Supplement of Geosci. Model Dev., 18, 5937–5969, 2025
https://doi.org/10.5194/gmd-18-5937-2025-supplement

Geoscientific Model Development
Open Access
EGU

Supplement of

The Atmospheric Potential Oxygen forward Model Intercomparison Project (APO-MIP1): evaluating simulated atmospheric transport of air-sea gas exchange tracers and APO flux products

Yuming Jin et al.

Correspondence to: Yuming Jin (yumingjin@ucar.edu)

Supplement

S1. APO spatial and temporal bias correction

Observed seasonal average APO cross-$M_{\theta e}$ gradients derived from airborne data are biased due to the limited spatial and temporal coverage of airborne sampling. To assess the spatial bias of $M_{\theta e}$-averaged APO used to calculate the cross-$M_{\theta e}$ gradients for each airborne campaign, we compare the true model APO with the average calculated by subsampling the model 3-D atmospheric field along flight tracks. We selected three ATMs (MIROC-ACTM and two versions of NICAM) and surface flux products (Jena and CESM), which are identified as more realistic models or products in our analysis. We correct for the mean model difference and include an uncertainty contribution on these corrections based on the 1σ spread in the six model-flux estimates of the correction.

The true model average is computed by averaging over all model tropospheric grid cells (troposphere defined as potential vorticity unit, PVU, smaller than 2) on the flight dates of each airborne campaign within the corresponding $M_{\theta e}$ box. The subsampled average is computed by subsampling model data along the flight tracks on the same flight dates and by trapezoidal integration of subsampled model data as a function of $M_{\theta e}$. Both calculations use $M_{\theta e}$ calculated from MERRA-2 and interpolated to the model grids. Prior to the trapezoidal integration, the subsampled data are extrapolated to $M_{\theta e}$ = 0 surface using the average of the 100 observations with the lowest $M_{\theta e}$ values, except for HIPPO4, in which we only extrapolate to $M_{\theta e}$ = 15 (10^{16} kg).

We also correct for temporal bias when we calculate seasonal average cross-$M_{\theta e}$ APO gradients. We use the same three ATMs and two flux products to compare the true seasonal average APO from 3D fields on all days and APO averaged using 3D fields only the airborne campaign mean dates that fall within the seasonal window. These corrections are described in more detail in Jin et al. (2024).

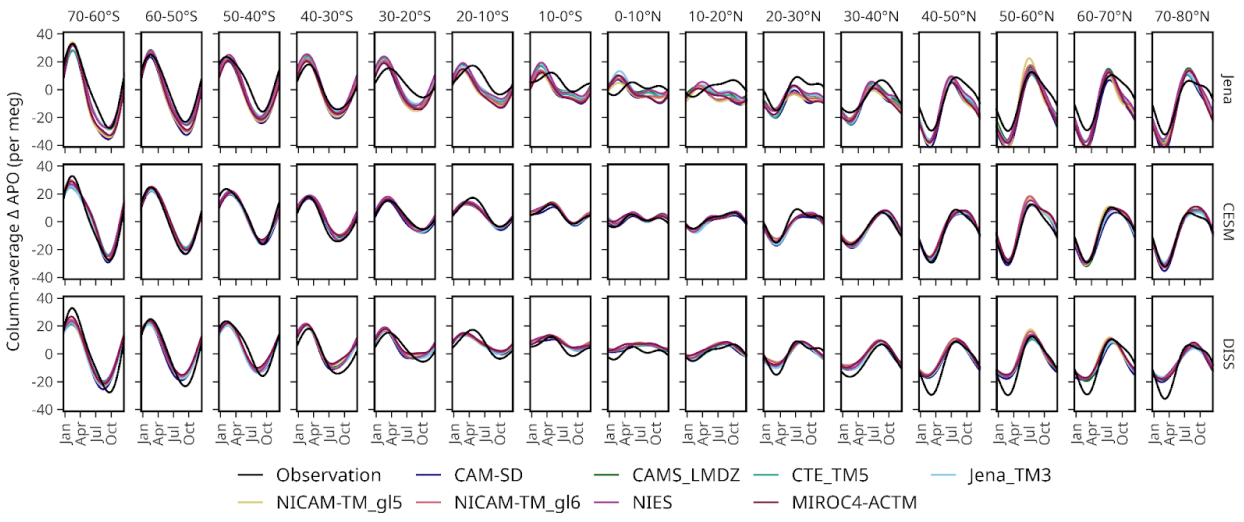


Figure S1: Comparison of simulated and observed column-average APO seasonal cycles, organized from southern high-latitudes (left) to northern high-latitudes (right). In each panel, the black line represents observations, while colored lines show simulations from different transport models using three different flux products (Jena, CESM, and DISS, shown in rows). The y-axes show APO anomalies in per meg units, and the x-axes show months from January to December. Observations and model simulations are grouped into 10° latitude bands, and are first detrended using a multiple-station weighted average trend. Seasonal cycles are calculated as 2-harmonic fits to each observation or model simulation.

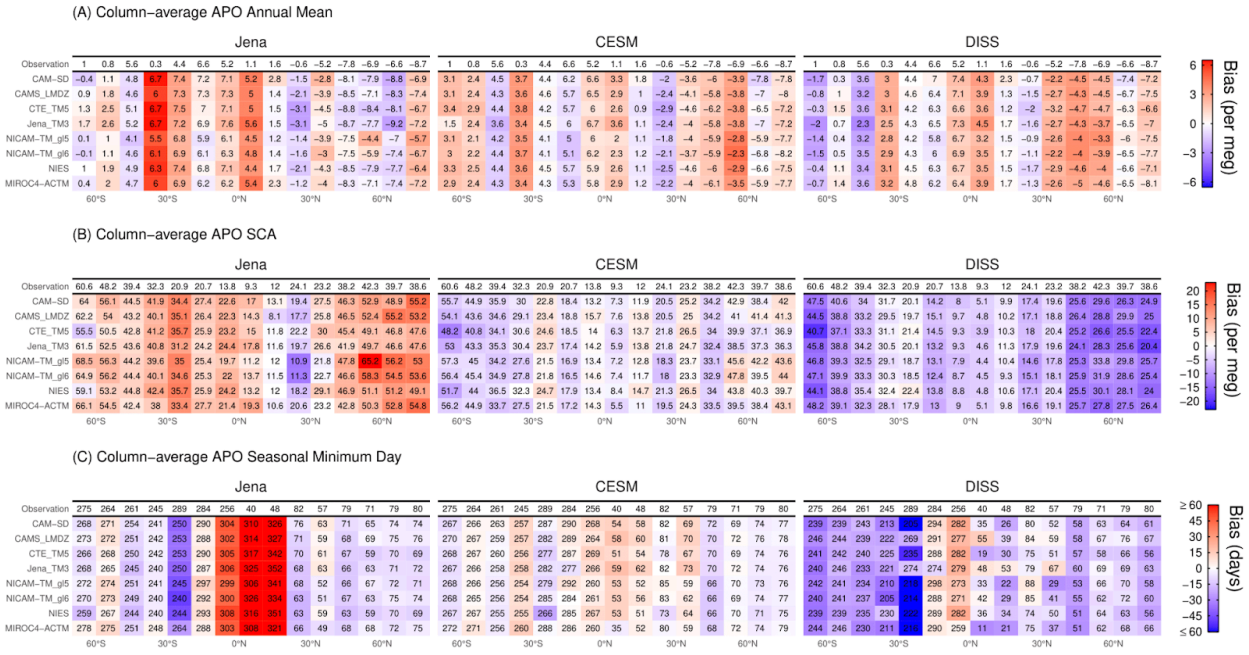


Figure S2: Similar to Fig. 4, but showing evaluation of column average APO (A) annual mean relative to a multi-station global mean, (B) seasonal cycle amplitude, and (C) seasonal minimum day across latitude bins using different flux-transport model combinations, derived from Fig. 6. For each panel, results are organized by flux products (Jena, CESM, DISS) in columns and transport models in rows, with observations on the top. The metrics are printed in black, with background colors indicating biases relative to observations. Positive bias is shown in red, and negative bias is shown in blue.

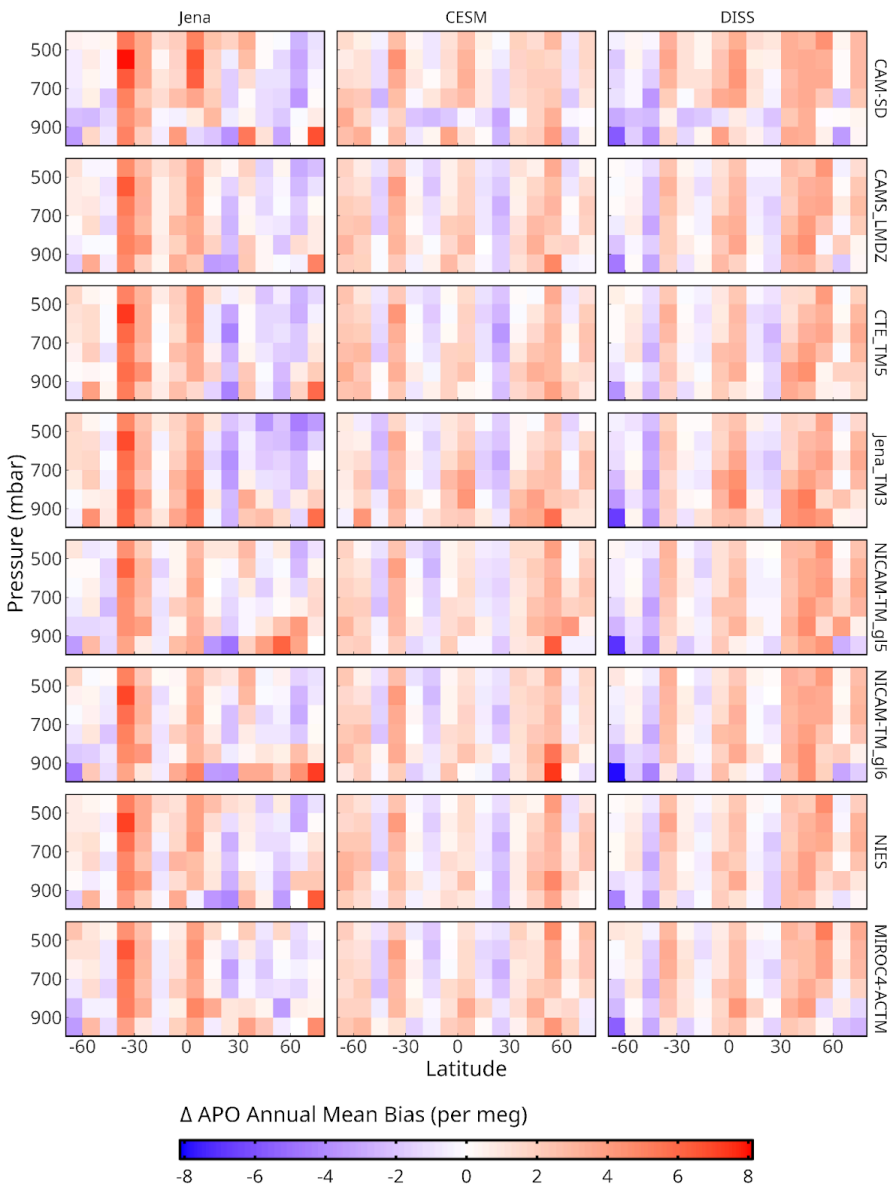
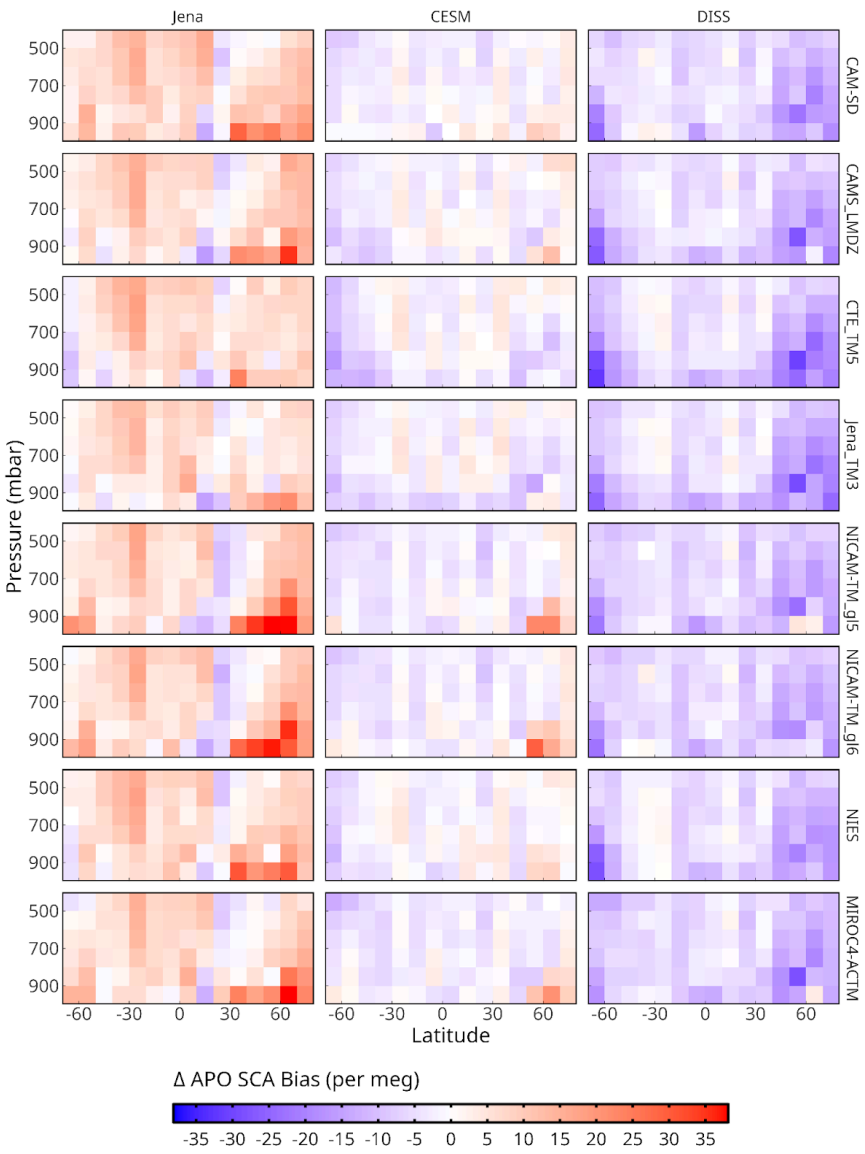
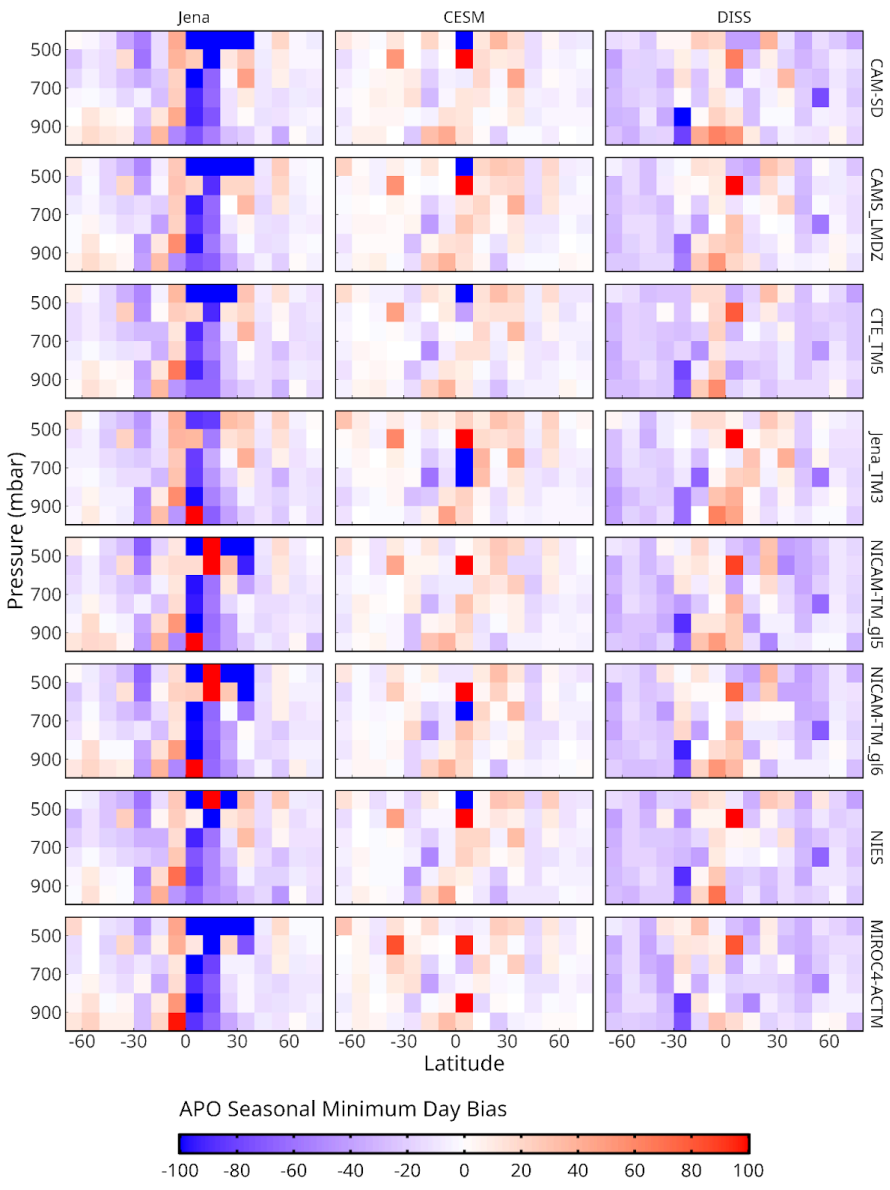


Figure S3: Model biases in APO annual mean compared to airborne observations. The latitude-pressure cross-sections show the difference between simulated and observed annual mean APO (model minus observation, in per meg). Columns represent different flux products (CESM, DISS, and Jena), while rows show different atmospheric transport models. Negative values indicate model underestimation of annual mean APO. Simulated annual mean APO is calculated using the same method as for observations.



51

52 Figure S4: Similar to Fig. S3 but showing model biases in APO SCA compared to airborne
53 observations.



54

55 Figure S5: Similar to Fig. S4 but showing model biases in seasonal minimum day.

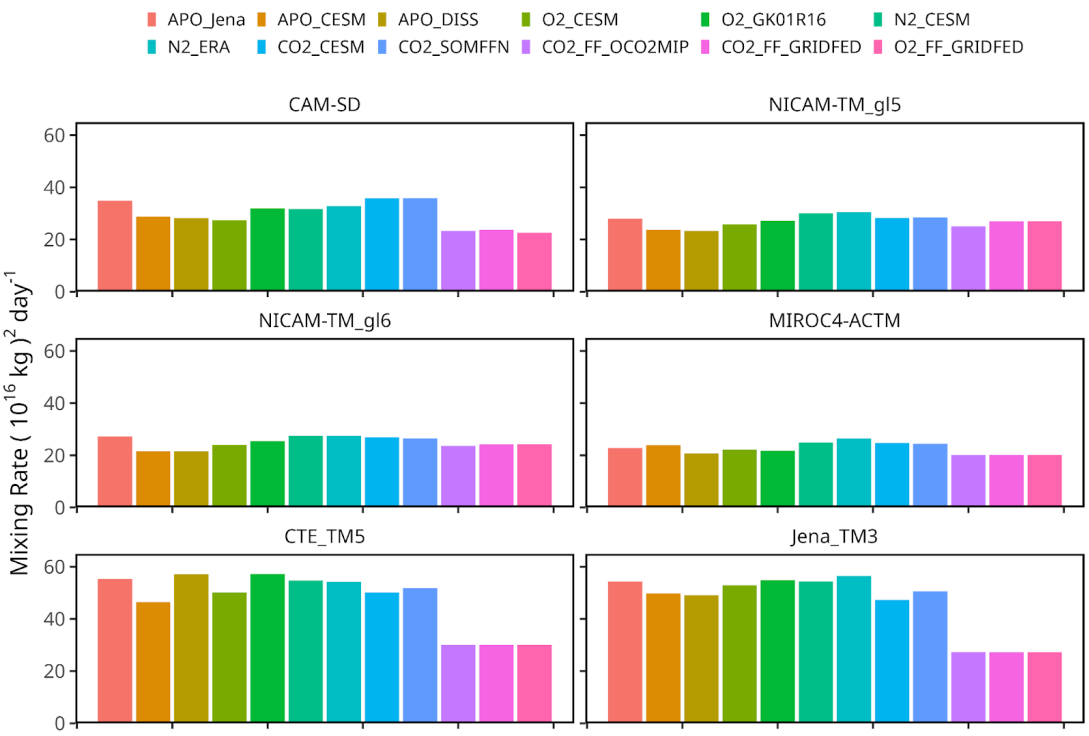


Figure S6: Comparison of January-February averaged diabatic mixing rates across the 30 (10^{16} kg) $M_{\theta e}$ surface in the Southern Hemisphere, calculated using different tracers across six ATMs. ATMs with generally low diagnosed diabatic mixing rates (i.e., CAM-SD, two versions of NICAM, and MIROC4-ACTM) show relatively consistent mixing rates across different tracers. In contrast, the two ATMs with rapid mixing (CTE_TM5 and Jena_TM3) demonstrate considerable variation in mixing rates across different tracers, with the three fossil fuel tracers showing notably reduced mixing rates.

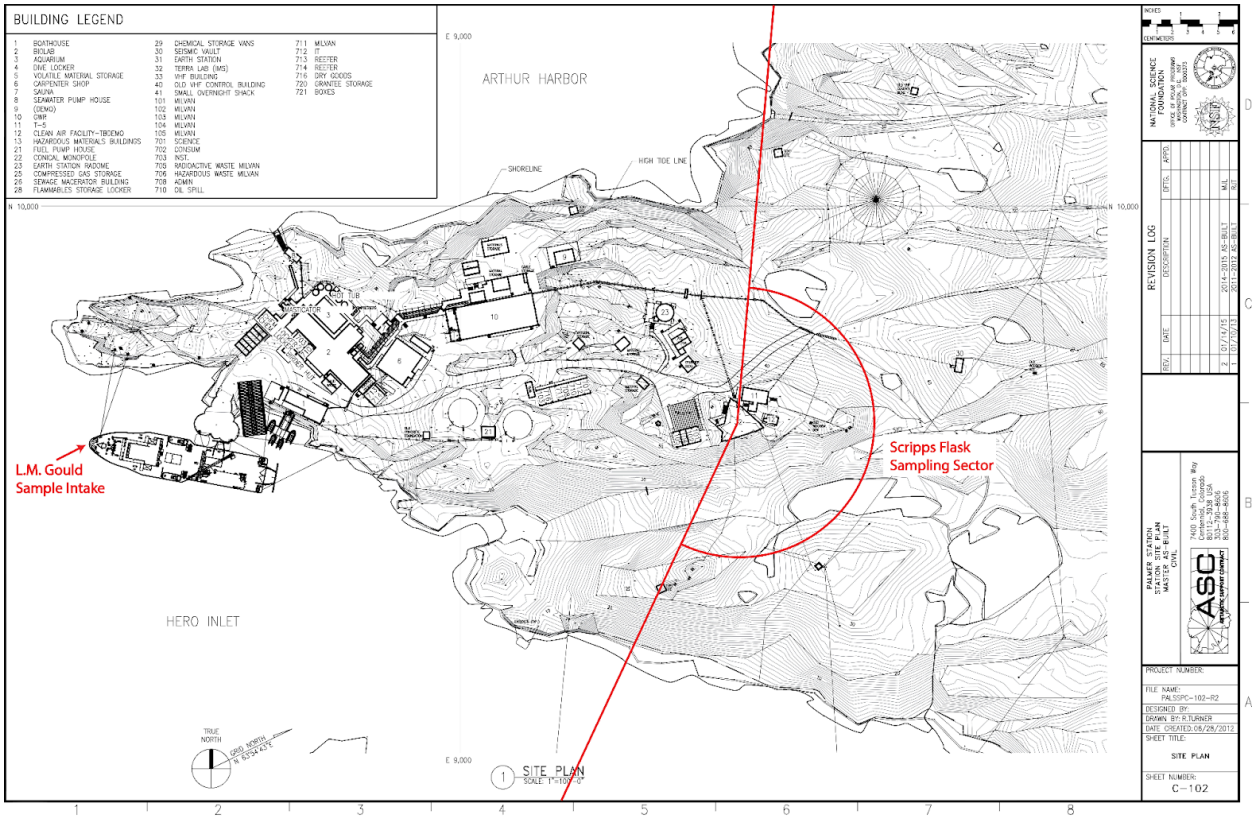


Figure S7: A map that shows PSA with sampling location (Terra Lab) and clean air wind sector for the Scripps O_2 Program flasks, as well as the location of the L.M. Gould bow intake while docked at the pier (adapted from https://www.usap.gov/scienceSupport/documents/PAL%20Station%20Map_Jan%202015.pdf)

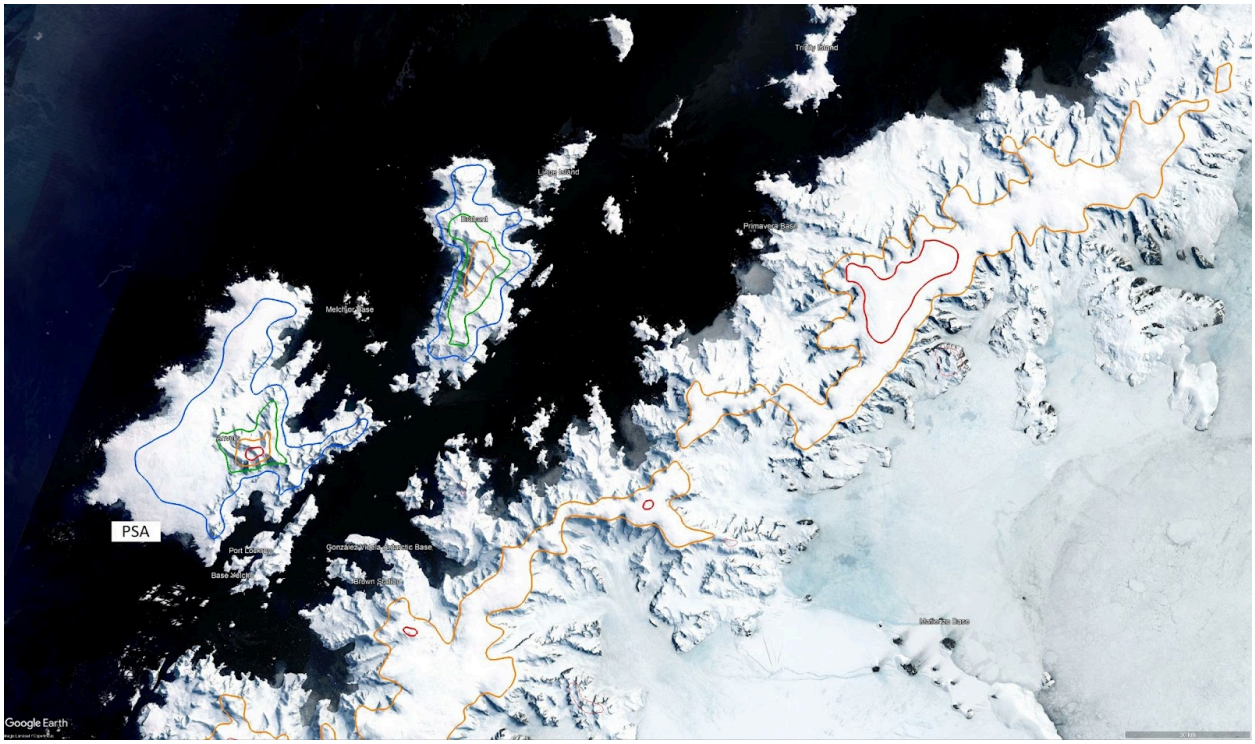


Figure S8: A map that shows the location of PSA within the Palmer Archipelago, including elevation contours (Gerrish et al., 2020) at 500 (blue), 1000 (green), 1500 (orange), and 2000 (red) m.

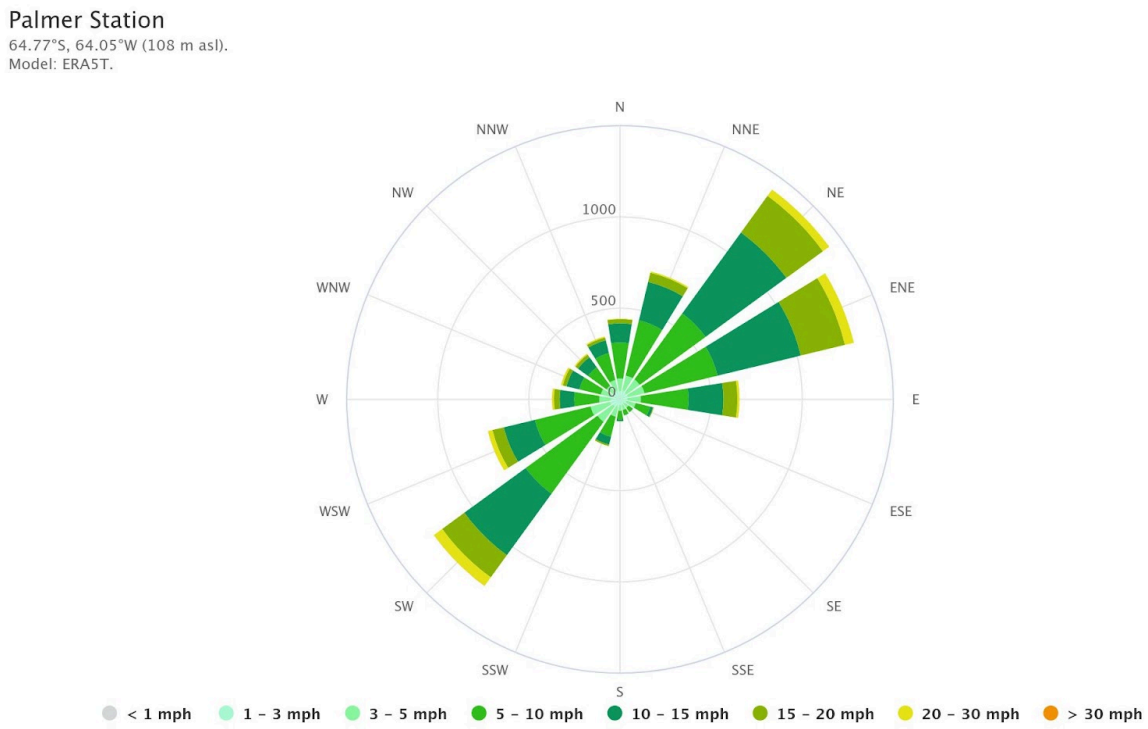


Figure S9: Wind rose for PSA (from MeteoBlue). SIO flasks primarily sample descending air from the northeast, whereas the ship measurements at the pier are filtered to exclude air influenced by the station.

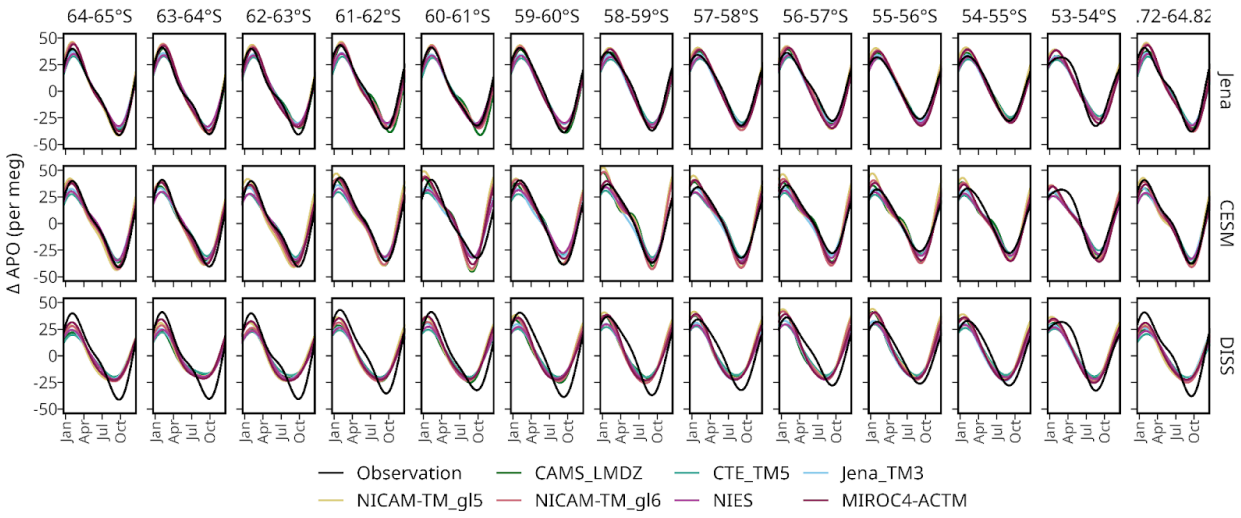


Figure S10: Comparison of simulated and observed APO seasonal cycles across the Drake Passage region (53°S-65°S). In each panel, the black line represents observations, while colored lines show simulations from different transport models using three different flux products (Jena, CESM, and DISS, shown in rows). The y-axes show APO anomalies in per meg units, and the x-axes show months from January to December. We calculate SCA (as in Fig. 11) from these resolved seasonal cycles. Both observational and model data were detrended using corresponding cubic smooth spline fits from SPO and binned into 1° latitude bands. Seasonal cycles are calculated as 2-harmonic fit to each observation or model simulation. However, this difference is not captured in the model simulations. We exclude CAM-SD because the simulated ship data is only available from 2012 to 2015 (i.e., missing 2016 to 2017 data).

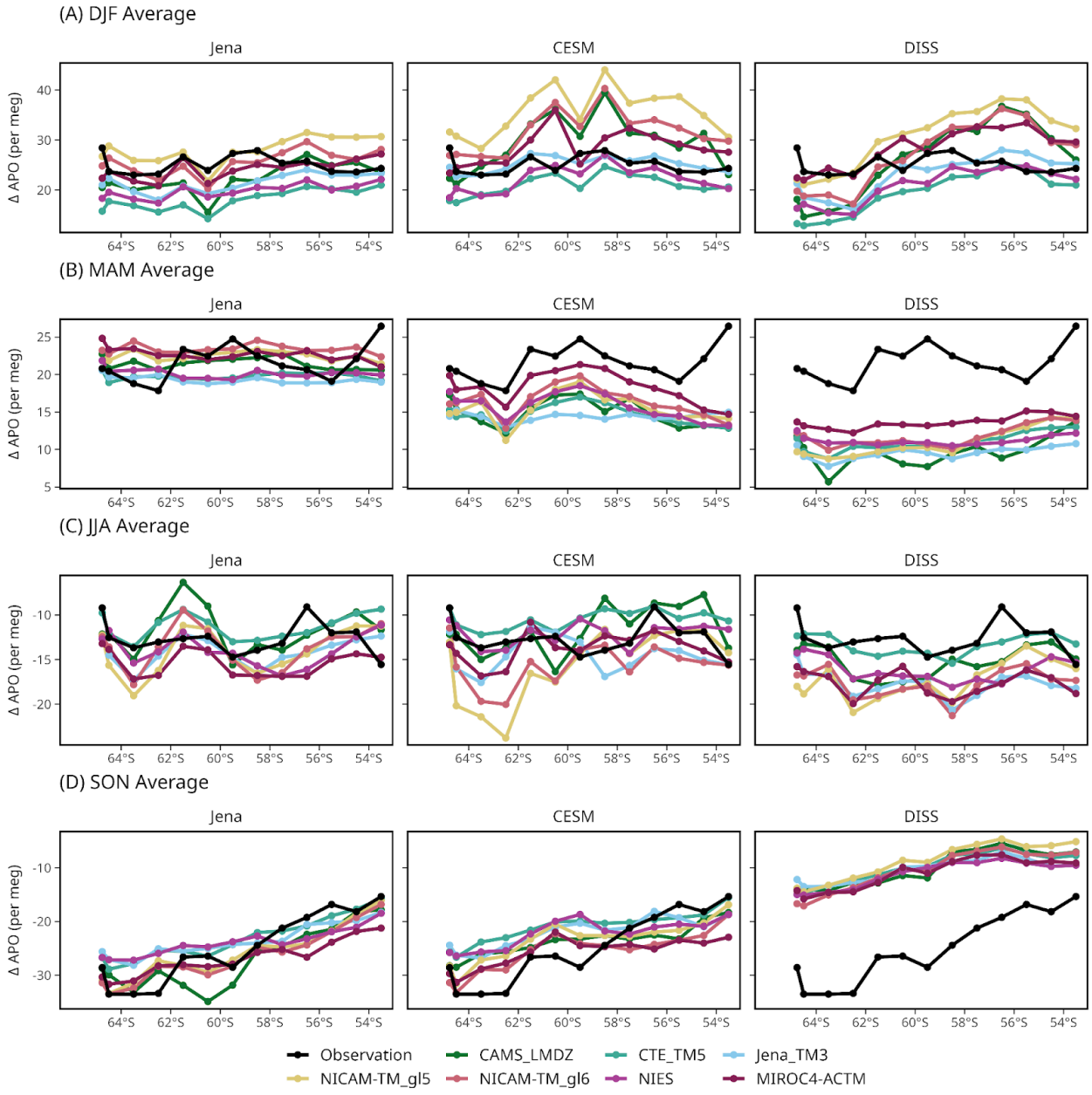


Figure S11: Comparison of simulated and observed APO seasonal anomalies across the Drake Passage region (53°S-65°S) during (A) DJF, (B) MAM, (C) JJA, and (D) SON. The APO seasonal cycles for each latitude band are calculated following the method in Figs. 11 and S10. The seasonal anomaly is calculated by averaging data from the corresponding months. We exclude CAM-SD because the simulated ship data is only available from 2012 to 2015 (i.e., missing 2016 to 2017 data).

Table S1: List of ObsPack files that were required for ATM simulation sampling.

Primary aircraft campaigns samples	co2_hip_aircraft-insitu_59_allvalid.nc
	co2_orc_aircraft-insitu_3_allvalid-merge10.nc
	co2_tom_aircraft-insitu_1_allvalid.nc
Scripps O_2 Program stations	co2_alt_surface-flask_4_representative.nc
	co2_brw_surface-flask_4_representative.nc
	co2_cba_surface-flask_4_representative.nc
	co2_cgo_surface-flask_4_representative.nc
	co2_kum_surface-flask_4_representative.nc
	co2_ljo_surface-flask_4_representative.nc
	co2_mlo_surface-flask_4_representative.nc
	co2_psa_surface-flask_4_representative.nc
	co2_smo_surface-flask_4_representative.nc
	co2_spo_surface-flask_4_representative.nc
ARSV Laurence M. Gould Ship	co2_gould_shipboard-insitu_1_allvalid.nc
AIST/JMA aircraft samples	co2_aoa_aircraft-flask_19_allvalid.nc

Table S2: Optional ATM output submitted. For additional shipboard and station data, two versions of NICAM provide hourly output, and NIES provides 3-hourly output.

Model	Full ObsPack	Additional Shipboard	Additional Stations	3D	Others
CAM-SD				X	ObsPack winds; 3D meteorology (T, P, Q) and boundary layer height
CAMS_LMDZ					
CTE_TM5	X		X	X	Winds, Boundary layer height, meteorology (T, P, Q), and CTE components including land, ocean, and fossil fuel CO_2 from ObsPack, additional stations, and 3D
Jena_TM3	X			X	
MIROC4-ACTM				X	Additional ObsPack Files, ACTM CO_2 inversion
NICAM-TM	X	X	X	X	Meteorology (T, P, Q) from ObsPack samplings, additional stations/shipboard, 2D surface fields
NIES		X	X		

References

Gerrish, L., Fretwell, P., and Cooper, P.: Medium resolution vector contours for Antarctica (7.3) [Data set]. UK Polar Data Centre, Natural Environment Research Council, UK Research & Innovation, 2020.

Jin, Y., Keeling, R. F., Stephens, B. B., Long, M. C., Patra, P. K., Rödenbeck, C., Morgan, E. J., Kort, E. A., and Sweeney, C.: Improved atmospheric constraints on Southern Ocean CO_2 exchange, Proc. Natl. Acad. Sci., 121, e2309333121, https://doi.org/10.1073/pnas.2309333121, 2024.