Paleoceanography and Paleoclimatology

Supporting Information for

Proxy-Model Comparison for the Eocene-Oligocene Transition in Southern High Latitudes

Emily J. Tibbett[1], Natalie J. Burls[2], David K. Hutchinson[3], Sarah J. Feakins[1]

[1]Department of Earth Science, University of Southern California, Los Angeles, CA, USA
[2]Atmospheric, Oceanic, and Earth Sciences Department, George Mason University, Fairfax, VA, USA
[3]Climate Change Research Centre, University of New South Wales, Sydney, Australia

Contents of this file

Introduction

The supporting information associated with the manuscript referenced above includes two supplementary text sections, five supplementary tables and eleven supplementary figures. Text S1 provides background information on the model simulations and text S2 discusses the impact of seasonality on the results. Table S1 contains the RMSE values and calculated pCO_2 decrease. Table S2 compares the RMSE when applying MAT versus MAF from each model. Table S3 reports the RMSE between the proxies and models for both MAT and MAF for each timeslice. Table S4 reports the RMSE and pCO_2 decrease for the paleogeographic runs combined with the ensemble mean. Table S5 reports the RMSE for the inter-model comparison. Table S6 evaluates the impact of the time window averaging for the MAT results and Table S7 does the same for the SST results. Table S8 and S9 provide the results of removing individual proxy records for MAT and SST respectively. Table S10 and S11 evaluate the removal of all records of the same proxy type on the results for MAT and SST respectively. Figure S1 shows the southern hemisphere high latitude MAT for each timeslice. Figure S2 is a detailed map (a zoom in of that mapped on Figure S1) to note localized heterogeneity. Figure S3 southern hemisphere high latitude SSTs. Figure S4 MAT from model runs with and without ice. Figure S5 SST from model runs with and without ice. Figure S6 MAT in paleogeographic model runs. Figure S7 SST in paleogeographic model runs. Figure S8 proxy-model comparison for models scaled to a 25% reduction in pCO_2. Figure S9 model seasonality for each timeslice. Figure S10 compares the best fit pCO_2 scaling for the MAT and MAF comparison. Figure S11 paleogeographic model runs with the model ensemble runs. Figure S12 shows the results of the inter-model comparison using the "perfect model" approach.

Text S1. Model boundary conditions

Here we review the initial grid information for each model from the published literature. For the proxy-model comparison performed here, the models were evaluated on the same uniform grid, that of NorESM-L (Hutchinson et al., 2021). We also summarize key details of each simulation.

CESM-B: Prior to the uniform grid the atmospheric resolution was 144x96x26 and the ocean resolution was 384x320x60 (Baatsen et al., 2020). The low and high pCO_2 experiments come from Baatsen 2020 and 2016 respectively. The spin up was 3600 model years for the low pCO_2 (560 ppm) experiment and 4600 for the high pCO_2 experiment (1120 ppm). The land topography has low resolution which leads to smoothing that underestimates local temperature heterogeneity (Baatsen et al., 2020). The model simulations were in equilibrium for sea surface temperatures and the deep ocean.

CESM-H: The initial resolution was 96x48x26 for the atmosphere with 122x100x25 for the ocean (Goldner et al., 2014). CESM_H contained model simulations that include high and low pCO_2, with and without ice, and changes in paleogeography. For the high (1120ppm) and low (560 ppm) pCO_2 simulations the model was run for 3300 and 3400 model years respectively. For the no ice vs ice simulations, the prescribed pCO_2 was 560 ppm. The prescribed ice volume for the ice simulation was 20.3×10^6 km^3. For the no ice simulation, the model years were 3400 and for the ice simulation the model years were 3000. The paleogeography runs were the closing of the Tasman and Drake Passage (pre-EOT) and opening of both passages (post-EOT) with both simulations using a pCO_2 of 1120ppm. The model years used were 1300 and 1000 model years for both gateways closed and open respectively. The topography on land reflected nearly modern-day levels for the glaciated EOT simulations while the unglaciated EOT simulations used the paleo-elevation reconstruction from Sewall et al., (2000) and discussed potential error introduced by uncertainty in topography. The model was in equilibrium for the deep and surface ocean.

NorESM-L: The initial resolution was 96x48x26 for the atmosphere with 100x116x32 for the ocean (Zhang et al., 2012,2014). The NorESM-L simulations used in this study include low (560 ppm) and high pCO_2 (980 ppm) simulations as well as a simulation with paleogeography changes. The change in pCO_2 between the NorESM-L simulations was not a decrease in pCO_2 of 50%; therefore, to be consistent with the other modelling studies, the high pCO_2 simulation was scaled to 1120ppm to reflect a pCO_2 decrease of 50%. For the paleogeography simulations, the pre-EOT uses the continental configuration of 35 Ma from Scotese et al. (2001) while the post-EOT uses the 33 Ma continental configuration. Both paleogeography simulations were prescribed a pCO_2 of 560 ppm. There are no changes in passageways or changes in the Antarctic continent above/below sea level. For all simulations the model ran for 2200 years.

GFDLCM 2.1: The initial resolution was 96x60x24 for the atmosphere with 240x175x50 for the ocean (Hutchinsons et al., 2018,2019). The GFDLCM 2.1 simulations used were low (400 ppm) vs high (800 ppm) pCO_2 and changes in paleogeography. For the paleogeography changes the pCO_2 was set at 800 ppm for both simulations. The pre-EOT used a 38 Ma reconstruction from Baatsen et al., (2016) while the post-EOT simulation closed the Arctic gateway. All simulations ran for 6500 model years. The topography applied was the 38 Ma reconstruction from Baatsen et al. (2016). The model simulations were determined to be in quasi-equilibrium with the deep ocean gradually cooling (Hutchinson et al., 2018). Although the deep ocean continued to cool, surface temperature and salinity in areas of deep water formation regions were stable along with the overturning circulation (Hutchinson et al., 2019)

FOAM: The initial resolution was 48x40x18 for the atmosphere with 128x128x24 for the ocean (Ladant et al., 2014a,b). The FOAM simulations used here were the low (560 ppm) and high (1120 ppm) pCO_2 runs, no ice vs ice runs, and a run with changes in paleogeography. The ice volume prescribed was 25.0×10^6 km^3. For the paleogeography simulations the pre-EOT was a continental configuration of 34 Ma and post-EOT was 30 Ma. These paleogeography simulations did not include changes in ocean

passages/gateways or changes in large areas of Antarctica below or above sea level. All model simulations used 2000 model years. The topography for the model simulations is based on the topography of Antarctica after isostatic adjustment form the removal of the present-day ice sheet (Ladant et al., 2014a). The model simulations used were in equilibrium. In Ladant et al., (2014b) the Antarctic topographic reconstruction is based on that of Wilson et al. (2012) which reconstructed the elevation based on erosion of Antarctica. This leads to a larger area than the isostatically adjusted present day Antarctica topography after removal of the ice sheet (Ladant et al., 2014b).

UViC: The initial resolution was 150x100x11 for the atmosphere with 150x140x0 for the ocean (Sijp et al., 2016). The UviC model simulations used here were for the paleogeography comparison (there were no low and high pCO_2 runs available). The pCO_2 was consistent between runs with a pCO_2 of 1600ppm. The pre-EOT and post-EOT paleogeography was set at 45 Ma (Sewall et al., 2000). For the pre-EOT the Drake passage was closed while in the post-EOT it was open. All simulations were run for 9000 model years. No information was available on the topography. There is also no indication of MAT data from the UviC model in Sijp et al. (2016). The model simulations were run to equilibrium including in the deep ocean (Sijp et al., 2016).

HadCM3BL: The initial resolution was 96x73x19 for the atmosphere with 96x73x20 for the ocean (Kennedy et al., 2015). For HadCM3BL the model simulations used here include low (560 ppm) and high (1120 ppm) for pCO_2, no ice versus ice, and changes in paleogeography. For the no ice and ice runs the pCO_2 was 560ppm for both with a prescribed ice volume of 17.0×10^6 km^3 for the post-EOT ice run. For the paleogeography the pCO_2 was set to 560 ppm. The pre-EOT runs were a continental configuration for the Priabonian and for the post-EOT the Chattian. All model simulations were run for 1422 years. The model simulations were determined to be in quasi-equilibrium as the atmospheric and surface ocean, down to 670 m, are stable.

Text S2. Accounting for proxy seasonality in proxy-model comparisons

To assess the impact of seasonality on the proxy model comparison, we considered a seasonally defined rather than mean annual comparison for land temperature proxies only. For land temperature proxies, both plants and soil bacteria are thought to be summer-active recorders, this motivates an effort to compare to seasonal climate from the models. At this time, however, only $BayMBT_0$ has been explicitly calibrated to MAF, and only one site, Prydz Bay, has MAF estimates for both the Eocene and Oligocene. Using this MAF estimate and the other constraints as before, we compared to the climate model results and recalculated the pCO_2 scaling. The MAF proxy model comparison yields lower RMSE for the individual timeslices (**Figure S8, Table S2**). However, across the EOT the RMSE does not change and yields a similar estimated pCO_2 decrease (**Table S3 and Figure S9**). Although this seasonality proxy-model comparison is limited in application here, it demonstrates the potential for future work to calibrate proxies to seasons and to compare to the seasonal output of climate models in future efforts.

Table S1. Proxy-model discrepancy (as root mean standard error; RMSE) for the MAT and SST proxy-model comparison for the best fit pCO_2 forcing. The proxy constraints are based on n sites where SST proxy n = 6 and MAT proxy n=6. The pCO_2 is expressed in concentration units for the Oligocene-Eocene climate model runs and as a % decrease in the Oligocene relative to the Eocene runs.

	MAT proxy-model comparison			SST proxy-model comparison		
Climate Model	RMSE (°C)	pCO_2 decrease (ppmv)	pCO_2 decrease (%)	RMSE (°C)	pCO_2 decrease (ppmv)	pCO_2 decrease (%)
CESM_B	1.98	129	18.8	0.90	179	33.1
CESM_H	1.68	221	28.3	0.92	221	35.8
GFDLCM 2.1	1.35	200	26.3	1.02	111	21.0
HadCM3BL	2.24	169	23.2	1.23	362	46.4
FOAM	1.58	232	29.3	1.07	159	27.3
NorESM-L	2.34	66	10.5	1.45	375	43.4
Ensemble	**1.88**	**243**	**30.3**	**1.01**	**200**	**33.1**

Table S2. Proxy-model discrepancy (as root mean standard error; RMSE) for land surface air temperature estimates for both the Eocene and Oligocene timeslices, and EOT (Oligocene-Eocene) difference for mean annual surface air temperature (MAT) and months above freezing (MAF). MAF were selected in model runs based on their seasonal climatology. The *br*GDGT calibrations include both MAT and MAF formal calibrations and both are available for Prydz Bay and are used as defined. The other proxies are maintained with their same temperature conversions for both scenarios in order to assess whether they better approximate mean annual or 'summer' (above freezing) conditions. As hypothesized, we find greater agreement in the MAF comparison.

	Eocene		Oligocene		EOT	
Model Run	MAT RMSE (°C)	MAF RMSE (°C)	MAT RMSE (°C)	MAF RMSE (°C)	MAT RMSE (°C)	MAF RMSE (°C)
CESM_B	4.6	4.5	5.7	5.1	3.6	2.4
CESM_H	4.6	3.8	5.6	4.6	2.5	2.4
GFDL CM2.1	4.6	3.1	5.7	3.6	2.8	2.6
HadCM3BL	8.0	3.8	6.8	3.8	2.5	1.8
FOAM	6.1	4.0	7.3	4.3	2.4	1.8
NorESM-L	7.5	3.7	5.7	3.2	3.1	1.9
Ensemble	**5.2**	**3.2**	**5.0**	**3.4**	**2.3**	**2.1**

Table S3. Estimated pCO_2 decrease across the EOT for scaling experiments based on land surface air temperatures comparison results of the MAT and MAF* model outputs and proxy** comparisons. Only Prydz Bay had proxy data for both the Eocene and Oligocene as MAF formally derived, all other proxy data were unchanged from the MAT case, but many may have warm-season recording bias

Model Run	MAT		MAF	
	pCO_2 decrease (%)	pCO_2 decrease (ppmv)	pCO_2 decrease (%)	pCO_2 decrease (ppmv)
CESM_B	18.8	129	27.3	210
CESM_H	28.3	221	26.3	200
GFDL CM2.1	26.3	200	24.2	179
HadCM3BL	23.2	169	33.1	277
FOAM	29.3	232	35.8	313
NorESM-L	10.5	66	31.2	254
Ensemble	**30.3**	**243**	**30.3**	**243**

Table S4. Comparison of the three, paleogeography, model runs which have significant gateway changes or changes in continent extent showing proxy-model SSTs discrepancy as RMSE and resulting pCO_2 scaling. SST proxies are from n=7 sites. The paleogeography estimates in this table can be compared with the ensemble mean in Table S3 for the scenarios without consideration of paleogeography changes.

	UVic	CESM_H	FOAM
RMSE (°C) Oligocene-Eocene	1.66	1.10	0.95
pCO_2 decrease (%)	19.9	30.3	23.2
pCO_2 decrease (ppmv)	139	243	169

Table S5. Inter-model temperature comparison using each model in turn as the true value ("perfect model approach") to assess model uncertainty for land surface mean air temperatures (MAT) and sea surface temperatures (SST). The "proxy sites" were used as sampling points and compared to sampling all suitable grid cells (i.e., all land/sea grid cells for MAT/SST respectively) to assess the potential reduction in uncertainty with a larger number of possible sampling locations. The RMSE and pCO_2 decrease are the mean of all other models when each model simulation is used as the true temperature values.

Model Run	MAT				SST			
	Proxy sites n=6		All land n=960		Proxy sites n=6		All marine n=710	
	RMSE (°C)	pCO_2 decrease (%)	RMSE (°C)	pCO_2 decrease (%)	RMSE (°C)	pCO_2 decrease (%)	RMSE (°C)	pCO_2 decrease (%)
CESM_B	2.80	25.6	1.65	41.3	1.39	37.7	1.32	38.9
CESM_H	1.79	34.7	1.64	41.1	1.16	41.3	1.41	36.9
GFDL CM2.1	2.56	26.4	2.04	26.4	1.74	24.4	1.94	23.9
HadCM3BL	1.50	47.4	1.83	49.3	0.80	54.3	1.50	44.7
FOAM	1.82	34.3	1.96	38.3	1.74	32.3	1.84	31.5
NorESM-L	2.14	32.0	1.77	43.4	0.95	53.6	1.12	51.2
Ensemble	**1.37**	**51.1**	**1.42**	**49.2**	**1.09**	**48.2**	**1.18**	**48.1**

Table S6. Effect of time window selection to characterize each time period (4, 2 or 1 Ma) on proxy-model MAT RMSE (°C). There is no notable improvement in the RMSE when comparing 4 Ma (time windows used in the paper) versus 2 or 1 Ma windows.

Model	4 Ma time averaging		2 Ma time averaging		1 Ma time averaging		ΔRMSE (2 Ma-4 Ma)	Δ RMSE (1 Ma-4 Ma)
	RMSE (°C)	pCO_2 decrease (%)	RMSE (°C)	pCO_2 decrease (%)	RMSE (°C)	pCO_2 decrease (%)		
CESM_B	2.01	18.8	2.01	19.9	2.10	21.0	0.00	0.09
CESM_H	1.68	28.3	1.78	28.3	1.58	35.8	0.10	-0.10
GFDLCM 2.1	1.38	26.3	1.38	27.3	1.64	24.2	0.00	0.26
HadCM3BL	2.23	23.2	2.34	21.0	2.18	40.1	0.11	-0.05
FOAM	1.60	29.3	1.69	29.3	1.75	31.2	0.09	0.15
NorESM-L	2.34	10.5	2.44	5.4	2.55	16.5	0.10	0.21
Ensemble	**1.88**	**30.3**	**2.04**	**29.3**	**2.08**	**40.1**	**0.16**	**0.20**

Table S7. Time interval impact on proxy-model SST comparison. Model SST RMSE (°C) from adjusted time window used for the proxies. The time windows used in this comparison include 4, 2, and 1 Ma do not have a significant difference, the 4 Ma is selected for this study.

Model	4 Ma time averaging		2 Ma time averaging		1 Ma time averaging		ΔRMSE (2 Ma-4 Ma)	Δ RMSE (1 Ma-4 Ma)
	RMSE (°C)	pCO_2 decrease (%)	RMSE (°C)	pCO_2 decrease (%)	RMSE (°C)	pCO_2 decrease (%)		
CESM_B	0.90	33.1	0.80	26.3	1.93	50.7	-0.10	1.03
CESM_H	0.92	35.8	0.65	29.3	1.96	48.6	-0.27	1.04
GFDLCM 2.1	1.02	21.0	0.79	17.6	2.19	32.2	-0.23	1.17
HadCM3BL	1.23	46.4	0.52	44.1	1.41	69.6	-0.71	0.18
FOAM	1.07	27.3	0.67	23.2	2.05	42.6	-0.41	0.98
NorESM-L	1.45	43.4	0.78	39.3	1.28	58.2	-0.67	-0.17
Ensemble	**1.01**	**33.1**	**0.67**	**28.3**	**1.98**	**48.6**	**-0.34**	**0.97**

Table S8. Testing the effect of excluding single MAT proxy records (n=6 reduced to n=5) on a) RMSE (°C) and b) pCO_2 decrease (%) for the pCO_2 scaling.

	Excluding the following 1 record from the proxy set													
Model	All proxies (n=6)		Prydz Bay bayMBT		Prydz Bay S-index		CIROS-1 /CRP S-index		WW7 BayMBT		Site 696 NLR		Site 1172 NLR	
	a	b	a	b	a	b	a	b	a	b	a	b	a	b
CESM_B	1.98	19	1.40	16	1.96	19	2.14	20	2.01	13	2.24	18	2.08	22
CESM_H	1.68	28	1.54	18	1.95	24	2.11	26	1.77	23	1.76	32	1.47	35
GDFLCm 2.1	1.35	26	1.22	19	1.51	25	1.56	28	1.38	24	1.27	30	0.95	32
HadCM3BL	2.24	23	1.80	0	2.38	9	2.57	0	2.20	7	2.44	26	2.38	31
FOAM	1.58	29	1.51	18	1.87	25	2.02	26	1.69	24	1.79	30	1.51	34
NorESM-L	2.34	11	1.80	0	2.39	0	2.57	0	2.20	0	2.57	8	2.55	13
Ensemble	**1.88**	**30**	**1.66**	**16**	**2.11**	**26**	**2.32**	**25**	**1.94**	**24**	**1.99**	**36**	**1.76**	**40**

Table S9. Effect of removing individual SST proxy records (n=6 reduced to n=5) on a) RMSE (°C) and b) pCO_2 decrease (%) for the pCO_2 scaling.

	Excluding the following 1 record from the proxy set													
Model	All proxies (n=6)		Prydz Bay BAYSPAR		Site 689 Δ_{47}		Site 511 BAYSPAR		Site 277 BAYSPAR		Site 277 BAYSPLINE		Site 1172 BAYSPAR	
	a	b	a	b	a	b	a	b	a	b	a	b	a	b
CESM_B	0.90	33	1.00	31	0.99	34	0.96	30	0.96	30	1.01	31	0.53	38
CESM_H	0.92	36	0.88	34	0.80	37	0.88	32	0.88	32	0.96	33	0.77	38
GDFLCm 2.1	1.02	21	0.89	21	0.87	23	0.99	20	0.99	20	1.05	20	0.85	24
HadCM3BL	1.23	46	1.07	50	0.92	55	0.76	51	0.76	51	0.85	52	0.71	58
FOAM	1.07	27	1.18	25	1.01	29	0.94	26	0.94	26	1.02	26	0.91	31
NorESM-L	1.45	43	0.92	49	0.94	51	1.18	43	1.18	43	1.28	43	1.53	46
Ensemble	**1.01**	**33**	**0.93**	**33**	**0.83**	**36**	**0.90**	**32**	**0.90**	**32**	**0.97**	**32**	**0.79**	**38**

Table S10. Impact on removing all records of the same proxy type for MAT to evaluate individual proxy impacts on the pCO_2 scaling experiment.

Model	All proxies (n=6)		bayMBT (n=4)		S-index (n=4)		NLR (n=4)	
	RMSE (°C)	pCO_2 decrease (%)	RMSE (°C)	pCO_2 decrease (%)	RMSE (°C)	pCO_2 decrease (%)	RMSE (°C)	pCO_2 decrease (%)
CESM_B	1.98	18.8	1.14	19.7	2.18	21.0	1.98	34.0
CESM_H	1.68	28.3	1.20	43.0	1.83	26.3	0.78	47.9
GDFLCM 2.1	1.35	26.3	1.03	37.8	1.56	26.3	0.63	35.8
HadCM3BL	2.24	23.2	1.29	35.7	2.01	41.0	2.65	34.0
FOAM	1.58	29.3	1.15	38.3	1.76	27.3	0.57	47.9
NorESM-L	2.34	10.5	1.29	8.7	2.07	40.1	2.85	11.7
Ensemble	**1.88**	**30.3**	**1.28**	**53.6**	**1.93**	**30.3**	**1.41**	**54.6**

Table S11. Effect of removing proxy types for MAT to evaluate individual proxy impacts on the pCO_2 scaling experiment.

Model	All proxies (n=6)		BAYSPAR (n=3)		BAYSPLINE (n=5)		Δ_{47} (n=5)	
	RMSE (°C)	pCO_2 decrease (%)	RMSE (°C)	pCO_2 decrease (%)	RMSE (°C)	pCO_2 decrease (%)	RMSE (°C)	pCO_2 decrease (%)
CESM_B	0.90	33.1	0.53	31.2	0.97	31.2	0.94	34.9
CESM_H	0.92	35.8	0.86	31.2	0.93	34.0	0.87	39.3
GDFLCM 2.1	1.02	21.0	0.89	17.6	1.05	19.9	0.93	25.3
HadCM3BL	1.23	46.4	0.59	62.1	1.22	43.4	1.34	45.7
FOAM	1.07	27.3	0.72	31.2	1.08	25.3	1.16	28.3
NorESM-L	1.45	43.4	1.48	29.3	1.34	41.0	1.36	58.2
Ensemble	**1.01**	**33.1**	**0.86**	**31.2**	**1.02**	**31.2**	**1.02**	**36.7**

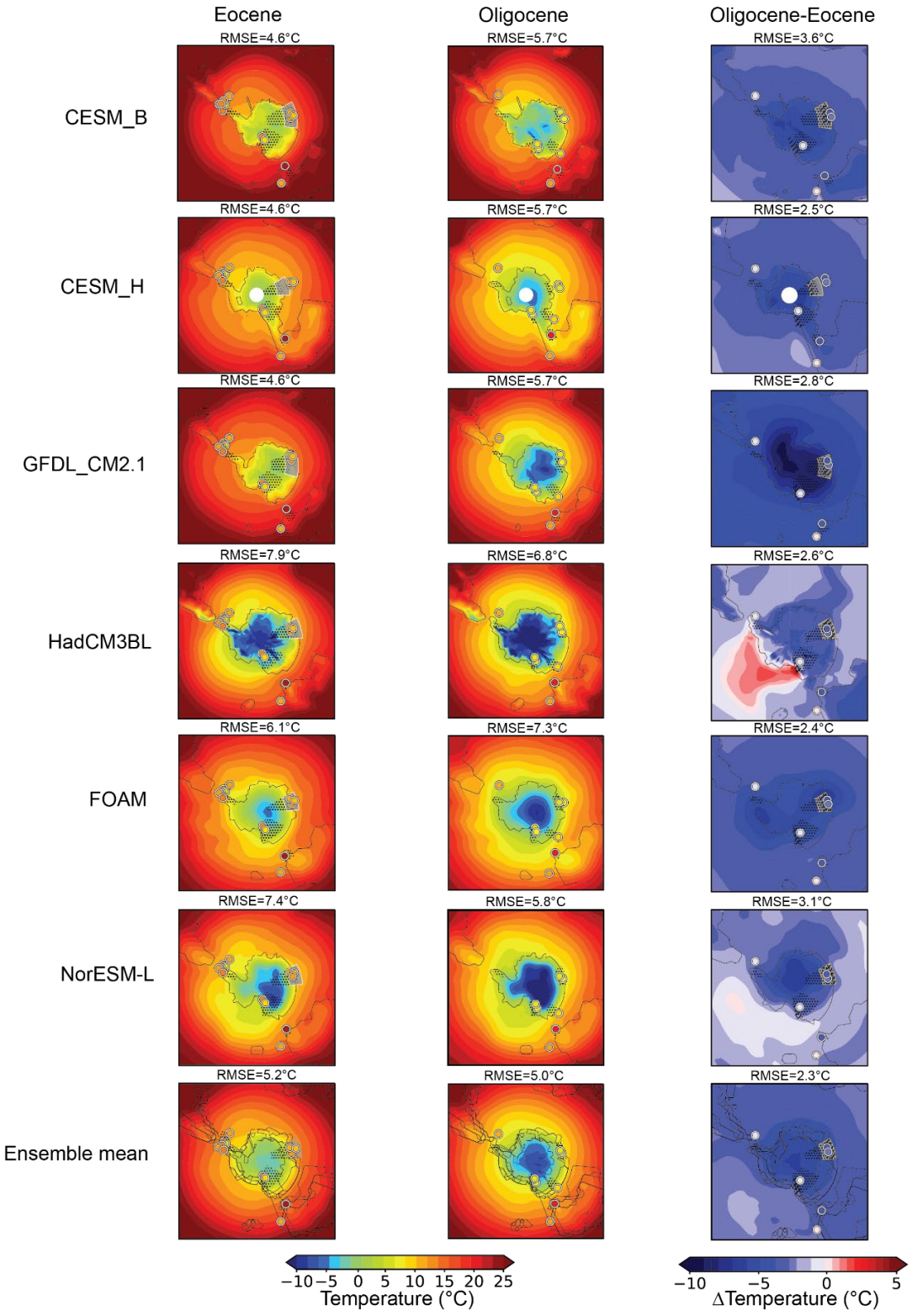

Figure S1. Southern hemisphere high latitude surface air temperatures for the Eocene (4x pCO_2 model runs), Oligocene (2x pCO_2), and the difference across the transition (2x-4x) for the various climate models and the ensemble mean. The circles correspond to proxy mean annual air temperature records

while the dotted areas show the source area used to compare the model temperature to the proxy record. The grey area in the Eocene timeslice identifies the separate source regions used in Prydz Bay based on the proxy type – a larger catchment is used for the rock weathering proxy (S-index) which almost certainly has more high-altitude erosional influence, whereas a smaller catchment is used for the soil biomarker proxy (brGDGTs) as these are presumably dominated by production in lower altitudes (Tibbett et al., 2021). In the Oligocene run the catchment is restricted for Prydz Bay and the Ross Sea on the basis of ice expansion reducing the effective catchment area. In the Oligocene-Eocene panel the grey panel represent the Eocene area used for Prydz Bay while the line for source region corresponds to proxy sourcing areas for the Oligocene where an ice sheet may have reached the coast.

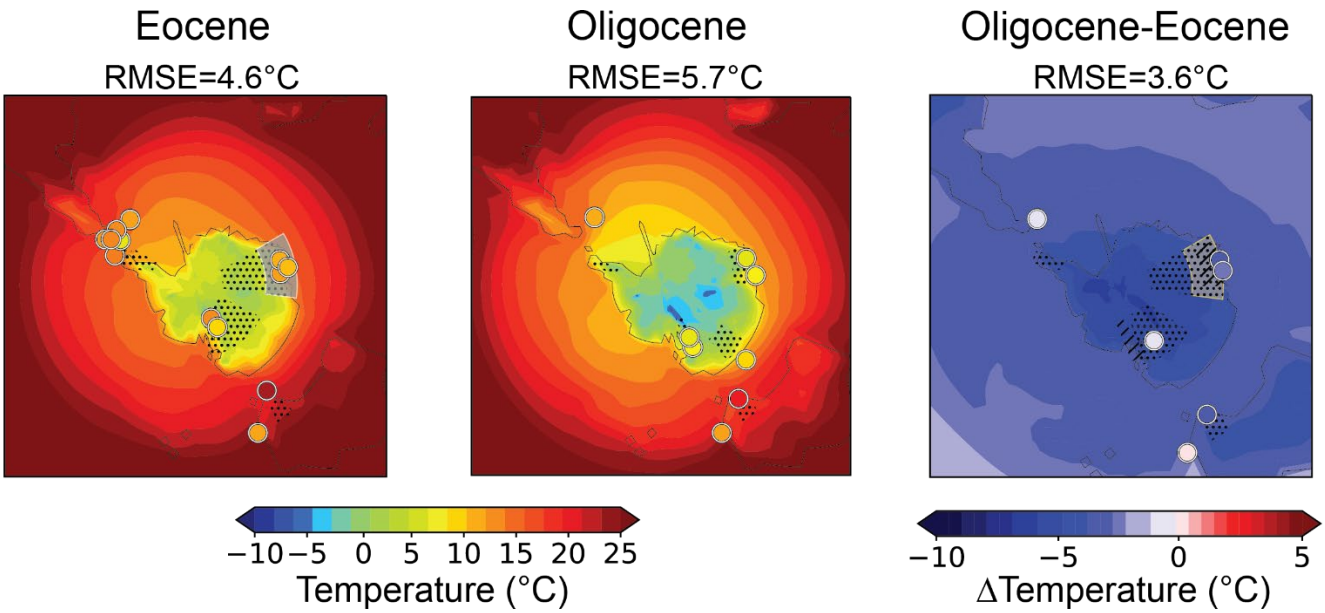

Figure S2. Zoom in from Figure S1 CESM_H simulations for source region comparisons. The grey area in the Eocene timeslice identifies the separate source regions used in Prydz Bay based on the proxy type – a larger catchment is used for the rock weathering proxy (S-index) which almost certainly has more high-altitude erosional influence, whereas a smaller catchment is used for the soil biomarker proxy (brGDGTs) as these are presumably dominated by production in lower altitudes (Tibbett et al., 2021). In the Oligocene run the catchment is restricted for Prydz Bay and the Ross Sea on the basis of ice expansion reducing the effective catchment area. In the Oligocene-Eocene panel the grey panel represents the Eocene area used for Prydz Bay while the line for source region corresponds to proxy sourcing areas for the Oligocene where an ice sheet may have reached the coast.

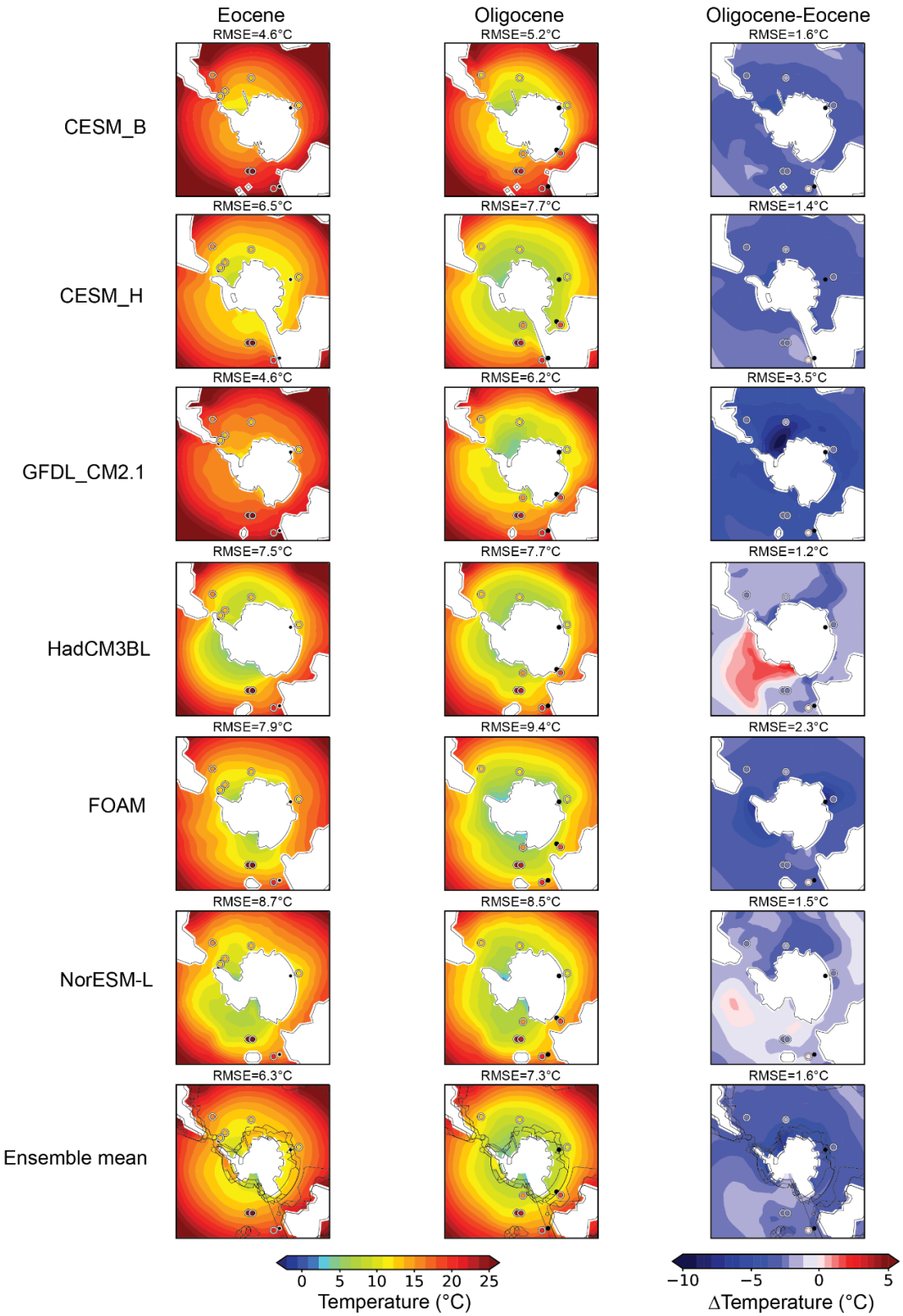

Figure S3. Modelled Southern Ocean SSTs for the Eocene (4x pCO_2 model runs), Oligocene (2x pCO_2), and the difference across the transition (2x-4x). The colored circles show proxy SSTs estimates, when these marine core sites appear to plot "on land", this is due to imprecision in modelled coastlines, and a black circle denotes the nearest marine location in the model used for proxy-model comparison.

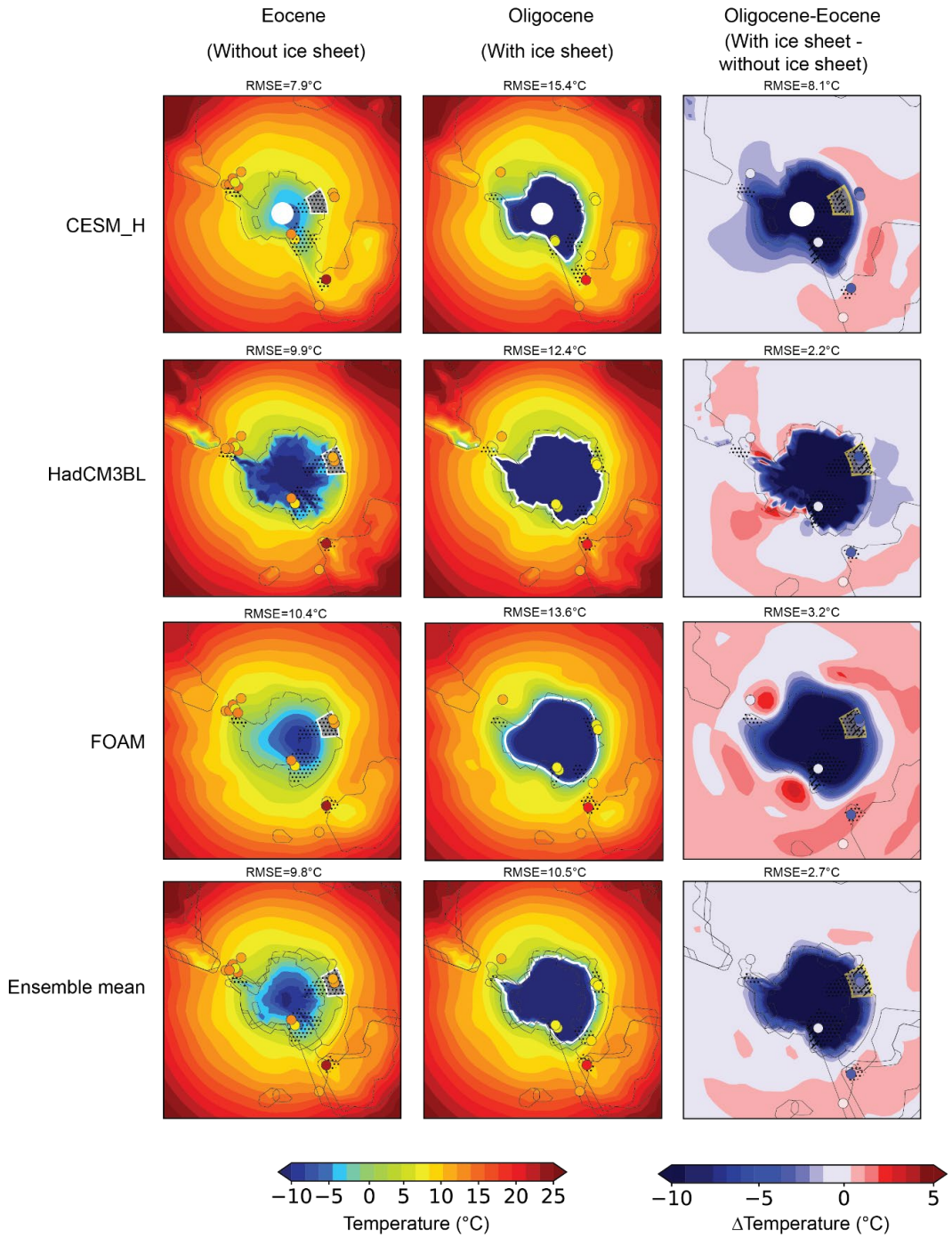

Figure S4. Climate model reconstructions for the Eocene (without ice), Oligocene (with ice) and Oligocene-Eocene showing the modelled surface air temperature (MAT) change associated with EOT glaciation. Circles show proxy evidence for MAT for comparison. White outline denotes -5°C to indicate ice sheet extent.

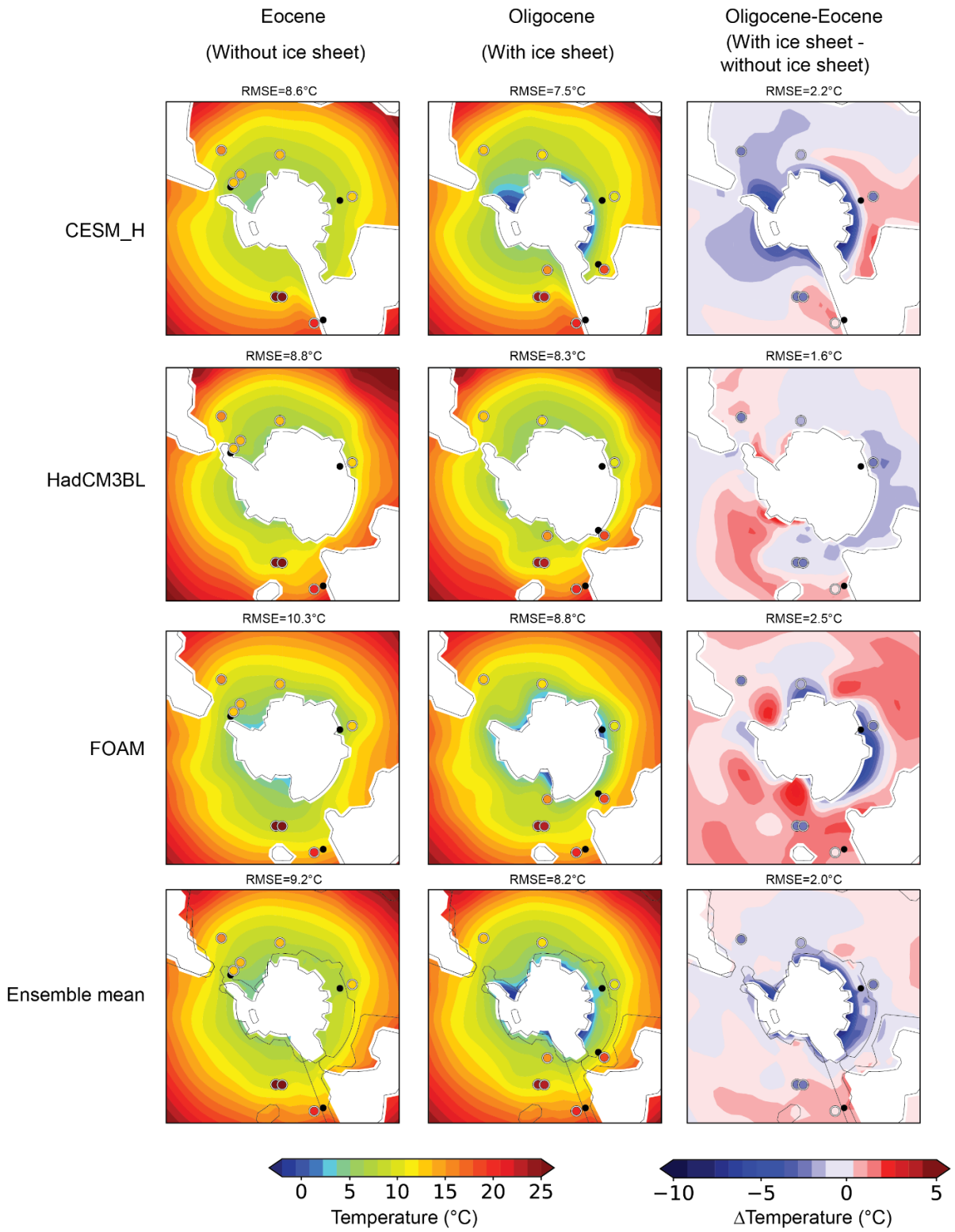

Figure S5. Climate model reconstructions for the Eocene (without ice), Oligocene (with ice) and Oligocene-Eocene showing the modelled sea surface temperature (SST) change associated with EOT glaciation. Circles show proxy evidence for SSTs for comparison.

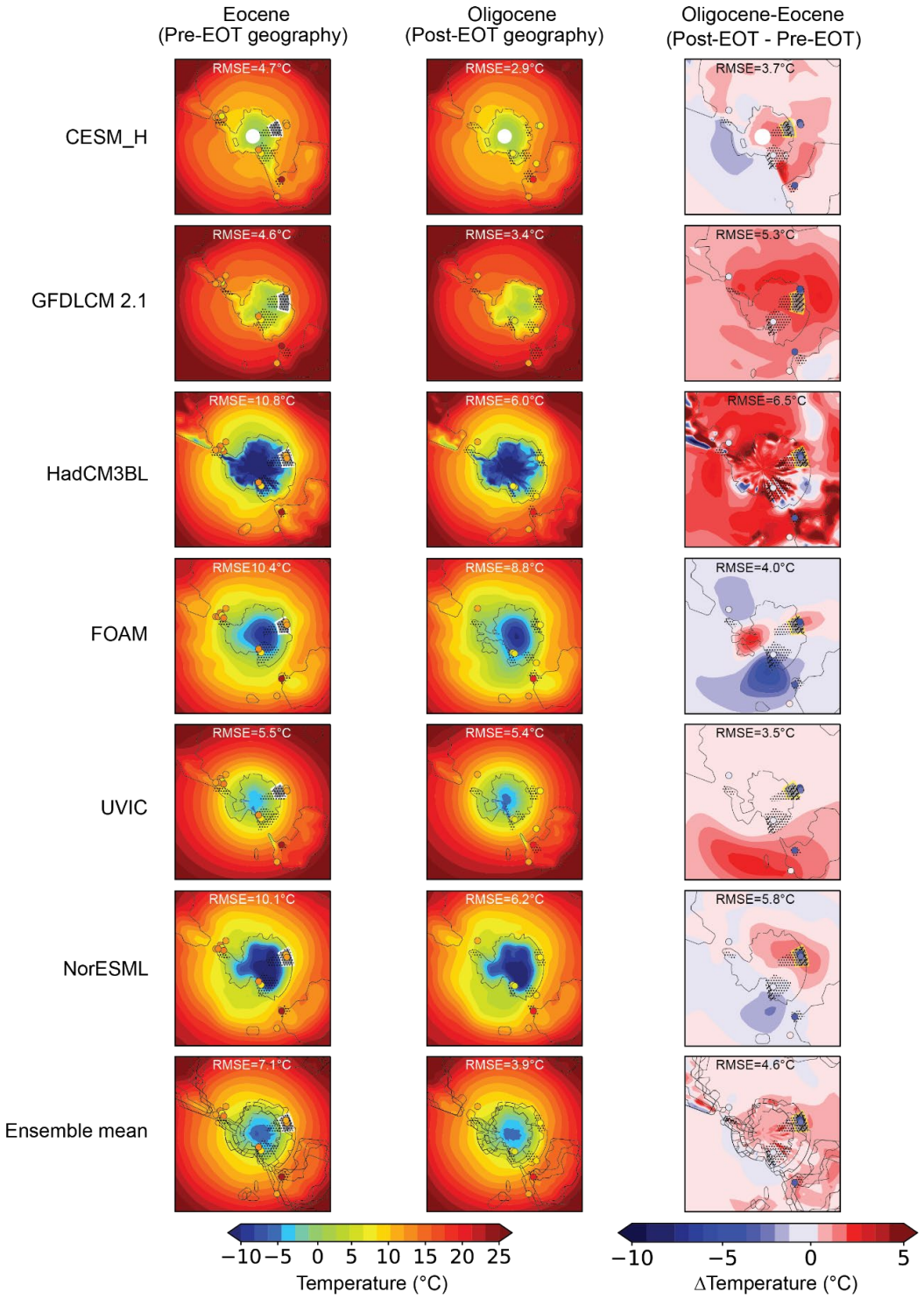

Figure S6. Climate model MAT reconstructions of Eocene, Oligocene and Oligocene-Eocene difference associated with changes in geography. CESM_H contrasts Tasman Gateway closed in the Eocene and open in the Oligocene. FOAM compares the geography of West Antarctica, being above sea level (Eocene) and below sea level (Oligocene). The UVIC experiments compare a closed (Eocene) and open (Oligocene) Drake Passage. Circles show surface air temperature values reconstructed from proxies.

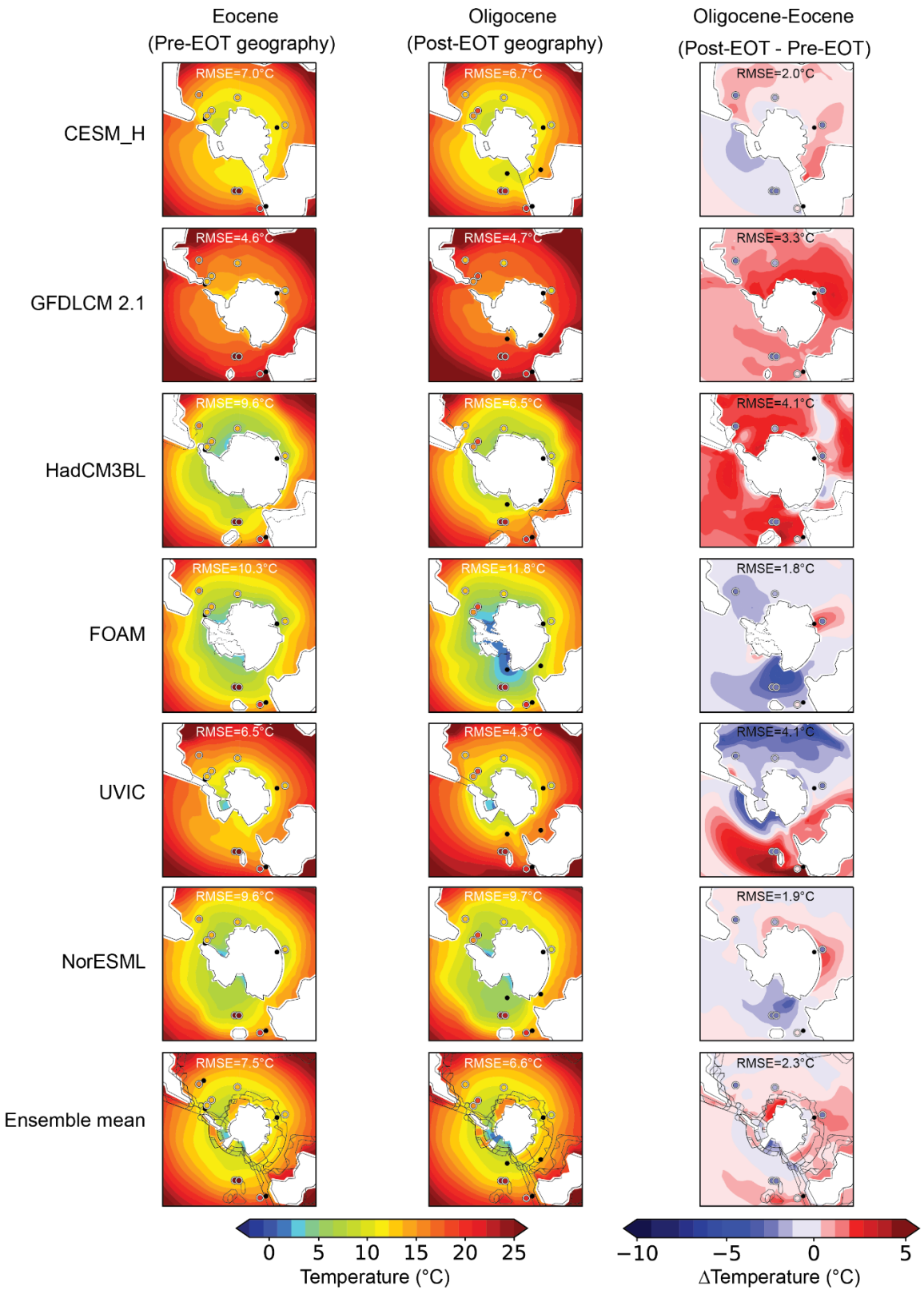

Figure S7. SST comparison between Eocene and Oligocene climate model scenarios with different paleogeographies, and proxy-model comparison. Experiments, as in Figure S3.

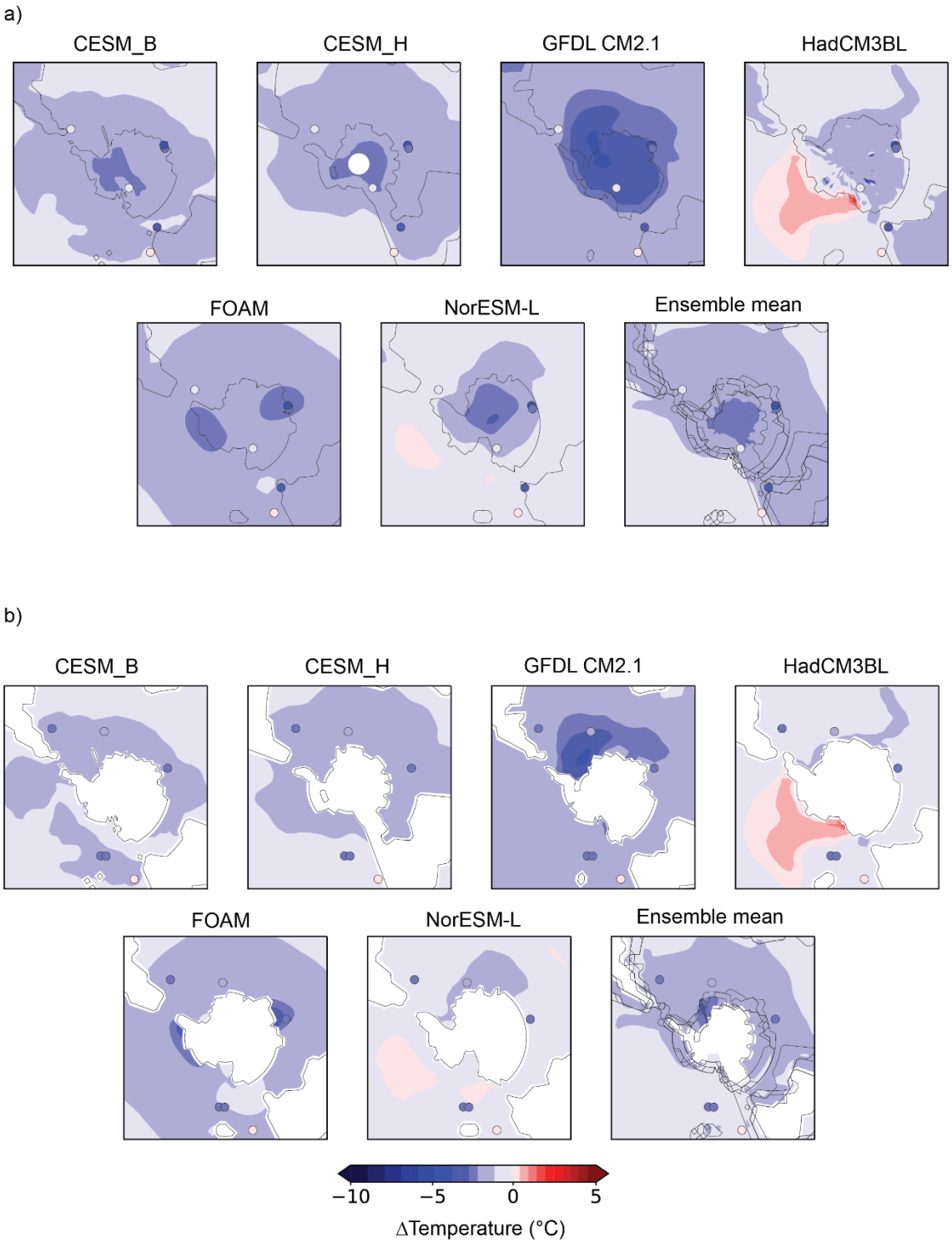

Figure S8. Proxy-model comparison for models scaled to a 25% reduction in pCO_2 for Oligocene-Eocene scenarios showing a) MAT b) SST. The colored circles show proxy SSTs estimates, when these marine core sites appear to plot "on land", this is due to imprecision in modelled coastlines, and a black circle denotes the nearest marine location in the model used for proxy-model comparison.

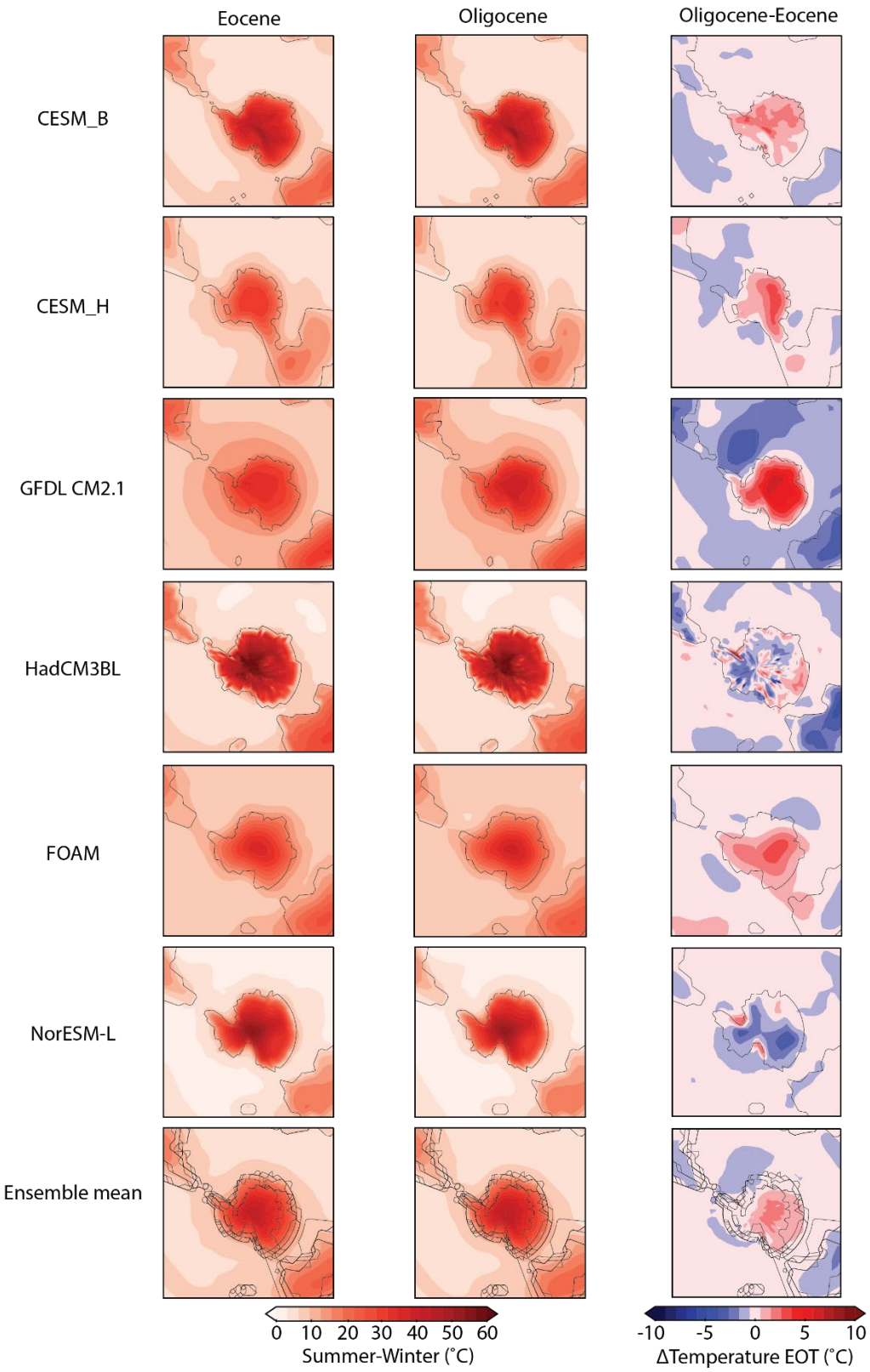

Figure S9. Model seasonality (summer-winter) for the Eocene and Oligocene timeslices and Oligocene-Eocene difference.

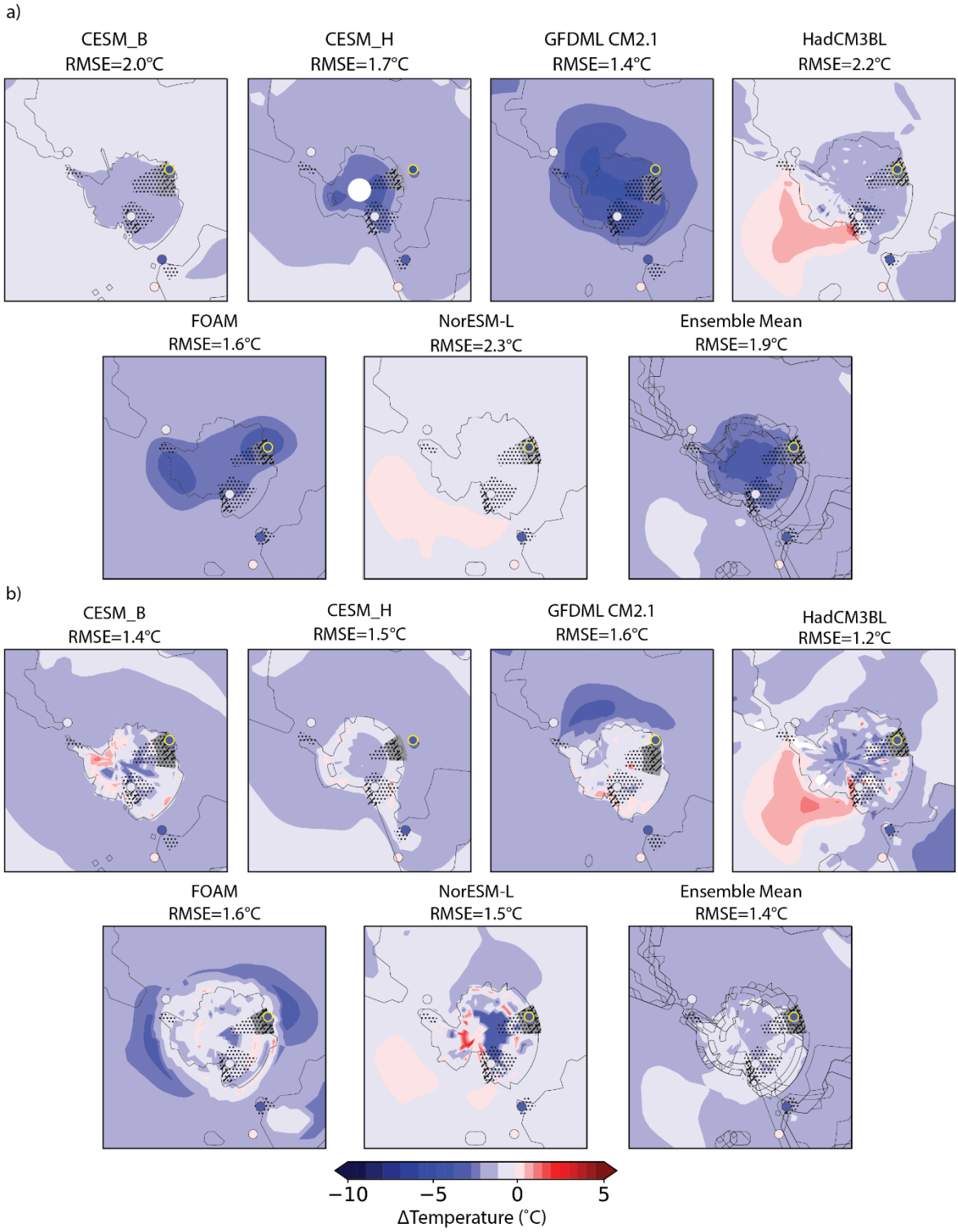

Figure S10. Best fit for pCO_2 across all models for a) MAT and b) MAF with RMSE for each model and the ensemble mean. The yellow outline marks the proxy site that was recalibrated to MAF.

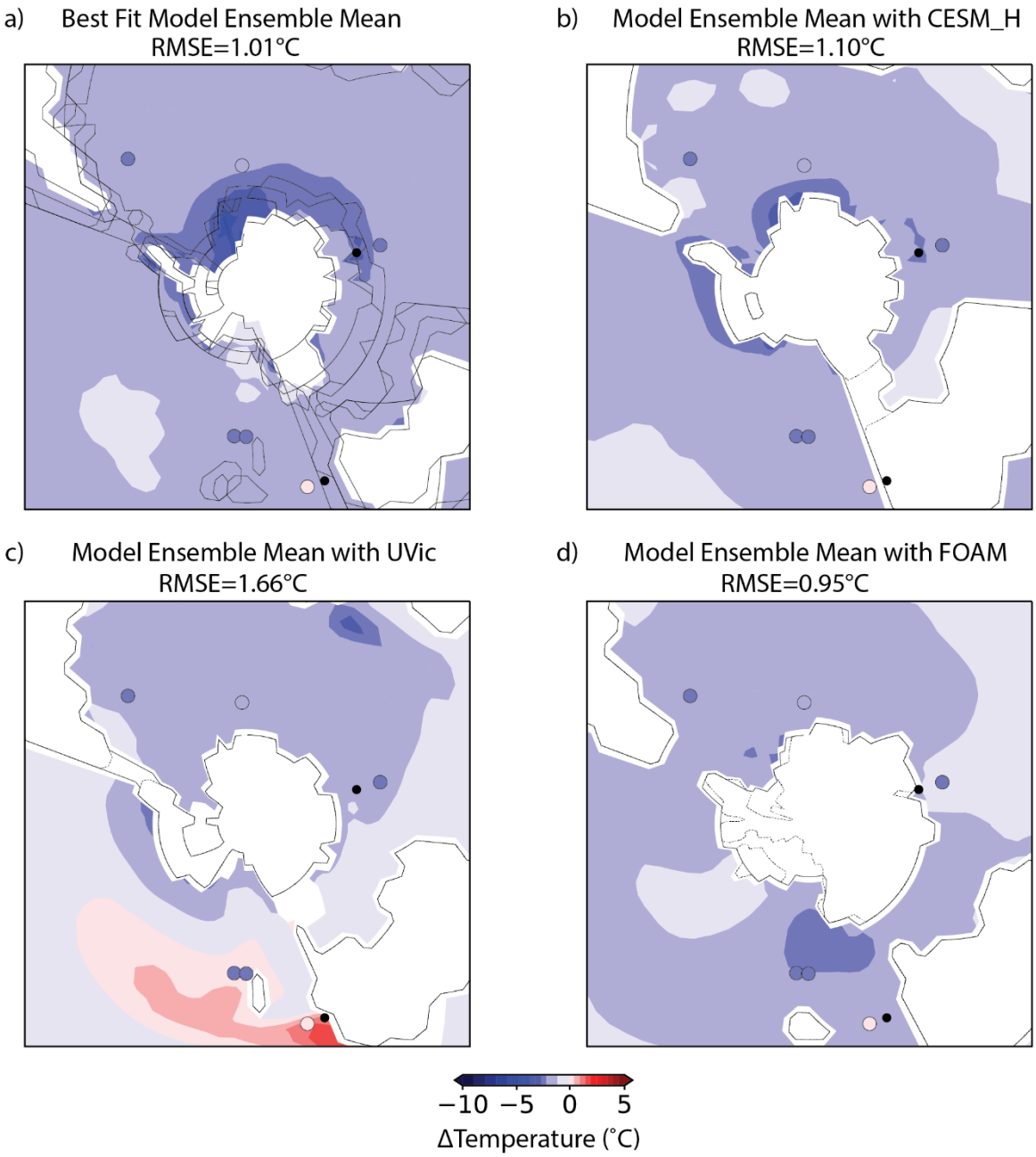

Figure S11. a) pCO_2 best fit model ensemble mean for Oligocene-Eocene b) pCO_2 from panel a + CESM_H ΔEOT paleogeography c) pCO_2 from panel a + pCO_2 from panel a + ΔEOT UVic paleogeography d) pCO_2 + FOAM ΔEOT paleogeography.

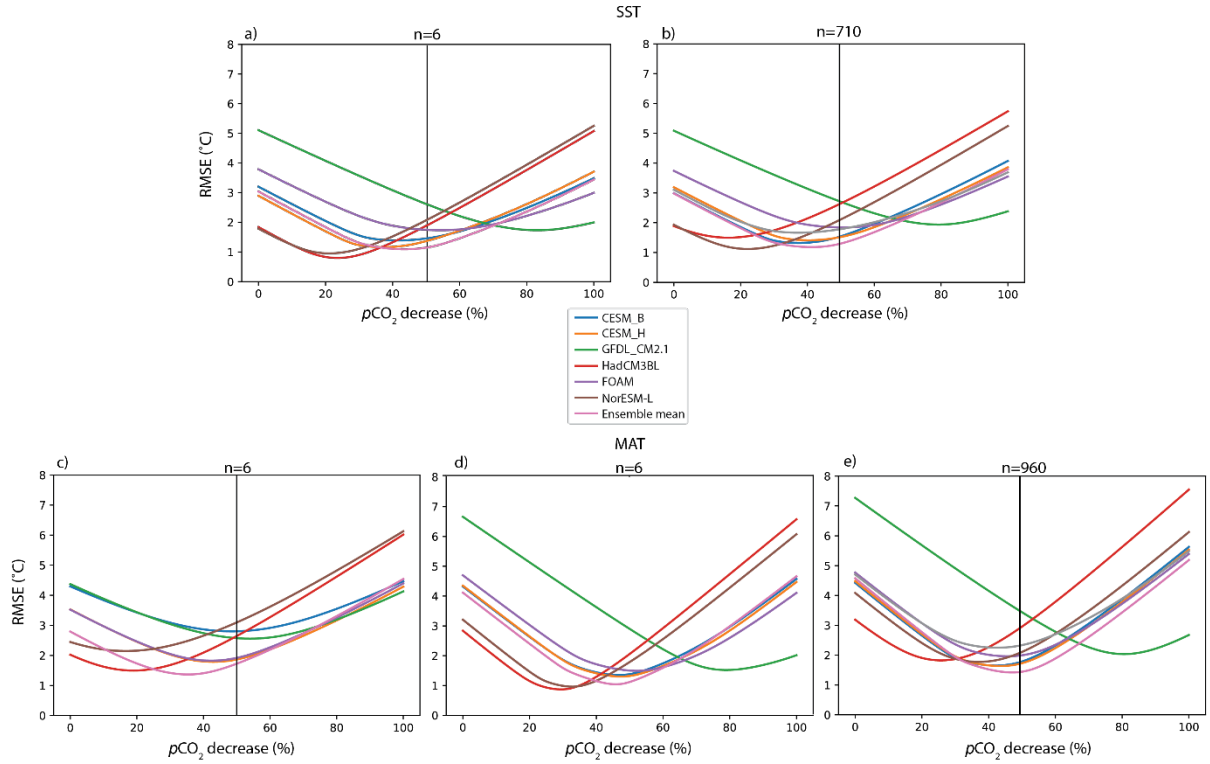

Figure S12. The "perfect model approach" assessment of inter-model temperature sensitivity to pCO_2 scaling. For each iteration one model or the ensemble mean (see legend) is used as the true temperature values ("perfect model" approach) with the additional models scaled to the "perfect mode" to assess how well the models replicate the 50% decrease in pCO_2. We show how the RMSE based on the remaining models changes across scaling factors for a) SSTs sampled at the proxy locations in the perfect model (n=6) and b) all possible marine grid cells within set parameters (n=710); c) MATs sampled at the proxy locations within the perfect model (n=6) using the adjusted source regions for the Eocene and Oligocene to account for ice sheet extent during the Oligocene, d) proxy location using the same source are for both the Eocene and Oligocene timeslices and e) then with all possible land grid cells (n=960). The black line marks the 50% decrease in pCO_2 which is the expected lowest RMSE given that a 50% decrease in pCO_2 was in fact imposed in the respective "perfect model" datasets. The RMSE does indeed minimize near 50% when the ensemble mean is used as the "prefect model" with only a slight improvement when more locations are added, indicating that the current sample size of proxy location should be adequate. The range RMSE minima in pCO_2 % decrease seen when using a given model as the "perfect model" is due to model differences in their regional climate sensitivity to the given pCO_2 change.