Review of "Southern Annular Mode Persistence and Westerly Jet: A Reassessment Using High-Resolution Global Models" by Ting-Chen Chen et al.

Overall Assessment

This study explores the atmospheric and oceanic influences on modelled SAM persistence and its relationship with the latitude of the mid-latitude jet. The authors note the longstanding issue that CMIP models (including the latest suite: CMIP6) overestimate SAM persistence (quantified using the decorrelation timescale), particularly in early austral summer, which appears to be much improved when using high-resolution, eddy-resolving simulations from the EERIE project. This appears to be in part due to more realistic simulation of the jet position/distribution (CMIP models have tended to be too equatorward biased) but the importance of accurate SST representation is also clear. In fact, the authors show that AMIP model experiments of the EERIE simulations perform better than coupled experiments in terms of more realistically representing the jet position and SAM decorrelation timescale.

Enhanced resolution of the EERIE simulations likely plays a role in the improvement relative to CMIP6 models but also improved model physics, particularly concerning ocean mesoscale eddies which appear to slightly enhance the SAM decorrelation timescale in early summer (at least for simulations run at 28 km resolution) according to sensitivity simulations performed. However, cancellation effects (e.g., atmospheric eddy feedback strength versus surface friction) make it difficult to ascertain which aspects help improve modelling of the SAM persistence. For instance, the decorrelation timescale is more realistic still in 9 versus 28 km AMIP simulations, yet the role of ocean mesoscale eddies in enhancing SAM persistence is not evident at this finer resolution.

I found the study to be very well written, organised and logically structured. The Figures are clearly presented and straightforward to understand and the step-by-step computation of the different diagnostics examined will I think be much appreciated by many readers. The conclusions drawn are supported by the results shown, so I can recommend this be accepted for publication in *Weather and Climate Dynamics*. I include just a few comments for the authors to consider prior to acceptance.

General Comments

I only had one main consideration for the authors which I found lacking in the paper. That is information of which CMIP6 models were considered (if supplementary information were to be provided, a table for this would be warranted). That is not to say that knowledge of which models lie where in the distributions shown (Figures 2 and 3) are important in understanding this work. But for others reading, it might be useful for them to know and the best way I feel to include this would be to show similar Figures with individual CMIP6 models indicated in supplementary information.

Nevertheless, knowledge of which CMIP6 models lie where could help others or even the authors to comment upon whether commonalities such as shared model components or known biases more widely within the climate system might influence the results. So, I would encourage the authors to think about providing this information.

Specific Comments

L15: "a critical driver" → "a leading mode"? I wouldn't consider the SAM to a driver itself, but more of a reflection of driving influences. Expressing it as a leading mode would be more scientifically accurate and consistent with earlier literature (e.g., Marshall, 2003; Marshall et al., 2022).

L340-342: How reliable is the SAM derived from ERA5 pre-satellite era when comparing with the EERIE coupled simulations? Presumably the SAM is more reliably reconstructed after 1979 from ERA5 but maybe difficult to quantify how much of an improvement there would be. It may be worthy of further comment or caveating, however?

L343-344: Suggesting that a relative minority of models are considerably worse in representing SAM decorrelation timescale than the rest of the pack. Did the authors investigate why this might be or were they at least able to note some commonalities in the most unrealistic models that might point to the cause(s)? For instance, could there be an association between too equatorward jet position and shared model components? Or factors that may give plausibly give rise to the issue of realistic eddy feedback strength? It may be beyond the scope of the paper to delve into this, but others reading might be encouraged to look into this.

L477: "…a key driver…" → Again I think '…a leading mode…" would be more technically correct.

Technical Corrections

L52-53: "spring (MAM) and summer (JJA)". → "autumn (MAM) and winter (JJA)".

L153: "observation" → "observations"

Table 1: Some font size inconsistencies noted.

L245: Tabulation before "Finally…"

Additional References

Marshall, G. J. (2003). Trends in the Southern Annular Mode from observations and reanalyses. *Journal of climate*, *16*(24), 4134-4143, 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.