Response to the reviews of "Current status of ocean observation, ensemble reanalysis and CMIP6 models in describing Antarctic Bottom Water" by S. Chen, Liu, J., and X. Han

Response to comments by Reviewer #2

We would like to thank the reviewer for the helpful comments on the paper.

Major Comments

1. Novelty and physical basis of the classification.

The manuscript presents the classification as a "new framework". However, the idea that AABW can be separated into regional source types (Weddell, Ross, Adélie/Prydz) based on thermohaline properties is already well established in the literature (e.g. Talley, Pardo et al., and many other studies). The present approach appears to formalise this using fixed salinity ranges and geographic masks, but it is not clear that this constitutes a genuinely new framework.

More importantly, the classification is defined using very narrow salinity ranges applied far downstream of formation regions. It is unclear whether such narrow thresholds can robustly distinguish source contributions in the interior ocean. The decision not to include conservative temperature also weakens the constraints. This raises the question of whether the method is identifying dynamically meaningful water masses, or simply segmenting the field based on arbitrary thresholds. As the previous reviewer noted, sensitivity analysis would help here.

Response:

Thanks for the reviewer's constructive comments. We agree that it is important to show our water mass classification is dynamically meaningful. In the revised manuscript, we have conducted a sensitivity analysis to demonstrate the physical robustness of the Antarctic Bottom Water (AABW) subtype classification, and thoroughly revised the text to contextualize our method within existing classical literature.

We agree that the existence of multiple origins for AABW is established in classical oceanography, primarily built upon pioneering observational studies (e.g., Talley, 2013; Pardo et al., 2012). In this revision, we removed the potentially ambiguous term 'new framework' and explicitly cited classical literature. The core of our work is to systematically evaluate whether the ensemble reanalysis and CMIP6 climate models can simulate AABW subtypes that are characterized by different thermohaline properties and spatial distributions.

First, we only used the neutral density ($\gamma_n \geq 28.27$ kg m^{-3}) as the threshold to serve as a physical reference for the AABW. Figure R1 outlines the area covered by the AABW. It is mainly distributed from the eastern Ross Sea to the Bellingshausen Sea, from the central Weddell Sea to the Indian Ocean sector, and north of the East Antarctic continental margin from 75°E to 150°E.

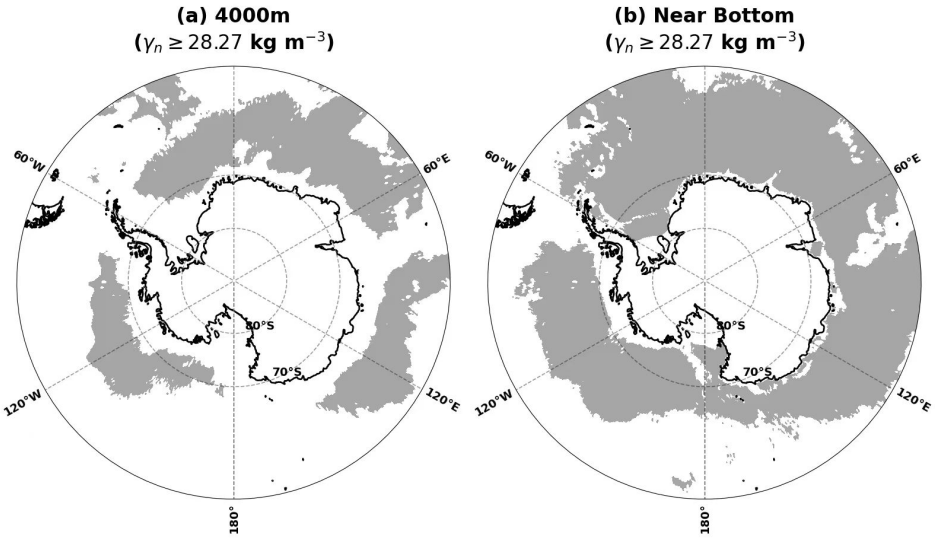

Figure R1. Spatial footprint of Antarctic Bottom Water (AABW) defined solely by the classical neutral density threshold ($\gamma_n \geq 28.27 kgm^{-3}$) across the Southern Ocean. The gray shaded regions illustrate the geographical extent of water masses satisfying this density criterion at (a) the 4000 m depth layer and (b) the near-bottom layer. The diagnostic results are derived from the World Ocean Atlas 2023 (WOA23) climatology.

Second, in cold polar waters, the thermal expansion coefficient of seawater plays a relatively minor role, making salinity the primary driver of density variation. Therefore, we adopted the central salinity values of the respective water masses, including the Weddell-Prydz Bottom Water (WPBW), the Ross Sea Bottom Water (RSBW), and Adélie Land Bottom Water (ALBW), as defined by Pardo et al. (2012). Here, we also conducted a sensitivity analysis of the classification to different salinity ranges relative to these central values. In our previous manuscript, we used a salinity range of $\Delta S = \pm 0.01$ psu relative to each central salinity value, corresponding to twice the uncertainty reported by Pardo et al. (2012). We subsequently tested a series of widened thresholds, including $\Delta S = \pm 0.02$, ± 0.03, and ± 0.04 psu.

Figure R2 shows the spatial distributions of the three AABW subtypes evaluated under these progressively widened salinity thresholds at both the 4000 m and near-bottom layers. At the 4000 m depth layer, the three AABW subtypes within their respective prescribed longitudinal sectors exhibit quite stable and similar spatial distributions when $\Delta S \geq 0.02$ psu. This feature also holds true for the near-bottom layer. Table R1 provides spatial area (10^6 km^2) of each AABW subtype calculated under different salinity thresholds. It demonstrates that the areas of the three AABW subtypes tend to slowly converge to the values based solely on the neutral density criterion as ΔS reaches or exceeds ± 0.02 psu (Table R1).

Given that a salinity threshold greater than ± 0.02 psu may overly incorporate diluted peripheral waters and the overlying Circumpolar Deep Water (CDW), we selected ± 0.02 psu as the revised threshold. This provides a reasonable balance, facilitating the identification of the main bottom water bodies while mitigating the potential over-inclusion of overlying CDW. Ultimately, this framework robustly delineates the bottom water distributions with distinct source-region characteristics and main advection pathways.

Third, in the first step of identifying water masses, we used the neutral density threshold ($\gamma_n \geq 28.27$ kg m^{-3}), in which the temperature effect is also included as a thermodynamic factor for calculating density. We removed the phrase 'excluding temperature' in the revised manuscript. In our response to the reviewer's major comment #3, we conducted a sensitivity analysis of the classification to different temperature ranges relative to these central values (see detailed discussion there).

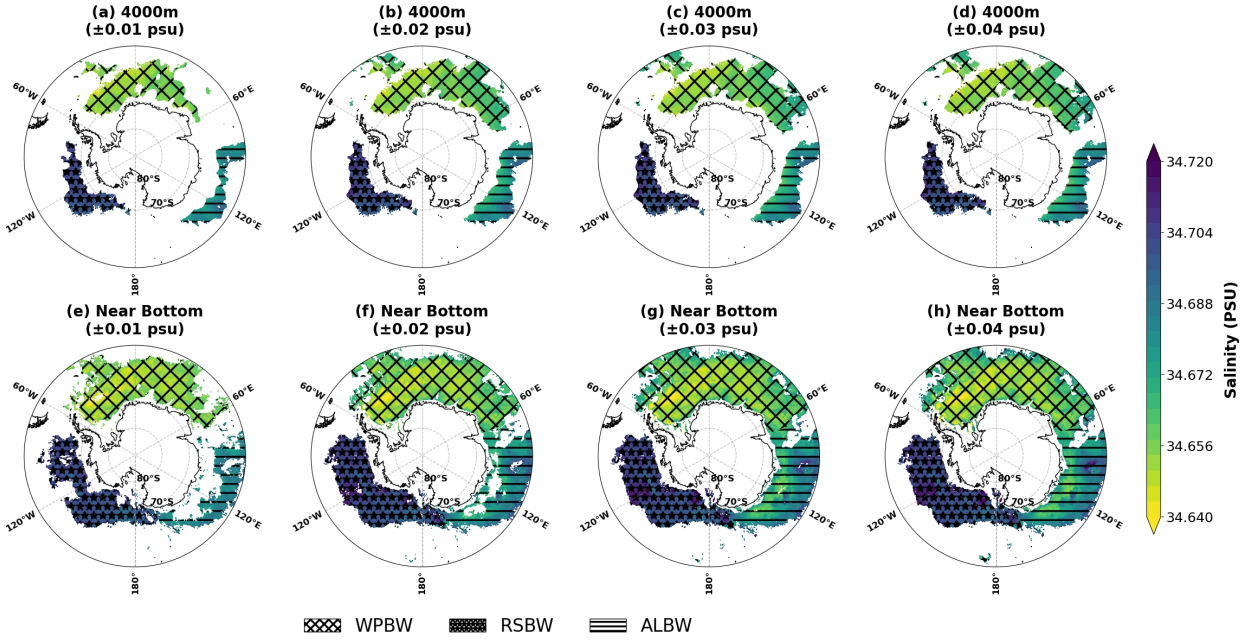

Figure R2. Spatial distributions of the identified Antarctic Bottom Water (AABW) subtypes under different salinity thresholds (ΔS). The top row (a–d) and bottom row (e–h) represent the 4000 m depth layer and the near-bottom layer, respectively. Columns from left to right show classifications based on applied ΔS values of ±0.01, ±0.02, ±0.03, and ±0.04 around the central salinity values (WPBW: 34.650; RSBW: 34.695; ALBW: 34.683). Note that they are also constrained by the neutral density threshold ($\gamma_n \geq 28.27$ kg m^{-3}). The respective AABW subtypes are distinguished by distinct hatching patterns, as detailed in the figure legend. The background color shading represents the practical salinity field.

Table R1. Calculated areas (10^6 km^2) of the three AABW subtypes under different salinity thresholds (ΔS) at the 4000 m layer and the near-bottom layer. These thresholds are defined around the central salinity values (WPBW: 34.650; RSBW: 34.695; ALBW: 34.683).

Layer	Subtype	Neutral Density Only	±0.010	±0.020	±0.030	±0.040
4000 m	WPBW	8.36	4.96	7.56	8.30	8.36

	RSBW	2.92	2.53	2.92	2.92	2.92
	ALBW	3.25	2.03	3.10	3.25	3.25
Near-Bottom	WPBW	13.96	9.05	12.19	13.28	13.60
	RSBW	7.71	5.73	7.27	7.45	7.50
	ALBW	6.88	3.50	5.59	6.62	6.70

2. Use of endmember properties without a mixing framework (OMP issue). This is probably my biggest concern. The salinity ranges are taken from Pardo et al. (2012), who use Optimum Multiparameter (OMP) analysis. Therefore the quoted values represent end members, rather than mean properties. OMP explicitly accounts for mixing and derives bulk water mass fractions based on unmixed source waters. In the present manuscript, these endmember values are used directly as classification thresholds without performing any mixing analysis. It is unclear to me whether this is appropriate.

Response:
Thank you for raising this important point. We agree that the hydrographic values reported by Pardo et al. (2012) represent source-water endmembers derived within an Optimum Multiparameter (OMP) framework, rather than the mean properties of mixed AABW in the ocean interior. We therefore clarified this distinction in the revised manuscript.

In our study, these endmember salinity values are not used to calculate water-mass fractions or to attribute individual water parcels quantitatively to a unique formation source. Instead, the reported endmember salinities are used only as central reference values for defining hydrographic subtypes associated with the Weddell–Prydz, Ross Sea, and Adélie Land source regions. The classification is additionally constrained by the neutral-density threshold $\gamma_n \geq 28.27$ kg m^{-3} and by prescribed geographical sectors. We revised the terminology accordingly and no longer imply that the method provides an OMP-type decomposition of source-water contributions.

The sensitivity analysis described in our response to major comment #1 evaluates the dependence of the classification on the adopted endmember values. Rather than requiring grid cells to match an exact endmember salinity, we tested progressively wider ranges of $\Delta S = \pm 0.01, \pm 0.02, \pm 0.03$, and ± 0.04 around each reference value. As shown in Figure R2 and Table R1, the three AABW subtypes within their respective prescribed longitudinal sectors exhibit quite stable and similar spatial distributions when $\Delta S \geq 0.02$ psu (Figure R2). The areas of the three AABW subtypes tend to slowly converge to the values based solely on the neutral density criterion as ΔS reaches or exceeds ± 0.02 psu (Table R1). No abrupt spatial reorganization occurs when the threshold is widened, indicating that the classification is not controlled by an isolated or arbitrarily precise endmember value.

On the basis of this analysis, we adopted $\Delta S = \pm 0.02$ in the revised manuscript. This range allows for moderate modification of the source-water properties through mixing and downstream spreading. We therefore interpret the resulting categories as hydrographic AABW subtypes with properties characteristic of the respective source regions, rather than as unmixed endmembers or quantitative source fractions. We have added this methodological qualification and its associated limitation in the revised manuscript.

3. *Exclusion of temperature. The manuscript explicitly excludes potential temperature from the classification on the basis that it is "too variable". This is difficult to justify physically, particularly as the manuscript also uses density (and therefore implicitly temperature). Also - excluding temperature while later discussing temperature patterns extensively leads to an inconsistency between the definition and the interpretation.*

Response:
Thank you for the reviewer's comment. We agree that our previous statement that potential temperature was "excluded" because it was "too variable" was insufficiently precise. In fact, temperature is incorporated into the calculation of the neutral density threshold ($\gamma_n \geq 28.27$ kg m^{-3}), which is used in the first step to identify Antarctic Bottom Water (AABW). We have therefore removed the phrase "excluding temperature" from the revised manuscript.

To assess whether potential temperature can provide an additional independent constraint for distinguishing the three AABW subtypes, we conducted a sensitivity analysis using progressively widened potential-temperature ranges relative to the central value reported by Pardo et al. (2012), including $\Delta T = \pm 0.10$, ± 0.20, ± 0.30, ± 0.40, and ± 0.50 °C (Figure R3 and Table R2). Under a narrow range (i.e., ± 0.10 °C), the identified water-mass areas are small and spatially fragmented. At 4000 m, a range of ± 0.10 °C identifies WPBW, RSBW, and ALBW areas of 0.46, 0.85, and 1.66×10^6 km², respectively, whereas a range of ± 0.30 °C identifies much larger areas, 4.92, 4.17, and 4.67×10^6 km². The identified distributions continue to expand and change in spatial continuity under wider ranges, with similar behavior at the near-bottom layer.

Although increasing the identified area is an expected consequence of widening the classification range, the potential-temperature analysis does not exhibit a clear stage at which the principal spatial patterns become relatively stable. By contrast, once the salinity range reaches $\Delta S = \pm 0.02$, the principal distributions of the three AABW subtypes remain broadly consistent, and further widening mainly produces a gradual expansion of their boundaries. A fixed potential-temperature range therefore provides a less stable standalone constraint for subtype classification.

Accordingly, we use neutral density (temperature effect partially considered), salinity characteristics, and geographical sectors to define the subtype distributions. Potential temperature is subsequently evaluated within these identified spatial footprints to diagnose thermodynamic biases in the Ensemble Reanalysis and CMIP6 models. We have clarified this distinction between water-mass classification and thermodynamic-property evaluation in the revised manuscript.

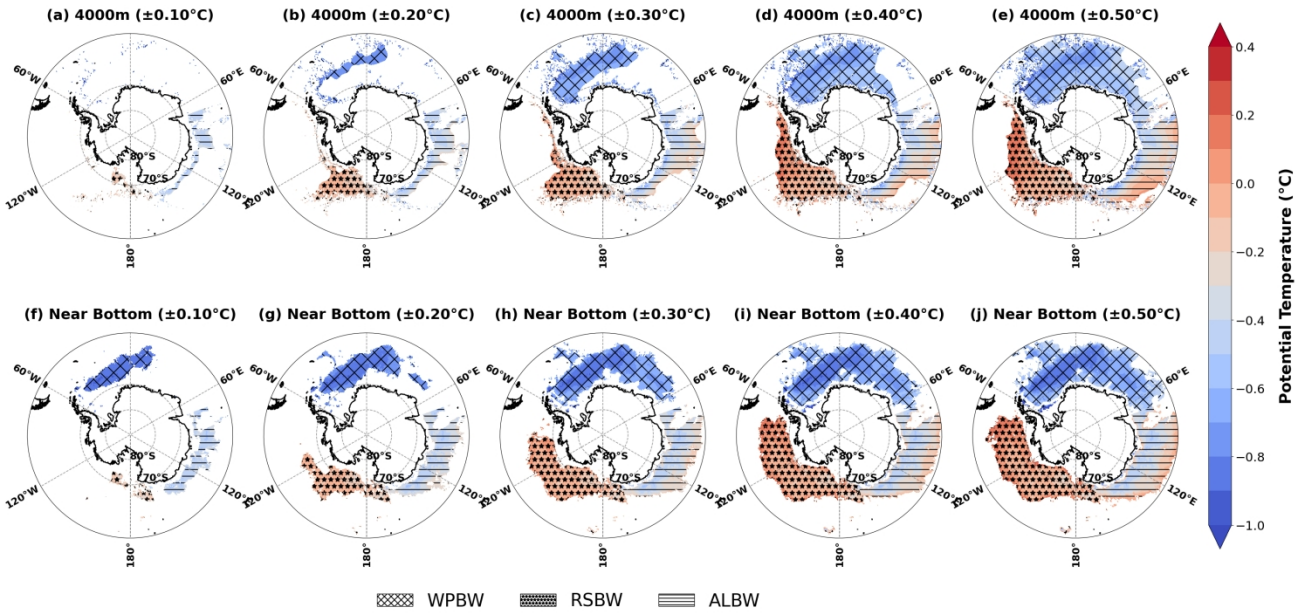

Figure R3. Sensitivity of the spatial distribution of Antarctic Bottom Water (AABW) subtypes to progressively widened potential temperature (PT) classification thresholds. The 10-panel layout illustrates the diagnostic results at the 4000 m depth layer (upper row, panels a–e) and the near-bottom layer (lower row, panels f–j) based on the World Ocean Atlas 2023 (WOA23) climatology. The classification bounds are symmetrically relaxed with specific values of ±0.10, ±0.20, ±0.30, ±0.40, and ±0.50 °C. The background color shading represents the background potential temperature field (°C). The respective AABW subtypes are distinguished by distinct hatching patterns, as detailed in the figure legend.

Table R2. Quantitative sensitivity of the identified areal extent (10^6 km^2) of AABW subtypes (WPBW, RSBW, and ALBW) to varying potential temperature (PT, °C) thresholds. The analysis is evaluated at both the 4000 m layer and the near-bottom layer based on the World Ocean Atlas 2023 (WOA23) climatology.

Layer	Subtype	±0.10 °C	±0.20 °C	±0.30 °C	±0.40 °C	±0.50 °C
4000 m	WPBW	0.46	2.07	4.92	8.21	10.52
	RSBW	0.85	2.81	4.17	5.80	6.61
	ALBW	1.66	3.16	4.67	5.74	6.72
Near-Bottom	WPBW	2.19	4.05	6.93	8.90	10.47
	RSBW	0.63	3.08	5.34	6.92	7.42
	ALBW	2.23	3.95	5.03	5.78	6.35

4. Bias correction.

The bias correction (Section 3.4) subtracts regional mean T/S differences between models and reanalysis. This approach effectively forces the models closer to the reference state and artificially improves agreement. To me, it is not clear that this procedure provides meaningful insight into model skill.

Response:

Thanks for the reviewer's comments. We agree that correcting regional mean temperature and salinity biases does bring the simulated temperature and salinity states closer to the reference dataset. In this study, bias correction is used only as a diagnostic method to examine whether some models contain systematic mean-state biases and whether, after removing these biases, their AABW spatial distributions can be recovered to some extent. At the same time, some models still perform poorly after correction, indicating that their problems cannot be explained solely by mean temperature and salinity biases.

Figure R4 compares the spatial pattern of potential temperature and salinity at 4000 m and the near-bottom layer in 16 CMIP6 models and the Ensemble Reanalysis relative to WOA23. In the Taylor diagrams, the spatial correlation coefficient (SCC) is used to measure the similarity of the spatial structures between the simulations and the reference dataset, the normalized standard deviation (NSTD) is used to measure the difference in the amplitude of simulated spatial variability relative to the observations, and the normalized centered root-mean-square error (NCRMSE) reflects the spatial pattern error between the two after removing their spatial means. The results show that, although some models cannot identify AABW distributions according to the neutral density and salinity thresholds, their temperature and salinity fields exhibit a certain degree of spatial correlation and relatively reasonable spatial variability. This indicates that these models may retain some spatial structures similar to the observations, and that their poor AABW identification may, to some extent, be affected by systematic mean-state biases. In contrast, some models exhibit low or even negative SCCs, NSTD values that clearly deviate from the reference dataset, and large NCRMSE values, indicating that their problems also involve the temperature and salinity spatial structures themselves.

Based on the differences revealed by the Taylor diagrams, we further compared the AABW identification results before and after bias correction to determine whether the AABW spatial distributions of the models could be recovered after removing the regional mean temperature and salinity biases. Tables R3 and R4 give the area of WPBW, RSBW, and ALBW identified by each model before correction at 4000 m and the near-bottom layer, respectively, together with the mean potential temperature and salinity within the area of each subtype. Tables R5 and R6 provide the corresponding results after bias correction. Therefore, by comparing Table R3 with Table R5 and Table R4 with Table R6, we can determine whether the AABW subtypes in each model can be recovered to some extent after removing the regional mean temperature and salinity biases. For example, before correction, ACCESS-ESM1-5 identifies only 0.03 × 10^6 km^2 of WPBW at 4000 m and does not identify RSBW or ALBW (Table R3). After correction, the identified areas of WPBW, RSBW, and ALBW increase to 6.20, 2.47, and 2.71 × 10^6 km^2, respectively (Table R5).

This indicates that, after removing the mean-state bias, part of the regional spatial structure simulated by this model can satisfy the AABW classification criteria. In contrast, some models cannot reasonably identify specific AABW subtypes either before or after correction. For example, before correction, CNRM-CM6-1-HR does not identify WPBW, RSBW, or ALBW at the near-bottom layer (Table R4); after the regional mean temperature and salinity bias correction, the model still does not identify WPBW or RSBW, and the identified ALBW area is only 0.01×10^6 km^2 (Table R6).

In summary, bias correction is not intended to artificially improve the agreement between the models and the reference dataset, but rather to distinguish models that are primarily affected by systematic mean-state biases from those that still exhibit clear spatial structural deficiencies after correction.

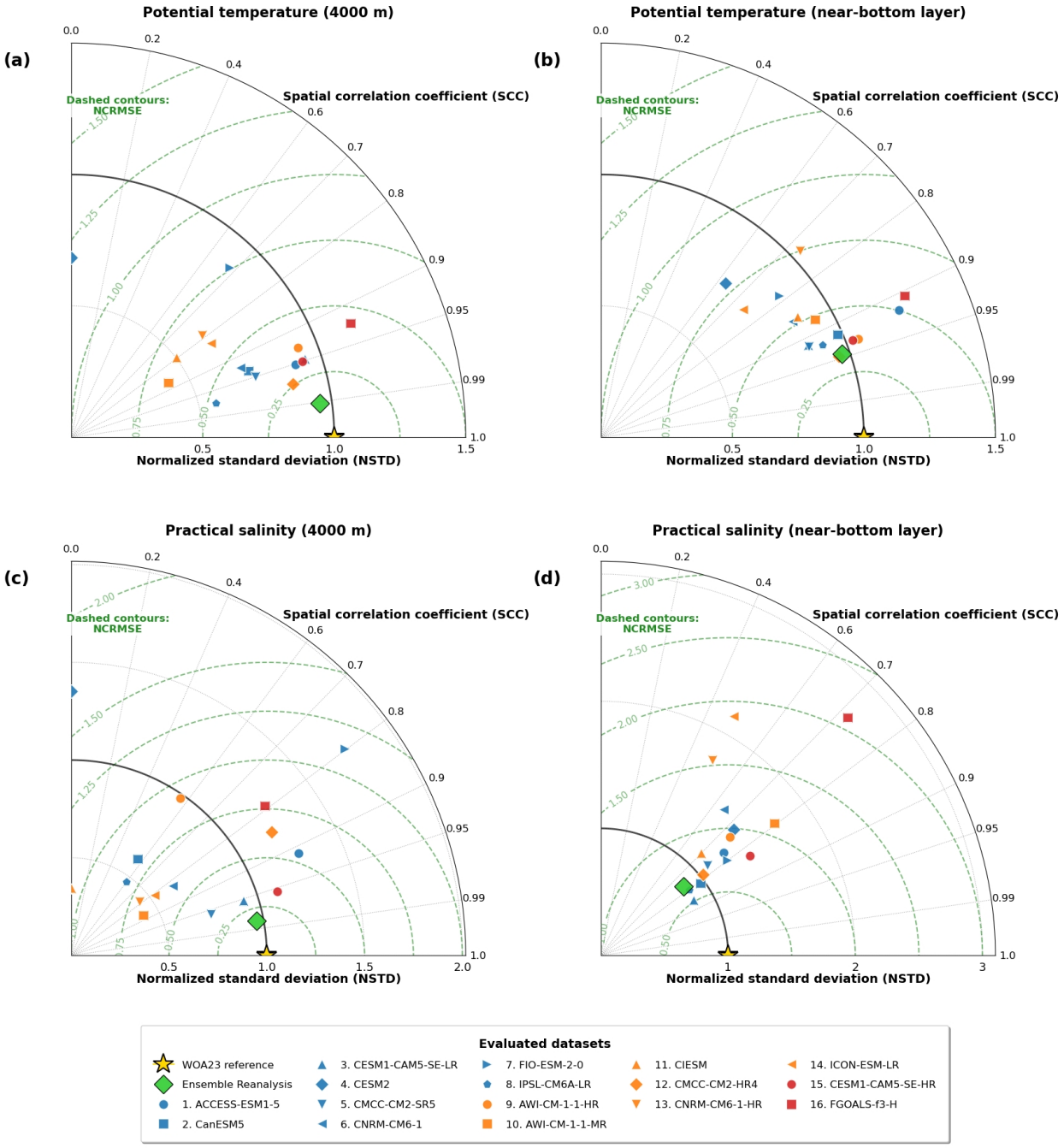

Figure R4. Taylor diagrams evaluating the ability of the simulations to reproduce the spatial distributions of thermohaline properties across the Southern Ocean south of 50°S. The upper row shows potential temperature at the 4000 m depth layer (a) and the near-bottom layer (b), while the lower row shows practical salinity at the 4000 m depth layer (c) and the near-bottom layer (d). The thermohaline spatial fields from the Ensemble Reanalysis and 16 CMIP6 models are evaluated against the WOA23 climatological reference. The gold star at SCC = 1 and NSTD = 1 represents the WOA23 reference. In each diagram, the angular coordinate represents the spatial correlation coefficient (SCC), which measures the similarity between the simulated and reference spatial distributions, while the radial coordinate represents the normalized standard deviation (NSTD), which indicates the amplitude of simulated spatial variability relative to WOA23. The green

dashed contours represent the normalized centered root-mean-square error (NCRMSE). The datasets are distinguished by color: the Ensemble Reanalysis is shown in green, low-resolution CMIP6 models in blue, moderate-resolution models in orange, and high-resolution models in red. Because the diagrams display only the non-negative-correlation quadrant, models with SCC ≤ 0 are plotted on the zero-correlation boundary. To prevent excessive stretching of the radial scale, data points with NSTD > 10 are not displayed.

Table R3. Quantitative metrics of the simulated Antarctic Bottom Water (AABW) subtypes at the 4000 m depth layer across the Southern Ocean (south of 50°S). The table details the spatial area (10^6 km^2), mean potential temperature (PT, °C), and mean practical salinity (S) for the Weddell–Prydz Bottom Water (WPBW), Ross Sea Bottom Water (RSBW), and Adélie Land Bottom Water (ALBW). Metrics are evaluated for the WOA23 climatological reference, the Ensemble Reanalysis, and 16 CMIP6 models. AABW subtype properties are averaged strictly within the spatial extent where each dataset successfully satisfies the prescribed density and salinity thresholds. A value of 0 in the area column indicates that no grid cells satisfy the prescribed criteria for the corresponding subtype. In such cases, the mean potential temperature and practical salinity cannot be calculated and are therefore denoted by dashes.

Dataset	WPBW Area	WPBW Mean PT	WPBW Mean S	RSBW Area	RSBW Mean PT	RSBW Mean S	ALBW Area	ALBW Mean PT	ALBW Mean S
[Observational & Reanalysis References]									
WOA23 (Reference)	7.56	-0.55	34.657	2.92	0.04	34.701	3.10	-0.21	34.677
Ensemble Reanalysis	7.51	-0.53	34.659	2.83	0.04	34.704	3.25	-0.19	34.682
[Low-Resolution CMIP6 Models]									
ACCESS-ESM1-5	0.03	-0.29	34.631	0	-	-	0	-	-
CanESM5	8.26	-0.96	34.652	3.19	-0.49	34.689	1.40	-0.72	34.664
CESM1-CAM5-SE-LR	6.39	-0.47	34.659	2.20	0.05	34.693	2.38	-0.13	34.680
CESM2	0	-	-	0	-	-	0	-	-
CMCC-CM2-SR5	3.16	-0.20	34.667	0	-	-	0.01	0.04	34.676
CNRM-CM6-1	0	-	-	0	-	-	0	-	-
FIO-ESM-2-0	0	-	-	0	-	-	0	-	-
IPSL-CM6A-LR	8.30	-1.14	34.657	3.53	-0.85	34.683	3.42	-0.90	34.672
[Moderate-Resolution CMIP6 Models]									
AWI-CM-1-1-HR	0	-	-	2.09	-0.41	34.710	1.89	-0.80	34.695
AWI-CM-1-1-MR	0	-	-	0	-	-	0	-	-
CIESM	0	-	-	0	-	-	0	-	-
CMCC-CM2-HR4	0	-	-	0	-	-	0	-	-

CNRM-CM6-1-HR	0	-	-	0	-	-	0	-	-
ICON-ESM-LR	0	-	-	0	-	-	0	-	-
[High-Resolution CMIP6 Models]									
CESM1-CAM5-SE-HR	8.42	-0.39	34.659	1.24	0.15	34.709	0.10	-0.19	34.667
FGOALS-f3-H	4.24	-0.92	34.641	2.59	-0.02	34.689	1.59	-0.25	34.677

Table R4. Quantitative metrics of the simulated Antarctic Bottom Water (AABW) subtypes at the near-bottom layer. The variables, evaluated datasets, and spatial averaging criteria are identical to those detailed in Table R3. A value of 0 in the area column indicates that no grid cells satisfy the prescribed criteria for the corresponding subtype.

Dataset	WPBW Area	WPBW Mean PT	WPBW Mean S	RSBW Area	RSBW Mean PT	RSBW Mean S	ALBW Area	ALBW Mean PT	ALBW Mean S
[Observational & Reanalysis References]									
WOA23 (Reference)	12.19	-0.62	34.654	7.27	-0.04	34.700	5.59	-0.26	34.678
Ensemble Reanalysis	10.57	-0.56	34.658	6.33	0.04	34.704	6.01	-0.18	34.682
[Low-Resolution CMIP6 Models]									
ACCESS-ESM1-5	0.16	-0.25	34.635	0.02	-1.75	34.694	0	-	-
CanESM5	14.55	-0.87	34.653	8.09	-0.48	34.690	2.95	-0.70	34.665
CESM1-CAM5-SE-LR	10.07	-0.53	34.657	7.82	-0.02	34.691	4.81	-0.15	34.679
CESM2	0	-	-	0.05	-1.93	34.697	0	-	-
CMCC-CM2-SR5	6.75	-0.28	34.665	0.08	-0.81	34.692	0.72	0.01	34.678
CNRM-CM6-1	0	-	-	0	-	-	0	-	-
FIO-ESM-2-0	0.26	-0.26	34.669	0.03	-1.11	34.693	0	-	-
IPSL-CM6A-LR	14.95	-1.10	34.656	8.68	-0.85	34.685	6.51	-0.85	34.675
[Moderate-Resolution CMIP6 Models]									
AWI-CM-1-1-HR	0.56	-1.09	34.658	3.63	-0.44	34.708	3.13	-0.78	34.692
AWI-CM-1-1-MR	0.08	-1.24	34.650	0.08	-0.88	34.698	0.04	-0.48	34.687
CIESM	0	-	-	0	-	-	0	-	-
CMCC-CM2-HR4	0.07	-1.15	34.652	0.16	-0.54	34.695	0.05	-0.51	34.684
CNRM-CM6-1-HR	0	-	-	0	-	-	0	-	-
ICON-ESM-LR	0.01	-0.31	34.658	0.01	-0.24	34.698	0.004	-0.13	34.683
[High-Resolution CMIP6 Models]									

CESM1-CAM5-SE-HR	11.61	-0.41	34.657	3.20	0.14	34.709	0.44	-0.15	34.668
FGOALS-f3-H	5.71	-0.91	34.643	4.35	-0.07	34.689	3.89	-0.26	34.676

Table R5. Quantitative metrics of the bias-corrected simulated Antarctic Bottom Water (AABW) subtypes at the 4000 m depth layer. The table details the spatial area (10^6 km^2), mean potential temperature (PT, °C), and mean practical salinity (S) for the Weddell–Prydz Bottom Water (WPBW), Ross Sea Bottom Water (RSBW), and Adélie Land Bottom Water (ALBW). A value of 0 in the area column indicates that no grid cells satisfy the prescribed criteria for the corresponding subtype despite the mean-state bias correction.

Dataset	WPBW Area	WPBW Mean PT	WPBW Mean S	RSBW Area	RSBW Mean PT	RSBW Mean S	ALBW Area	ALBW Mean PT	ALBW Mean S
[Low-Resolution CMIP6 Models]									
ACCESS-ESM1-5	6.20	-0.54	34.657	2.47	0.07	34.690	2.71	-0.20	34.680
CanESM5	8.17	-0.48	34.661	0.34	0.21	34.705	3.11	-0.18	34.685
CESM1-CAM5-SE-LR	6.54	-0.55	34.657	2.64	0.05	34.698	2.62	-0.21	34.680
CESM2	6.40	-0.52	34.657	1.14	0.05	34.683	1.37	-0.19	34.681
CMCC-CM2-SR5	6.38	-0.55	34.657	1.24	0.12	34.702	2.58	-0.19	34.681
CNRM-CM6-1	7.63	-0.49	34.660	2.32	0.15	34.703	3.22	-0.19	34.681
FIO-ESM-2-0	2.79	-0.61	34.646	2.15	-0.14	34.681	1.42	-0.19	34.678
IPSL-CM6A-LR	8.29	-0.47	34.661	0.64	0.21	34.714	3.21	-0.17	34.685
[Moderate-Resolution CMIP6 Models]									
AWI-CM-1-1-HR	5.93	-0.44	34.655	0.47	0.12	34.692	2.54	-0.18	34.680
AWI-CM-1-1-MR	7.41	-0.50	34.659	0.88	0.22	34.712	2.99	-0.17	34.683
CIESM	6.88	-0.53	34.659	1.37	0.12	34.699	2.74	-0.18	34.682
CMCC-CM2-HR4	7.03	-0.51	34.659	1.40	0.07	34.694	2.72	-0.17	34.684
CNRM-CM6-1-HR	7.65	-0.49	34.660	0	-	-	3.77	-0.12	34.685
ICON-ESM-LR	6.44	-0.55	34.658	1.90	0	34.695	2.60	-0.18	34.682
[High-Resolution CMIP6 Models]									
CESM1-CAM5-SE-HR	8.11	-0.51	34.659	1.80	0.08	34.703	3.42	-0.15	34.683
FGOALS-f3-H	5.89	-0.60	34.654	1.60	0	34.683	1.72	-0.18	34.680

Table R6. Quantitative metrics of the bias-corrected simulated Antarctic Bottom Water (AABW) subtypes at the near-bottom layer. The evaluated variables, bias-correction methodology, and spatial averaging criteria are identical to those detailed in Table R5. A value of 0 in the area column

indicates that no grid cells satisfy the prescribed criteria for the corresponding subtype despite the mean-state bias correction.

Dataset	WPBW Area	WPBW Mean PT	WPBW Mean S	RSBW Area	RSBW Mean PT	RSBW Mean S	ALBW Area	ALBW Mean PT	ALBW Mean S
[Low-Resolution CMIP6 Models]									
ACCESS-ESM1-5	9.71	-0.69	34.653	8.53	0	34.701	5.50	-0.26	34.686
CanESM5	12.55	-0.51	34.661	3.53	0.16	34.703	6.35	-0.17	34.692
CESM1-CAM5-SE-LR	10.58	-0.57	34.655	5.56	0.09	34.703	5.70	-0.22	34.677
CESM2	4.17	-0.43	34.666	5.17	0.01	34.691	1.70	-0.14	34.681
CMCC-CM2-SR5	11.8	-0.54	34.651	4.04	0.15	34.708	5.93	-0.14	34.675
CNRM-CM6-1	0	-	-	0.30	0.13	34.710	5.64	-0.15	34.694
FIO-ESM-2-0	2.79	-0.45	34.648	6.51	-0.10	34.692	1.77	-0.10	34.679
IPSL-CM6A-LR	13.53	-0.43	34.659	3.43	0.17	34.712	5.64	-0.13	34.688
[Moderate-Resolution CMIP6 Models]									
AWI-CM-1-1-HR	9.47	-0.61	34.650	2.52	0.07	34.697	3.90	-0.20	34.678
AWI-CM-1-1-MR	9.34	-0.54	34.665	0.07	-0.48	34.699	6.61	-0.17	34.694
CIESM	7.19	-0.64	34.668	4.95	0.17	34.709	5.29	-0.23	34.686
CMCC-CM2-HR4	12.68	-0.43	34.653	3.20	0.10	34.703	5.41	-0.09	34.685
CNRM-CM6-1-HR	0	-	-	0	-	-	0.01	0.19	34.702
ICON-ESM-LR	8.39	-0.63	34.664	5.71	0.03	34.710	6.49	-0.14	34.688
[High-Resolution CMIP6 Models]									
CESM1-CAM5-SE-HR	10.01	-0.52	34.66	0.14	-0.24	34.706	6.27	-0.03	34.689
FGOALS-f3-H	8.69	-0.88	34.654	0.03	-0.54	34.699	4.06	-0.29	34.677

Minor Comments

1. Introduction

- *L47: repeat transects remain the primary constraint on AABW properties and underpin climatologies such as WOA.*

Response:

Thanks for the reviewer's comment. We have rewritten the relevant descriptions in the Introduction to explicitly emphasize the irreplaceable, foundational role of repeat hydrographic section observations in

constraining the physical properties of AABW and in underpinning climatological datasets such as the WOA.

- *L86: exhibits*

Response:
Thanks for the reviewer's comment. We have corrected the spelling error to "exhibits".

- *L87–90: please provide a reference for the stated claim - it does not obviously follow in my opinion.*

Response:

Thanks for the reviewer's comments. We agree that the original statement was too broad and insufficiently supported. We have therefore revised it and cited Purkey and Johnson (2010), Gunn et al. (2023), Marinov et al. (2006) and Zhang et al. (2023). The revised text reads:

"A source-region perspective on AABW is relevant because abyssal warming varies spatially across ocean basins (Purkey and Johnson, 2010), while bottom-water transport and ventilation can undergo substantial regional changes (Gunn et al., 2023). In addition, regional Southern Ocean circulation influences air–sea CO_2 exchange (Marinov et al., 2006), and dense-shelf-water export transports anthropogenic carbon from the Antarctic shelf into the deep ocean (Zhang et al., 2023)."

References:
Gunn, K. L., Rintoul, S. R., England, M. H., and Bowen, M. M.: Recent reduced abyssal overturning and ventilation in the Australian Antarctic Basin, Nat. Clim. Change, 13, 537–544, https://doi.org/10.1038/s41558-023-01667-8, 2023.

Marinov, I., Gnanadesikan, A., Toggweiler, J. R., and Sarmiento, J. L.: The Southern Ocean biogeochemical divide, Nature, 441, 964–967, https://doi.org/10.1038/nature04883, 2006.

Purkey, S. G. and Johnson, G. C.: Warming of global abyssal and deep Southern Ocean waters between the 1990s and 2000s: Contributions to global heat and sea level rise budgets, J. Clim., 23, 6336–6351, https://doi.org/10.1175/2010JCLI3682.1, 2010.

Zhang, S., Wu, Y., Cai, W.-J., Cai, W., Feely, R. A., Wang, Z., Tanhua, T., Wang, Y., Liu, C., Li, X., Yang, Q., Ding, M., Xu, Z., Kerr, R., Luo, Y., Cheng, X., Chen, L., and Qi, D.: Transport of anthropogenic carbon from the Antarctic shelf to the deep Southern Ocean triggers acidification, Global Biogeochem. Cy., 37, e2023GB007921, https://doi.org/10.1029/2023GB007921, 2023.

2. Methods

- *L155: Pardo et al. (2012) is based on OMP analysis. Are the quoted values intended as endmembers? This should be clarified explicitly.*

Response:
Thanks for the reviewer's comment. We agree that the values reported by Pardo et al. (2012) represent source-water endmembers derived from an Optimum Multiparameter analysis rather than mean properties of fully mixed AABW. As explained in our response to major comment #2, we now use these salinity values only as central hydrographic reference values, rather than as unmixed endmembers for calculating source-water fractions. The classification is additionally constrained by $\gamma_n \geq 28.27$ kg m^{-3} and geographical sectors. The sensitivity analysis in Figure R2 and Table R1 further shows that the principal subtype distributions remain broadly stable when the salinity range reaches $\Delta S = \pm 0.02$. We have clarified this interpretation and its limitation in the revised Methods section.

- *L163: please justify the exclusion of temperature more clearly.*

Response:
Thanks for the reviewer's comment. We agree that our previous statement that potential temperature was "excluded" was insufficiently precise. As explained in our response to major comment #3, temperature is incorporated into the neutral-density criterion used to identify AABW. We have therefore removed the phrase "excluding temperature." We also conducted an additional potential-temperature sensitivity analysis (Figure R3 and Table R2), which shows that a fixed potential-temperature range provides a less stable standalone constraint than the adopted salinity criterion. Potential temperature is therefore retained as an important diagnostic variable for evaluating thermodynamic properties after the subtype distributions have been identified.

- *L174–176: the analysis appears to focus only on two depth levels (4000 m and near-bottom). Please justify this choice and discuss its limitations. Is it possible that the properties of AABW remain similar, but the models just form less? This distinction is needed and is entirely missed by focusing on 2 discrete depth levels.*

Response:
Thank you for this important comment. We agree that the original analysis based only on the 4000 m and near-bottom layers provided relatively limited information on the vertical extent of the AABW subtypes. We have now added salinity diagnostics at two additional depth levels, 3500 m and 4500 m, as shown in Figures R5 and R6, respectively.

The new results show that WOA23 and the Ensemble Reanalysis retain spatially coherent WPBW, RSBW, and ALBW characteristics at both 3500 m and 4500 m. Different CMIP6 models, however, exhibit clearly different levels of vertical persistence. For example, CESM1-CAM5-SE-LR and CanESM5 identify relatively extensive AABW subtype distributions at both 3500 m and 4500 m, indicating that the bottom-water characteristics simulated by these models maintain relatively good vertical continuity between the two depth levels, although their specific spatial extents still differ from those in WOA23.

In contrast, CMCC-CM2-SR5 and FGOALS-f3-H identify some AABW subtype distributions at 4500 m, but the corresponding spatial extents shrink markedly or largely disappear at 3500

m. This rapid contraction toward shallower depths is consistent with limited vertical extension or a smaller overall volume of bottom water.

On the other hand, ACCESS-ESM1-5, CESM2, CNRM-CM6-1, CIESM, AWI-CM-1-1-MR, CMCC-CM2-HR4, and ICON-ESM-LR identify only very small areas of AABW subtypes at both 3500 m and 4500 m, or no grid cells at all that satisfy the identification criteria adopted in this study. Therefore, the biases in these models cannot be explained solely by insufficient AABW vertical thickness or reduced formation volume, but may also involve biases in the simulated deep-ocean thermohaline structure or regional spatial distribution.

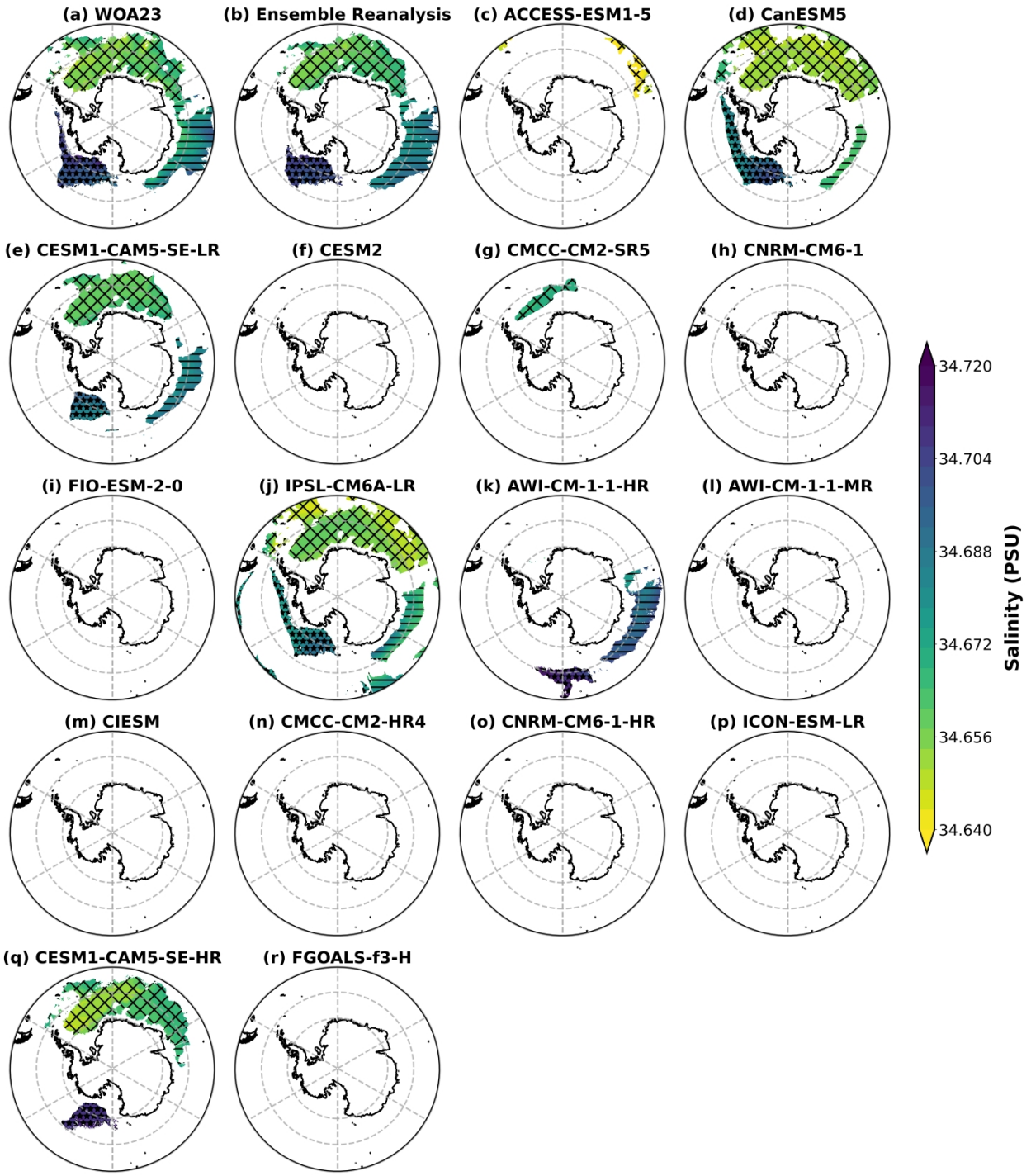

Figure R5. Spatial distributions of practical salinity and the identified Antarctic Bottom Water (AABW) subtypes at the 3500 m depth layer in the Southern Ocean (south of 50°S). The figure displays the WOA23 observational reference, the Ensemble Reanalysis, and 16 CMIP6 models, arranged within a 5 × 4 layout, with 18 populated panels labeled from (a) to (r). The respective AABW subtypes are distinguished by distinct hatching patterns, as detailed in the figure legend.

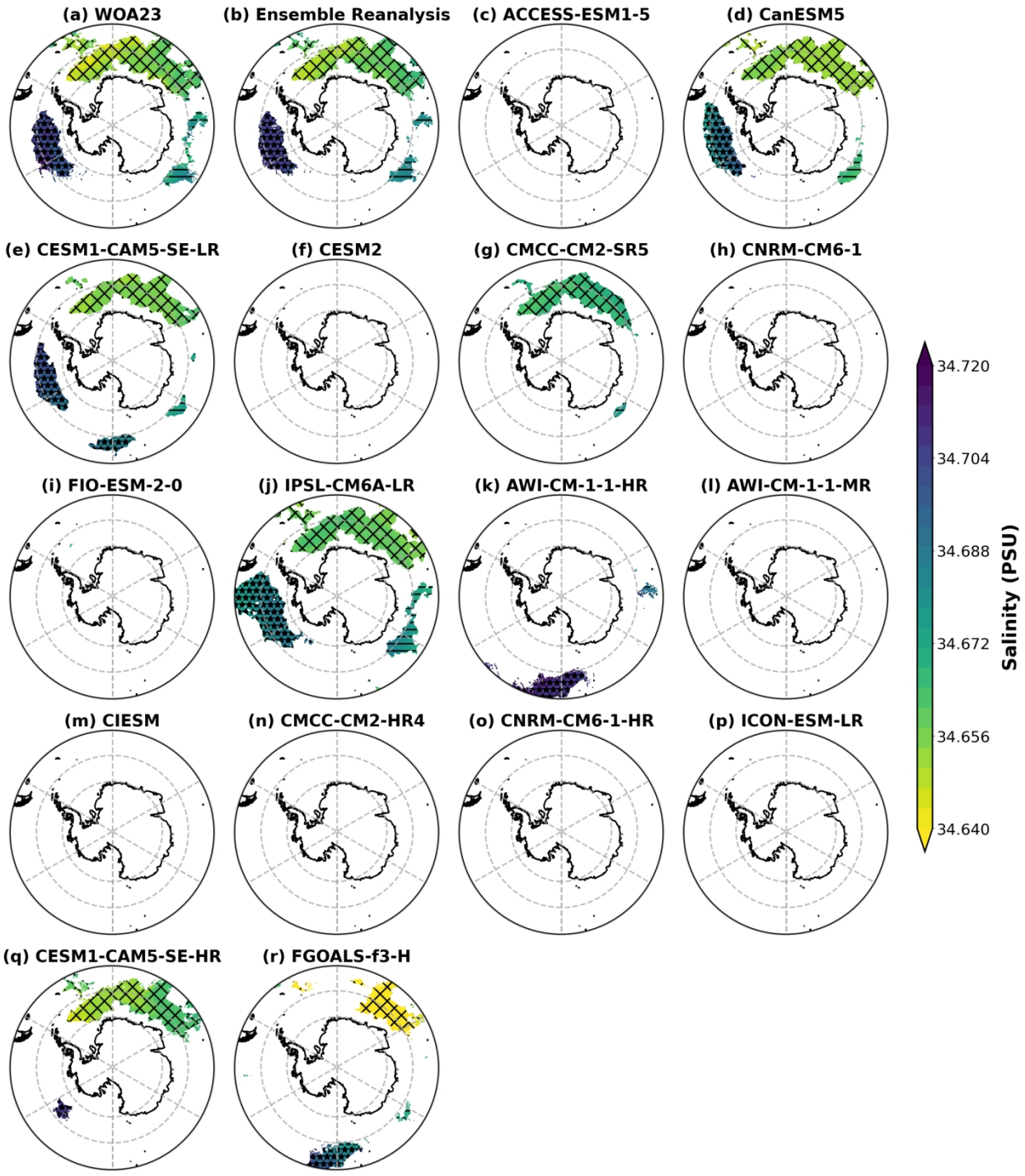

Figure R6. Same as Figure R5, but for the 4500 m depth layer.

3. Results

- *L211: capitalisation issue.*

Response:

Thanks for the reviewer's comment. We have corrected the capitalization error.

- *Figures 1–2 and throughout: there is extensive discussion of potential temperature despite it not being part of the classification.*

Response:
Thanks for the reviewer's comment. Potential temperature is not used as an additional independent threshold for distinguishing the three AABW subtypes. However, temperature is incorporated into the neutral-density criterion ($\gamma_n \geq 28.27$ kg m^{-3}) used to identify AABW.

After the subtype distributions are identified using neutral density, salinity characteristics, and geographical sectors, potential temperature is examined as a diagnostic property within those identified regions. The purpose of Figures 1–2 and the Results section is to describe the thermohaline characteristics of the identified water masses and to evaluate temperature biases in the Ensemble Reanalysis and CMIP6 models, rather than to determine subtype membership.

- *L239–241: The relevance of this discussion to AABW formation is unclear.*

Response:
Thanks for the reviewer's comment. We agree that the original discussion did not clearly explain the relationship of mesoscale eddies and model resolution to AABW formation.

We have revised the paragraph to clarify the important roles of mesoscale eddies in cross-shelf exchanges of heat and salt, the transport and mixing of dense shelf water, and its subsequent export and spreading into the deep ocean. Therefore, the extent to which mesoscale eddies are resolved may affect the simulated thermohaline properties and spatial distribution of AABW.

In addition, AABW formation and export are jointly controlled by other physical processes and their associated parameterizations. We therefore classify the CMIP6 models into low-, moderate-, and high-resolution groups to compare their representation of the thermohaline properties and spatial distribution of AABW, rather than to assume in advance that higher-resolution models necessarily have better simulation skill.

- *Figures 3–6: These figures are difficult to read due to small size and dense layout. Consider enlarging panels or splitting figures.*

Response:
Thank you for this suggestion. We agree that the original figures were difficult to read because too many model panels were presented in a dense layout. We have therefore reorganized the results and added new Figures R7 and R8, which present the original and bias-corrected simulations, respectively, using an enlarged 5 × 4 matrix layout. These figures focus on five representative CMIP6 models spanning different resolutions and levels of performance, allowing the spatial distributions and thermohaline properties of the AABW subtypes to be examined more clearly. The quantitative results for all models are retained in the Taylor diagrams and summary tables.

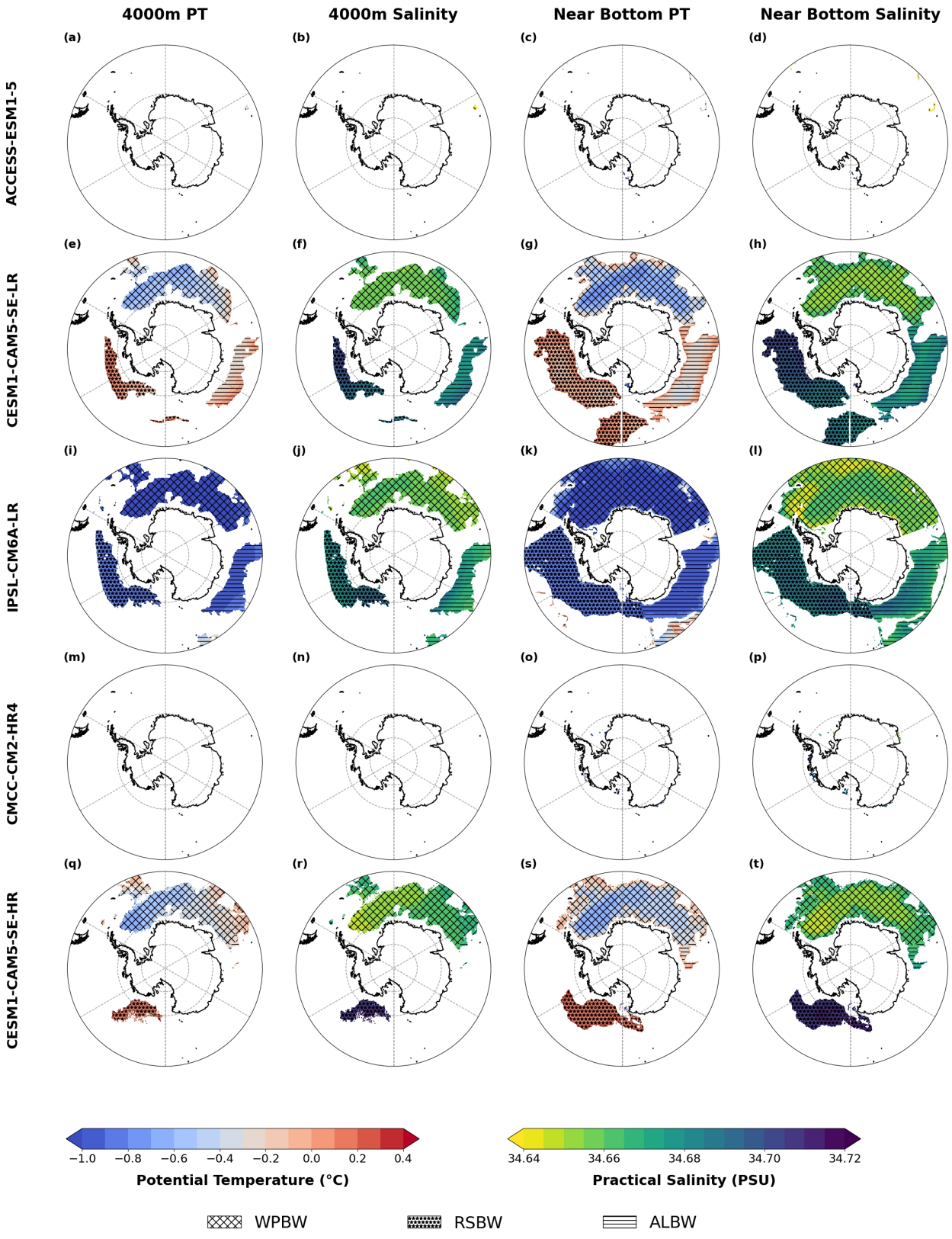

Figure R7. Spatial distributions of the native (uncorrected) simulated potential temperature and practical salinity for Antarctic Bottom Water (AABW) subtypes across five representative CMIP6

models in the Southern Ocean (south of 50°S). The figure is organized into a 5 × 4 matrix. The rows represent the strategically selected models: ACCESS-ESM1-5, CESM1-CAM5-SE-LR, IPSL-CM6A-LR, CMCC-CM2-HR4, and CESM1-CAM5-SE-HR. The columns display the evaluated thermohaline properties: potential temperature at 4000 m, practical salinity at 4000 m, potential temperature at the near-bottom layer, and practical salinity at the near-bottom layer. Subpanels are labeled from (a) to (t). The respective AABW subtypes are distinguished by distinct hatching patterns, as detailed in the figure legend.

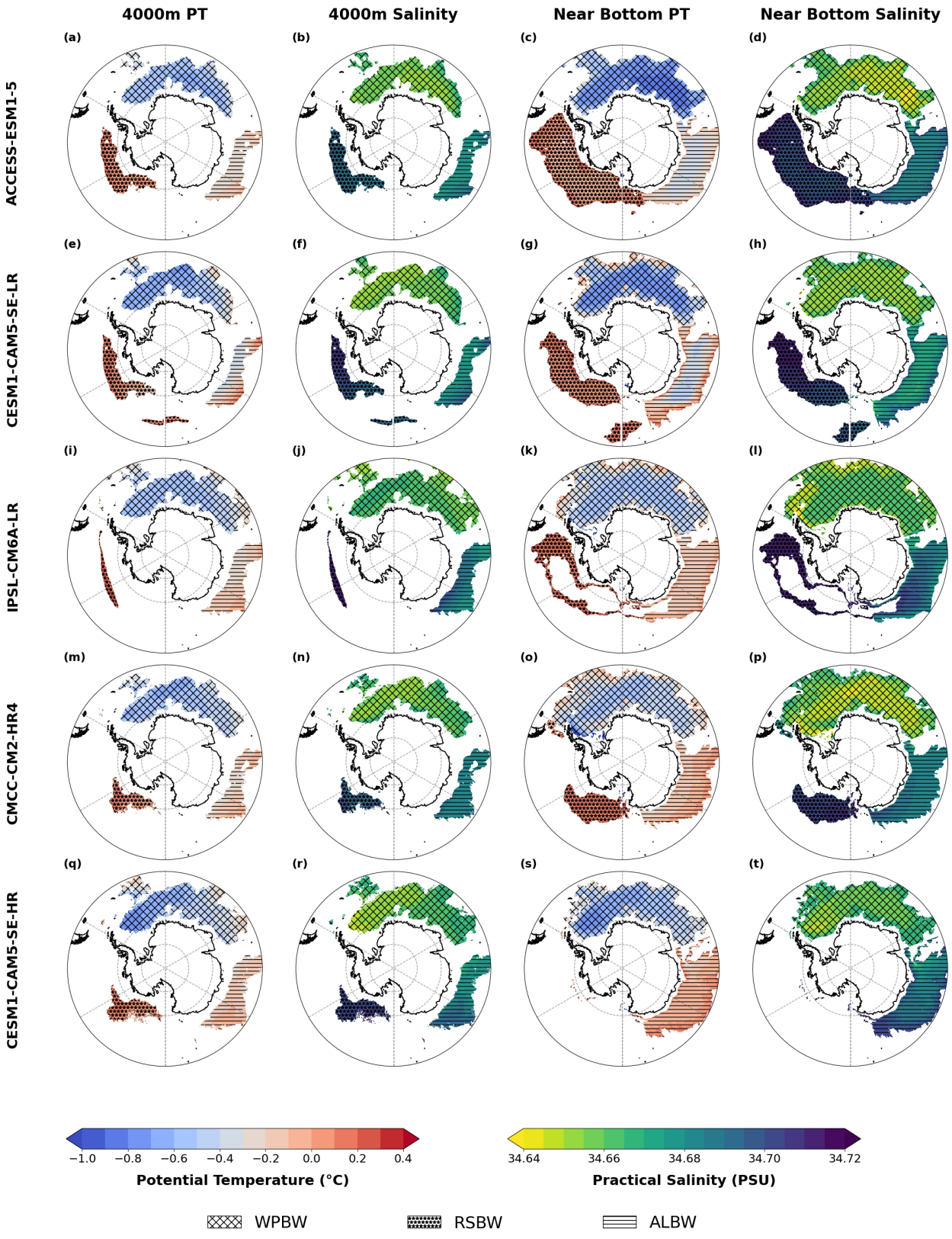

Figure R8. Spatial distributions of the bias-corrected potential temperature and practical salinity for AABW subtypes across the same five representative CMIP6 models shown in Figure R7. The

matrix layout, including the models in the rows and the thermohaline properties and depth levels in the columns, is identical to Figure R7. Subpanels are labeled from (a) to (t). The respective AABW subtypes are represented by distinct hatching patterns, as detailed in the figure legend.

- *please quantify how well or otherwise each product describes AABW relative to the baseline*

Response:
Thank you for this important suggestion. We agree that relying solely on visual inspection of spatial distribution maps makes it difficult to objectively assess how well different data products represent AABW. Therefore, in the revised manuscript, we have added quantitative evaluations of the Ensemble Reanalysis and all 16 CMIP6 models, using WOA23 consistently as the reference baseline.

First, the Taylor diagrams in Figure R4 use the spatial correlation coefficient (SCC), normalized standard deviation (NSTD), and normalized centered root-mean-square error (NCRMSE) to quantitatively evaluate the simulated spatial patterns of potential temperature and salinity at the 4000 m and near-bottom layers. The SCC measures the agreement between the simulated spatial pattern and that of WOA23; the NSTD indicates whether the magnitude of simulated spatial variability is weaker or stronger than that in WOA23; and the NCRMSE provides an integrated measure of the difference between the simulated and WOA23 spatial structures after removal of the regional mean.

The results show that the Ensemble Reanalysis reproduces the large-scale abyssal thermohaline structure in WOA23 reasonably well. For example, its SCC for salinity at 4000 m is approximately 0.98, indicating strong agreement with the regional salinity contrasts and major spatial gradients in WOA23. In contrast, several CMIP6 models exhibit pronounced structural biases. CESM2 and CIESM yield SCC values of approximately −0.55 and −0.10, respectively, for salinity at 4000 m. The negative correlation for CESM2 indicates that its simulated salinity anomalies are spatially opposed to those in WOA23 over a substantial part of the domain; that is, regions that are relatively saline in WOA23 may be represented as relatively fresh in the model, and vice versa. The SCC of CIESM is slightly below zero, indicating almost no correspondence between its simulated salinity pattern and that of WOA23.

Second, Tables R3 and R4 quantitatively summarize the identified areas and mean potential temperature and salinity of WPBW, RSBW, and ALBW before bias correction at the 4000 m and near-bottom layers, respectively, while Tables R5 and R6 present the corresponding results after bias correction. These statistics allow us to determine whether each product identifies the respective AABW subtypes and to quantify differences in their simulated spatial extent and thermohaline properties relative to WOA23.

At the 4000 m layer, for example, WOA23 identifies WPBW, RSBW, and ALBW areas of 7.56, 2.92, and 3.10 × 10^6 km^2, respectively, whereas the corresponding areas identified by the Ensemble Reanalysis are 7.51, 2.83, and 3.25 × 10^6 km^2. These values differ from WOA23 by approximately −0.7%, −3.1%, and +4.8%, respectively. This indicates that the Ensemble Reanalysis not only identifies all three AABW subtypes, but also reproduces their spatial extents reasonably closely.

Substantial differences are evident among the CMIP6 models. CESM1-CAM5-SE-LR identifies all three subtypes at 4000 m, but its WPBW, RSBW, and ALBW areas are 6.39, 2.20, and 2.38 × 10^6 km^2, respectively, all smaller than those in WOA23. CanESM5 identifies WPBW and RSBW areas of 8.26 and 3.19 × 10^6 km^2, respectively, both larger than the WOA23 values, whereas its identified ALBW area is only 1.40 × 10^6 km^2. ACCESS-ESM1-5 identifies only approximately 0.03 × 10^6 km^2 of WPBW at 4000 m and does not identify RSBW or ALBW. CESM2, CNRM-CM6-1, and CIESM identify none of the three AABW subtypes satisfying the prescribed criteria at this depth.

These additional analyses allow us to quantitatively compare the ability of each data product to represent the spatial patterns, subtype extents, and thermohaline properties of AABW relative to WOA23.

- *some references appear in inconsistent fonts throughout.*

Response:
Thanks for the reviewer's comments. We have carefully checked the entire manuscript and standardized the font, font size, and formatting of all in-text citations and entries in the reference list.