

Southern Annular Mode Persistence and Westerly Jet: A Reassessment Using High-Resolution Global Models

Ting-Chen Chen¹, Hugues Goosse¹, [Cécile Davrinche¹](#), [Stephy Libera¹](#), [Christopher Roberts²](#), Matthias Aengenheyster², Kristian Strommen³, [Christopher Roberts²](#), Malcolm Roberts⁴, Rohit Ghosh⁵, Jin-Song von Storch⁶, [Stephy Libera¹](#)

¹Earth and Life Institute, Université Catholique de Louvain, [Ottignies](#)-Louvain-la-Neuve, Belgium

²European Centre for Medium Range Weather Forecasting (ECMWF), Reading, United Kingdom

³Department of Atmospheric, Oceanic, and Planetary Physics, University of Oxford, Oxford, UK

⁴Met Office Hadley Centre, Exeter, United Kingdom

10 ⁵Alfred Wegener Institute Helmholtz Centre for Polar and Marine Research, Bremerhaven, Germany

⁶Max-Planck Institute for Meteorology, Hamburg, Germany

Correspondence to: Stephy Libera (stephy.libera@uclouvain.be)

Abstract. This study evaluates the performance of high-resolution (grid sizes of 9–28 km for the atmosphere; 5–13 km for the ocean) global simulations from the EERIE project in representing the persistence of the Southern Annular Mode (SAM), a [critical driver of Southern Hemisphere climate variability. Using the decorrelation timescale of the SAM index \(\$\tau\$ \), we compare EERIE coupled and atmosphere-only \(AMIP\) simulations with CMIP6 and ERA5 datasets. EERIE coupled simulations improve the long-standing biases in SAM persistence, especially in early summer, with \$\tau\$ values of 9–17 days compared to CMIP6's 9–32 days. This improvement generally correlates with a more accurate climatological jet latitude \(\$\lambda_0\$ \) distribution in EERIE simulations than in CMIP6, but such a correlation is not robust within EERIE AMIP simulations with a well-represented jet location, suggesting other factors in play. With prescribed SSTs, EERIE AMIP show even smaller biases in both \$\tau\$ and \$\lambda_0\$ than EERIE coupled runs, highlighting the critical role of SST representation. Using the same AMIP model, finer grids \(9 km vs. 28 km\) can further reduce \$\tau\$, but the underlying cause remains unclear, likely because of potential compensation between different processes. Sensitivity experiments filtering ocean mesoscale features in SST boundary conditions suggest that mesoscale processes enhance SAM persistence by \$\sim 2\$ days in early summer, though this effect is clear in ensemble means at 28 km but not in the single 9 km](#)

15

20

25

runs. We also show that the atmospheric eddy feedback strength is a better indicator than λ_0 to infer the SAM persistence, although the metric alone does not fully explain the τ differences across SST scenarios. These findings underscore the interplay of dynamic processes influencing SAM persistence and offer insights for advancing global climate model performance. leading mode of Southern Hemisphere climate variability. Using the decorrelation timescale of the SAM index (τ), we compare EERIE simulations with CMIP6 models and ERA5 reanalysis.

EERIE simulations reduce long-standing biases in SAM persistence, especially in early summer, with τ values of 9–20 days compared to CMIP6’s 9–32 days and ERA5’s 11 days. This improvement correlates with a more accurate climatological jet latitude (λ_0). EERIE atmosphere-only AMIP runs outperform the coupled simulations in both τ and λ_0 , showing smaller biases and ranges of variability, underscoring the critical role of SST representation in shaping atmospheric circulation. In these AMIP experiments, the atmospheric eddy feedback strength, combined with the damping timescale estimated via friction, correlates more strongly with τ than λ_0 . We speculate that the well-captured jet position (biases $<1^\circ$ relative to ERA5), due to prescribed SSTs, limits λ_0 ’s explanatory power for τ differences, allowing other processes to dominate. Using a finer model grid (9 km vs. 28 km) of the same AMIP model results in reduced τ , though the mechanism remains unclear. Finally, motivated by the importance of oceanic eddies in the Southern Ocean, we conducted sensitivity experiments that filter transient mesoscale features from the SST boundary conditions. The results suggest that oceanic eddies may enhance summertime SAM persistence (by ~ 2 days), though this signal is not statistically significant and is absent in the single 9-km run, pointing to a subtle role of mesoscale ocean-atmosphere interaction that remains to be explored.

1 Introduction

Over the extratropical Southern Hemisphere, the daily- to decadal climate variability is dominated by the Southern Annular Mode (SAM), a mode of natural variability manifested in the large-scale oscillation of atmospheric mass between mid- and high-latitudes and hence ~~the north-southward shift and intensity changes~~ of the eddy-driven jet in the midlatitudes (e.g., Fogt and Marshall, 2020). This internal variability both influences and is influenced by the atmospheric circulation, affecting regional temperatures and precipitation patterns, sea ice extent, and ocean circulation, with consequences for

global heat and carbon redistribution (e.g., Doddridge and Marshall, 2017; Gillett et al., 2006; Lefebvre and Goosse, 2005; Lenton and Matear, 2007; Lovenduski and Gruber, 2005).

As inferred by its name “annular”, the spatial structure of SAM is approximately “ring-shaped” when viewed from above the South Pole and is nearly barotropic in the vertical direction (Gerber et al., 2010).

60 During the positive phase of SAM, lower air pressure anomalies overlay Antarctica while higher pressure anomalies spread over the mid-latitudes, and such anomalous pressure distribution indicates a strengthening and poleward shifting of the westerly jet that climatologically sits at around 50°S (Lim et al., 2013). While the SAM can, to a first approximation, be described from a zonal-mean perspective, its structure can deviate from the zonal mean and vary across different timescales, affected by factors such
65 as the seasonal cycle of midlatitude jet (atmospheric eddy activity), sea surface temperature (SST) variability, tropical oscillations such as the El Niño-Southern Oscillation (ENSO), stratosphere-troposphere interactions and so on (e.g., Campitelli et al., 2022; Ding et al., 2012; Fogt and Marshall, 2020; Karoly, 1989). On the seasonal scale, SAM is overall more zonally symmetric in austral summer (DJF) but exhibits asymmetric wavenumber 3 components when entering [springautumn](#) (MAM) and
70 [summerwinter](#) (JJA). Readers interested in a comprehensive review of the SAM literature are encouraged to consult Fogt and Marshall (2020) and Thompson et al. (2011).

~~An important~~[A key](#) characteristic of SAM is its temporal persistence, referring to how long a given phase of the SAM (positive, negative or neutral) tends to last before transitioning. ~~‡This long persistence is important as it provides a source of predictability at a timescale longer than the one associated with~~
75 [synoptic variability \(e.g., Robinson 2000; Lorenz and Hartmann 2001, Simpson and Polvani 2016\). SAM persistence](#) is often measured as the decorrelation timescale (e-folding timescale) which indicates the average duration over which the SAM index remains strongly correlated with its past values. [A standard explanation attributes the extended SAM persistence to the reinforcement of westerly flow anomalies by atmospheric eddy momentum fluxes, which are generated by changes in the mean flow and act to](#)
80 [counteract dissipation from surface friction. Several mechanisms can be at the origin of this eddy-mean flow feedback that reinforces the shifted jet, including barotropic processes related to anomalous wave propagation and breaking and baroclinic processes associated with eddy generation and enhanced baroclinicity in the lower troposphere in response to shift in the westerly flow \(e.g., Robinson 2000,](#)

[Lorenz and Hartmann 2001](#), [Zurita-Gotor et al. 2014](#), [Hassanzadeh and Kuang, 2019](#)). The westerly flow anomalies also induce changes in the diabatic heating and cooling due to latent heat release and cloud radiative effect that modify the temperature gradients, affecting SAM persistence ([Xia and Chang 2014](#), [Smith et al. 2024](#), [Vishny et al. 2024](#)). In addition to this eddy-mean flow feedback, SAM persistence can have an origin from the stratosphere, which introduces some non-stationary forcing to SAM. The main influence is likely in late spring and summer at the time of the seasonal breakdown of the stratospheric vortex ([Simpson et al. 2011](#), [Byrne et al. 2016](#), [Byrne et al. 2017](#), [Saggioro and Shepherd 2019](#)). Furthermore, interactions between a stationary mode and a propagating mode of the zonal variability could also affect SAM persistence ([Lubis and Hassanzadeh 2021](#), [Sheshadri and Plumb 2017](#), [Smith et al. 2024](#)).

While global climate models (GCMs) have shown good skills in capturing the spatial structure of SAM variability, a long-standing ~~common~~ challenge for GCMs is that they tend to overestimate the SAM persistence during the austral summer. Based on global reanalysis data, the SAM decorrelation timescale is found to be approximately 10 days on annual mean and is a couple of days higher in early summer (November–January; NDJ), during which period GCMs typically show values that are two to three-times larger ([Bracegirdle et al., 2020](#)). Many studies have found a strong dependency between the SAM persistence bias and the bias in the climatological westerly jet location (e.g., [Kidston and Gerber, 2010](#); [Simpson et al., 2013](#)[2013a, b](#); [Simpson and Polvani, 2016](#); [Son et al., 2010](#)), that is, GCMs showing too persistent SAM tend to be associated with a too equatorward-placed tropospheric westerly jet. A possible explanation for such a relationship is that the structure of the climatological jet can affect the tropospheric eddy-mean flow feedback, ~~the process by which small-scale atmospheric eddies interact with large-scale climate anomalies to amplify them~~, and models with lower latitude jets exert stronger feedback to maintain SAM ([Codron, 2005](#); [Simpson and Polvani, 2016](#)).

However, the climatological position of the midlatitude jet is not the only factor for the overly persistent SAM variability in GCMs. [Simpson et al. \(2013a\)](#) performed a series of experiments with nudging and bias correcting procedures using a stratosphere-resolving GCM, the Canadian Middle Atmosphere Model (CMAM). They found that the SAM persistence bias remains even when the representation of the climatological tropospheric winds is artificially improved. Similar conclusions are obtained when another

common bias for the overly-persistent summertime SAM —~~specifically,~~ the delayed breakdown of the stratospheric vortex— was manually nudged toward the reanalysis-based seasonal climatology. Based on these results, they suggested that a substantial proportion of the SAM timescale bias arises from “internal” tropospheric dynamics, specifically the atmospheric eddy-mean flow feedback.

~~Such a feedback mechanism assumes that the shifted midlatitude jet associated with the positive/negative SAM acts as a source of eddies. When eddies propagate away from the source region before breaking, convergence of eddy vorticity flux is produced in the upper troposphere. Such eddy vorticity flux convergence can reinforce the shifted jet by enhancing the baroclinicity through adiabatic heating/cooling associated with the induced secondary circulation, thus extending the SAM persistence against dissipation from surface friction. Within this framework, Barnes and Hartmann (2010) performed a budget analysis of the relative vorticity tendency equation in a global reanalysis. Their analysis showed that this feedback is present across the hemisphere in austral summer but lacking over the western Pacific in winter due to a weaker climatological midlatitude jet in that region. Such a seasonal variation appears to explain the longer SAM timescale in summer than in winter. Following a similar concept, Simpson et al. (2013) performed momentum budget analysis of the vertically averaged and zonally averaged zonal wind associated with SAM and confirmed that the maintenance of SAM anomalous wind is dominated by the eddy momentum flux convergence that compensates the negative contribution by friction. They also showed~~Simpson et al. (2013b) examined the Coupled Model Intercomparison Project Phase 5 (CMIP5) climate models, and found that the overly persistent SAM is highly correlated (coefficient of 0.83) with the too strong total eddy feedback during the summer season ~~in the Coupled Model Intercomparison Project Phase 5 (CMIP5) climate models.~~

As GCMs improve in their representation of physics, resolution, and overall complexity, some advancements have been made in reducing biases associated with SAM persistence and the climatological jet latitude. Compared to earlier versions of CMIP models (e.g., CMIP3, CMIP5), noticeable reductions in these biases have been reported. Bracegirdle et al. (2020) found that the ensemble-mean bias in the westerly jet latitude decreased from 1.9° in CMIP5 to 0.4° in CMIP6 on an annual mean basis. Consistently, the early-summertime SAM persistence was reduced from approximately 30 days in CMIP5 to 20 days in CMIP6. Nevertheless, the SAM decorrelation timescale remains systematically biased.

140 While higher resolution is generally regarded as beneficial, it is worth exploring whether additional improvements are achievable by further increasing the resolution or if other factors become increasingly significant when the resolution has reached beyond a sufficiently high level.

This potential role of an increase in resolution in reducing the biases in SAM persistence and the relationship with the mean-state westerly jet is investigated here in new experiments conducted as part of
145 the Horizon Europe project European Eddy-Rich Earth System Models (EERIE) (M. J. Roberts et al., 2024a). A distinctive feature of the ~~high-resolution~~ atmosphere-ocean coupled Earth System Models (ESMs) built under EERIE is their adoption of high oceanic resolutions to explicitly represent ocean mesoscale processes, which have been increasingly recognized as critical for weather and climate simulation ~~in recent studies (e.g., Busecke and Abernathy, 2019; Chassignet and Xu, 2021; Hewitt et al.,~~
150 ~~2020).~~ (e.g., Busecke and Abernathy, 2019; Chassignet and Xu, 2021; Hewitt et al., 2020). Mesoscale oceanic features can influence SAM persistence by strongly affecting surface heat fluxes and surface stress in the Southern Ocean—a hotspot of mesoscale activity (Frenger et al., 2013; Bishop et al., 2017). These ocean-atmosphere interactions can alter atmospheric temperature gradients and boundary layer structure, modifying diabatic heating and low-level baroclinicity, both of which have been linked to SAM persistence (Xia and Chang, 2014; Smith et al., 2024; Robinson, 2000; Zurita-Gotor et al., 2014). Furthermore, surface stress also plays a role as it tends to damp the westerly winds but also to enhance baroclinicity and the baroclinic feedback (Robinson 2000, Zurita-Gotor 2014, Vishny et al. 2024).

In addition to the development of new coupled models, EERIE also includes a suite of atmosphere-only simulations and idealized experiments to facilitate exploration of the atmosphere response to the ocean
160 mesoscales by excluding effects attributed to the air-sea coupling and SST biases. Those experiments will allow disentangling the role of the explicit resolution of the eddies compared to the one of increasing the model resolution. Using those experiments, we specifically analyze the potential role of the mesoscale oceanic eddies on SAM persistence, a contribution that has been studied to a minimum to date. The data sources and diagnostics are detailed in Sections 2 and 3, respectively, followed by the results in Section
165 4 and the conclusions in Section 5.

2 Data

2.1 EERIE models & simulations

Running from January 2023 to December 2026, the EERIE project aims to build new generations of ESMs
run at “eddy-rich resolution” (note that the “eddy” here refers to ocean eddies), which explicitly resolve
ocean mesoscale processes with scales of 10–100 km. Crucial components at this scale include the ocean
eddies (analogous to cyclones in the atmosphere) and boundary/frontal currents. EERIE will deliver
simulations over multi-centennial timescales centered on four global coupled ESMs and two atmosphere-
only models, with an overarching objective to reveal and to quantify the role of ocean mesoscales in
shaping the climate trajectory over seasonal to centennial time scales, regionally and globally (European
Commission, 2022). ~~The simulations are organized in two phases, with the Phase 1 providing guidance
for the planning of Phase 2.~~

2.1.1 Coupled simulations

This study evaluates the preliminary EERIE Phase 1 simulations ~~based on their~~(Wachsmann et al., 2024).
~~To facilitate direct comparison across experiments, all outputs were~~ regridded ~~output onto~~ a uniform
0.25° × 0.25° grid ~~(Wachsmann et al., 2024). prior to analysis, except for the westerly jet location~~
~~identification (Section 3.2).~~ A detailed description of the EERIE models can be found in M. J. Roberts et
al., (2024a), and ~~here we~~Table 1 briefly summarizes the simulations used in the current study ~~(Table 1).~~
~~We analyze five coupled, IFS-FESOM2 and ICON model~~ simulations ~~in total: four~~are conducted
following a protocol similar to the CMIP6 HighResMIP (High Resolution Model Intercomparison
Project; Haarsma et al., 2016), ~~using either the IFS-FESOM2 or ICON models, and one following~~
~~similar).~~ The HadGEM3-GC5-EERIE model simulation follows protocol similar as CMIP6 DECK
(Diagnostic, Evaluation and Characterization of Klima; Eyring et al., 2016) ~~with the HadGEM3-GC5-~~
~~EERIE model).~~ HighResMIP differs from CMIP6 DECK primarily in its use of 1950s’ climate conditions
instead of 1850s’ as the initial state and a shorter spin-up (~50 years instead of >= 200 years; which have
been discarded and not counted in the simulation length shown in Table 1) due to the computational
demands of high-resolution models. Using the IFS-FESOM2 model, we analyze ~~selected segments of the~~
~~spin-up and a 65-year~~ control ~~runs (referred to as 1950spinup and 1950control), both simulation~~

195 conducted under fixed 1950 forcings. ~~The analyzed segments vary in length depending on data availability, as detailed in Table 1. Additionally, we examine a portion of the (referred to as 1950control), along with a~~ historical simulation covering the period from 1950–1969, to 2014. For the ICON model, ~~only the 22-year 1950control and the 1950spinup~~ historical run ~~is available~~ are analyzed. For the HadGEM3 model, we ~~investigate the~~ examine a 30-year pre-industrial control simulation (*piControl*) forced by 1850 conditions.

200 ~~While~~ As these simulations cover different time periods, ~~we do not expect these differences to significantly impact the evaluation of specific model performance in SAM persistence, as~~ and some of them include transient forcing, linear and low-frequency nonlinear trends and a slowly varying climatology are removed as standard procedures in the SAM-related diagnostics. ~~However, it~~ This should ~~be borne in mind that~~ reduce the impact of the difference in experimental design on the evaluation of the model

205 performance. However, this removal ~~may~~ does not fully eliminate ~~all the~~ non-stationary features ~~in the second moments. Therefore, when evaluating EERIE runs against~~ that could have a clear influence on the evaluation of SAM persistence and of the eddy feedbacks (Byrne et al. 2016). We therefore adopt a bootstrapping procedure (Section 3.1) to provide partial quantification of the influence of non-stationarity and uncertainty due to the short period of some simulations. As will be shown later, through bootstrapping

210 resampling, different results can be obtained with the same model even after de-trending. We also provide results using two different periods of the ERA5 reanalysis (Section 2.2), a fairer comparison would be based on a similar earlier period, but the best available option starts from 1958,) as observation in the Southern Hemisphere were sparse before then references for comparisons. Note that the difference can be partly attributed to the larger data coverage after 1979 in ERA5.

Table 1. EERIE simulations analyzed in the current study.

Institution	Alfred Wegener Institute (AWI)	Max Planck Institute (MPI-M)	Met Office (MO)	European Centre for Medium-range Weather Forecasting (ECMWF)
	Coupled atmosphere-ocean models (eddy-rich)			Atmospheric model
System name	IFS-FESOM2	ICON	HadGEM3-GC5-EERIE	IFS

Model components	IFS CY48R1, FESOM2, FESIM2	ICON-A, ICON-O	UM, NEMO4.0.4, SI3	IFS CY48R1
Atmos. grid (km)	Tco1279 (~9 km)	R2B8 (~10 km)	N640 (~20 km at 50N)	Tco1279 (~9 km) Tco399 (~28 km)* *five ensemble members
Atmos. vertical levels (model top)	137 (0.01 hPa)	90 (0.01 hPa)	85 (85 km)	137 (0.01 hPa)
Ocean grid (km)	NG5 (~13-5 km)	R2B9 (~5 km)	eORCA12 (~8 km)	-
Ocean vertical levels	70	72	75	-
Protocol	CMIP6 HighResMIP		CMIP6 DECK	HighResMIP2
Simulations (analyzed segment lengths)	1950control (65 yrs) Historical (1950–2014)	1950control (22 years) Historical (1950–2014)	piControl (30 yrs)	Historical (1980–2023)

2.1.2 Atmosphere-only simulations & sensitivity experiments

The EERIE AMIP simulations were performed for the historical period of 1980–2023 following the HighResMIP2 highresSST-present experimental design (M. J. Roberts et al., 2024b). [We analyze](#) the simulations ~~were~~ produced with the IFS model, ~~tested with~~ [in](#) two model grid sizes (~28 km and ~9 km; [both with convection parameterization](#)), and the higher-resolution configuration is identical to the atmosphere component of the coupled IFS-FESOM2 (Table 1). One member has been performed at the 9-km resolution, but the 28-km simulations are supplemented with five ensemble members to represent a range of model uncertainty or noise. These ensembles are generated by perturbing the atmospheric initial conditions for January 1, 1980, using the same methodology employed in operational ECMWF ensemble forecasts (C. Roberts et al., 2024a).

The prescribed boundary conditions ~~include~~ [are taken from](#) the daily-mean SST reanalysis from the European Space Agency Sea Surface Temperature Climate Change Initiative (ESA CCI SST v3) and the daily-mean sea-ice concentration from the European Organisation for the Exploitation of Meteorological Satellites (EUMETSAT) Ocean and Sea Ice Satellite Application Facility (OSI-SAF), both retrieved on

a $0.05^\circ \times 0.05^\circ$ grid. External radiative forcings are generally specified following CMIP6/HighResMIP protocols, and the specificity can be found in C. Roberts et al. (2024a).

235 ~~Taking the IFS-AMIP simulations as the reference (ObsSST), a series of idealized experiments with the SST boundary condition modified are carried out as part of the EERIE project~~ To enable exploration of the response of the atmosphere to the extratropical SST ocean mesoscale features ~~(C. Roberts et al., 2024b). Two types of modifications are designed. One is, EERIE project also conducted idealized experiments with modified SST boundary conditions. Taking the NoFronts experiment, in which IFS-AMIP simulations as the quasi-stationary features in the extratropics are smoothed out by applying a spatial low-pass filter to the climatological mean field of SST. The other is the control experiments~~
240 ~~(denoted as ObsSST), NoEddies experiment, with experiments have~~ the transient oceanic eddy features removed from their SST boundary conditions with the ~~same a spatial~~ low-pass filter ~~but~~ applied to the SST anomaly field (difference from the climatological mean). Sea ice cover remains unchanged in ~~all experiments. NoEddies.~~ We emphasize that such ~~a design of NoEddies or NoFronts experiments~~ only allows us to test the direct thermodynamic impact from ocean mesoscales (as reflected in SSTs) but not
245 their relative winds-currents effects (wind stress feedback; or called current feedback).

The ~~filter being~~ employed filter is a Gaussian filter from the GCM-Filters Python package (Loose et al., 2022). The filter length scale is set to be $20L_R$ ~~with a lower and higher limit of 30 km and 700 km, respectively,~~ where L_R is the spatially varying climatological Rossby radius in the ocean. ~~with a lower and higher limit of 30 km and 700 km, respectively.~~ While ~~a filter depending on~~ a smaller L_R at high
250 latitudes can effectively remove the smaller oceanic eddies there, it also removes the larger-scale tropical instability waves near the equator as L_R reaches maximum. The latter consequence may obscure the interpretation on the impact of the targeted extratropical ocean mesoscales due to the tropical-extratropical teleconnections. To avoid this, low-latitude areas are masked out from the filtering with ~~the a~~ function ~~with values~~ ranging from 0 to 1: $M(\lambda) = \frac{1}{2} \left(\tanh \left(\frac{|h-\lambda|}{s} \right) + 1 \right)$, where $h=10$ determines the latitude where
255 the M value is halved (0.5) and $s=3$ scales the steepness of the masking function. Like the ~~reference simulations~~ ObsSST, the NoEddies experiment is run with two model grid sizes of ~ 28 km (five ensemble

members) and ~9 km (one member), ~~and the NoFronts setup is only performed for the 28 km ensemble.~~
For more details of the experimental design, we refer readers to ~~(C. Roberts et al., (2024b).~~

2.2 CMIP6 models & ERA5 reanalysis

260 For the diagnostics of SAM persistence and westerly jet characteristics, the CMIP6 models are used to
compare with EERIE models. We ~~utilize the~~analyze 31 CMIP6 historical simulations from the first
ensemble member that provide outputs of daily geopotential at 500-hPa level and monthly zonal wind at
850 hPa. ~~A total of 31 simulations are available, and only the period of 1980-2014 is extracted to ensure~~
~~a uniform data length.~~ All CMIP6 outputs are regridded to a uniform $1^\circ \times 1^\circ$ grid with the bilinear
265 interpolation before performing the analysis. ~~and only the period of 1980-2014 is extracted to ensure a~~
~~uniform data length.~~ As a proxy of observation, we use the global reanalysis dataset ERA5 (Hersbach et
al., 2020) for the same variables and a total period from 1958 to 2023. ~~The chosen period of ERA5-based~~
~~analysis will be adjusted to align with the corresponding period of the target simulation for comparison.~~
~~to cover the earlier period included in some EERIE simulations. Among the reanalysis products that~~
270 ~~extend backwards in time beyond 1979 (ERA5, 20CRv3, JRA-55), ERA5 is found to agree best with~~
~~station observations and produces good representation of SAM, both before and after the advent of~~
~~satellite sounder data (Marshall et al., 2022).~~ While we analyze ERA5 on the commonly distributed
 $0.25^\circ \times 0.25^\circ$ grid, we have tested the impact with regridding it to the $1^\circ \times 1^\circ$ grid and ~~fi~~ound no notable
changes in our results.

275 3 Diagnostics

For the overall assessment of model performance, the diagnostics described in subsections 3.1 and 3.2
below are applied to all available CMIP6 historical and EERIE simulations. Due to the limited
accessibility of the EERIE data at the time of writing, diagnostics in subsections 3.3 and 3.4 are only
performed on the EERIE atmosphere-only sensitivity experiments to provide ~~more insights~~deeper
280 investigation on the tropospheric mechanisms critical to the SAM persistence.

3.1 SAM persistence timescale

Some variations exist in the definition of the SAM across the literature (Ho et al., 2012), and its persistence estimation may be sensitive to the methods employed. While many studies adopt similar methodological concepts, the details are often not fully transparent. To ensure clarity, we provide a step-
285 by-step explanation of our approach. Note that SAM is a rather barotropic feature, so even though some traditional definitions consider the vertical averaged field, we have chosen to follow Bracegirdle et al. (2020) using a single level for simplicity.

We define the SAM as the first empirical orthogonal function (EOF) of daily zonal-mean geopotential anomalies on the 500-hPa level for the region south of 20°S (Bracegirdle et al., 2020). The anomalies are
290 calculated based on Gerber et al. (2010). First, a time series of 500-hPa zonally mean $\bar{\Phi}(\lambda, t)$ is taken, where λ and t refer to latitude and time at daily intervals, respectively, and the bar indicates zonal average. Then, for each day, we subtract the global mean of 500-hPa geopotential from $\bar{\Phi}(\lambda, t)$ at each latitude, and the resulting data is linearly detrended. Lastly, a slowly varying climatology $\tilde{\Phi}(\lambda, t)$ is subtracted to remove the seasonal cycle and the low-frequency nonlinear trends associated with known external
295 forcings such as the ozone hole formation/recovery and global warming signal. The $\tilde{\Phi}(\lambda, t)$ field is derived in two steps: 1) Applying. First, a 60-day low-pass filter is applied to the detrended $\bar{\Phi}(\lambda, t)$ along the t axis, which leaves us the to retain only seasonal-scale variations. 2) Then, for Second, the same date in a time axis (t) is reindexed into calendar day (d) and year, (y). For each calendar day (e.g., Jan 1st, Jan 2nd, etc.), a 30-year low-pass filter is applied. along the y axis to capture long-term variations. If the data spans
300 less than 30 years, the averaging of average across all available years for that same calendar date is performed, and hence $\tilde{\Phi}(\lambda, t)$ is used, resulting in a fixed, repeating annual cycle that repeats across years.

The resultant anomalies $\bar{\Phi}'(\lambda, t)$ reflect the internal/natural variability. We can then obtain SAM as the first EOF of $\bar{\Phi}'(\lambda, t)$ over 20–90°S. For the computation of EOFs, $\bar{\Phi}'(\lambda, t)$ is weighted by $\sqrt{\cos(\lambda)}$ to
305 account for the decreasing distance between meridians toward the pole. The resultant leading EOF $\mathbf{e}(\lambda)$ represents the spatial patterns of SAM, and its corresponding principal component time series PC(t) is referred to “SAM index”, expressed in normalized form with zero mean and unit variance (Fig. 1a-b).

To quantify the ~~temporal~~SAM persistence ~~of SAM~~, the decorrelation time scale is computed ~~as the e-~~
~~folding timescale of~~based on the autocorrelation function of the SAM index. ~~Following the procedure in~~
 310 ~~following~~ Simpson et al. (2013), ~~we first calculate the autocorrelation function as~~(2013a):

$$\text{ACF}(d, l) = \frac{\sum_{y=1}^{N-1} \text{PC}(d, y) \text{PC}(d+l, y)}{\sqrt{\sum_{y=1}^{N-1} \text{PC}(d, y)^2 \sum_{y=1}^{N-1} \text{PC}(d+l, y)^2}}, \frac{\sum_{y=1}^{N-1} \text{PC}(d, y) \text{PC}(d+l, y)}{\sqrt{\sum_{y=1}^{N-1} \text{PC}(d, y)^2 \sum_{y=1}^{N-1} \text{PC}(d+l, y)^2}}. \quad (1)$$

~~where~~Here, the daily time series $\text{PC}(t)$ ~~has been converted to~~is reindexed as a function of calendar day
 of the year (d (e.g., Jan 1st to Dec 31st) and lag days (l), year y , and N ~~is~~denotes the total number of years
 315 ~~of the data~~. Equation (1) computes the autocorrelation of PC between a given day d and a lagged day $d+l$,
averaged over all available years. The $\text{ACF}(d, l)$ is then smoothed over a 181-day window along the d
 axis (to smoothen daily fluctuations) using a Gaussian filter with a full width at half maximum of 42 days
 (standard deviation of 8 days). Finally, for each d , an exponential curve is fitted to the smoothed $\text{ACF}(l)$
 up to a lag of 50 days using the least squares method. ~~An, and the~~ e-folding time scale (τ) is then derived
 320 ~~for a given d , representing the time at which the exponential fit of the ACF decreases to e^{-1} (Fig. 1c).~~

~~We found that the estimation of $\tau(d)$ exhibits some sensitivity to the length of the data record.~~ To provide
 a measure of sampling uncertainty ~~in the ERA5 of τ~~ , we perform 1,000 times of bootstrap resampling,
 each time redrawing all yearly $\text{PC}(d, y)$ with replacement to form a new sample as large as the original
 sample size ~~(same number of total years). N~~ . Repeating the above ACF calculation for all bootstrap
 325 samples leads us to 1,000 values of τ for a given day (Fig. 1c). ~~The distribution of these values provides~~
~~the), showing its possible range of the estimated $\tau(d)$.~~

Note that ~~we perform~~ the above EOF analysis is performed separately for each datasetall datasets to
 identify SAM as the leading mode within each simulation, allowing for potential differences in its spatial
 structure across models.

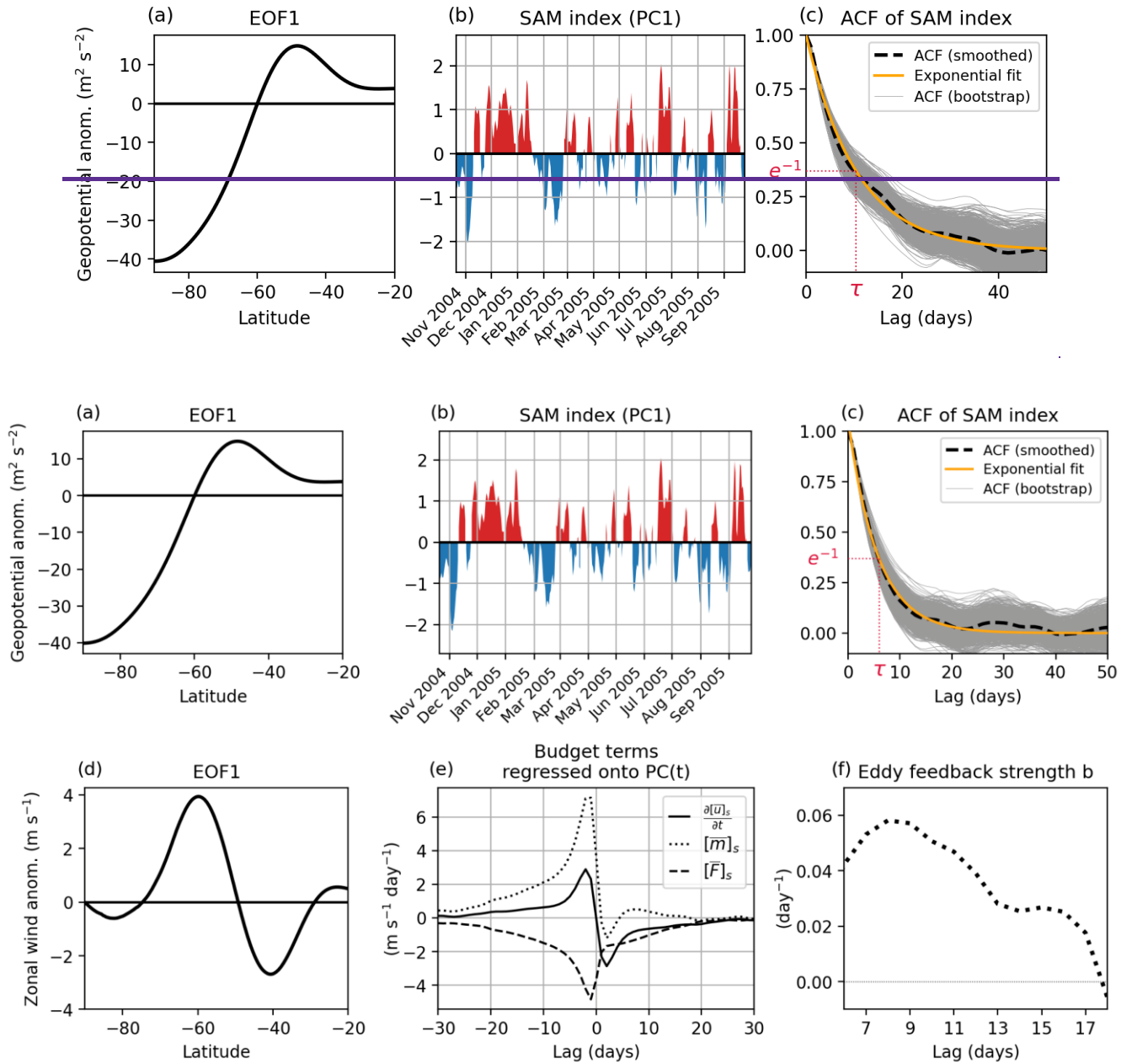


Figure 1. Example of the SAM decorrelation timescale and eddy feedback strength calculation based on ERA5: (a) The first EOF pattern based on 500-hPa geopotential; (b) The associated first PC1 time series (only a partial segment is shown here); (c) Autocorrelation function (ACF) of the SAM index (smoothed with a Gaussian filter) shown for a given day of the year (black dashed), and an exponential fit (yellow). The e-folding timescale is denoted as τ_e . The calculation of ACF is repeated 1,000 times (gray)

~~withfor the bootstrap sampling with replacement samples (gray). (d) Same as (a) but based on vertically averaged zonal wind. (e) Lagged regression of the budget terms in Eqn. (3) onto the SAM index. (f) Eddy feedback strength b for lags 7–14 days.~~

3.2 Tropospheric westerly jet position

The westerly jet position is diagnosed following ~~Bracegirdle~~Menzel et al. ~~(2020)~~based(2019) and Barnes and Polvani (2015) using the output on the native model grid. We apply a quadratic fit on the monthly mean zonally averaged 850-hPa zonal wind, at the latitude where the maximum value is found between 75°S and 10°S and the four adjacent latitudes of the model. The latitude where the maximum value of the quadratic fit is found defines the position of the tropospheric westerly jet, ~~λ_{Q}~~ .

3.3 Contribution of atmospheric eddy feedback strength to SAM persistence

~~In the literature,~~ Various methods have been proposed to assess the strength of tropospheric eddy-mean flow feedback. We adopt the approach of Simpson et al. (2013b), as it has been ~~successfully~~ applied to ~~potentially explain~~CMIP5 model evaluation and showed a high correlation with the SAM persistence bias ~~in CMIP5 models~~. This approach estimates the contribution of eddy momentum flux convergence to the tendency of SAM-associated westerly wind anomalies. Therefore, within this framework, SAM is alternatively ~~(but physics consistently)~~ described by the first EOF of vertically averaged (pressure weighted) zonal-mean zonal wind anomalies, deseasonalized and detrended, over 20-90°S. The resultant EOF latitudinal pattern (**e**) and associated PC time series are defined such that the former has units of m s^{-1} , (Fig. 1d), the latter has unit variance, and their multiplication reconstructs the SAM-associated ~~(EOF1)~~ zonal wind anomaly fields in latitude and time space. This shift from a definition of the SAM persistence timescale using geopotential height to the zonal wind for the estimation of the eddy-mean flow feedback is based on the standard assumption that geostrophic equilibrium provides a good approximation of the relevant variables. However, ageostrophic terms can also contribute to SAM persistence, introducing limitations to this hypothesis (Vishny et al. 2024; Smith et al. 2024). For simplicity and consistency with Simpson et al. (2013b) in their CMIP5 assessment, only three pressure levels of 850, 500, 250 hPa are utilized for ~~calculating this diagnostic analysis.~~

A quantity or a forcing term (denoted as X as an example) associated with the SAM is derived by projecting it onto the EOF pattern (**e**) with the operator:

$$[\bar{X}]_s = \frac{[\bar{\mathbf{X}}]\mathbf{W}\mathbf{e}}{\sqrt{\mathbf{e}^\top \mathbf{W}\mathbf{e}}}, \quad (2)$$

365 where the overbars denote the zonal mean, brackets indicate the vertical average, $[\bar{\mathbf{X}}]$ is a vector form of $[\bar{X}](\lambda, t)$, where λ and t are latitude and time, and $\mathbf{W}$ is a matrix with diagonal elements equal to the $\cos(\lambda)$ weighting when defining the EOF in Simpson et al. (2013b). The resultant $[\bar{X}]_s$ is a time series. How strongly the eddy forcing sustains the SAM wind anomalies is then estimated by projecting the vertically and zonally averaged zonal momentum tendency equation onto $\mathbf{e}$:

$$370 \quad \frac{\partial [\bar{u}]_s}{\partial t} = [\bar{m}]_s + [\bar{F}]_s, \quad (3)$$

$$[\bar{m}]_s = - \left[\frac{1}{a \cos^2 \lambda} \frac{\partial (\overline{u'v' \cos^2 \lambda})}{\partial \lambda} \right]_s, \quad (4)$$

where $[\bar{m}]_s$ is the eddy momentum flux convergence attributed to SAM, u' and v' are the deviation of the zonal and meridional velocities from their zonal means, respectively, and are calculated based on the instantaneous fields at 6-hourly intervals before being converted to daily means, a is the Earth radius, and
 375 $[\bar{F}]_s$ represents all the residual momentum forcing associated with SAM. Note that Equation (3) assumes that the sum of individual projected forcing terms on the right-hand side is in balance with the tendency of the SAM anomalies. While this assumption may not be strictly ~~mathematically~~ valid, Simpson et al. (2013b) demonstrated that it holds in their simulations.

Lorenz and Hartmann (2001) hypothesized that the eddy forcing of the SAM consists of a random
 380 component and a feedback component that depends linearly on the pre-existing state of SAM, $[\bar{m}]_s = \tilde{m} + b[\bar{u}]_s$, where b denotes the eddy feedback strength. To obtain b , Simpson et al. (2013b) performed the lagged linear regressions of $[\bar{m}]_s$ and $[\bar{u}]_s$ onto the SAM index $PC(t)$, such that for a lag day l , $[\bar{m}]_s(t+l) \approx \beta_m(l)PC(t)$ and $[\bar{u}]_s(t+l) \approx \beta_u(l)PC(t)$, where β_m and β_u are the regression coefficients- (Fig. 1e). Accordingly, the eddy forcing of SAM at lag l , $[\bar{m}]_s(t+l)$, can be expressed as
 385 $\beta_m(l)PC(t) = \beta_{\tilde{m}}(l)PC(t) + b\beta_u(l)PC(t)$. Assuming that at sufficiently large positive lags, the feedback component dominates the eddy forcing, i.e., $\beta_{\tilde{m}} \approx 0$, we can estimate the eddy feedback strength as a function of lag days (l) by

$$b(l) = \frac{\beta_m(l)}{\beta_u(l)}.$$

In Simpson et al. (2013b), the b averaged over lags from 7 to 14 days is used to denote the eddy feedback strength contributing to the SAM for the intercomparison of the models: (Fig. 1f).

The approach followed here assumes that analyzing only the first PC is a good approximation to study SAM persistence. However, although the PCs are uncorrelated by construction on short timescale, this is not the case at longer lags and the coupling between the first two components influences SAM persistence (Sheshadri and Plumb 2017, Lubis and Hassanzadeh 2021, and Lubis and Hassanzadeh 2023). Analyzing only the first PC brings thus clear limitations in our analysis of the model spread in simulated SAM persistence. Furthermore, positive regression coefficients could be caused by non-stationarity of the series and in particular by interaction with the stratosphere and not just by eddy mean flow interactions. This introduces biases in the estimate of eddy feedback, particularly in late spring and summer (Byrne et al. 2016, Byrne et al. 2017), although this does not necessarily prevent using the regression method (Ma et al. 2017). The methodology is thus imperfect, but it provides an interpretative framework for the difference between the simulations and allows a comparison with earlier studies.

3.4 Contribution of surface ~~wind stress~~friction to SAM persistence

While the eddy momentum flux convergence primarily contributes positively to the persistence of SAM, it is counteracted by the negative impacts, predominantly ~~driven~~ by the surface friction (~~appears as a subcomponent of $[\bar{F}]_s$ in Equation (3),~~ which acts to dissipate the SAM anomalies. Since the friction term for the zonal wind tendency is not directly forcing is not a standard output of EERIE simulations, we estimate it from the available from the phase 1 EERIE simulation outputs, we infer its impact from the surface variable: the turbulent wind stress in the eastward direction. This variable, *metss* (or *avg_iews* as renamed in the ECMWF encoding), represents (in units of N m^{-2}). By assuming the turbulent wind stress is zero at the model top, we can estimate the friction as $\frac{0 - \rho_0^{-1} WS_0}{H_0}$, where WS_0 indicates the daily-mean eastward turbulent ~~surface~~ stress near the surface, resulting from turbulent atmospheric eddies ~~near the surface~~ (due to the roughness of the surface) and turbulent orographic form drag, ~~represented in units of N m^{-2} . Positive.~~ For simplicity, we assume fixed values of *metss* indicate that the air flowing

over the surface exerts stress directed toward the East, which translates into a westward frictional force acting on the wind, slowing down the eastward wind tendencies.

density $\rho_0 = 1.204 \text{ kg/m}^3$ and the atmosphere column depth $H_0 = 8,464$ meters here. Following a similar approach for calculating $[\bar{m}]_s$, we multiplied metss by -1 and $[\bar{m}]_s$, we projected ~~it~~ the result onto the EOF pattern (**e**) as in Equation (2) to ~~estimate~~ obtain the frictional forcing for the SAM zonal wind anomalies, denoted as $[\bar{f}]_s$. Note that $[\bar{f}]_s$ does not have the same physical units as $[\bar{m}]_s$ or $\frac{\partial[\bar{u}]_s}{\partial t}$ (unit: m s^{-2}), and thus cannot be quantitatively compared with $[\bar{m}]_s$. Still, it is useful for the intercomparison among IFS-AMIP idealized experiments to examine whether the surface friction causing the SAM dissipation mechanism is affected by the presence of ocean mesoscales features (specifically their direct thermodynamic forcing) in the SST field: $[\bar{f}]_s$. To provide an alternative measure of friction forcing and verify the estimation, the residual term of Eq. (3), $[\bar{F}]_s$, is also computed based on the estimates of the acceleration and eddy momentum flux convergence, given the dominance of friction in this residual as shown in Simpson et al. (2013b).

It is important to note that the projection values of all budget terms are resolution (number of data points)-dependent, as defined by Eq. (2). Therefore, their magnitudes are not directly comparable across datasets with differing resolutions unless regridded to a common grid, as done here.

4. Results

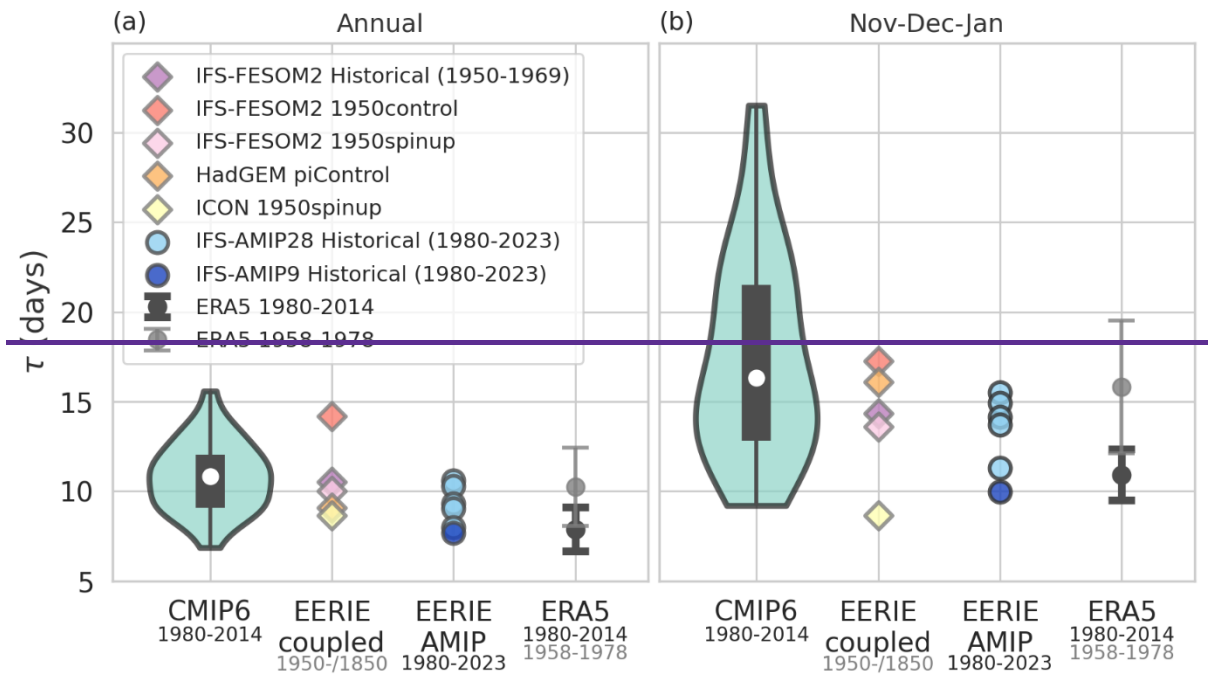
4.1 Model performance for SAM persistence

Figure 2 compares the performance of EERIE and CMIP6 models in representing SAM persistence, measured by the decorrelation time scale (τ). Consistent with Bracegirdle et al. (2020), CMIP6 models tend to overestimate SAM persistence compared to the reanalysis data analyzed over the same historical period (1980–2014). On the annual mean, CMIP6 presents a median value of 11 days, while the ERA5 shows a τ of 8 days. Systematically positive biases in τ seems to persist in the A reduced bias is found for EERIE coupled simulations, with a median value of 10 days. However, it should be noted that most of 9 days, although the distribution spread is still large, suggesting a large inter-model variability. Among these simulations, positive biases persist in the IFS-FESOM2 1950control ($\tau=13$) and historical ($\tau=11$)

440 runs, and ICON historical simulation show negative bias ($\tau=6$). Meanwhile, HadGEM3 *piControl* ($\tau=9$)
 and ICON *1950control* ($\tau=8$) runs are closer to ERA5. Given that some of these simulations are run under
 a pre-industrial 1850s' or 1950s' forcing (~~even the IFS-FESOM2 historical simulation is only available~~
~~for 1950–1969~~). The Southern Hemisphere circulation and the SAM can change over time, influenced by
 multiple factors such as the stratospheric ozone depletion and recovery, ENSO variability, changes in sea-
 445 ice extent and more. While our analysis of the SAM has removed the seasonal cycle and long-term trends,
 some non-stationary features may remain, leading to a varying τ depending on the analysis periods.
 Indeed, the EERIE median value is much closer to the we also examine the result based on an earlier-
 period ERA5 (1958–1978) ERA5 result of 10.3 days, for which τ increases to 10 days. Note, however,
 that there is relatively less confidence in the accuracy of the value of the SAM in ERA5 prior to the
 450 satellite era. Nevertheless, EERIE still show an improved agreement with biases that are smaller and more
evenly distributed on ERA5 as their τ fully cover the uncertainty ranges of ERA5 for both sides periods.
 During the austral early summer (NDJ), the overestimation of SAM persistence in CMIP6 is more
 pronounced with a longer tail of τ distribution. The maximum and median τ in CMIP6 is 32 days and
 4617 days, respectively, compared to the ERA5 value of 11 days for the same historical period. Compared
 455 to CMIP6, EERIE coupled simulations exhibit some ~~noticeable~~ improvement with the maximum and
 median ~~value of 17~~ values dropping to 20 days and ~~14~~16 days, respectively. ~~Again, such a distribution is~~
~~even less~~ However, the spread among different EERIE simulations remains large; while the positively
 biased ~~when compared to the earlier period~~ are mostly captured by IFS-FESOM2, ICON tends to exhibit
much smaller τ than ERA5 (~~τ of 16 at 6–7 days~~).
 460 Interestingly, the atmosphere-only EERIE simulations (IFS-AMIP) ~~ov~~generally outperform the ocean-
 coupled runs ~~with, exhibiting~~ a reduced positive bias in τ compared to their coupled versions (IFS-
FESOM2) and a much smaller spread ~~of τ , suggesting~~. This suggests that the prescribed historical SST
 boundary condition serves a ~~good~~strong physical constraint on the SAM persistence. With all five
 members considered, the simulated τ at 28 km is still positively biased for both annual and austral-summer
 465 means, but the biases do not exceed more than 4.5 days and at least one member presents almost identical
 values (8 days annually and 11 days in NDJ) to ERA5 ~~7~~ (1980–2014). Refining the atmospheric resolution
 from 28 km to 9 km ~~lowers~~ suggests a lowering of the SAM decorrelation timescale ~~to become slightly~~

negatively biased (, with τ of 8 days annually and 10 days in NDJ) although they are still within the measurement uncertainty ranges of ERA5.

470



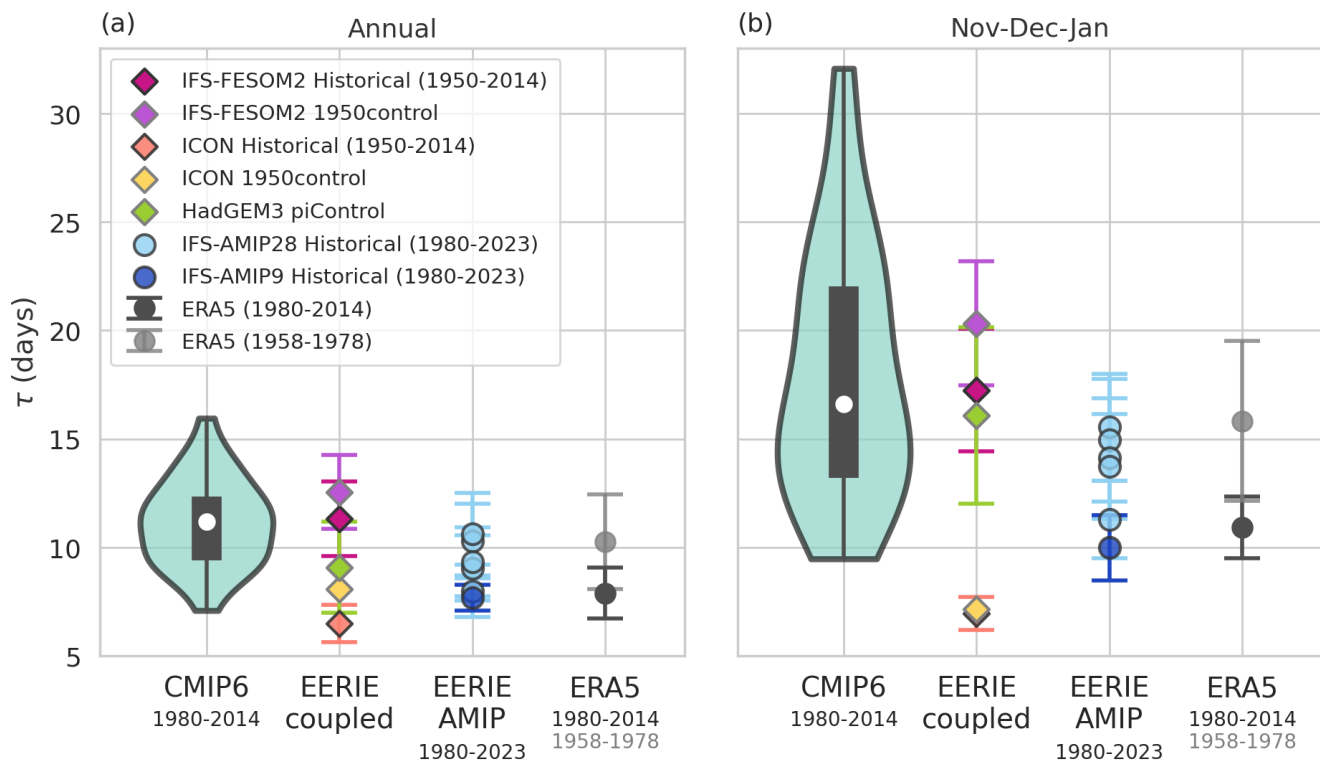


Figure 2. Distribution of τ (days) in CMIP6, EERIE coupled, and EERIE atmosphere-only (AMIP) simulations. The simulations from CMIP6 and EERIE AMIP are both historical simulations, with their fixed period indicated in the x-axis labels, and the simulations from EERIE coupled vary as simulations cover varied periods as indicated in Table 1. ERA5 is analyzed for two time periods as references for. CMIP6 results from 31 experiments are presented in violin plot, in which the historical width indicates the density of the data points, the thin gray vertical box in the middle shows the 25th – 75th quantiles, and pre-industrial periods, both expressed in the white dot presents the median. For the rest, error bars showing are added wherever applicable to show the ± 1 standard deviation of the results from the 1,000 bootstrap resampling.

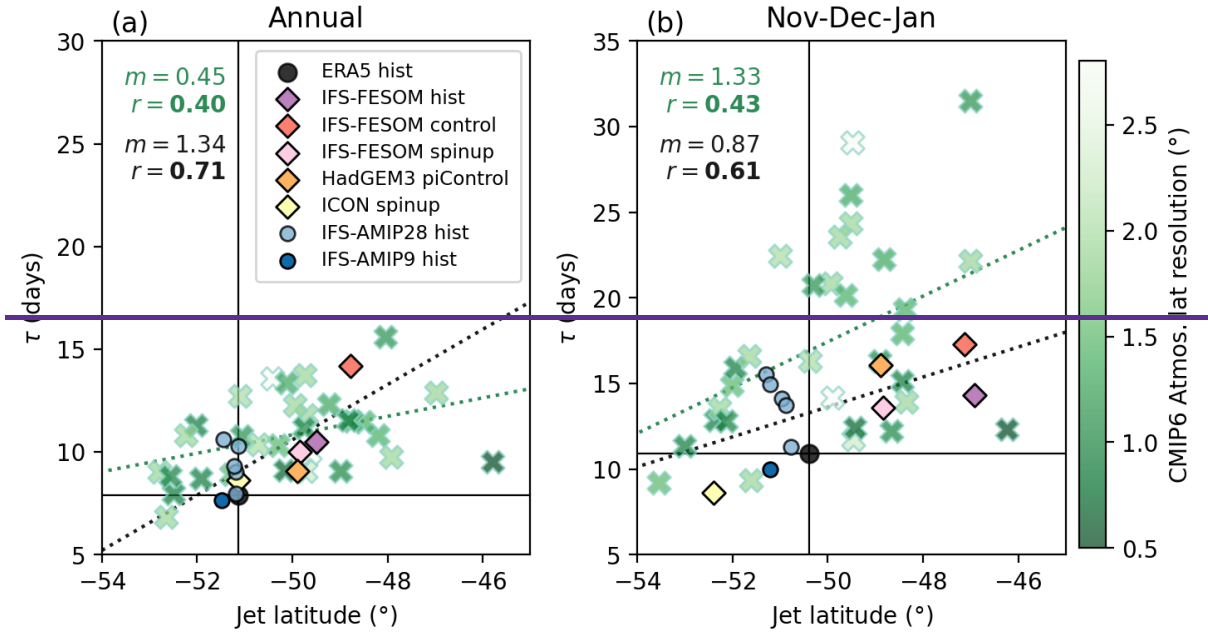
4.2 The relationship between jet location and τ

The bias relationship between westerly jet location (λ_0) and SAM decorrelation timescale (τ) is then revisited. Similar to their predecessors, CMIP6 models show a positive correlation between λ_0 and τ , that is, models with a more equatorward jet location tend to exhibit a more persistent SAM. Consistent with Simpson & Polvani's (2016) result based on CMIP5 models, the slope of the linear fit is larger during NDJ, indicating a larger variation in τ given the same variation in λ_0 during this season. Examining the

model resolution of each CMIP6 simulation, there appears no strong or clear relationship between the model resolution and the model biases in these two quantities (same for both latitudinal and longitudinal resolutions and for both atmosphere and ocean models, although only the atmospheric latitudinal resolution is expressed in Fig. 3). A potential dependency on resolution could be obscured in the CMIP6 ensemble by other cancelling factors, which vary from simulations incorporating different configurations and model systems. However, it is also possible that the ~~importance of resolution (or the improvements typically attributed to a higher resolution) in determining on~~ the performance of ~~the~~ large-scale SAM variability and ~~the~~ mean jet has ~~reached~~ a plateau ~~withat~~ the ~~typical grid size reached~~ sizes used in the current GCMs (~~e.g., CMIP6) and other factors are becoming more critical.~~). For instance, based on simplified atmospheric GCMs with idealized forcing, Gerber et al. (2008) found that the decorrelation timescale of the annular mode is unrealistically large at a coarse resolution of T21 (5.6°). While such a bias was notably reduced by refining the model resolution to T42 (2.8°), no further improvement was shown with a ~~model higher~~ resolution ~~higher than T42 of T85 (1.4°)~~ and the τ converges to a still positively biased value. No test was performed in this study to determine if τ is improved again at even higher resolution or if the plateau continues.

On the annual mean, EERIE simulations generally fall within a region smaller than that covered by CMIP6, with the IFS-FESOM ~~1950~~ control being the worst performing experiment among the EERIE simulations (Fig. 3a), showing both the greatest positive bias in τ and λ_0 . For NDJ, a clear improvement of EERIE models in representing the SAM persistence is again shown as the spread of EERIE clearly shifts toward a lower τ , closer to ERA5's τ compared to other CMIP6 exhibiting a similar jet location. In all, a positive λ_0 - τ relationship remains and appears stronger in summertime across EERIE models (Fig. 3). ~~However, the slope of the linear fit is greater for annual means than for early summer, mainly because of the specific behavior of the IFS-FESOM control.~~ The most skillful EERIE simulations for the SAM persistence, IFS-AMIP, all well capture the jet location (with a bias $< 1^\circ$). This highlights again the importance of well-represented sea surface features to the large-scale atmospheric circulation and variability. Still, even with the same IFS model and the same 28-km grid size, the five IFS-AMIP ensemble members generated by perturbing the initial conditions (the only difference is the internal atmospheric variability) exhibits a spread in τ of about 5 days, which is not positively correlated with the

515 corresponding simulated jet location. This result suggests that the documented bias relationship between τ and λ_0 in the literature does not ~~always hold, particularly in our IFS-AMIP setup~~this configuration with prescribed SSTs. It is also possible that when the jet location ~~is relatively~~has already been well-captured, other factors become increasingly important to influence the persistence of SAM, and we explore some of these potential factors in the next section using idealized sensitivity experiments.



520

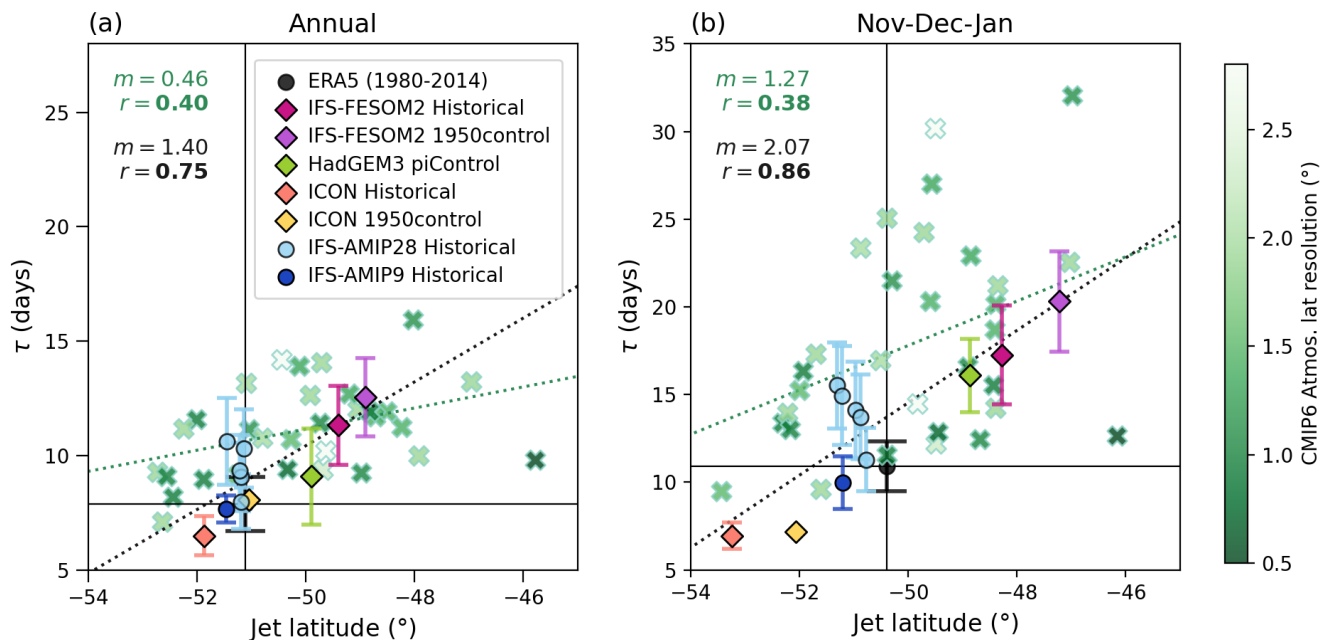


Figure 3. Scatter plot of climatological jet latitude ($^{\circ}$) versus SAM decorrelation timescale τ (days; error bar indicates ± 1 standard deviation from the bootstrapping) for (a) annual and (b) early-summer (NDJ) means in the Southern Hemisphere. Green crosses are based on CMIP6 historical simulations (colored by their latitudinal atmospheric resolution). Model names are not labeled here for visual clarity, but details are provided in their atmospheric model-Supplementary Table 1. ERA5 reanalysis and EERIE simulations are indicated as in the legend. Vertical and horizontal black lines are the ERA5 values. The green dotted straight line is the linear least-squares regression fit for CMIP6 models (slope is denoted as m , and Pearson correlation coefficient r is expressed in bold if statistically significant with the p value < 0.05 in green in the top left corner). Similarly, the black dotted fitted line is ~~the linear fit~~ for all EERIE simulations.

4.3 Sensitivities to varying SST boundary conditions

EERIE simulations demonstrate a reduced bias in summer-timesummertime SAM persistence compared to CMIP6, but identifying the cause is challenging due to variability in model systems. Although CMIP6 results show no clear link between model resolution and performance in τ and λ_0 , the higher resolution in EERIE remains one possible contributing factor to such an improvement. One piece of evidence is the reduction in τ when transitioning from a 28-km to a 9-km model grid size using a consistent IFS model. Another possibility is that the new generation of models in EERIE improves model physics, reducing the biases in processes that resulted in a too overly-persistent SAM in earlier CMIP-like GCMs. In addition,

EERIE begins to explicitly resolve the ocean mesoscales, which are parameterized in CMIP6, though the resulting impacts on SAM persistence have not been investigated. To explore these possibilities, this section focuses on EERIE atmosphere-only sensitivity experiments with varied SST boundary conditions, aiming to investigate the influence of ocean mesoscales and model resolution in the observed improvement within a controlled framework.

We first focus on the 28-km simulations. Regarding the seasonal variation of τ (Fig. 4a), the NoEddies experiments ~~in which ocean mesoscales are removed (either NoEddies or NoFronts)~~ exhibit intermingled patterns overlapping with those of ObsSST. ~~However, Although~~ their ensemble means suggest a slight reduction in τ (by approximately 2 days) ~~during NDJ. This result indicates that the presence in the absence~~ of ocean eddies — hinting that mesoscale SST features, ~~whether quasi-stationary or transient, helps may help~~ sustain SAM persistence ~~(hence—this difference is not statistically significant at 95% confidence level. For the larger 9-km configuration, the subtle impact of ocean eddies is not observed as NoEddies shows no clear changes in τ in the ensemble mean off from ObsSST) during this period., and both show smaller τ than the 28-km counterparts (Fig. 4b).~~

All these 28-km sensitivity experiments show a slightly poleward biased jet latitude compared to ERA5 (within 1°) during NDJ, and ~~the ensemble mean of~~ ObsSST is the least generally less biased ~~compared to those of~~ than NoEddies ~~and NoFronts~~ (Fig. 4b4c). While this seems to be in agreement with the literature that a more southward-shifted jet is associated with a longer SAM persistence, ~~there is an overall small and even negative the~~ correlation between ~~the~~ λ_0 and τ is weak (with a correlation coefficient of -0.22 considering 03) across all IFS-AMIP simulations in the IFS-AMIP configurations.

~~In contrast, Compared to λ_0 , the metric eddy feedback strength b appears a better indicator for~~ shows a much stronger correlation with SAM persistence ~~, exhibiting a correlation coefficient with τ of 0.47,~~ with a higher correlation coefficient of 0.52 and a lower p-value of 0.08 (Fig. 4c). While the p-value does not meet the conventional threshold for statistical significance ($p < 0.05$), it is notably lower than the examined λ_0 - τ relationship (p-value 0.42). However, the difference of b among different sets of SST boundary condition experiments cannot explain solely the observed reduction (4d), suggesting it may be a more informative indicator of SAM persistence in ~~both NoEddies and NoFronts. Specifically, while the ensemble mean of NoEddies shows a slightly weaker b than this configuration. Meanwhile, the surface~~

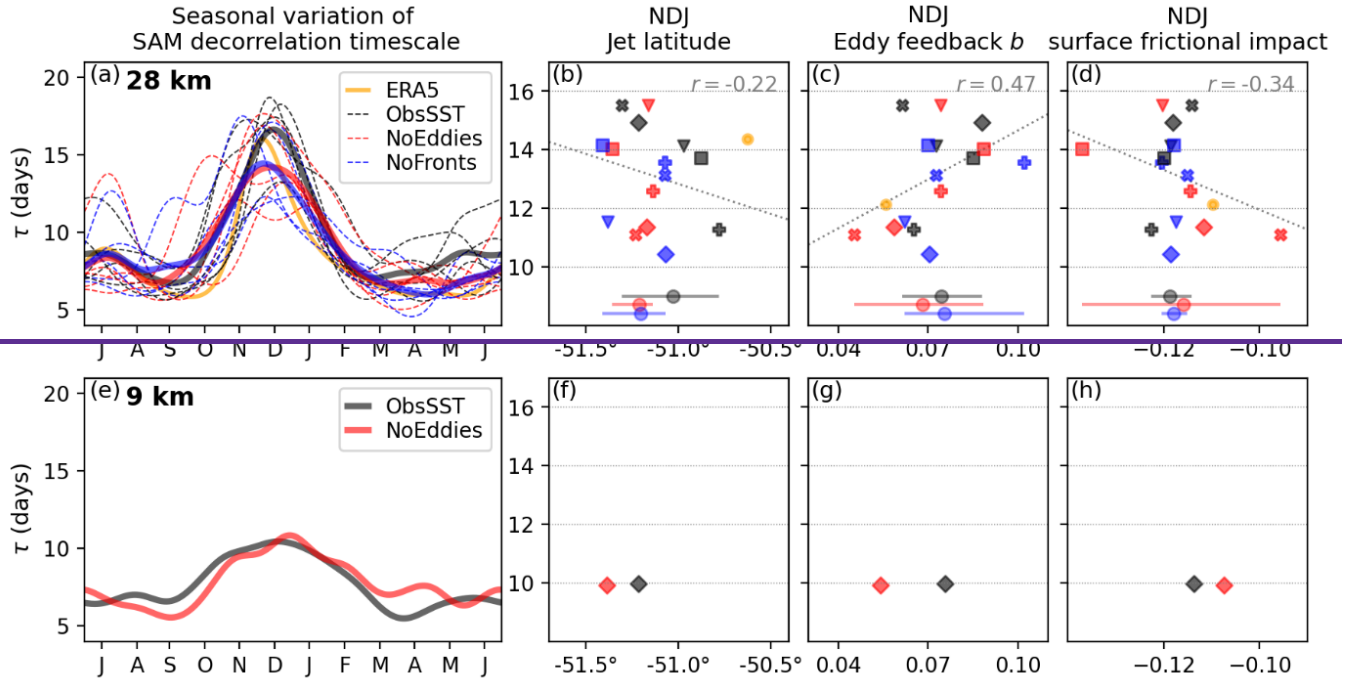
friction and τ exhibit a negative correlation (Fig. 4e) with a moderate correlation coefficient of -0.48 and p-value of 0.11. It is worth noting that of ObsSST, NoFronts our results using $[\bar{f}]_s$ based on surface wind stress show values similar with or even larger than ObsSST.

We then examine the surface frictional effects on SAM in these qualitatively consistent patterns with those using the residual estimates, $[\bar{F}]_s$, across simulations (Fig. 4d), and no clear difference in $[\bar{f}]_s$ is found among ObsSST, NoEddies, and NoFronts. Experiments ObsSST and NoFronts exhibit highly similar values, while NoEddies exhibit a large ensemble spread and a slightly larger ensemble mean value despite some differences in the absolute values (Fig. S1a, b). A closer examination shows that the member in NoEddies with the largest value in $[\bar{f}]_s$ is accompanied by the weakest eddy feedback b (red cross markers in Fig. 4c-d, e) and vice versa (red square markers). The opposite shifts of these two dominant mechanisms indicates an offset between each other, which explains why these members with the largest magnitudes of $[\bar{f}]_s$ do not appear as outliers in the distribution of SAM persistence. However, the quantification of the net effects of eddy forcing and surface friction requires a comprehensive budget analysis leading to subtle combined effects on the SAM persistence.

The potential impact from ocean mesoscale on the SAM persistence is, however, not observed in the 9 km simulations, as NoEddies shows almost no change from ObsSST in τ (Fig. 4e). That said, one ensemble member may be not enough to identify the influence of the ocean eddies as the τ fluctuation is quite large among the 28 km ensemble members, even given the same SST boundary condition. For the other metrics, the single runs of 9 km qualitatively align with the ensemble means of 28 km simulations: The suppression/filtering of ocean eddies in SST boundaries leads to a slight poleward shift (more biased than ObsSST) of the mean state jet, a weaker Lorenz and Hartmann (2001) proposed that the eddy feedback sustaining the SAM anomalies, and a slightly reduced surface can interact with the frictional impact dissipating SAM. The same signs of changes in the latter two counteracting mechanisms render their net impact on sustaining SAM uncertain. It is possible that the changes in these two processes, influenced to lengthen the effective timescale of SAM by the presence of ocean eddies, counterbalance each other at 9 km, resulting in a comparable $\tau \frac{t_f}{(1-bt_f)}$, where t_f is the damping timescale. Here, we estimate t_f by taking the ratio between NoEddies the regressed $[\bar{u}]_s$ (in unit of m/s) and ObsSST. the

regressed $[\bar{F}]_s$ (unit of m/s^2) averaged over the 7-14 lag days, which gives a value of 8.6 days for ERA5 (close to the 8.9 days in Lorenz and Hartmann (2001)). We found that this metric indeed correlates with τ more strongly than b or $[\bar{f}]_s$ ($[\bar{F}]_s$) alone with a higher correlation coefficient of 0.61 and a lower p-value of 0.03 (Fig. S1d), pointing to its superior usefulness to assess the joint impact of the two competing dominant mechanisms.

However, ~~these analyses do not~~ although those metrics explain some of the differences between individual experiments, none of them shows systematic differences between ObsSST and NoEddies and none clearly accounts for the significant reduction in τ when the model grid size is refined from 28 km to 9 km, ~~as all three diagnostics indicate that the 9 km simulation falls within the range of values covered by the 28 km simulations.~~ Considering the large variability in the 28-km ensemble members, one member at 9 km may be not enough to identify the influence of the resolution. Additional simulations and diagnostics ~~would~~ different experimental approaches may be required to confirm the underlying cause for the observed model grid spacing dependency.



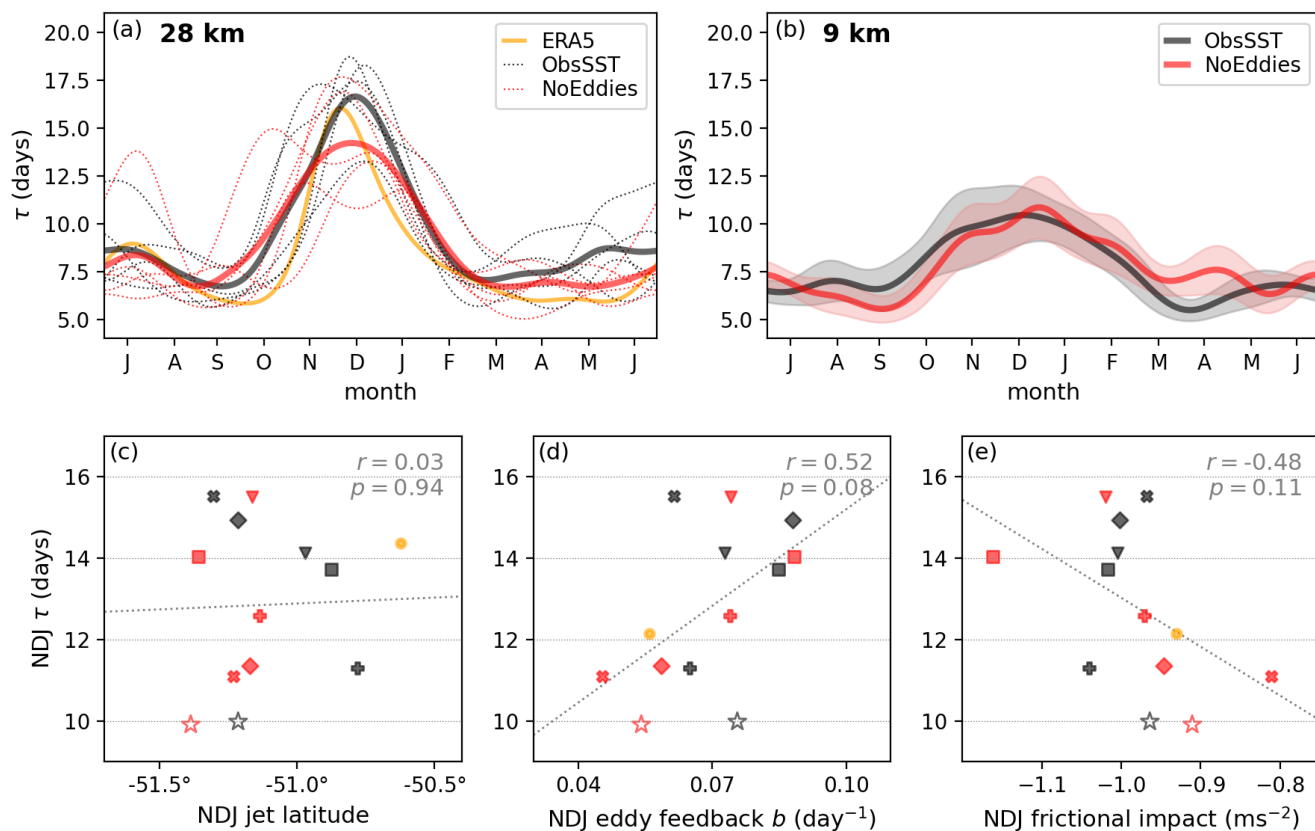


Figure 4. (a) SAM decorrelation timescale (τ) as a function of month for IFS-AMIP 28km simulations (dashed lines for each ensemble members and solid lines for the ensemble means; black for ObsSST and red for NoEddies experiments indicated in the legend) and ERA5 (yellow). (b) Similar to (a) but for 9 km experiments (shades for the ± 1 standard deviation of τ from the 1,000 bootstrap resampling). (c) Scatter plot of τ (days; y-axis) and westerly jet latitude (each marker x-axis; filled-color markers for one simulation; the horizontal bar in the bottom shows the minimum to maximum range with the circle 28 km; hollow stars for the ensemble mean of each set of the SST experiments). (e) (9 km simulations). (d)–(e) Similar to (b)–(c) but for with x-axis variable replaced with the eddy feedback strength and frictional effects associated with SAM impact, respectively. In (b)–(d), the gray dotted line represents the linear regression fit, and the correlation coefficient r and p-value are indicated in the top-right corner. (e)–(h) Same as (a)–(d), but for the 9-km simulations.

5 Discussion and conclusions

This study assesses the performance of new high-resolution global model simulations developed under the EERIE project in capturing the persistence of the Southern Annular Mode (SAM), a key driver leading mode of climate variability in the Southern Hemisphere. EERIE simulations are conducted with a model

grid size of 9–28 km for the atmosphere and 5–13 km for the ocean. The persistence of the SAM is
625 assessed using the decorrelation timescale of the SAM index (τ), for which CMIP GCMs have historically
exhibited a systematic positive bias (overly persistent) in austral summer, often correlated with a
climatological westerly jet that is too equatorward. Our conclusions and discussion based on the phase 1
preliminary simulations of the EERIE models are organized into two subsections: (1) the performance of
coupled simulations, and (2) the performance of atmosphere-only (AMIP) simulations and insights
630 obtained from the sensitivity experiments with varied SST boundary conditions under the AMIP setup.

5.1 EERIE Coupled simulations

Compared to CMIP6, the EERIE coupled simulations show ~~a slight~~ improvement in representing the
SAM persistence ~~on an~~. Although the inter-model variability remains large, the annual-mean basis and a
notable advancement in early summer (NDJ). The annual-mean τ distribution in the of EERIE coupled
635 simulations (clearly shifts to lower biases with a median: 10 value of 9 days) ~~is similar to that in the~~
~~CMIP6 historical simulations (median: 11 days), both are positively biased compared, closer~~ to the ERA5
value of 8 days ~~for 1980–2014. Notably, the EERIE simulations primarily represent~~ than the earlier 1850s
~~or 1950s period, for which reanalysis suggest a higher τ CMIP6's median of 1011 days. For~~ During early
summer ~~(NDJ), the EERIE simulations demonstrate τ values~~ pronounced long tail of τ in CMIP6
640 simulations is also noticeably reduced in EERIEs, with the former ranging from 9 to 32 days (median: 16
days) and the latter ranging from 9 to 17 days (median: 14 days), in contrast to CMIP6's overestimated
range of 9) closer to 32 days (median: 16 days). This clear improvement of the EERIE simulations for
early summer τ holds regardless of the two reference periods used for ERA5 (ERA5's 11 days for the
historical period as in CMIP6 and 16 days for an earlier period of 1958–1978).
645 . The relationship between biases in the westerly jet location (λ_o) and τ remains positively correlated in
EERIE simulations as has been ~~found~~ documented for CMIP-like models. Consistently, the smaller bias
for τ in EERIE simulations ~~during the austral summer~~ is accompanied by ~~a smaller range~~ improved
representation of λ_o ~~(closer to ERA5)~~ compared to CMIP6. However, the outperformance for τ in EERIE
models compared to the CMIP6 runs that capture similar jet locations indicates that other factors ~~in~~ are at
650 play. Among While the improvement of EERIE models compared to CMIP6, indicates that increased

resolution can offer benefits, the varied skills within CMIP6 in representing either λ_0 or τ do not show a clear dependency on the model resolution ~~shows weak dependency on biases in either λ_0 or τ , but,~~ It is possible that the impact of resolution is outweighed by other factors varying in CMIP6 simulations incorporating different configurations and model systems, or that the benefit requires the resolution exceeding a certain threshold to emerge.

5.2 EERIE Atmosphere-only simulations

Among EERIE simulations, the IFS-AMIP runs with prescribed historical SST and sea ice boundary conditions show the optimal performance in both SAM persistence and westerly jet location, with smaller spreads and closer values to ERA5 than the coupled runs. This highlights the importance of accurately representing sea surface thermal conditions to improve the simulation of these large-scale atmospheric quantities. While ~~some studies have demonstrated~~ Sen Gupta and England (2006) showed that ~~the~~ air-sea coupling is critical ~~infor~~ for modulating the SAM ~~(although they focused—albeit focusing~~ on interseasonal scales, timescales, which are longer than the intraseasonal scale ~~being~~ investigated here; ~~Sen Gupta and England, 2006),—~~ our results suggest that ~~the~~ atmosphere-ocean coupling plays a secondary role ~~relative to the.~~ Instead, SST biases introduced by the coupling, ~~—an ongoing challenge common to ocean-atmospherein~~ coupled GCMs (Zhang et al., 2023). ~~Notably, the IFS-AMIP historical simulations significantly reduce biases compared to CMIP6, with errors of less than 4.5 days in τ and 1° in λ_0 with respect to ERA5.)—~~ appear to be more influential.

For the AMIP historical simulations, the λ_0 - τ bias relationship ~~does not apply, as the approximately 5-day ensemble spread in τ shows no positive correlation with the minimally varying λ_0 .~~ It is possible is ~~virtually absent.~~ We speculate that when the jet ~~location~~ is already well -captured (all AMIP runs are with $<1^\circ$ bias) and SSTs are prescribed, other ~~factors become increasingly important for second-order processes may come into play to affect τ .~~ Indeed, we find that the metrics of atmospheric eddy-mean feedback strength, surface friction and their combination correlate more strongly with τ than with λ_0 in the AMIP configurations, highlighting the importance of these two competing dominant mechanisms on SAM persistence ~~of SAM.~~ Especially, ~~However, these metrics cannot fully explain the impact of clear reduction of τ when the model resolution stands out and can be more clearly interpreted using the same~~

model. Refining the IFS model grid size is refined from 28 km to 9 km lowers the simulated τ to even lower than the ERA5 value, although it is worth noting that ERA5 has a resolution closer to 28 km using the same atmospheric model.

The potential importance of Finally, the thermodynamic impact from the ocean mesoscale features is explored via idealized AMIP experiments. Taking the historical runs as reference (ObsSST experiment), two experiments are carried out by filtering out the quasi-stationary ocean fronts (NoFronts experiment) or the transient ocean eddies (NoEddies) in the SST boundary conditions. While three experiments (each with five ensemble members) at 28 km show intertangled seasonal variations of τ , their the difference between the 28-km ensemble means suggest of ObsSST and NoEddies imply that the presence of ocean mesoscales, regardless of fronts or eddies, mesoscale SST features may help to maintain the SAM anomalies increasing the persistence (increase τ by roughly 2 days) in early summer. However, such an effect impact is not statistically significant and is not captured in the 9-km simulations, although there is only one ensemble member for each experiment. Among the 28-km members, we also do not see a systematic change of eddy feedback or surface friction due to the presence or absence of ocean eddies in the SST field. The critical role of oceanic mesoscale eddies in the Southern Ocean climate system is well documented. While their local impact on the atmospheric boundary layer is well established, their direct influence in modulating large-scale modes such as the SAM appears limited under our AMIP setup without air-sea coupling. A similar conclusion was obtained by Purich et al. (2021) performed similar experiments but with a coarser coupled GCM, (model resolution of ~ 130 km), ACCESS1.0, to explore the role of oceanic variability on SAM by restoring the SST to the monthly mean patterns. They found that suppressing Southern Ocean SST variability by restoring the SST to the monthly mean patterns does not impact SAM persistence in their simulations, but their analysis was on monthly rather than daily means and with a coarser model resolution of ~ 130 km. As they also concluded, that eddy-resolving models are required to properly capture the air-sea feedbacks in the Southern Hemisphere.

Using the idealized AMIP experiments, we conduct further exploration on the “internal” atmospheric dynamical mechanisms contributing to the SAM persistence, eddy feedback strength and surface friction. It is worth noting that the eddy feedback strength appears to be a better indicator than the mean state jet latitude λ_0 , linking positively and more statistically significantly to the simulated summertime τ among

the IFS-AMIP simulations. Still, by how much it is counteracted by the surface friction requires a more comprehensive budget analysis. Likely because of those compensating effects, the metrics of jet latitude, eddy feedback and frictional impact, do not provide a clear answer as to what contribute to the observed differences in the simulated τ when ocean mesoscales are removed.

710 This result highlights the complexity of mechanisms contributing to the SAM persistence in GCMs. For instance, the interplay between eddy forcing and surface friction suggests that even small errors in one process may result in notable uncertainty in their net impact. Additionally Between EERIE coupled and AMIP simulations, the superior performance of the latter seems to suggest that model skills in representing SAM persistence does not clearly benefit from the two-way ocean-atmosphere coupling or

715 from the explicit inclusion of ocean mesoscale features. Our hypothesis is that while coupled models offer a more physically consistent representation of the climate system, they also tend to introduce SST biases—potentially due to under-tuning in high-resolution configurations or imbalances in the coupling process. In fact, previous studies have shown that eddy-permitting models can exhibit larger SST biases than either coarser models with parameterized eddy fluxes or fully eddy-rich models (e.g., Storkey et al.

720 2025). Reducing SST biases remains essential for advancing the representation of SAM and Southern Hemisphere climate variability. The large variability among ensemble members with the same model configuration also highlights the complexity of mechanisms contributing to the SAM persistence in GCMs and call for further investigation or different approaches to address the outstanding questions. For example, this study only considers the zonally averaged properties, but non-zonal components likely play

725 important roles in shaping SAM characteristics ~~in climate systems~~ and hence their representation in GCMs (e.g., Barnes and Hartmann, 2010; Sen Gupta and England, 2006). Nevertheless, the general improvements seen in the phase 1 simulations of the EERIE coupled models present ~~a~~ promising ~~path forward~~ results in addressing the long-standing GCM biases in SAM persistence, especially considering the challenges in optimally configuring high-resolution models (i.e., tuning) and the lack of community

730 experience in doing so. Furthermore, the controlled framework of the IFS-AMIP idealized eddy-rich experiments offers significant potential for enhancing our understanding of atmospheric responses to ocean mesoscales.

Acknowledgments

This publication is part of the EERIE project funded by the European Union. Views and opinions
735 expressed are however those of the author(s) only and do not necessarily reflect those of the European
Union or the European Climate Infrastructure and Environment Executive Agency (CINEA). Neither the
European Union nor the granting authority can be held responsible for them. This work has received
funding from the Swiss State Secretariat for Education, Research and Innovation (SERI) under contract
#22.00366. This work was funded by UK Research and Innovation (UKRI) under the UK government's
740 Horizon Europe funding guarantee (grant number 10057890, 10049639, 10040510, 10040984). Hugues
Goosse is Research Director with the Fonds de la Recherche Scientifique (F.R.S. -FNRS). The authors
also thank Francesco Ragone, Isla Simpson, Matthew Patterson, Thomas Bracegirdle for providing
constructive suggestions or clarification about the methodologies employed in their works.

Data Availability Statement

745 All EERIE simulation outputs are publicly accessible at <https://eerie.cloud.dkz.de> and Wachsmann et al.
(2024). The calculation of EOFs was performed using the publicly available Python package by Dawson
(2016). All scripts used for the analysis and figure generation will be shared in the GitHub repository
under the EERIE project (<https://github.com/eerie-project>) upon acceptance of the publication.

References

- 750 Barnes, E.A., Hartmann, D.L., 2010. Dynamical Feedbacks of the Southern Annular Mode in Winter and
Summer. J. Atmospheric Sci. 67, 2320–2330.
<https://doi.org/10.1175/2010JAS3385.1>
Barnes, E. A., and L. M. Polvani, 2015. CMIP5 Projections of Arctic Amplification, of the North
American/North Atlantic Circulation, and of Their Relationship. J. Climate, 28, 5254–5271,
755 <https://doi.org/10.1175/JCLI-D-14-00589.1>.
Bishop, S. P., R. J. Small, F. O. Bryan, and R. A. Tomas, 2017. Scale Dependence of Midlatitude Air–
Sea Interaction. J. Climate, 30, 8207–8221, <https://doi.org/10.1175/JCLI-D-17-0159.1>.
Bracegirdle, T.J., Holmes, C.R., Hosking, J.S., Marshall, G.J., Osman, M., Patterson, M., Rackow, T.,
2020. Improvements in Circumpolar Southern Hemisphere Extratropical Atmospheric Circulation
760 in CMIP6 Compared to CMIP5. Earth Space Sci. 7, e2019EA001065.
<https://doi.org/10.1029/2019EA001065>

- Busecke, J.J.M., Abernathey, R.P., 2019. Ocean mesoscale mixing linked to climate variability. *Sci. Adv.* 5, eaav5014. <https://doi.org/10.1126/sciadv.aav5014>
- Byrne, N. J., Shepherd, T. G., Woollings, T., & Plumb, R. A., 2016. Annular modes and apparent eddy feedbacks in the Southern Hemisphere. *Geophysical Research Letters*, 43(8), 3897–3902. <https://doi.org/10.1002/2016GL068851>
- Byrne, N. J., Shepherd, T. G., Woollings, T., & Plumb, R. A., 2017. Nonstationarity in Southern Hemisphere Climate Variability Associated with the Seasonal Breakdown of the Stratospheric Polar Vortex. *Journal of Climate*, 30(18), 7125–7139. <https://doi.org/10.1175/JCLI-D-17-0097.1>
- Campitelli, E., Díaz, L.B., Vera, C., 2022. Assessment of zonally symmetric and asymmetric components of the Southern Annular Mode using a novel approach. *Clim. Dyn.* 58, 161–178. <https://doi.org/10.1007/s00382-021-05896-5>
- Chassignet, E.P., Xu, X., 2021. On the Importance of High-Resolution in Large-Scale Ocean Models. *Adv. Atmospheric Sci.* 38, 1621–1634. <https://doi.org/10.1007/s00376-021-0385-7>
- Codron, F., 2005. Relation between Annular Modes and the Mean State: Southern Hemisphere Summer. *J. Clim.* 18, 320–330. <https://doi.org/10.1175/JCLI-3255.1>
- Dawson, A., 2016. eofs: A Library for EOF Analysis of Meteorological, Oceanographic, and Climate Data. *J. Open Res. Softw.* 4, 14. <https://doi.org/10.5334/jors.122>
- Ding, Q., Steig, E.J., Battisti, D.S., Wallace, J.M., 2012. Influence of the Tropics on the Southern Annular Mode. *J. Clim.* 25, 6330–6348. <https://doi.org/10.1175/JCLI-D-11-00523.1>
- Doddridge, E.W., Marshall, J., 2017. Modulation of the Seasonal Cycle of Antarctic Sea Ice Extent Related to the Southern Annular Mode. *Geophys. Res. Lett.* 44, 9761–9768. <https://doi.org/10.1002/2017GL074319>
- European Commission, 2022. European Eddy-Rich ESMs [WWW Document]. URL <https://doi.org/10.3030/101081383> (accessed 2.2.25).
- Eyring, V., Bony, S., Meehl, G.A., Senior, C.A., Stevens, B., Stouffer, R.J., Taylor, K.E., 2016. Overview of the Coupled Model Intercomparison Project Phase 6 (CMIP6) experimental design and organization. *Geosci. Model Dev.* 9, 1937–1958. <https://doi.org/10.5194/gmd-9-1937-2016>
- Fogt, R.L., Marshall, G.J., 2020. The Southern Annular Mode: Variability, trends, and climate impacts across the Southern Hemisphere. *WIREs Clim. Change* 11, e652. <https://doi.org/10.1002/wcc.652>
- Frenger, I., Gruber, N., Knutti, R., Münnich, M., 2013. Imprint of Southern Ocean eddies on winds, clouds and rainfall. *Nature Geosci* 6, 608–612. <https://doi.org/10.1038/ngeo1863>
- Gerber, E.P., Baldwin, M.P., Akiyoshi, H., Austin, J., Bekki, S., Braesicke, P., Butchart, N., Chipperfield, M., Dameris, M., Dhomse, S., Frith, S.M., Garcia, R.R., Garny, H., Gettelman, A., Hardiman, S.C., Karpechko, A., Marchand, M., Morgenstern, O., Nielsen, J.E., Pawson, S., Peter, T., Plummer, D.A., Pyle, J.A., Rozanov, E., Scinocca, J.F., Shepherd, T.G., Smale, D., 2010. Stratosphere-troposphere coupling and annular mode variability in chemistry-climate models. *J. Geophys. Res. Atmospheres* 115, 2009JD013770. <https://doi.org/10.1029/2009JD013770>
- Gerber, E.P., Voronin, S., Polvani, L.M., 2008. Testing the Annular Mode Autocorrelation Time Scale in Simple Atmospheric General Circulation Models. *Mon. Weather Rev.* 136, 1523–1536. <https://doi.org/10.1175/2007MWR2211.1>

- Gillett, N.P., Kell, T.D., Jones, P.D., 2006. Regional climate impacts of the Southern Annular Mode. *Geophys. Res. Lett.* 33, 2006GL027721. <https://doi.org/10.1029/2006GL027721>
- 805 Haarsma, R.J., Roberts, M.J., Vidale, P.L., Senior, C.A., Bellucci, A., Bao, Q., Chang, P., Corti, S., Fučkar, N.S., Guemas, V., Von Hardenberg, J., Hazeleger, W., Kodama, C., Koenigk, T., Leung, L.R., Lu, J., Luo, J.-J., Mao, J., Mizielinski, M.S., Mizuta, R., Nobre, P., Satoh, M., Scoccimarro, E., Semmler, T., Small, J., Von Storch, J.-S., 2016. High Resolution Model Intercomparison Project (HighResMIP v1.0) for CMIP6. *Geosci. Model Dev.* 9, 4185–4208. <https://doi.org/10.5194/gmd-9-4185-2016>
- 810 <https://doi.org/10.5194/gmd-9-4185-2016>
Hassanzadeh, P., and Kuang, Z., 2019. Quantifying the Annular Mode Dynamics in an Idealized Atmosphere. *J. Atmos. Sci.*, 76, 1107–1124, <https://doi.org/10.1175/JAS-D-18-0268.1>.
- Hersbach, H., Bell, B., Berrisford, P., Hirahara, S., Horányi, A., Muñoz-Sabater, J., Nicolas, J., Peubey, C., Radu, R., Schepers, D., Simmons, A., Soci, C., Abdalla, S., Abellan, X., Balsamo, G., 815 Bechtold, P., Biavati, G., Bidlot, J., Bonavita, M., De Chiara, G., Dahlgren, P., Dee, D., Diamantakis, M., Dragani, R., Flemming, J., Forbes, R., Fuentes, M., Geer, A., Haimberger, L., Healy, S., Hogan, R.J., Hólm, E., Janisková, M., Keeley, S., Laloyaux, P., Lopez, P., Lupu, C., Radnoti, G., De Rosnay, P., Rozum, I., Vamborg, F., Villaume, S., Thépaut, J., 2020. The ERA5 global reanalysis. *Q. J. R. Meteorol. Soc.* 146, 1999–2049. <https://doi.org/10.1002/qj.3803>
- 820 Hewitt, H.T., Roberts, M., Mathiot, P., Biastoch, A., Blockley, E., Chassignet, E.P., Fox-Kemper, B., Hyder, P., Marshall, D.P., Popova, E., Treguier, A.-M., Zanna, L., Yool, A., Yu, Y., Beadling, R., Bell, M., Kuhlbrodt, T., Arsouze, T., Bellucci, A., Castruccio, F., Gan, B., Putrasahan, D., Roberts, C.D., Van Roekel, L., Zhang, Q., 2020. Resolving and Parameterising the Ocean Mesoscale in Earth System Models. *Curr. Clim. Change Rep.* 6, 137–152. <https://doi.org/10.1007/s40641-020-00164-w>
- 825 Ho, M., Kiem, A.S., Verdon-Kidd, D.C., 2012. The Southern Annular Mode: a comparison of indices. *Hydrol. Earth Syst. Sci.* 16, 967–982. <https://doi.org/10.5194/hess-16-967-2012>
- Karoly, D.J., 1989. Southern Hemisphere Circulation Features Associated with El Niño-Southern Oscillation Events. *J. Clim.* 2, 1239–1252. [https://doi.org/10.1175/1520-0442\(1989\)002<1239:SHCF&W>2.0.CO;2](https://doi.org/10.1175/1520-0442(1989)002<1239:SHCF&W>2.0.CO;2)
- 830 Kidston, J., Gerber, E.P., 2010. Intermodel variability of the poleward shift of the austral jet stream in the CMIP3 integrations linked to biases in 20th century climatology. *Geophys. Res. Lett.* 37, 2010GL042873. <https://doi.org/10.1029/2010GL042873>
- Lefebvre, W., Goosse, H., 2005. Influence of the Southern Annular Mode on the sea ice-ocean system: the role of the thermal and mechanical forcing. *Ocean Sci.* 1, 145–157. <https://doi.org/10.5194/os-1-145-2005>
- 835 Lenton, A., Matear, R.J., 2007. Role of the Southern Annular Mode (SAM) in Southern Ocean CO₂ uptake. *Glob. Biogeochem. Cycles* 21, 2006GB002714. <https://doi.org/10.1029/2006GB002714>
- Lim, E.-P., Hendon, H.H., Rashid, H., 2013. Seasonal Predictability of the Southern Annular Mode due to Its Association with ENSO. *J. Clim.* 26, 8037–8054. <https://doi.org/10.1175/JCLI-D-13-00006.1>
- 840 Loose, N., Abernathey, R., Grooms, I., Busecke, J., Guillaumin, A., Yankovsky, E., Marques, G., Steinberg, J., Ross, A., Khatri, H., Bachman, S., Zanna, L., Martin, P., 2022. GCM-Filters: A

Python Package for Diffusion-based Spatial Filtering of Gridded Data. *J. Open Source Softw.* 7, 3947. <https://doi.org/10.21105/joss.03947>

Lorenz, D.J., Hartmann, D.L., 2001. Eddy–Zonal Flow Feedback in the Southern Hemisphere. *J. Atmospheric Sci.* 58, 3312–3327. [https://doi.org/10.1175/1520-0469\(2001\)058<3312:EZFFIT>2.0.CO;2](https://doi.org/10.1175/1520-0469(2001)058<3312:EZFFIT>2.0.CO;2)

Lovenduski, N.S., Gruber, N., 2005. Impact of the Southern Annular Mode on Southern Ocean circulation and biology. *Geophys. Res. Lett.* 32, 2005GL022727. <https://doi.org/10.1029/2005GL022727>

Lubis, S. W., and P. Hassanzadeh, 2021. An Eddy–Zonal Flow Feedback Model for Propagating Annular Modes. *J. Atmos. Sci.*, 78, 249–267, <https://doi.org/10.1175/JAS-D-20-0214.1>.

Lubis, S. W., & Hassanzadeh, P., 2023. The Intrinsic 150-Day Periodicity of the Southern Hemisphere Extratropical Large-Scale Atmospheric Circulation. *AGU Advances*, 4(3), e2022AV000833. <https://doi.org/10.1029/2022AV000833>

Ma, D., Hassanzadeh, P., & Kuang, Z., 2017. Quantifying the Eddy–Jet Feedback Strength of the Annular Mode in an Idealized GCM and Reanalysis Data. *Journal of the Atmospheric Sciences*, 74(2), 393–407. <https://doi.org/10.1175/JAS-D-16-0157.1>

Marshall, G. J., 2003: Trends in the Southern Annular Mode from Observations and Reanalyses. *J. Climate*, 16, 4134–4143, [https://doi.org/10.1175/1520-0442\(2003\)016<4134:TITSAM>2.0.CO;2](https://doi.org/10.1175/1520-0442(2003)016<4134:TITSAM>2.0.CO;2).

Marshall, G. J., Fogt, R. L., Turner, J., & Clem, K. R., 2022. Can current reanalyses accurately portray changes in Southern Annular Mode structure prior to 1979?. *Climate Dynamics*, 59(11), 3717–3740, <https://doi.org/10.1007/s00382-022-06292-3>.

Menzel, M. E., Waugh, D., & Grise, K., 2019. Disconnect Between Hadley Cell and Subtropical Jet Variability and Response to Increased CO₂. *Geophysical Research Letters*, 46(12), 7045–7053. <https://doi.org/10.1029/2019GL083345>

Purich, A., Boschat, G., Liguori, G., 2021. Assessing the impact of suppressing Southern Ocean SST variability in a coupled climate model. *Sci. Rep.* 11, 22069. <https://doi.org/10.1038/s41598-021-01306-2>

Roberts, C., Aengenheyster, M., Roberts, M., 2024a. Description of EERIE atmosphere-only reference simulations. <https://doi.org/10.5281/ZENODO.14334612>

Roberts, C., Aengenheyster, M., Roberts, M., 2024b. Description of EERIE idealised atmosphere-only simulations. <https://doi.org/10.5281/ZENODO.14514510>

Roberts, M.J., Jung, T., Ortega, P., Ghosh, R., Wachsmann, F., von Storch, J.-S., Kroeger, J., Aengenheyster, M., Roberts, C., 2024a. Phase 1 simulations, including HighResMIP and CMIP6-PI-control simulations. <https://doi.org/10.5281/ZENODO.14288726>

Roberts, M.J., Reed, K.A., Bao, Q., Barsugli, J.J., Camargo, S.J., Caron, L.-P., Chang, P., Chen, C.-T., Christensen, H.M., Danabasoglu, G., Frenger, I., Fučkar, N.S., Hasson, S.U., Hewitt, H.T., Huang, H., Kim, D., Kodama, C., Lai, M., Leung, L.-Y.R., Mizuta, R., Nobre, P., Ortega, P., Paquin, D., Roberts, C.D., Scoccimarro, E., Seddon, J., Treguier, A.M., Tu, C.-Y., Ullrich, P.A., Vidale, P.L., Wehner, M.F., Zarzycki, C.M., Zhang, B., Zhang, W., Zhao, M., 2024b. High Resolution Model Intercomparison Project phase 2 (HighResMIP2) towards CMIP7. <https://doi.org/10.5194/egusphere-2024-2582>

- Robinson, W. A., 2000. A Baroclinic Mechanism for the Eddy Feedback on the Zonal Index. *J. Atmos. Sci.*, 57, 415–422. [https://doi.org/10.1175/1520-0469\(2000\)057<0415:ABMFTE>2.0.CO;2](https://doi.org/10.1175/1520-0469(2000)057<0415:ABMFTE>2.0.CO;2).
- Saggioro, E., & Shepherd, T. G., 2019. Quantifying the Timescale and Strength of Southern Hemisphere Intraseasonal Stratosphere-troposphere Coupling. *Geophysical Research Letters*, 46(22), 13479–13487. <https://doi.org/10.1029/2019GL084763>
- Sen Gupta, A., England, M.H., 2006. Coupled Ocean–Atmosphere–Ice Response to Variations in the Southern Annular Mode. *J. Clim.* 19, 4457–4486. <https://doi.org/10.1175/JCLI3843.1>
- Sheshadri, A., & Plumb, R. A., 2017. Propagating Annular Modes: Empirical Orthogonal Functions, Principal Oscillation Patterns, and Time Scales. *Journal of the Atmospheric Sciences*, 74(5), 1345–1361. <https://doi.org/10.1175/JAS-D-16-0291.1>
- Simpson, I. R., Hitchcock, P., Shepherd, T. G., & Scinocca, J. F., 2011. Stratospheric variability and tropospheric annular-mode timescales. *Geophysical Research Letters*, 38(20). <https://doi.org/10.1029/2011GL049304>
- Simpson, I.R., Hitchcock, P., Shepherd, T.G., Scinocca, J.F., 2013a. Southern Annular Mode Dynamics in Observations and Models. Part I: The Influence of Climatological Zonal Wind Biases in a Comprehensive GCM. *J. Clim.* 26, 3953–3967. <https://doi.org/10.1175/JCLI-D-12-00348.1>
- Simpson, I.R., Polvani, L.M., 2016. Revisiting the relationship between jet position, forced response, and annular mode variability in the southern midlatitudes. *Geophys. Res. Lett.* 43, 2896–2903. <https://doi.org/10.1002/2016GL067989>
- Simpson, I.R., Shepherd, T.G., Hitchcock, P., Scinocca, J.F., 2013b. Southern Annular Mode Dynamics in Observations and Models. Part II: Eddy Feedbacks. *J. Clim.* 26, 5220–5241. <https://doi.org/10.1175/JCLI-D-12-00495.1>
- Smith, S., Lu, J., & Staten, P. W., 2024. Diabatic Eddy Forcing Increases Persistence and Opposes Propagation of the Southern Annular Mode in MERRA-2. *J. Atmos. Sci.*, 81, 743–764. <https://doi.org/10.1175/JAS-D-23-0019.1>
- Son, S. -W., Gerber, E.P., Perlwitz, J., Polvani, L.M., Gillett, N.P., Seo, K. -H., Eyring, V., Shepherd, T.G., Waugh, D., Akiyoshi, H., Austin, J., Baumgaertner, A., Bekki, S., Braesicke, P., Brühl, C., Butchart, N., Chipperfield, M.P., Cugnet, D., Dameris, M., Dhomse, S., Frith, S., Garny, H., Garcia, R., Hardiman, S.C., Jöckel, P., Lamarque, J.F., Mancini, E., Marchand, M., Michou, M., Nakamura, T., Morgenstern, O., Pitari, G., Plummer, D.A., Pyle, J., Rozanov, E., Scinocca, J.F., Shibata, K., Smale, D., Teyssède, H., Tian, W., Yamashita, Y., 2010. Impact of stratospheric ozone on Southern Hemisphere circulation change: A multimodel assessment. *J. Geophys. Res. Atmospheres* 115, 2010JD014271. <https://doi.org/10.1029/2010JD014271>
- Storkey, D., Mathiot, P., Bell, M. J., Copsey, D., Guiavarc'h, C., Hewitt, H. T., Ridley, J., and Roberts, M. J., 2024. Resolution dependence of interlinked Southern Ocean biases in global coupled HadGEM3 models, *EGUsphere* [preprint], <https://doi.org/10.5194/egusphere-2024-1414>.
- Thompson, D.W.J., Solomon, S., Kushner, P.J., England, M.H., Grise, K.M., Karoly, D.J., 2011. Signatures of the Antarctic ozone hole in Southern Hemisphere surface climate change. *Nat. Geosci.* 4, 741–749. <https://doi.org/10.1038/ngeo1296>

- Vishny, D. N., Wall, C. J., & Lutsko, N. J., 2024. Impact of Atmospheric Cloud Radiative Effects on Annular Mode Persistence in Idealized Simulations. *Geophysical Research Letters*, 51(15), e2024GL109420. <https://doi.org/10.1029/2024GL109420>
- 930 Wachsmann, F., Matthias Aengenheyster, Wickramage, C., Seddon, J., Nikolay Koldunov, Ziemen, F., 2024. Data access to EERIE ESM output via intake catalogues: Phase1 plus Hackathon simulations. <https://doi.org/10.5281/ZENODO.14243677>
- 935 Xia, X., & Chang, E. K. M., 2014. Diabatic Damping of Zonal Index Variations. *Journal of the Atmospheric Sciences*, 71(8), 3090–3105. <https://doi.org/10.1175/JAS-D-13-0292.1>
- Zhang, Q., Liu, B., Li, S., Zhou, T., 2023. Understanding Models’ Global Sea Surface Temperature Bias in Mean State: From CMIP5 to CMIP6. *Geophys. Res. Lett.* 50, e2022GL100888. <https://doi.org/10.1029/2022GL100888>
- 940 Zhang, X., He, B., Liu, Y., Bao, Q., Zheng, F., Li, J., Hu, W., & Wu, G., 2022. Evaluation of the seasonality and spatial aspects of the Southern Annular Mode in CMIP6 models. *International Journal of Climatology*, 42(7), 3820-3837. <https://doi.org/10.1002/joc.7447>
- 945 Zurita-Gotor, P., J. Blanco-Fuentes, and E. P. Gerber, 2014. The Impact of Baroclinic Eddy Feedback on the Persistence of Jet Variability in the Two-Layer Model. *J. Atmos. Sci.*, 71, 410–429, <https://doi.org/10.1175/JAS-D-13-0102.1>