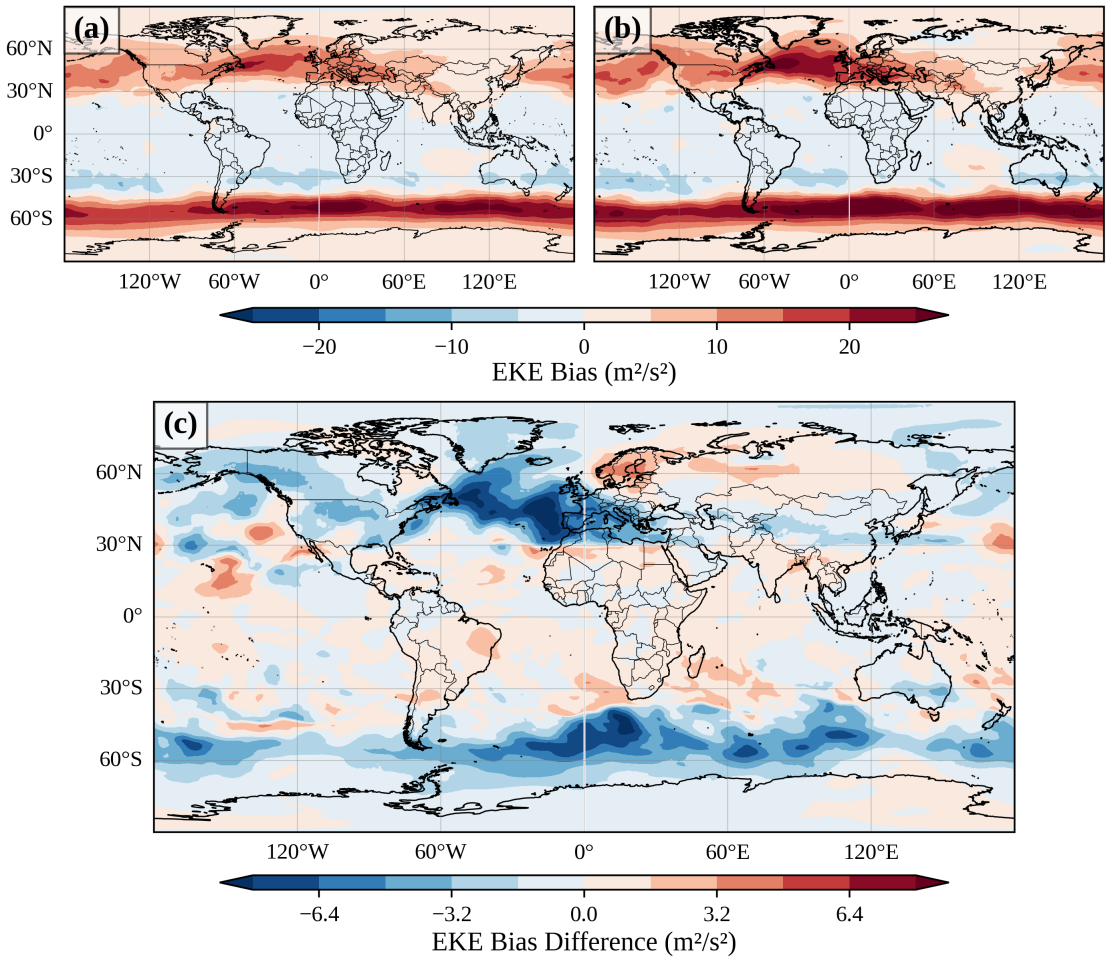

Figure 3.10: Mean EKE bias (m^2s^{-2}) at 200 hPa globally during DJF 1980–2023 for TCo399 ensemble mean **(a)** and TCo1279 **(b)** relative to ERA5. **(c)** Shows the difference in bias magnitude, with blue indicating smaller TCo1279 bias and red indicating smaller TCo399 bias.

The enhanced bias magnitudes at 200 hPa compared to 500 hPa in both configurations suggest that the models overestimate the efficiency of upward energy transfer during baroclinic development, producing excessively energetic upper-tropospheric eddies. The systematic over-estimation at upper levels with increased resolution indicates that TCo1279 poorly regulates this vertical energy transfer.

Global Context

Analysis of global EKE distributions (Figure 3.10) confirms that the North Atlantic bias patterns form part of a hemisphere-wide tendency towards excessive storm track activity in both configurations. The North Pacific exhibits similar positive biases to the North Atlantic, whilst the Southern Ocean storm track shows a more uniform band of enhanced EKE. The resolution dependence observed in the North Atlantic — increased biases in TCo1279—extends to these other storm track regions as well, though with varying magnitudes across different basins.

The EKE analysis reveals systematic positive biases in both model configurations, with excessive eddy activity along the primary Northern Hemisphere storm tracks. These biases are more pronounced at 200 hPa than at 500 hPa, suggesting overestimation of vertical energy transfer during baroclinic development. The TCo1279 high-resolution configuration shows substantial over-estimation of the EKE at upper levels, with 30–40% bias across the North Atlantic. At 500 hPa, the resolution impact is more spatially complex, with improvements in the central Atlantic storm track core but some degradation in downstream regions. The subsequent analysis of eddy feedback parameters will examine whether these changes in eddy energetics have an impact in the representation of eddy-mean flow coupling.

3.3.2 Eddy Feedback Parameter

The EFP quantifies the degree to which transient eddies reinforce large-scale atmospheric anomalies through momentum transport. As explain in Chapter 2, the EFP is calculated as the squared correlation coefficient between DJF-mean zonal wind and horizontal Eliassen-Palm flux divergence at each latitude. This diagnostic aims to capture the fundamental coupling between synoptic-scale eddies and the mean flow: positive feedback occurs when eddy momentum fluxes systematically accelerate regions of strong westerlies and decelerate regions of weak westerlies, thereby amplifying existing circulation anomalies rather than damping them towards climatology. Values closer to unity indicate stronger, more realistic dynamical coupling, whilst lower values suggest insufficient eddy-mean flow interaction.

Recent research has identified deficient eddy feedback as a primary physical mechanism underlying the signal-to-noise problem in seasonal predictions (Hardiman et al. 2022). Models with stronger EFP values approaching observations show enhanced teleconnection strength and larger predictable signals. The analysis here examines whether increased atmospheric resolution from ∼27 km to ∼9 km improves the representation of this critical dynamical process at two key levels: 500 hPa, representative of mid-tropospheric dynamics, and 200 hPa, capturing upper-tropospheric jet-eddy interactions.

500 hPa Eddy Feedback

Figure 3.11 presents the latitudinal structure of the EFP at 500 hPa, comparing ERA5 (black) with TCo1279 (blue), the TCo399 ensemble mean (red), and individual TCo399

members (thin red lines). The vertical dashed lines at 25°N and 72°N delineate the Northern Hemisphere extratropical region over which area-weighted EFP averages are computed in the literature.

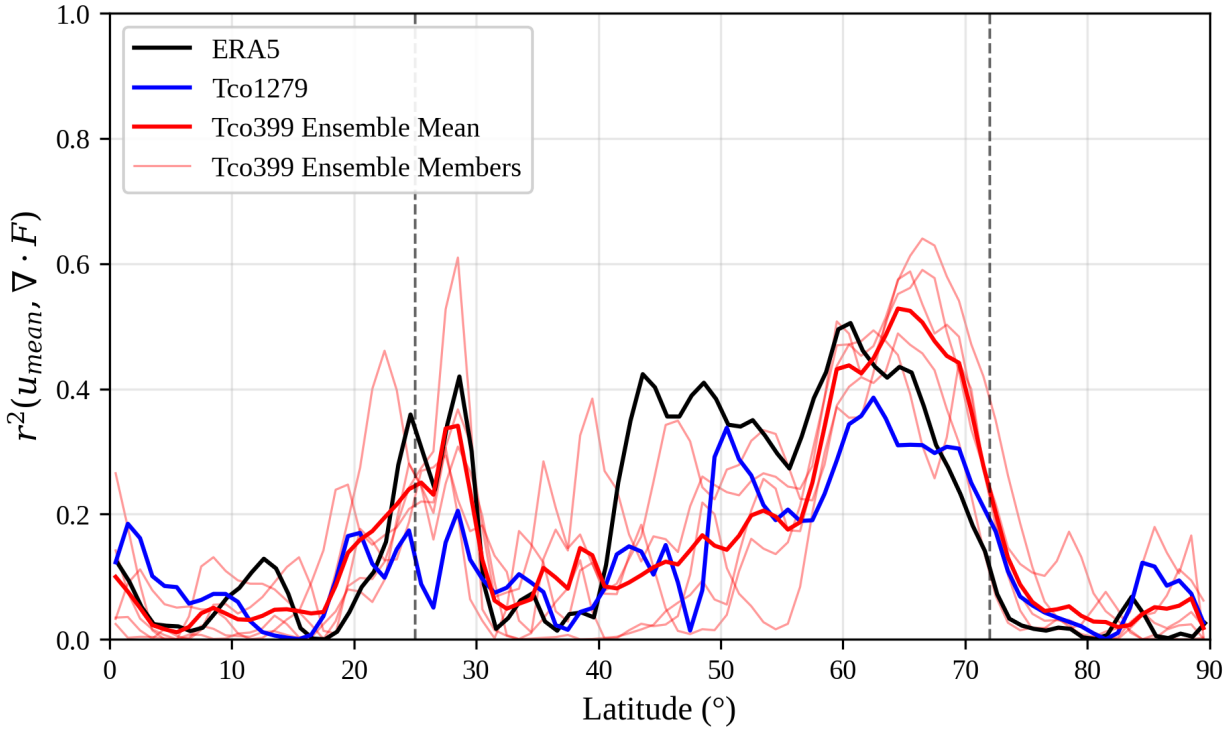

Figure 3.11: Eddy feedback parameter (r^2) at 500 hPa as a function of latitude for DJF 1980–2023. Black line shows ERA5, blue shows TCo1279, red shows TCo399 ensemble mean, and thin red lines show individual TCo399 members. Vertical dashed lines at 25°N and 72°N indicate the extratropical averaging domain. Higher values indicate stronger eddy-mean flow coupling.

The ERA5 reanalysis exhibits substantial latitudinal structure, with EFP values ranging from near zero in the tropics and high latitudes to peak values of 0.45–0.50 in the mid-latitude storm track regions around 35–65°N. The maximum occurs near ≈60°N, coinciding with the climatological position of the North Atlantic eddy-driven jet. This latitudinal structure reflects the geographical distribution of eddy-mean flow coupling strength, with the strongest feedback occurring where synoptic eddies are most active and where the background flow provides optimal conditions for eddy amplification.

Both model configurations capture the general latitudinal structure of the EFP, with peaks in the mid-latitudes and reduced values in the tropics and polar regions. However, both systematically underestimate the EFP magnitude across the mid latitudes. The TCo399 ensemble mean (red line) shows peak values of only ≈0.2 in the critical 40–55°N band, representing approximately 35–50% reduction compared to ERA5. In the high latitudes, the ensemble mean is greater than ERA5 with the peak offset to the north and dropping off later. Individual ensemble members (thin red lines) exhibit considerable spread, with some members achieving EFP values exceeding 0.6, exceeding observations especially in the high latitudes. This ensemble spread indicates substantial sampling variability in the EFP calculation arising from the finite 43-winter analysis period and the chaotic nature of atmospheric variability.

The TCo1279 high-resolution configuration (blue line) shows a complex resolution dependence. TCo1279 exhibits slightly lower EFP values than the TCo399 ensemble mean and individual simulations, suggesting that increased resolution does affect the representation of the mid-tropospheric eddy feedback. However, it is hard to conclude whether it represents a more realistic EFP, as the best performing configuration depends on the latitude. In particular, we see potential relative improvements of TCo1279 in high-latitude regions where diabatic processes, orographic effects, or ice-edge dynamics may be better resolved at finer scales.

Critically, the deficient EFP in the models compared to ERA5—particularly in the 40–60°N band that encompasses the North Atlantic storm track and NAO centres of action—confirms the systematic underrepresentation of eddy-mean flow coupling identified in previous studies. The fact that individual ensemble members occasionally match or exceed observed EFP values whilst the ensemble mean remains deficient suggests that sampling uncertainty contributes to the apparent bias, though the systematic tendency towards lower values across most members indicates a genuine model deficiency rather than purely statistical fluctuation.

200 hPa Eddy Feedback

Figure 3.12 presents the corresponding analysis at 200 hPa, revealing markedly different characteristics compared to the mid-troposphere. The upper-tropospheric EFP exhibits slightly larger magnitudes and a distinctly different latitudinal structure.

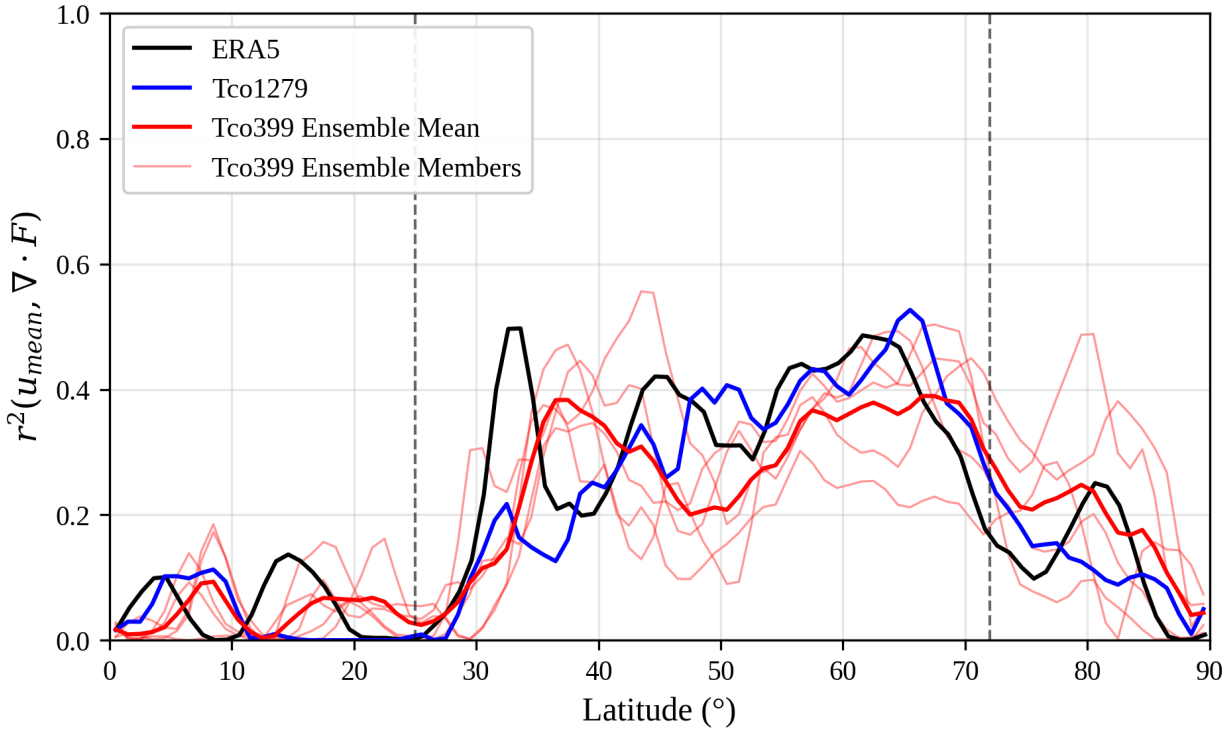

Figure 3.12: Eddy feedback parameter (r^2) at 200 hPa as a function of latitude for DJF 1980–2023. Black line shows ERA5, blue shows TCo1279, red shows TCo399 ensemble mean, and thin red lines show individual TCo399 members. Vertical dashed lines at 25°N and 72°N indicate the extratropical averaging domain. Higher values indicate stronger eddy-mean flow coupling.

ERA5 shows peak EFP values of ≈0.50 concentrated in the subtropical jet region around

30–40°N, with a secondary maximum of 0.45–0.50 in the mid-latitudes around 65°N. The dual-peak structure reflects the distinct dynamical regimes of the subtropical jet (thermally driven) and the eddy-driven polar jet, each associated with different eddy-mean flow interaction characteristics. Between these peaks, around 35–50°N, the EFP exhibits a local minimum of 0.20–0.35, corresponding to the transition region between the two jet regimes.

The model configurations show quite different behaviour at this level compared to 500 hPa. The TCo399 ensemble mean exhibits consistently high EFP values throughout the mid-latitudes (40–70°N), with individual members frequently matching or exceeding ERA5 values. Peak ensemble mean values reach ≈ 0.40 around 55–65°N, slightly deficient compared to observations, and remain higher than ERA5 poleward of 66°N. Ensemble members show significant spread at this level, with some members achieving EFP values of 0.55 — higher than observed in ERA5.

The TCo1279 high-resolution configuration also shows a different behaviour at 200 hPa than at 500 hPa. Between approximately 25–40°N, TCo1279 shows notably reduced EFP values compared to both the ensemble mean and observations, with a pronounced minimum around 32°N where the EFP is significantly lower than both. However, poleward of 40°N, TCo1279 shows enhanced EFP values approaching or reaching 0.50–0.55, comparable to the ensemble mean and observations. The differences between both configurations, however, might be due to internal variability because TCo1279 values are generally within the spread of the 5 TCo399 members.

Vertical Structure and Resolution Dependence

Comparing the 500 hPa and 200 hPa results reveals fundamental differences in eddy feedback representation across the tropospheric column. At mid-levels, both configurations systematically underestimate eddy feedback in the critical 40–60°N storm track region, with TCo1279 showing no systematic improvement over the ensemble mean. At upper levels, the models achieve or exceed observed EFP magnitudes in the mid-to-high latitudes (50–70°N), with the TCo1279 proving a very close result, suggesting that upper-tropospheric eddy-jet coupling is better represented—or potentially overrepresented—compared to mid-tropospheric processes.

Table 3.3: Comparison of area weighted EFP values for ERA5, TCo1279 and TCo399 Ensemble at 200 and 500 hPa.

Pressure Level	**Calculated EFP**							
	ERA5	**TCo1279**	**TCo399**					
			Mean	**r01**	**r02**	**r03**	**r04**	**r05**
200 hPa	0.308	0.268	0.257	0.246	0.301	0.199	0.258	0.281
500 hPa	0.264	0.161	0.208	0.169	0.249	0.179	0.252	0.192

This vertical discrepancy has important dynamical implications. Mid-tropospheric eddy feedback at 500 hPa primarily reflects the barotropic component of eddy-mean flow interaction, where momentum fluxes directly reinforce or weaken the large-scale wind field. Upper-tropospheric feedback at 200 hPa additionally involves baroclinic processes and the vertical propagation of eddy momentum fluxes from lower levels. The fact that models better capture upper-level feedback whilst underestimating mid-level feedback suggests that the vertical

structure of eddy forcing may be distorted, with consequences for how effectively eddies can drive tropospheric circulation anomalies.

The resolution dependence differs markedly between levels. At 500 hPa, increased resolution produces a negative systematic change in EFP magnitude across most latitudes, with modest results only appearing poleward of 60°N. At 200 hPa, resolution impacts are larger but spatially complex: degradation in the subtropical transition zone (25–45°N) but maintenance of strong feedback in the higher latitudes (50–70°N). This suggests that the dynamical processes governing eddy feedback respond differently to resolution changes depending on both altitude and latitude, potentially reflecting different sensitivities of baroclinic versus barotropic processes, or differences in how subtropical versus polar jet dynamics are affected by resolution.

The large ensemble spread observed at both levels, particularly pronounced at 200 hPa, highlights the substantial sampling uncertainty inherent in EFP calculations from 43-winter climatologies. Individual ensemble members can deviate substantially from the ensemble mean, with some members exhibiting EFP values that match or exceed observations even when the ensemble mean shows deficiencies. This sampling variability complicates the assessment of resolution impacts: differences between TCo1279 and the ensemble mean may partially reflect sampling fluctuation rather than systematic resolution-dependent changes. Longer simulation periods would thus be required to robustly quantify the magnitude of resolution-induced EFP changes.

Implications for NAO Dynamics

The deficient mid-tropospheric eddy feedback documented here, particularly in the 40–60°N latitude band encompassing the NAO centres of action, provides a mechanistic link to potential limitations in NAO representation. The 500 hPa level captures the core of the baroclinic zone where eddies interact most directly with the surface pressure field that defines the NAO (Hurrell et al. 2009; Scaife et al. 2019; Wanner et al. 2001). Insufficient coupling at this level means that transient eddies provide weaker forcing to the NAO dipole structure than observed, potentially affecting both the amplitude and persistence of NAO anomalies.

The fact that increased resolution from TCo399 to TCo1279 does not systematically improve mid-tropospheric EFP suggests that horizontal resolution alone — at least within the range examined here — is insufficient to address the eddy feedback deficiency. This finding aligns with Scaife et al. (2019), who found that doubling atmospheric resolution showed limited improvements in NAO signal-to-noise ratios despite incremental eddy feedback enhancements. The implication is that fundamental aspects of how models represent eddy-mean flow coupling—potentially involving parametrisation schemes, numerical diffusion, or the representation of diabatic processes—require attention beyond simply increasing resolution.

3.3.3 Teleconnections

Teleconnection patterns provide crucial pathways through which NAO predictability can arise from remote forcing sources (Hardiman et al. 2019). The ENSO-NAO teleconnection, whilst weaker and less spatially coherent than the ENSO-PNA connection, represents a potentially important source of seasonal predictability for the Euro-Atlantic sector. Recent research has demonstrated that seasonal forecast models substantially underrepresent the strength of this teleconnection, with the Pacific-to-Atlantic link appearing particularly deficient (O'Reilly et al.

2025). This section examines whether the AMIP-EERIE simulations accurately reproduce the observed teleconnection patterns and whether increased resolution improves the representation of these remote influences.

Upper-Tropospheric Jet Response

Figure 3.13 presents the regression of 200 hPa zonal wind onto the standardised Niño 3.4 index, revealing the jet stream response to ENSO forcing. The polar projection encompasses the Northern Hemisphere extratropics, enabling assessment of both the Pacific and Atlantic jet responses within a single view.

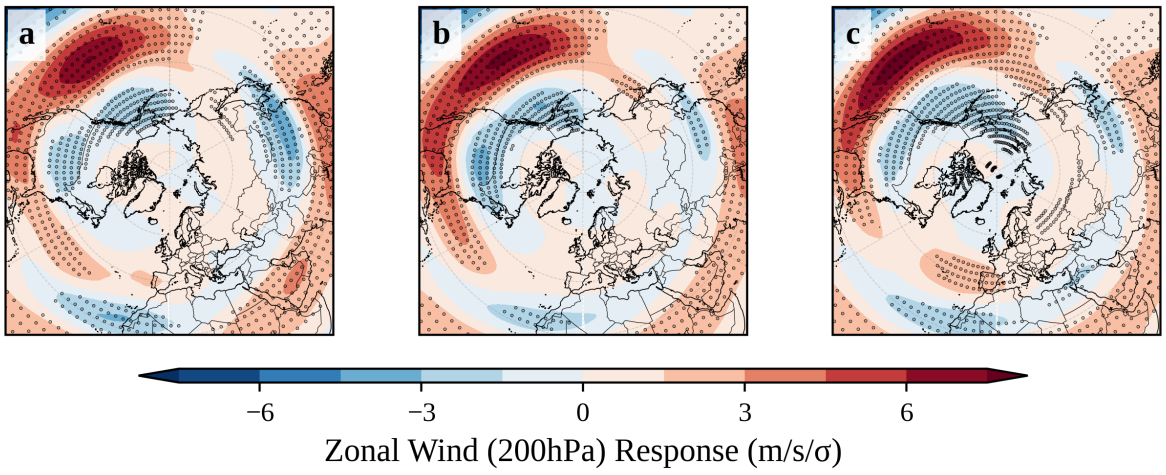

Figure 3.13: Regression of DJF 200 hPa zonal wind ($ms^{-1}\sigma^{-1}$) onto the Niño 3.4 index for 1980–2023. **(a)** ERA5, **(b)** TCo399 ensemble mean, **(c)** TCo1279. Positive values (red) indicate strengthened westerlies during El Niño; negative values (blue) indicate weakened westerlies. Polar stereographic projection centred on the North Pole, with stippling for statistical significance at the 95% interval.

ERA5 (panel a) exhibits a robust Pacific jet response, with positive zonal wind anomalies of 4–6 m/s/σ extending across the central and eastern North Pacific around 30–40°N. This represents a strengthening and equatorward extension of the Pacific subtropical jet during El Niño events, consistent with the atmospheric response to enhanced tropical convection and the associated Hadley cell modifications (Trenberth et al. 1998). Negative anomalies appear over Alaska and northwestern North America, reflecting the wave-like structure of the extratropical response. Critically, the Atlantic sector shows a coherent though weaker response, with negative zonal wind anomalies of 2–3 $ms^{-1}\sigma^{-1}$ over the subpolar North Atlantic around 50–60°N and positive anomalies over the subtropical Atlantic around 30–40°N. This dipole structure indicates a southward shift or weakening of the North Atlantic jet during El Niño conditions.

Both model configurations (panels b and c) capture the Pacific jet response with impressive fidelity. The TCo399 ensemble mean reproduces the magnitude and spatial structure of the positive Pacific jet anomalies, with values of 4–6 $ms^{-1}\sigma^{-1}$ in similar locations to observations. TCo1279 shows comparable Pacific jet response, confirming that both resolutions adequately represent the direct atmospheric circulation response to tropical Pacific forcing. The wave-like structure over North America is also present in both configurations, though with some spatial differences in the exact positioning of positive and negative centres.

In the Atlantic sector, the main dipole between the subtropical easterly flow off western Africa and the westerly flow in the Atlantic mid-latitudes is generally well presented. This suggests that the stratospheric link between the ENSO and the Atlantic has been captured in the models, although there are some spatial inconsistencies in the representation over Europe. Comparing TCo1279 with the TCo399 ensemble mean reveals at best a very minor improvement in Atlantic jet response with increased resolution.

Mid-Tropospheric Circulation Response

Figure 3.14 presents the 500 hPa geopotential height regression onto the Niño 3.4 index, providing a comprehensive view of the hemispheric teleconnection pattern including the canonical PNA structure.

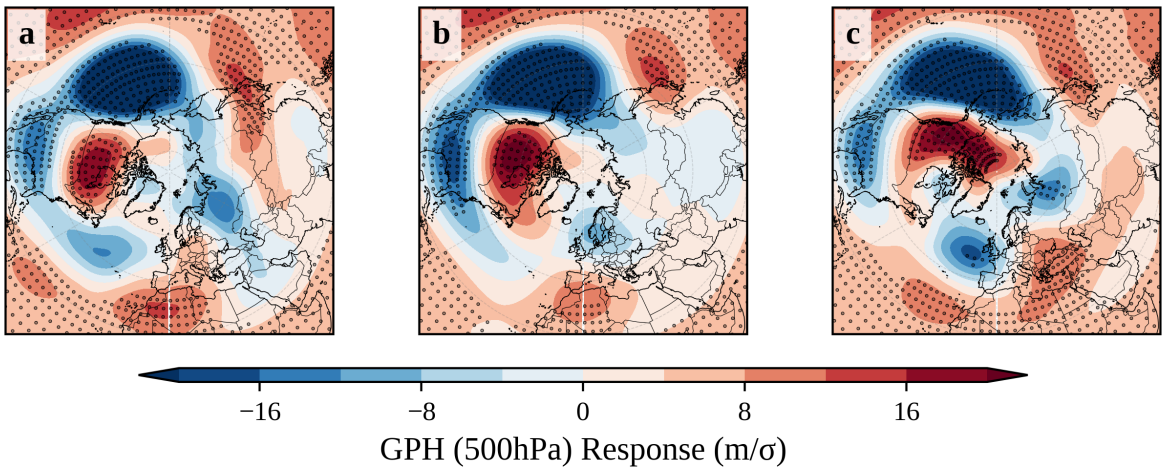

Figure 3.14: Regression of DJF 500 hPa geopotential height (m /σ) onto the Niño 3.4 index for 1980–2023. **(a)** ERA5, **(b)** TCo399 ensemble mean, **(c)** TCo1279. Positive values (red) indicate ridging during El Niño; negative values (blue) indicate troughing. Polar stereographic projection centred on the North Pole, with stippling for statistical significance at the 95% interval.

ERA5 (panel a) displays the classic PNA teleconnection pattern with remarkable clarity. A pronounced negative geopotential height anomaly of $>$ 16 m /σ appears over the Aleutians and North Pacific, representing the deepened Aleutian Low characteristic of El Niño winters. Positive anomalies of 8–12 m /σ extend over western North America and the eastern Pacific, whilst negative anomalies appear over the southeastern United States and western Atlantic. This wave train structure reflects the atmospheric Rossby wave response to tropical Pacific heating anomalies, propagating along a great circle route towards the pole and subsequently equatorward.

Crucially for NAO dynamics, ERA5 exhibits coherent negative geopotential height anomalies of 6–10 m /σ over the midlatitude North Atlantic, a small positive anomally over Iceland and positive anomalies over the subtropical Atlantic and Mediterranean. This spatial structure resembles the negative phase of the NAO or a pattern similar to the EA pattern, suggesting that El Niño conditions tend to suppress the Icelandic Low whilst enhancing subtropical high pressure—broadly consistent with a weakened meridional pressure gradient and reduced Atlantic storm track activity.

The model configurations again demonstrate accurate representation of the Pacific and

North American components of the teleconnection. Both TCo399 (panel b) and TCo1279 (panel c) reproduce the Aleutian Low deepening with magnitudes of >16 m /σ, closely matching observations. The positive anomalies over western North America and the wave train structure extending across the continent are similarly well captured in both configurations. This confirms that the models correctly simulate the tropical-to-extratropical teleconnection mechanisms operating over the Pacific basin and the generation of the PNA pattern.

However, the continuation of this wave train into the Atlantic sector reveals the same deficiency observed in the jet stream analysis. The coherent negative geopotential height anomalies over the mid-latitude North Atlantic and North Easst Europe present in ERA5 are displaced in both model configurations. TCo399 shows only weak, fragmented anomalies over the Nordic Seas, with magnitudes reduced to 2–4 m /σ. TCo1279, on the other hand, exhibits different weaknesses in the Atlantic sector, with an greater negative response in the mid-latitudes, however shifted to the east and poleward from the comparable centre of action in ERA5. The subtropical Atlantic positive anomalies are slightly improved in TCo1279.

The contrast between the robust PNA representation and the mixed Atlantic response illuminates the nature of the teleconnection deficiency. The problem is not that models fail to respond to tropical forcing—the Pacific atmospheric response demonstrates they can generate realistic extratropical circulation anomalies when directly forced. Rather, the deficiency lies in the propagation or maintenance of these circulation anomalies as they progress downstream from the Pacific into the Atlantic sector. The degreadation in teleconnection realism at the mid-troposphere (500hPa) compared to the high troposphere (200hPa) is strongly suggestive that the defficiency is related to parametrisation or vertical processes - which explains the lack of improvement from horizontal resolution alone. **Surface Pressure Response** Figure 3.15 presents the MSLP regression onto Niño 3.4, directly relevant to NAO dynamics as the NAO is defined by surface pressure gradients. ERA5 (panel a) shows pronounced negative pressure anomalies of >2.0 hpa/σ centred over the Aleutians, the surface manifestation of the upper-level Aleutian Low deepening. Over the Atlantic, a dipole structure emerges with negative anomalies of 0.5–1.0 hpa/σ over the subpolar region and weak positive anomalies over parts of the subtropical Atlantic. Whilst this Atlantic surface pressure response is weaker than the Pacific response—reflecting the generally weaker ENSO-NAO connection compared to ENSO-PNA—it nonetheless represents a coherent atmospheric response that projects onto NAO variability.

Both model configurations reproduce the Pacific surface pressure response accurately, with Aleutian pressure reductions of >2.0 hpa/σ matching observations. However, the Atlantic dipole structure is again poorly represented. The subpolar Atlantic pressure anomalies are reduced to less than 0.5 hpa/σ in both TCo399 and TCo1279, and lack the spatial coherence evident in ERA5. This surface pressure deficiency directly affects the models' ability to capture ENSO-forced NAO variability, as the pressure gradient between the Icelandic and Azores pressure centres shows minimal response to tropical Pacific forcing. **Composite Analysis: El Niño minus La Niña**

Figure 3.16 presents composite mean MSLP differences between El Niño and La Niña winters, providing a complementary perspective that emphasises the atmospheric response to extreme ENSO events whilst allowing for potential nonlinearities not captured by linear regression.

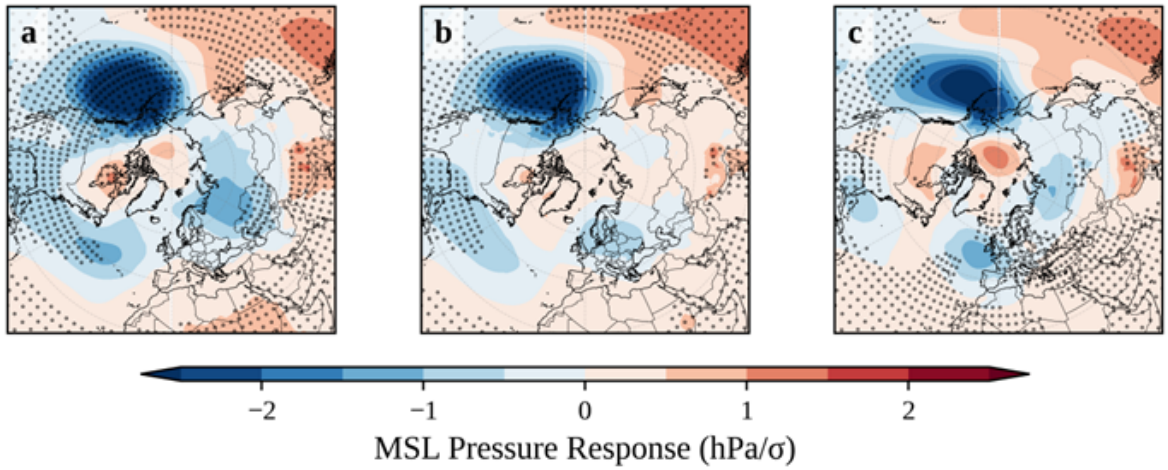

Figure 3.15: Regression of DJF mean sea level pressure (hpa/σ) onto the Niño 3.4 index for 1980–2023. **(a)** ERA5, **(b)** TCo399 ensemble mean, **(c)** TCo1279. Negative values (blue) indicate reduced pressure during El Niño. Polar stereographic projection centred on the North Pole, with stippling for statistical significance at the 95% interval.

ERA5 composites (panel a) show pressure differences of >3 hPa over the Aleutians, slightly larger than the regression-based estimates due to the nonlinear amplification during extreme events. The Atlantic sector exhibits pressure differences of approximately 1–1.5 hPa over the subpolar region, again confirming that extreme ENSO phases produce measurable NAO-like surface pressure responses. The composite approach, by focusing on strong ENSO events where teleconnection forcing is maximised, provides evidence that the Atlantic response is not merely statistical noise but represents a genuine physical connection.

The model composites tell a consistent story with the regression analysis. Pacific composite differences closely match ERA5, with both configurations showing 2–3 hPa Aleutian pressure reductions during El Niño. Atlantic composite differences remain weak in both TCo399 and TCo1279, with magnitudes less than 1.0 hPa and lacking spatial coherence. The fact that both regression and composite methodologies reveal the same deficiency—strong Pacific response, weak Atlantic response, minimal resolution dependence—confirms this as a robust feature of the model behaviour rather than an artefact of the analysis approach.

Interpretation and Implications

The teleconnection analysis reveals a clear and consistent pattern across all diagnostics: both model configurations accurately represent the ENSO-Pacific teleconnection (PNA pattern, Aleutian Low response, Pacific jet modifications) but struggle to represent continuation of this signal into the Atlantic-European sector at sea level. This "broken link" between Pacific and Atlantic circulation responses appears at first slightly in the 200 hPa jets, more prominently in the 500 hPa mid-tropospheric circulation and is most obvious in surface pressure fields. This trend is found independently of the variable in question, as the 850hPa zonal wind regression (Appendix A) shows a much weaker teleconnection than at 200hPa.

This finding has important implications for understanding the source of weak NAO predictability in atmosphere-only simulations. The models demonstrate that they can respond to remote tropical forcing when the forcing is sufficiently strong and direct, as evidenced by the robust Pacific response. The Atlantic deficiency therefore does not reflect a general inability

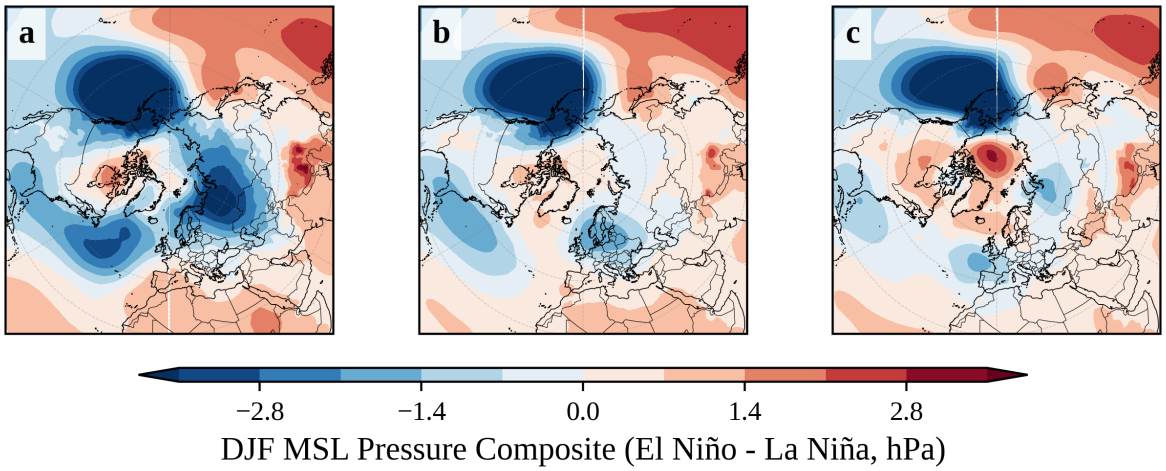

Figure 3.16: Composite difference in DJF mean sea level pressure (hPa) between El Niño and La Niña winters for 1980–2023. **(a)** ERA5, **(b)** TCo399 ensemble mean, **(c)** TCo1279. Negative values (blue) indicate lower pressure during El Niño compared to La Niña. Polar stereographic projection.

to simulate forced atmospheric responses, but rather points to specific deficiencies in the mechanisms that communicate circulation anomalies downstream through the teleconnection mechanisms.

Several physical mechanisms could contribute to this teleconnection breakdown. First, the deficient mid-tropospheric eddy feedback documented in the previous section could prevent adequate amplification of the teleconnection signal as it propagates downstream. Weak eddy-mean flow coupling would allow the circulation anomalies to be damped by dissipative processes before they can establish coherent Atlantic responses. Second, errors in stationary wave propagation characteristics could deflect or attenuate the Rossby wave energy before it reaches the Atlantic. Third, excessive damping through diabatic processes over the Atlantic basin could overwhelm the teleconnection forcing, preventing the establishment of persistent circulation anomalies. As shown above, the models struggle to realise sufficient eddy feedback in the mid-troposphere, however are much closer to ERA5 values towards to the top of the troposphere. This aligns with the finding that the teleconnection pathway is less evident closer to the surface.

The lack of clear improvements when resolution is enhanced suggests that these deficiencies are not primarily related to inadequate representation of small-scale features that higher resolution would better resolve. Rather, they likely involve fundamental aspects of large-scale atmospheric dynamics, parametrisation schemes, or the balance between forced responses and internal variability. The implications for NAO seasonal prediction are significant: even with perfect tropical Pacific forecasts, current atmospheric models would fail to translate ENSO information into skilful Atlantic sector predictions due to these teleconnection deficiencies.

Chapter 4

Discussion & Conclusion

Summary of Key Findings

Both model configurations reproduce the observed wintertime climatology with minimal systematic biases over the North Atlantic sector. Temperature, precipitation, and MSLP fields in the NAO centres of action show no statistically significant differences from ERA5. However, both exhibit a modest northward shift in the eddy-driven jet, with TCo1279 showing slightly more pronounced displacement. Substantial stratospheric westerly biases indicate a systematically weaker polar vortex than observed in the low-resolution configuration.

The NAO characteristics analysis confirmed that both configurations generate realistic interannual variability comparable to observations, with no clear mean state biases identified. However, the simulations fail to reproduce the observed NAO variability, as evidenced by negligible temporal correlations with observed NAO sequences ($r \approx 0.11$). This indicates that the signal-to-noise problems characteristic of initialised climate predictions with standard resolution coupled models are also present at these resolutions, even within the AMIP framework where oceanic forcing is constrained with observations and atmospheric evolution is dominated by internal dynamics.

The EKE analysis revealed systematic positive biases in both configurations, particularly at upper levels. At 200 hPa, TCo1279 demonstrates substantial over-representation with 30–40% bias across the North Atlantic. At 500 hPa, the resolution impact is spatially complex, with improvements in some regions but degradation in others.

Critically, the EFP analysis showed that both configurations systematically underestimate mid-tropospheric (500 hPa) eddy feedback. TCo1279 shows lower values that even the range of the TCo399 ensemble spread - ruling out any improvement from resolution. At 200 hPa, both configurations were in a similar range to ERA5, with the TCo1279 configuration performing better in the high latitudes and slightly worse in the mid latitudes. While a significant improvement from 500hPa, TCo1279 was within the spread of the TCo399 ensemble values again ruling out systematic improvement from resolution alone.

The teleconnection analysis corroborated these findings, with a subtle but noticeable degradation in representation from the higher pressure levels to the surface. Both configurations accurately reproduce the ENSO-Pacific connection with realistic PNA patterns and jet responses. However, moving through the pressure levels, the realism of the compared to observations across all levels examined. Crucially, increased resolution produces no systematic improvement in Atlantic teleconnection strength.

Interpretation

These findings provide a clear answer to the research question: within the resolution range examined, horizontal resolution increases do not improve the representation of eddy feedback or teleconnections. Horizontal resolution alone fails to address the signal to noise problem in NAO predictions. To address the stated aims:

1. Eddy feedback strength is deficient in both resolutions at both pressure levels studied. There was no improvement at 9km compared to 27km.

2. Higher horizontal resolution (9km) shows mixed results yet ultimately fails to systematically produce more realistic teleconnection patterns.

3. The failure of both configurations to predict NAO variability despite benefiting from observed SST forcing strongly suggests that realistic representation of eddy dynamics and teleconnection patterns is critical to representing the NAO.

The contrast between robust Pacific responses and weak surface-level Atlantic signals highlights the nature of the deficiency. Models correctly respond to tropical forcing when sufficiently strong and direct, but fail to propagate this signal downstream into the Atlantic. The deficient mid-tropospheric EFP provides a mechanistic explanation: insufficient eddy amplification prevents circulation anomalies from establishing persistent Atlantic responses, allowing dissipative processes to damp teleconnection signals.

The implications for seasonal prediction are significant. Simply increasing computational resources to run higher-resolution forecasts may not yield proportionate improvements in NAO skill without additional efforts to fine tune those model configurations. Efforts may also prove more effective when directed towards parametrisation improvements, particularly for processes controlling eddy-mean flow coupling strength, rather than resolution increases alone.

Research Limitations

The most significant limitation is the asymmetric ensemble design. TCo399 comprises 10 members for monthly data and 5 members for daily data enabling robust ensemble statistics, whilst TCo1279 consists of a single realisation. This fundamentally constrains comparison: differences between TCo1279 and the TCo399 ensemble mean could partially reflect sampling variability rather than systematic resolution effects. The large EFP ensemble spread—with individual members sometimes approaching or exceeding observations whilst the ensemble mean shows deficiencies—demonstrates this sampling sensitivity. Without a TCo1279 ensemble, we cannot definitively separate resolution impacts from internal variability.

Additionally, using only 43 DJF seasons introduces non-negligible sampling uncertainty for correlation-based metrics. The findings of Saffin et al. (2024) that the EFP is sensitive to sampling was tested (see Appendix B) and highlights the importance of finding other methods for quantifying the eddy feedback effects.

The atmosphere-only design, whilst valuable for isolating atmospheric processes, prevents assessment of ocean-atmosphere feedbacks identified as important for NAO variability (Patrizio et al. 2025). The lack of atmospheric initialisation means stratosphere-troposphere coupling

pathways available in operational seasonal forecasts (Scaife et al. 2016) are not fully represented. However, the persistence of eddy feedback and teleconnection deficiencies in this simplified framework suggests these problems are intrinsic to atmospheric model physics rather than arising solely from coupled system biases.

Using ERA5 as validation reference introduces complications, as reanalyses represent model-data synthesis rather than pure observations. Variables like precipitation and momentum fluxes retain substantial model dependence. ERA5's known stratospheric biases may affect validation of upper-level dynamics, and temporal inhomogeneities in the observing system — particularly major satellite transitions — can introduce subtle discontinuities affecting derived statistics.

Computational resources and sustainability represent fundamental constraints in climate modelling. The TCo1279 configuration required approximately three times the computational resources of TCo399, yet yielded limited improvements in the mid-tropospheric processes critical for NAO predictability. This has important implications for resource allocation: a 10-member TCo399 ensemble provides robust sampling of internal variability at one-third the cost of a single TCo1279 realisation. Within the examined resolution range, ensemble size may offer better return on investment than resolution increases. Such findings underscore the importance of evaluating modelling strategies within their broader social, environmental, and ethical context — ensuring that computational investments serve societal needs for improved climate predictions while minimising environmental impact.

Future Work

Several promising directions could advance understanding in this area:

- **Extended Ensembles:** A multi-member TCo1279 ensemble (5-10 members) would enable robust comparison with TCo399, clarifying whether observed differences reflect resolution or sampling. Century-scale integrations would reduce sampling uncertainty and reveal multi-decadal variability in teleconnection strength.
- **Higher Resolution:** Exploring resolutions at 3–5 km where convection is resolved would test whether a critical threshold exists for eddy feedback improvements.
- **Coupled Experiments:** Fully coupled configurations would assess whether atmospheric resolution improvements enhance ocean-atmosphere feedbacks, reduce coupled drift, and improve initialised forecast skill. Interactive coupling enables two-way processes that AMIP simulations exclude.
- **Parametrisation Sensitivity:**Many of the teleconnection mechanisms rely on stratosphere/troposphere coupling. Targeted improvements in relevant parametrisations might yield larger gains than resolution increases.

The diversity of possible research directions underscore that understanding NAO predictability requires creative combinations of higher resolution, improved parametrisations, and process-based understanding beyond traditional model development approaches. Improvements in resolution must be accompanied by improvements in the simulation of underlying physical processes in order to yield substantial improvements.

Appendix A

Zonal Wind ENSO Linear Regression at 850 hPa

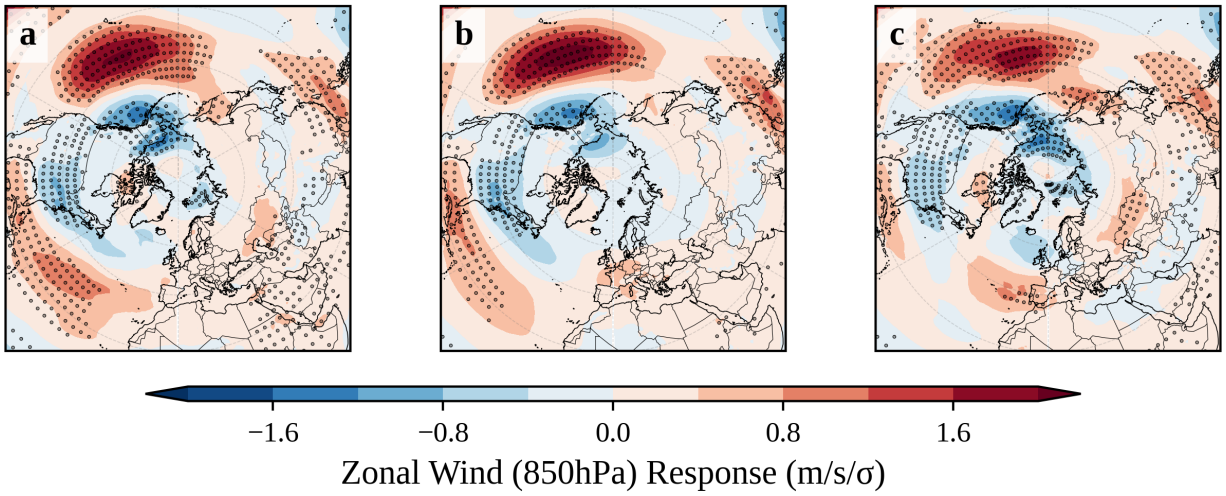

Figure A.1: Regression of DJF 850 hPa zonal wind ($ms^{-1}\sigma^{-1}$) onto the Niño 3.4 index for 1980–2023. **(a)** ERA5, **(b)** TCo399 ensemble mean, **(c)** TCo1279. Positive values (red) indicate strengthened westerlies during El Niño; negative values (blue) indicate weakened westerlies. Polar stereographic projection centred on the North Pole, with stippling for statistical significance at the 95% interval.

Appendix B

Temporal Sampling Analysis of the EFP

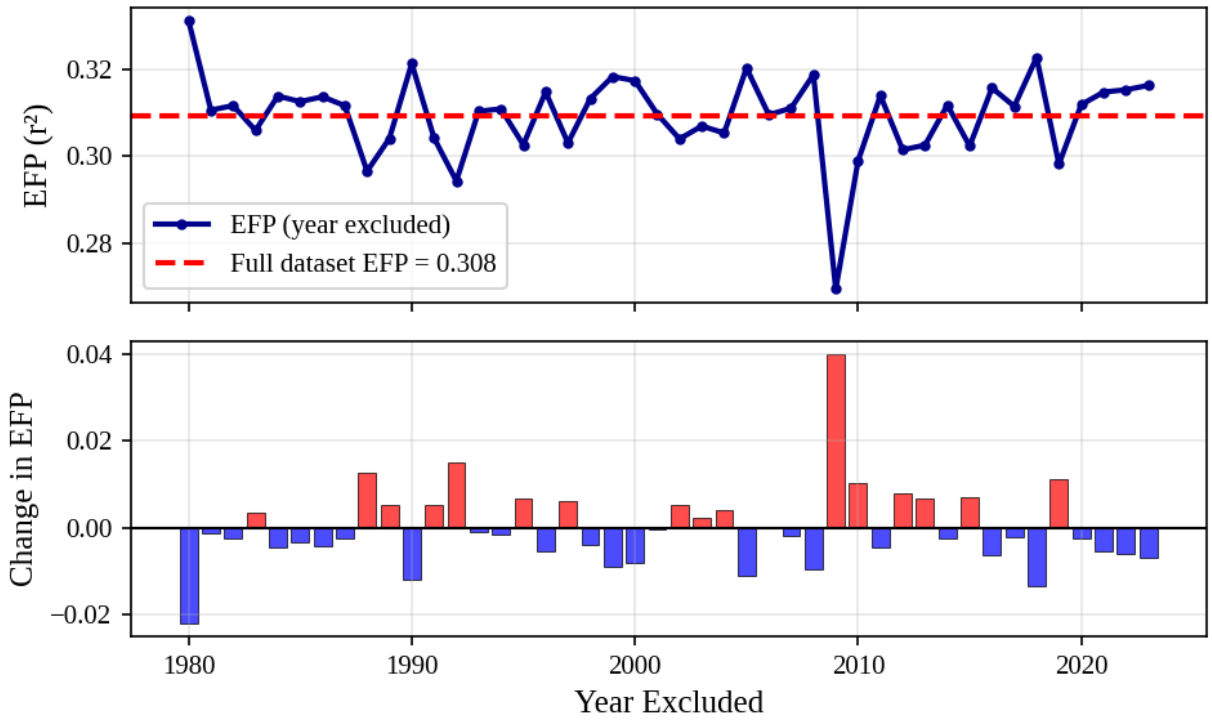

Figure B.1: A time series of the calculated EFP at 200 hPa for the ERA5 reanalysis for DJF 1980–2023, leaving out one year from the calculation (top). The affect of removing that year is shown underneath (bottom). The year indicated on the time series is the year removed from the calculation. This demonstrates the sampling uncertainties associated with the parameter.

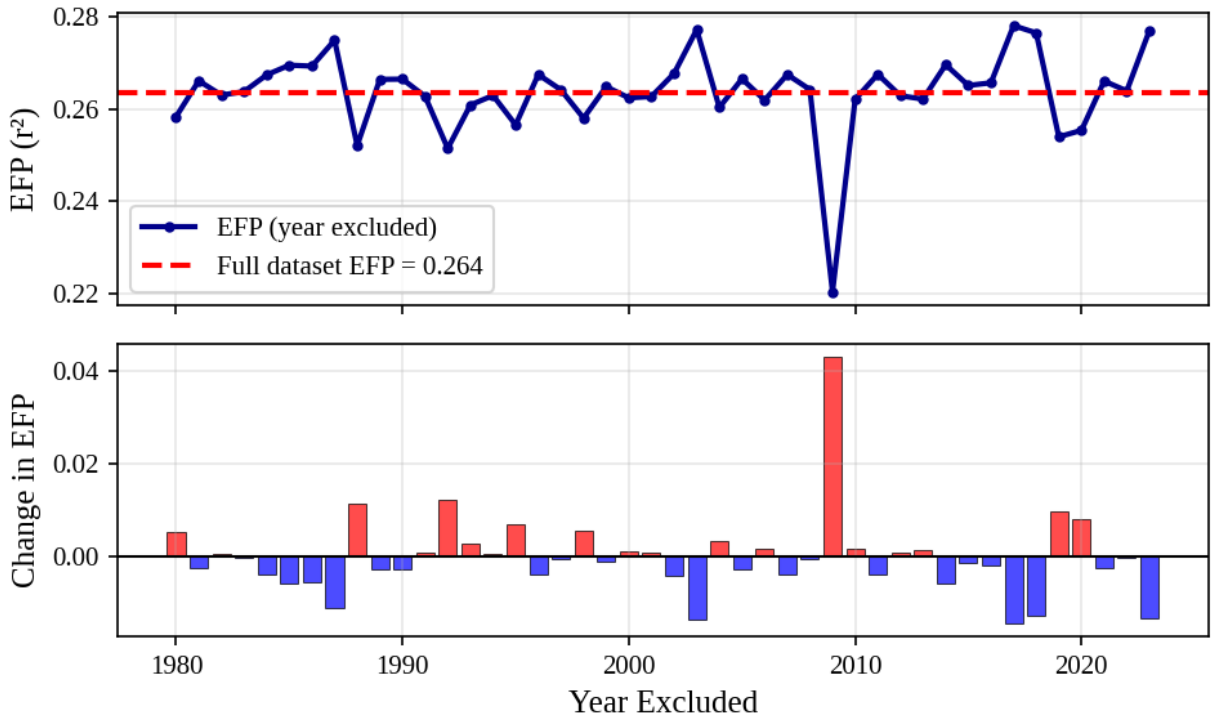

Figure B.2: A time series of the calculated EFP at 500 hPa for the ERA5 reanalysis for DJF 1980–2023, leaving out one year from the calculation (top). The affect of removing that year is shown underneath (bottom). The year indicated on the time series is the year removed from the calculation. This demonstrates the sampling uncertainties associated with the parameter.

Bibliography

Andrews, M. B., J. R. Knight, A. A. Scaife, Y. Lu, T. Wu, L. J. Gray, and V. Schenzinger (2019). "Observed and simulated teleconnections between the stratospheric Quasi-Biennial Oscillation and Northern Hemisphere winter atmospheric circulation". In: *Journal of Geophysical Research: Atmospheres* 124, pp. 1219–1232. DOI: 10.1029/2018JD029368.

Athanasiadis, Panos J., Fumiaki Ogawa, Nour-Eddine Omrani, Noel Keenlyside, Reinhard Schiemann, Alexander J. Baker, Pier Luigi Vidale, Alessio Bellucci, Paolo Ruggieri, Rein Haarsma, Malcolm Roberts, Chris Roberts, Lenka Novak, and Silvio Gualdi (2022). "Mitigating Climate Biases in the Midlatitude North Atlantic by Increasing Model Resolution: SST Gradients and Their Relation to Blocking and the Jet". In: *Journal of Climate* 35.21, pp. 6985–7006. DOI: 10.1175/JCLI-D-21-0515.1.

Baker, L. H., L. C. Shaffrey, R. T. Sutton, A. Weisheimer, and A. A. Scaife (2018). "An intercomparison of skill and overconfidence/underconfidence of the wintertime North Atlantic Oscillation in multimodel seasonal forecasts". In: *Geophysical Research Letters* 45, pp. 7808–7817. DOI: 10.1029/2018GL078838.

Baldwin, M. P. and T. J. Dunkerton (2001). "Stratospheric harbingers of anomalous weather regimes". In: *Science* 294, pp. 581–584. DOI: 10.1126/science.1063315.

Barnes, Elizabeth A and Dennis L Hartmann (2010). "Dynamical Feedbacks and the Persistence of the NAO". In: *Journal of the Atmospheric Sciences* 67.3, pp. 851–865. DOI: 10.1175/2009JAS3193.1.

Barnston, A. G. and R. E. Livezey (1987). "Classification, seasonality and persistence of low-frequency atmospheric circulation patterns". In: *Monthly Weather Review* 115, pp. 1083–1126.

Blackmon, M. L., J. M. Wallace, N.-C. Lau, and S. L. Mullen (1977). "An observational study of the Northern Hemisphere wintertime circulation". In: *Journal of the Atmospheric Sciences* 34, pp. 1040–1053.

Bojariu, R. and G. Reverdin (2002). "Large-scale variability modes of freshwater flux and precipitation over the Atlantic". In: *Climate Dynamics* 18, pp. 369–381. DOI: 10.1007/s003820100182.

Bruno Soares, Marta, Meaghan Daly, and Suraje Dessai (2018). "Assessing the value of seasonal climate forecasts for decision-making". In: *Wiley Interdisciplinary Reviews: Climate Change* 9.4, e523. DOI: 10.1002/wcc.523.

Davini, P., C. Cagnazzo, S. Gualdi, and A. Navarra (2012). "Bidimensional diagnostics, variability, and trends of Northern Hemisphere blocking". In: *Journal of Climate* 25, pp. 6496–6509. DOI: 10.1175/JCLI-D-12-00032.1.

DelSole, T. and J. Shukla (2009). "Artificial skill due to predictor screening". In: *Journal of Climate* 22, pp. 331–345. DOI: 10.1175/2008JCLI2414.1.

DeWeaver, E. and S. Nigam (2000). "Linearity and time scale dependence of ENSO in a coupled ocean–atmosphere model". In: *Journal of Climate* 13, pp. 3632–3645.

Diamantakis, M. and F. Váňa (2022). "A fast converging and concise algorithm for computing the departure points in semi-Lagrangian weather and climate models". In: *Quarterly Journal of the Royal Meteorological Society* 148.743, pp. 670–684. DOI: 10.1002/qj.4224.

Dunstone, N., D. Smith, A. Scaife, L. Hermanson, R. Eade, N. Robinson, M. Andrews, and J. Knight (2016). "Skilful predictions of the winter North Atlantic oscillation one year ahead". In: *Nature Geoscience* 9.11, pp. 809–814. DOI: 10.1038/ngeo2824.

Eade, R., D. Smith, A. Scaife, E. Wallace, N. Dunstone, L. Hermanson, and N. Robinson (2014). "Do seasonal-to-decadal climate predictions underestimate the predictability of the real world?" In: *Geophysical Research Letters* 41, pp. 5620–5628. DOI: 10.1002/2014GL061146.

Edmon Jr, Harold J, Brian J Hoskins, and Michael E McIntyre (1980). "Eliassen-Palm cross sections for the troposphere". In: *Journal of the Atmospheric Sciences* 37.12, pp. 2600–2616. DOI: 10.1175/1520-0469(1980)037<2600:EPCSFT>2.0.CO;2.

Embury, Owen, Christopher J. Merchant, Simon A. Good, Nick A. Rayner, Jacob L. Høyer, Chris Atkinson, Tom Block, Emy Alerskans, Katie J. Pearson, Matthew Worsfold, Niall McCarroll, and Craig Donlon (2024). "Satellite-based time-series of sea-surface temperature since 1980 for climate applications". In: *Scientific Data* 11.326. DOI: 10.1038/s41597-024-03147-w.

Eyring, V., S. Bony, G. A. Meehl, C. A. Senior, B. Stevens, R. J. Stouffer, and K. E. Taylor (2016). "Overview of the Coupled Model Intercomparison Project Phase 6 (CMIP6) experimental design and organization". In: *Geoscientific Model Development* 9, pp. 1937–1958. DOI: 10.5194/gmd-9-1937-2016.

Feldstein, S. B. (2003). "The dynamics of NAO teleconnection pattern growth and decay". In: *Quarterly Journal of the Royal Meteorological Society* 129.589, pp. 901–924. DOI: 10.1256/qj.02.76.

Flato, G., J. Marotzke, B. Abiodun, P. Braconnot, S. C. Chou, W. Collins, P. Cox, F. Driouech, S. Emori, V. Eyring, C. Forest, P. Gleckler, E. Guilyardi, C. Jakob, V. Kattsov, C. Reason, and M. Rummukainen (2013). "Evaluation of Climate Models". In: *Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*. Ed. by T. F. Stocker, D. Qin, G.-K. Plattner, M. Tignor, S. K. Allen, J. Boschung, A. Nauels, Y. Xia, V. Bex, and P. M. Midgley. Cambridge, United Kingdom and New York, NY, USA: Cambridge University Press, pp. 741–866.

Gastineau, G., F. Lia, and M. L. R. Liberato (2013). "Influence of the phase of the Quasi-Biennial Oscillation on the Madden–Julian Oscillation". In: *Quarterly Journal of the Royal Meteorological Society* 139, pp. 1816–1825.

Gelbrecht, Maximilian, Alistair White, Sebastian Bathiany, and Niklas Boers (2023). "Differentiable Programming for Earth System Modeling". In: *Geoscientific Model Development* 16, pp. 3123–3135. DOI: 10.5194/gmd-16-3123-2023.

Gerber, Edwin P and Geoffrey K Vallis (2007). "Eddy–zonal flow interactions and the persistence of the zonal index". In: *Journal of the Atmospheric Sciences* 64.9, pp. 3296–3311. DOI: 10.1175/JAS4006.1.

Hannachi, A. (1997). "Low-Frequency Variability in a GCM: Three-Dimensional Flow Regimes and Their Dynamics". In: *Journal of Climate* 10, pp. 1357–1379.

Hardiman, S. C., N. J. Dunstone, A. A. Scaife, D. M. Smith, D. Copsey, H.-L. Ren, and S. Yukimoto (2022). "Missing eddy feedback may explain weak signal-to-noise ratios in climate predictions". In: *npj Climate and Atmospheric Science* 5.1, pp. 1–11.

Hardiman, S. C., N. J. Dunstone, A. A. Scaife, D. M. Smith, S. Ineson, J. Lim, and D. Fereday (2019). "The effect of QBO phase on the atmospheric response to El Niño". In: *Journal of Geophysical Research: Atmospheres* 124, pp. 6924–6939.

Harris, Charles R, K Jarrod Millman, Stéfan J van der Walt, et al. (2020). "Array programming with NumPy". In: *Nature* 585.7825, pp. 357–362. DOI: 10.1038/s41586-020-2649-2.

Hersbach, H., B. Bell, P. Berrisford, S. Hirahara, A. Horányi, J. Muñoz-Sabater, J. Nicolas, C. Peubey, R. Radu, D. Schepers, A. Simmons, C. Soci, S. Abdalla, X. Abellan, G. Balsamo, P. Bechtold, G. Biavati, J. Bidlot, M. Bonavita, G. De Chiara, P. Dahlgren, D. Dee, M. Diamantakis, R. Dragani, J. Flemming, R. Forbes, M. Fuentes, A. Geer, L. Haimberger, S. Healy, R. J. Hogan, E. Hólm, M. Janisková, S. Keeley, P. Laloyaux, P. Lopez, C. Lupu, G. Radnoti, P. de Rosnay, I. Rozum, F. Vamborg, S. Villaume, and J.-N. Thépaut (2020). "The ERA5 global reanalysis". In: *Quarterly Journal of the Royal Meteorological Society* 146.730, pp. 1999–2049. DOI: 10.1002/qj.3803.

Hoskins, B. J., I. N. James, and G. H. White (1983). "Eliassen-Palm cross sections for the troposphere". In: *Journal of the Atmospheric Sciences* 40, pp. 1595–1615.

Hoskins, B. J. and D. J. Karoly (1981). "The steady linear response of a spherical atmosphere to thermal and orographic forcing". In: *Journal of the Atmospheric Sciences* 38, pp. 1179–1196.

Hoyer, Stephan and Joseph Hamman (2017). "xarray: ND labeled arrays and datasets in Python". In: *Journal of Open Research Software* 5.1, p. 10. DOI: 10.5334/jors.148.

Hunter, John D (2007). "Matplotlib: A 2D graphics environment". In: *Computing in Science & Engineering* 9.3, pp. 90–95. DOI: 10.1109/MCSE.2007.55.

Hurrell, J. W. (1995). "Decadal trends in the North Atlantic Oscillation: regional temperatures and precipitation". In: *Science* 269.5224, pp. 676–679. DOI: 10.1126/science.269.5224.676.

Hurrell, J. W., J. J. Hack, D. Shea, J. M. Caron, and J. Rosinski (2009). "A unified modeling approach to climate system prediction". In: *Journal of Marine Systems* 78.1, pp. 28–41. DOI: 10.1016/j.jmarsys.2008.11.026.

Hurrell, J. W., Y. Kushnir, G. Ottersen, and M. Visbeck (2003). "An overview of the North Atlantic Oscillation". In: *The North Atlantic Oscillation: Climatic Significance and Environmental Impact*. Ed. by J. W. Hurrell, Y. Kushnir, G. Ottersen, and M. Visbeck. Vol. 134. Geophysical Monograph Series. Washington, DC: American Geophysical Union, pp. 1–35.

Kang, D., M.-I. Lee, J. Im, D. Kim, H.-M. Kim, H.-S. Kang, S. D. Schubert, A. A. Arribas, and C. MacLachlan (2014). "Prediction of the Arctic Oscillation in boreal winter by dynamical seasonal forecasting systems". In: *Geophysical Research Letters* 41, pp. 3577–3585. DOI: 10.1002/2014GL060011.

King, M. P., A. Dobler, O. Alves, G. Boer, C. Deser, F. Driouech, B. Kirtman, A. Kumar, R. Saravanan, and M. Sigmond (2018). "The role of atmosphere-ice-ocean interactions in seasonal prediction of Arctic sea ice". In: *Philosophical Transactions of the Royal Society A* 376.2129, p. 20170153.

Knight, J. R., A. A. Scaife, and A. Maidens (2022). "n extratropical contribution to the signal-to-noise paradox in seasonal climate prediction". In: *Geophysical Research Letters* 49. DOI: 10.1029/2022GL100471.

Kumar, A. (2009). "Finite samples and uncertainty estimates for skill measures for seasonal prediction". In: *Monthly Weather Review* 137, pp. 2622–2631.

Lang, S. T. K., A. Dawson, M. Diamantakis, P. Dueben, S. Hatfield, M. Leutbecher, K. Mogensen, F. Pappenberger, D. S. Richardson, I. Tsonevsky, N. Wedi, and F. Váňa (2021). "More accuracy with less precision". In: *Quarterly Journal of the Royal Meteorological Society* 147.741, pp. 4358–4370. DOI: 10.1002/qj.4181.

Lau, N.-C. (1988). "Variability of the observed midlatitude storm tracks". In: *Journal of the Atmospheric Sciences* 45, pp. 2718–2743.

Limpasuvan, V. and D. L. Hartmann (2000). "Wave-maintained annular modes of climate variability". In: *Journal of Climate* 13, pp. 4414–4429.

Lock, S. J., S. T. K. Lang, M. Leutbecher, R. J. Hogan, and F. Vitart (2019). "Treatment of model uncertainty from radiation by the Stochastically Perturbed Parametrization Tendencies (SPPT) scheme and associated revisions in the ECMWF ensembles". In: *Quarterly Journal of the Royal Meteorological Society* 145.719, pp. 75–89. DOI: 10.1002/qj.3570.

Loeb, N. G., D. R. Doelling, H. Wang, W. Su, C. Nguyen, J. G. Corbett, L. Liang, C. Mitrescu, F. G. Rose, and S. Kato (2018). "Clouds and the Earth's Radiant Energy System (CERES) Energy Balanced and Filled (EBAF) Top-of-Atmosphere (TOA) Edition-4.0 Data Product". In: *Journal of Climate* 31.2, pp. 895–918. DOI: 10.1175/JCLI-D-17-0208.1.

Lorenz, David J and Dennis L Hartmann (2003). "Eddy–zonal flow feedback in the Northern Hemisphere winter". In: *Journal of Climate* 16.8, pp. 1212–1227. DOI: 10.1175/1520-0442(2003)16<1212:EFFITN>2.0.CO;2.

Maciel, P., T. Quintino, U. Modigliani, P. Dando, B. Raoult, W. Deconinck, F. Rathgeber, and C. Simarro (2017). *The new ECMWF interpolation package MIR*. Technical Memorandum 828. European Centre for Medium-Range Weather Forecasts. DOI: 10.21957/H20RZ8.

McKinney, Wes et al. (2010). "Data structures for statistical computing in Python". In: *Proceedings of the 9th Python in Science Conference*. Vol. 445. Austin, TX, pp. 51–56.

Mellado-Cano, J., D. Barriopedro, R. García-Herrera, R. M. Trigo, and A. Galán (2018). "Reconciling North Atlantic climate modes: revised monthly indices for the East Atlantic and the Scandinavian patterns beyond the 20th century". In: *Earth System Science Data* 10.4, pp. 2329–2344. DOI: 10.5194/essd-10-2329-2018.

Met Office (2015). *Cartopy: a cartographic python library with a Matplotlib interface*. URL: https://scitools.org.uk/cartopy.

O'Reilly, C. H., D. J. Befort, and A. Weisheimer (2025). "Challenges for seasonal prediction of summer European rainfall". In: *npj Climate and Atmospheric Science* 8, p. 146. DOI: 10.1038/s41612-025-01027-7.

O'Reilly, C. H., T. Woollings, and L. Zanna (2017). "The dynamical influence of the Atlantic multidecadal oscillation on continental climate". In: *Journal of Climate* 30, pp. 7213–7230.

Osborn, T. J. and K. R. Briffa (2006). "The spatial extent of 20th-century warmth in the context of the past 1200 years". In: *Science* 311.5762, pp. 841–844.

Palmeiro, Froila M., Javier García-Serrano, Mario Rodrigo, Marta Abalos, Bo Christiansen, and Shuting Yang (2023). "Boreal winter stratospheric climatology in EC-EARTH: CMIP6

version". In: *Climate Dynamics* 60, pp. 883–898. DOI: 10.1007/s00382-022-06368-0. URL: https://doi.org/10.1007/s00382-022-06368-0.

Patrizio, C. R. and D. W. J. Thompson (2025). "The influence of Atlantic multidecadal variability on the climate of northwest Europe". In: *npj Climate and Atmospheric Science* 8, p. 146. DOI: 10.1038/s41612-025-01027-7.

Rackow, T., X. Pedruzo-Bagazgoitia, T. Becker, S. Milinski, I. Sandu, R. Aguridan, L. F. Borchert, S. Danilov, H. Haak, S. Juricke, J. Kjellsson, N. Koldunov, P. Korn, G. Krokos, L. Loewenstein, C. Mehlmann, L. Mu, D. Mulholland, D. Sidorenko, J. Streffing, Q. Wang, and F. Ziemen (2024). "Multi-year simulations at kilometre scale with the Integrated Forecasting System coupled to FESOM2.5". In: *EGUsphere*. preprint. DOI: 10.5194/egusphere-2024-1492.

Roberts, C. D., R. Senan, F. Molteni, S. Boussetta, M. Mayer, and S. P. E. Keeley (2018a). "Climate model configurations of the ECMWF Integrated Forecasting System (ECMWF-IFS cycle 43r1) for HighResMIP". In: *Geoscientific Model Development* 11.9, pp. 3681–3712. DOI: 10.5194/gmd-11-3681-2018.

Roberts, M. J., P. L. Vidale, C. Senior, et al. (2018b). "The benefits of global high resolution for climate simulation: process understanding and the enabling of stakeholder decisions at the regional scale". In: *Bulletin of the American Meteorological Society* 99, pp. 2341–2359. DOI: 10.1175/BAMS-D-15-00320.1.

Robinson, W. A. (1991). "The dynamics of the zonal index in a simple model of the atmosphere". In: *Tellus A: Dynamic Meteorology and Oceanography* 43.5, pp. 295–305.

— (2006). "On the self-maintenance of midlatitude jets". In: *Journal of the Atmospheric Sciences* 63.8, pp. 2109–2122. DOI: 10.1175/JAS3732.1.

Saffin, Leo, Christine M. McKenna, Remy Bonnet, and Amanda C. Maycock (2024). "Large Uncertainties When Diagnosing the "Eddy Feedback Parameter" and Its Role in the Signal-To-Noise Paradox". In: *Geophysical Research Letters* 51.11, e2024GL108861. DOI: 10.1029/2024GL108861. URL: https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2024GL108861.

Sarmany, D., M. Valentini, P. Maciel, P. Geier, S. Smart, R. Aguridan, J. Hawkes, and T. Quintino (2024). "MultIO: A Framework for Message-Driven Data Routing for Weather and Climate Simulations". In: *Proceedings of the Platform for Advanced Scientific Computing Conference (PASC '24)*. New York, NY, USA: Association for Computing Machinery, pp. 1–12. DOI: 10.1145/3659914.3659938.

Scaife, A. A., A. Arribas, E. Blockley, A. Brookshaw, R. T. Clark, N. Dunstone, R. Eade, D. Fereday, C. K. Folland, M. Gordon, L. Hermanson, J. R. Knight, D. J. Lea, C. MacLachlan, A. Maidens, M. Martin, A. K. Peterson, D. Smith, M. Vellinga, E. Wallace, J. Waters, and A. Williams (2014). "Skilful long-range prediction of European and North American winters". In: *Geophysical Research Letters* 41, pp. 2514–2519. DOI: 10.1002/2014GL059637.

Scaife, A. A., J. Camp, R. Comer, P. Davis, N. Dunstone, M. Gordon, L. Hermanson, C. MacLachlan, N. Martin, Y. Nie, H.-L. Ren, M. Roberts, W. Robinson, D. Smith, and P. L. Vidale (2019). "Does increased atmospheric resolution improve seasonal climate predictions?" In: *Atmospheric Science Letters* 20.8, e922. DOI: 10.1002/asl.922.

Scaife, A. A., D. Copsey, C. Gordon, C. Harris, T. Hinton, S. Keeley, A. O'Neill, M. Roberts, and K. Williams (2016). "Improved Atlantic winter blocking in a climate model". In: *Atmospheric Science Letters* 17.1, pp. 51–56. DOI: 10.1002/asl.598.

Scaife, Adam A. and Doug Smith (2018). "A signal-to-noise paradox in climate science". In: *npj Climate and Atmospheric Science* 1.28, pp. 1–8. DOI: 10.1038/s41612-018-0038-4.

Siegert, S., D. B. Stephenson, P. G. Sansom, A. A. Scaife, R. Eade, and A. Arribas (2016). "A Bayesian framework for verification and recalibration of ensemble forecasts: How uncertain is NAO predictability?" In: *Journal of Climate* 29, pp. 995–1012.

Sigmond, M., J. F. Scinocca, V. V. Kharin, and T. G. Shepherd (2013). "Enhanced seasonal forecast skill following stratospheric sudden warmings". In: *Nature Geoscience* 6, pp. 98–102. DOI: 10.1038/ngeo1698.

Simmons, A. J., D. M. Burridge, M. Jarraud, C. Girard, and W. Wergen (1989). "The ECMWF medium-range prediction models development of the numerical formulations and the impact of increased resolution". In: *Meteorology and Atmospheric Physics* 40.1-3, pp. 28–60. DOI: 10.1007/BF01027467.

Siqueira, Leo and Ben P. Kirtman (2016). "Atlantic near-term climate variability and the role of a resolved Gulf Stream". In: *Geophysical Research Letters* 43.8, pp. 3964–3972. DOI: 10.1002/2016GL068694.

Smith, D. M., N. J. Dunstone, A. A. Scaife, et al. (2022). "Robust but weak winter atmospheric circulation response to future Arctic sea ice loss". In: *Nature Communications* 13, p. 727.

Smith, D. M., A. A. Scaife, R. Eade, et al. (2020). "North Atlantic climate far more predictable than models imply". In: *Nature* 583, pp. 796–800.

Stockdale, T., R. Senan, and R. Bilbao (2022). *Harmonized CAMS and CMIP6 datasets for aerosols*. Deliverable D2.1. CONFESS Project. URL: https://confess-h2020.eu/wp-content/uploads/2022/12/confess-d2-1-v1-0.pdf.

Stockdale, T., R. Senan, R. Hogan, Z. Kipling, and J. Flemming (2024). *A new time-varying tropospheric aerosol climatology for the IFS*. Technical Memorandum 914. European Centre for Medium-Range Weather Forecasts. DOI: 10.21957/BA371F56TS.

Stronmen, I. H., S. Yang, and M. Bentsen (2020). "Seasonal prediction of Arctic sea ice from summer onwards". In: *Climate Dynamics* 54, pp. 2283–2299.

Thompson, D. W. J. and J. M. Wallace (1998). "The Arctic Oscillation signature in the wintertime geopotential height and temperature fields". In: *Geophysical Research Letters* 25.9, pp. 1297–1300.

— (2000). "Annular modes in the extratropical circulation. Part I: Month-to-month variability". In: *Journal of Climate* 13, pp. 1000–1016.

— (2001). "Regional climate impacts of the Northern Hemisphere annular mode". In: *Science* 293.5527, pp. 85–89. DOI: 10.1126/science.1058958.

Trenberth, K. E., G. W. Branstator, D. Karoly, A. Kumar, N.-C. Lau, and C. Ropelewski (1998). "Progress during TOGA in understanding and modeling global teleconnections associated with tropical sea surface temperatures". In: *Journal of Geophysical Research: Oceans* 103.C7, pp. 14291–14324. DOI: 10.1029/97JC01444.

Trigo, R. M., I. F. Trigo, C. C. DaCamara, and T. J. Osborn (2004). "Climate impact of the European winter blocking episodes from the NCEP/NCAR Reanalyses". In: *Climate Dynamics* 23, pp. 17–28.

Virtanen, Pauli, Ralf Gommers, Travis E Oliphant, et al. (2020). "SciPy 1.0: fundamental algorithms for scientific computing in Python". In: *Nature Methods* 17.3, pp. 261–272. DOI: 10.1038/s41592-019-0686-2.

Wallace, J. M. and D. S. Gutzler (1981). "Teleconnections in the geopotential height field during the Northern Hemisphere winter". In: *Monthly Weather Review* 109, pp. 784–812.

Wanner, Heinz, Stefan Brönnimann, Carlo Casty, Dimitrios Gyalistrasm, Jürg Luterbacher, Christoph Schmutz, David B. Stephenson, and Eleni Xoplaki (2001). "North Atlantic Oscillation – Concepts and Studies". In: *Surveys in Geophysics* 22, pp. 321–382.

Weisheimer, A. and T. N. Palmer (2014). "On the reliability of seasonal climate forecasts". In: *Journal of the Royal Society Interface* 11.96, p. 20131162.

Weisheimer, A., N. Schaller, C. O'Reilly, D. A. MacLeod, and T. Palmer (2020). "Addressing model error through atmospheric stochastic physical parametrizations: impact on the coupled ECMWF seasonal forecasting system". In: *Philosophical Transactions of the Royal Society A* 378.2164, p. 20190290.

Willison, J., W. A. Robinson, and G. M. Lackmann (2015). "The importance of resolving mesoscale latent heating in the North Atlantic storm track". In: *Journal of the Atmospheric Sciences* 72, pp. 2234–2250.

Woollings, T., D. Barriopedro, J. Methven, et al. (2018). "Blocking and its response to climate change". In: *Current Climate Change Reports* 4, pp. 287–300.