S1 More detailed information for the individual CMIP6 models

Model	$MAAT$ (°C)	area $MAAT < 0°C$ (10^6km^2)	$PF_{benchmark}$ (10^6km^2)	$PF_{benchmark}$ / area $MAAT < 0°C$	$MAGT$ (°C)	median ALT (m)	$S_{depth,eff}$ (m)
BCC-CSM2-MR	-7.4	28.2	18.7	0.66	-0.7	1.68	0.42
BCC-ESM1	-7.6	26.7	17.7	0.66	-0.2	1.48	0.38
CESM2	-5.1	20.9	12.0	0.58	-2.1	1.08	0.30
CESM2-WACCM	-5.6	21.8	13.0	0.59	-2.6	1.02	0.30
CNRM-CM6-1	-6.8	23.8	14.8	0.62	-4.2	1.54	0.31
CNRM-ESM2-1	-6.1	21.6	13.3	0.62	-3.5	1.60	0.30
CanESM5	-5.4	21.6	13.1	0.61	0.5	3.43	0.33
EC-Earth3	-7.6	25.5	17.1	0.67	-5.1	1.70	0.41
GFDL-CM4	-7.8	25.7	16.6	0.65	-5.1	0.76	0.26
GISS-E2-1-G	-8.2	25.0	17.0	0.68	-5.6	0.65	0.38
GISS-E2-1-H	-6.0	21.9	14.1	0.65	-3.8	1.10	0.36
HadGEM3-GC31-LL	-7.8	25.8	17.0	0.66	-0.3	1.92	0.26
IPSL-CM6A-LR	-6.6	25.3	15.9	0.63	-0.7	4.49	0.38
MIROC6	-4.7	21.9	13.4	0.61	-0.9	1.45	0.23
MIROC-ES2L	-4.7	21.3	13.4	0.63	-2.5	1.17	-
MPI-ESM1-2-HR	-5.0	21.9	12.4	0.57	-3.8	3.87	-
MRI-ESM2-0	-4.9	21.8	12.5	0.57	-0.1	1.02	0.27
NorESM2-LM	-5.0	19.5	11.0	0.56	-2.0	0.94	0.29
UKESM1-0-LL	-8.6	26.6	17.2	0.65	-0.9	1.83	0.29
Observations	-6.8	24.4	15.1	0.62	-2.7	-	0.25

Table S1.1. Climate and land surface characteristics of the individual CMIP6 models. $PF_{benchmark}$ is the permafrost extent derived using the Chadburn et al. (2017) relationship and the model $MAAT$. These are climatologies for the period 1995–2014. The $MAAT$, $MAGT$ and $S_{depth,eff}$ are for the PF_{aff} region defined by the CCI-PF data. ALT is only calculated for the grid cells where the model simulates permafrost. The observations of $MAAT$ are for the same period (1986-2005) but the $MAGT$ from the CCI-PF is only available as the mean for 2000-2016 and the $S_{depth,eff}$ is for the period 1998-2016.
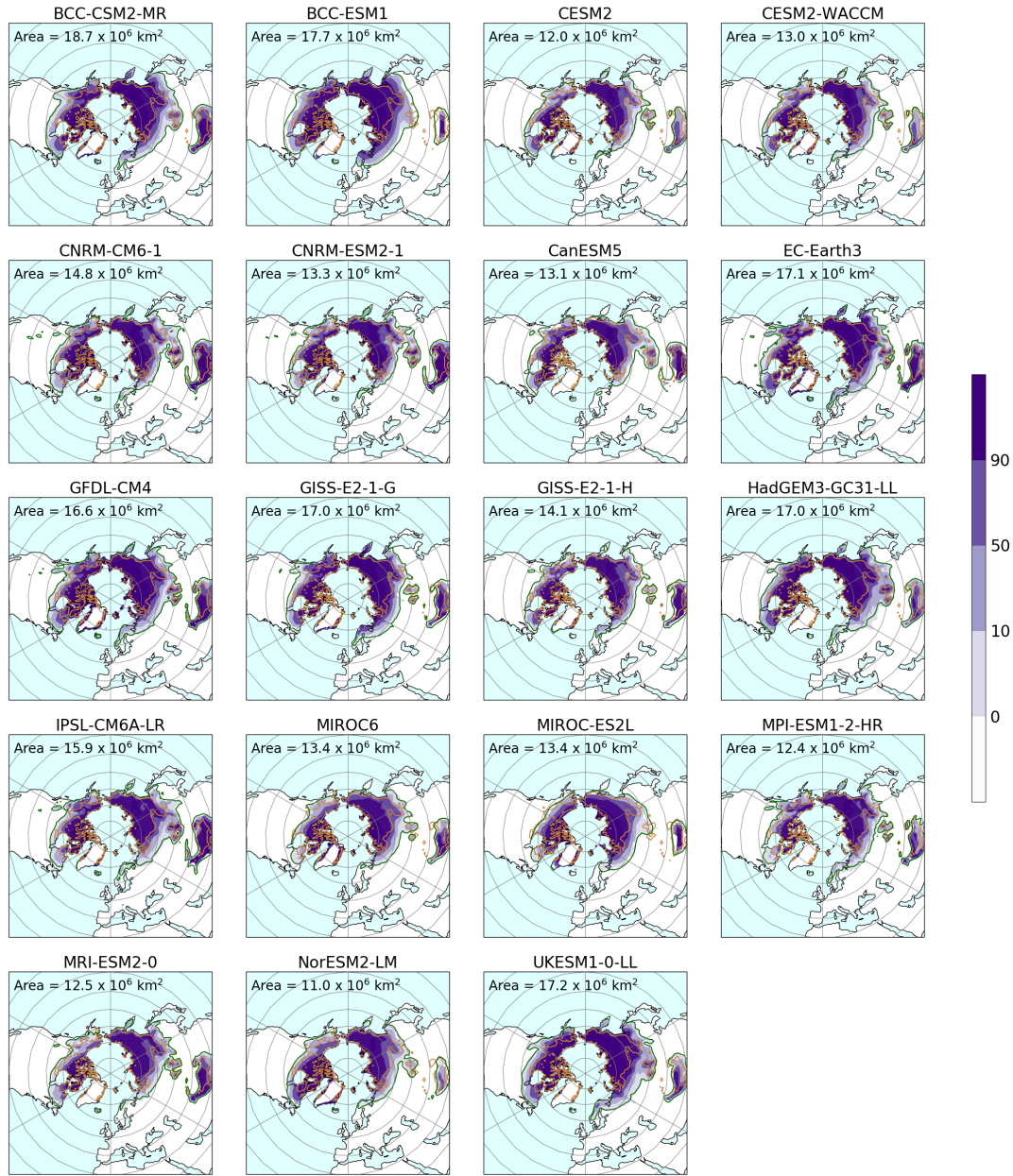
Figure S1.1. The probability of permafrost for the individual CMIP6 models derived using the mean $MAAT$ from each model for the period 1995–2014 and the Chadburn et al. (2017) relationship. This is denoted $PF_{benchmark}$ in the text. Any region where there is permafrost using this definition is shaded in purple with the continuous permafrost shaded dark purple. The orange line shows the threshold for the 50% probability of permafrost derived from the CCI-PF observations re-gridded to the same resolution as UKESM1-0-LL (2.5° latitude x 3.75° longitude). The green line show the threshold for $MAAT < 0°C$.
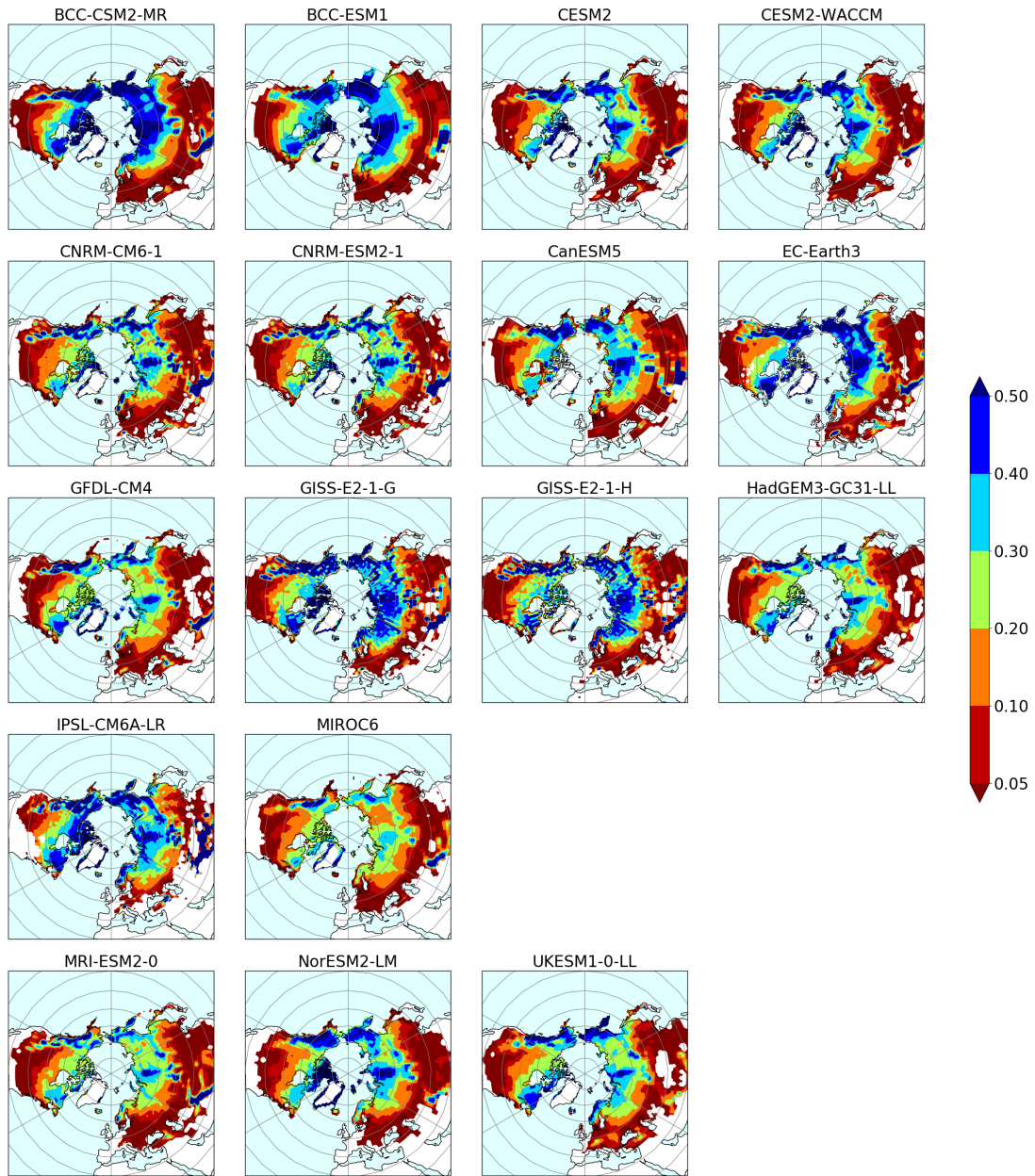
Figure S1.2. The climatology of the effective snow depth ($S_{depth,eff}$) for the period 1995–2014 for each individual CMIP6 model with snow depth data available on the archive. All grid cells with $S_{depth,eff}$ less than 2 cm are masked.
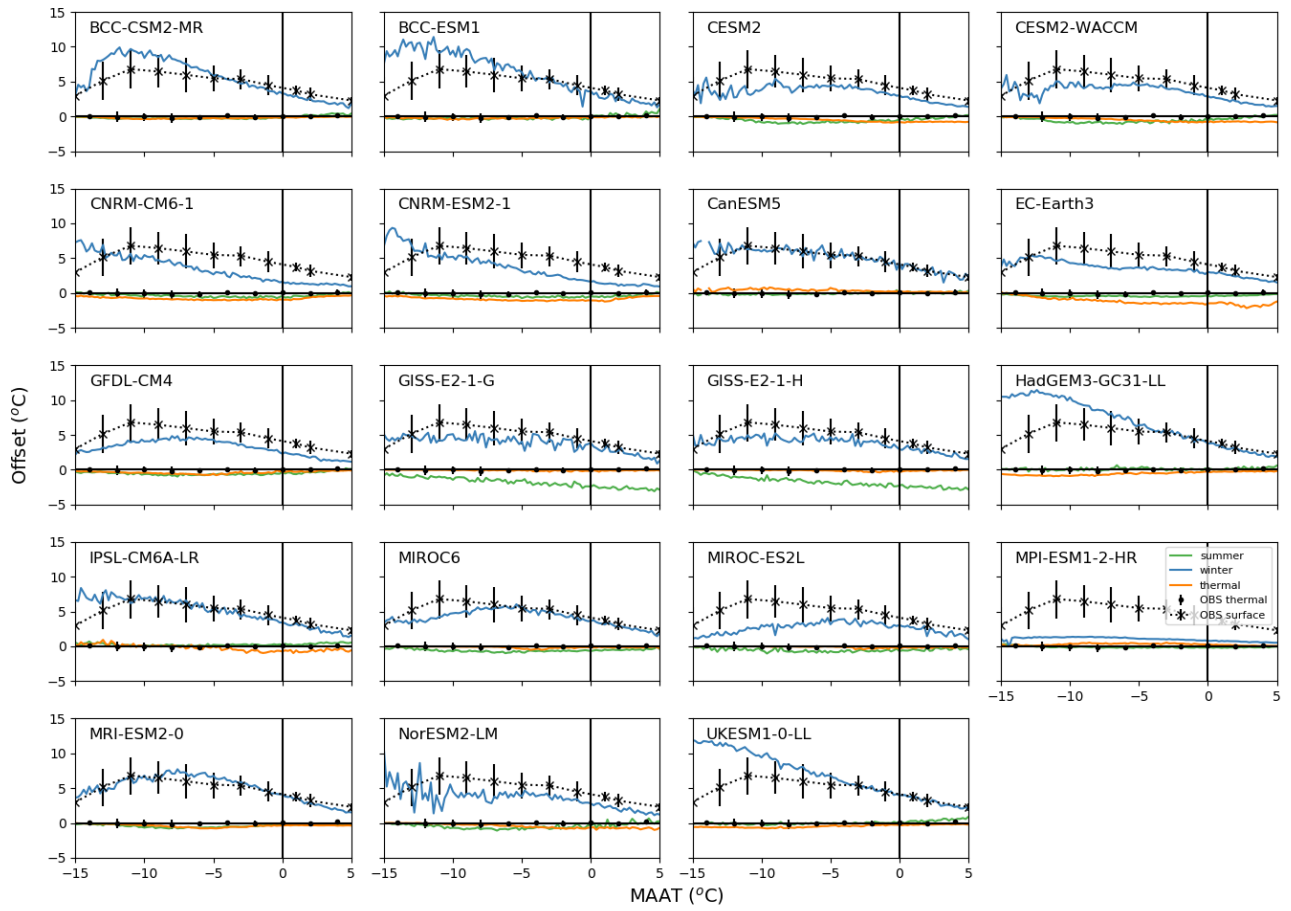
Figure S1.3. The winter, summer and thermal offsets for the CMIP6 models as a function of $MAAT$ for the climatology period of 1995–2014. These offset data are binned into $0.5°C$ bins and the median value of each offset taken. The observed surface and thermal offsets are added in black.
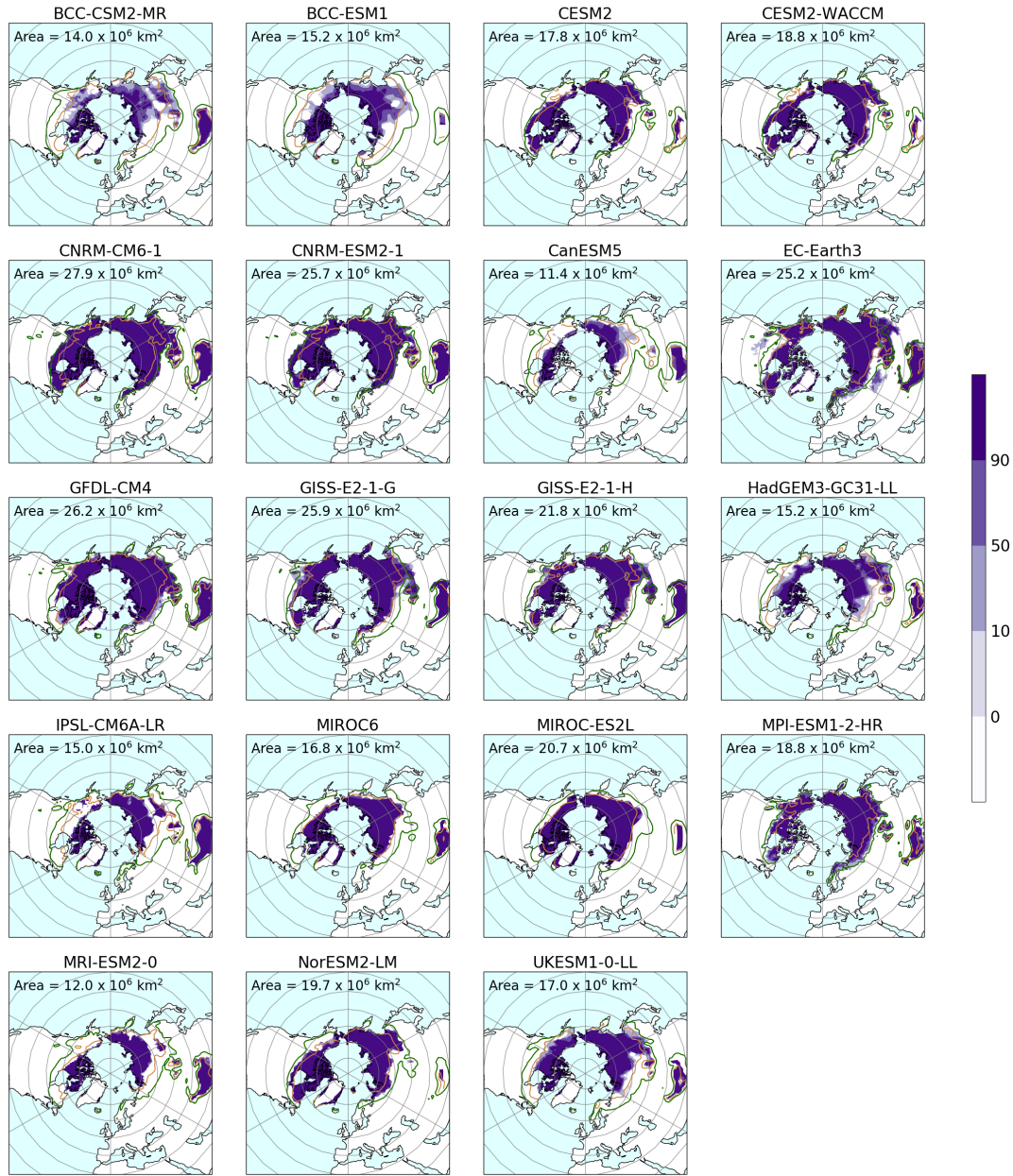
Figure S1.4. Permafrost extent (PF_{ex}) derived using the temperature at D_{zaa} or the lowest model level if the soil profile is too shallow. This is for the CMIP6 models for the period 1995–2014. Each model grid cell has either a 0% or 100% chance of finding permafrost in any particular year and the figure shows the mean for 20 years. Any region where there is permafrost is shaded in purple. Superimposed as an orange line on each plot is the 50% chance of finding permafrost using the model specific $PF_{benchmark}$ derived from the Chadburn et al. (2017) observed relationship. The green lines show the threshold where $MAAT < 0^{\circ}$C

Model	PF_{ex}/area $MAAT<0°C$	$\tilde{D}_{tot}$/area $MAAT<0°C$ (m)	snow off. (°C)	veg. off. (°C)	surface off. (°C)	thermal off. (°C)	$MAGT/MAAT$ (R²)	ALT (m; -12 < $MAAT$ <-10°C)	ALT (m; -6 < $MAAT$ <-4°C)
BCC-CSM2-MR	0.14	**0.22**	7.3	-0.3	**7.0**	-0.3	0.62 (46)	2.1	2.9
BCC-ESM1	0.10	**0.23**	7.7	-0.2	7.5	-0.3	0.53 (51)	2.2	2.9
CESM2	**0.72**	0.36	4.2	-0.7	3.5	-0.4	**0.95 (64)**	**0.6**	**1.7**
CESM2-WACCM	**0.72**	0.34	4.3	-0.7	3.6	-0.4	0.97 (66)	0.7	**1.7**
CNRM-CM6-1	0.96	0.62	4.1	-0.4	3.7	-0.8	0.55 (50)	1.3	**1.9**
CNRM-ESM2-1	0.97	0.65	4.4	-0.3	4.0	-0.9	0.50 (43)	1.3	2.1
CanESM5	0.08	0.10	5.8	-0.1	**5.7**	0.5	**0.87 (84)**	4.0	4.1
EC-Earth3	0.48	0.76	4.1	-0.4	3.7	-1.1	0.70 (72)	1.6	**1.8**
GFDL-CM4	0.77	**0.25**	4.0	-0.6	3.4	-0.5	1.01 (89)	0.7	**1.2**
GISS-E2-1-G	0.48	**0.23**	4.2	-1.5	2.7	**-0.1**	0.81 (80)	**0.6**	**1.6**
GISS-E2-1-H	0.32	0.28	4.2	-1.5	2.7	**-0.1**	0.83 (81)	0.7	2.6
HadGEM3-GC31-LL	0.47	0.47	8.1	0.1	8.2	-0.7	0.47 (55)	2.0	2.0
IPSL-CM6A-LR	0.49	0.47	5.9	0.2	**6.1**	**-0.1**	0.64 (59)	3.6	5.9
MIROC6	**0.66**	**0.24**	4.9	-0.7	**4.2**	**-0.1**	1.08 (90)	1.2	2.2
MIROC-ES2L	0.77	0.30	3.0	-0.6	2.3	**-0.1**	1.13 (97)	0.9	**1.5**
MPI-ESM1-2-HR	**0.72**	0.65	1.2	-0.1	1.1	0.4	0.98 (97)	2.0	4.5
MRI-ESM2-0	0.45	0.17	6.4	-0.6	**5.8**	-0.5	**0.94 (77)**	0.8	8.5
NorESM2-LM	**0.69**	0.35	4.1	-0.7	3.4	-0.4	**0.87 (56)**	**0.5**	**1.5**
UKESM1-0-LL	0.52	0.42	7.8	-0.1	7.7	-0.6	0.42 (58)	2.0	2.0
Mean observations	0.62	0.23	-	-	5.75	0.03	0.91	0.42	1.15
Min. observations	0.55	0.20	-	-	4.2	-0.15	0.86	0.49	0.64
Max. observations	0.77	0.25	-	-	7.1	0.15	0.95	0.65	1.98

Table S1.2. Evaluation metrics for the CMIP6 land surface modules. All of the offsets are calculated for the $MAAT$ range between -14°C and -2°C. The values within the range of the observations are highlighted in bold.

S2 CMIP5 equivalent of plots and tables

This section assesses the CMIP5 multi-model ensemble and creates comparable tables and plots so that the CMIP5 and CMIP6 ensembles can be directly compared in detail.

Model	Citation	No. soil layers	Soil depth (m)	D_{zaa} (m)
bcc-csm1-1	BCC	10	2.9	-
ACCESS1-0	CSIRO-BOM	4	2.0	-
CCSM4	NCAR	15	35.2	21.4
CESM1-CAM5	NSF-DOE-NCAR	15	35.2	21.2
CMCC-CM	CMCC	5	4.8	-
CanESM2	CCCma	3	2.2	-
FGOALS-g2	LASG-IAP	10	2.9	-
GFDL-ESM2M	NOAA-GFDL	20	8.8	-
GISS-E2-H-CC	NASA-GISS	6	2.7	-
GISS-E2-R	NASA-GISS	6	2.7	-
HadGEM2-CC	MOHC	4	2.0	-
HadGEM2-ES	MOHC	4	2.0	-
IPSL-CM5A-MR	IPSL	7	3.9	-
MIROC-ESM	MIROC	6	9.0	-
MIROC5	MIROC	6	9.0	-
MPI-ESM-MR	MPI-M	5	7.0	-
MRI-CGCM3	MRI	14	8.5	-
MRI-ESM1	MRI	14	8.5	-
NorESM1-M	NCC	15	35.2	21.2

Table S2.1. A summary of the CMIP5 models used in this study including the number of soil layers and the depth of the middle of the bottom soil layer. Also show is D_{zaa} for the models where the difference in the annual maximum and minimum soil temperatures at the maximum soil depth is less than $0.1°C$..

Model	$MAAT$ (°C)	area $MAAT < 0$°C (10^6km^2)	$PF_{benchmark}$ (10^6km^2)	$PF_{benchmark}$ / area $MAAT < 0$°C	$MAGT$ (°C)	median ALT (m)	$S_{depth,eff}$ (m)
bcc-csm1-1	-6.6	25.3	15.4	0.61	0.8	1.58	0.39
ACCESS1-0	-7.6	24.5	15.9	0.65	-4.5	1.88	0.23
CCSM4	-7.3	23.1	15.2	0.66	-0.5	1.07	0.41
CESM1-CAM5	-9.0	25.5	17.6	0.69	-1.7	1.08	0.37
CMCC-CM	-8.7	29.0	20.1	0.69	-7.9	1.61	0.21
CanESM2	-5.0	21.9	13.5	0.62	-4.8	1.75	0.23
FGOALS-g2	-10.5	33.3	24.4	0.73	-1.4	0.99	0.41
GFDL-ESM2M	-5.9	24.7	13.9	0.56	-6.2	0.83	-
GISS-E2-H-CC	-6.8	23.8	15.3	0.64	-4.2	1.50	0.34
GISS-E2-R	-6.7	24.1	14.9	0.62	-4.0	1.54	0.33
HadGEM2-CC	-9.3	28.0	18.9	0.68	-7.0	1.16	-
HadGEM2-ES	-8.1	25.3	16.9	0.67	-5.8	1.31	-
IPSL-CM5A-MR	-5.4	24.5	14.4	0.59	-3.5	2.68	-
MIROC-ESM	-4.7	20.5	12.1	0.59	-2.3	1.35	0.27
MIROC5	-6.4	22.8	14.8	0.65	-4.7	1.42	0.24
MPI-ESM-MR	-5.9	23.5	14.4	0.61	-5.2	3.04	-
MRI-CGCM3	-7.7	28.2	18.7	0.67	-2.0	0.97	0.30
MRI-ESM1	-7.3	27.2	17.7	0.65	-1.7	1.00	0.30
NorESM1-M	-8.4	24.8	16.5	0.66	-2.0	0.79	0.40
Observations	-7.1	24.8	15.7	0.61	-2.7	-	0.25

Table S2.2. Climate and land surface characteristics of the individual CMIP6 models. $PF_{benchmark}$ is the permafrost extent derived using the Chadburn et al. (2017) relationship and the model $MAAT$. These are climatologies for the period 1995–2014. The $MAAT$, $MAGT$ and $S_{depth,eff}$ are for the PF_{aff} region defined by the CCI-PF data. The ALT is only calculated for the grid cells where the model simulates permafrost. The observations of $MAAT$ are for the same period (1986-2005) but the $MAGT$ from the CCI-PF is only available as the mean for 2000-2016 and the $S_{depth,eff}$ is for the available period 1998-2016.
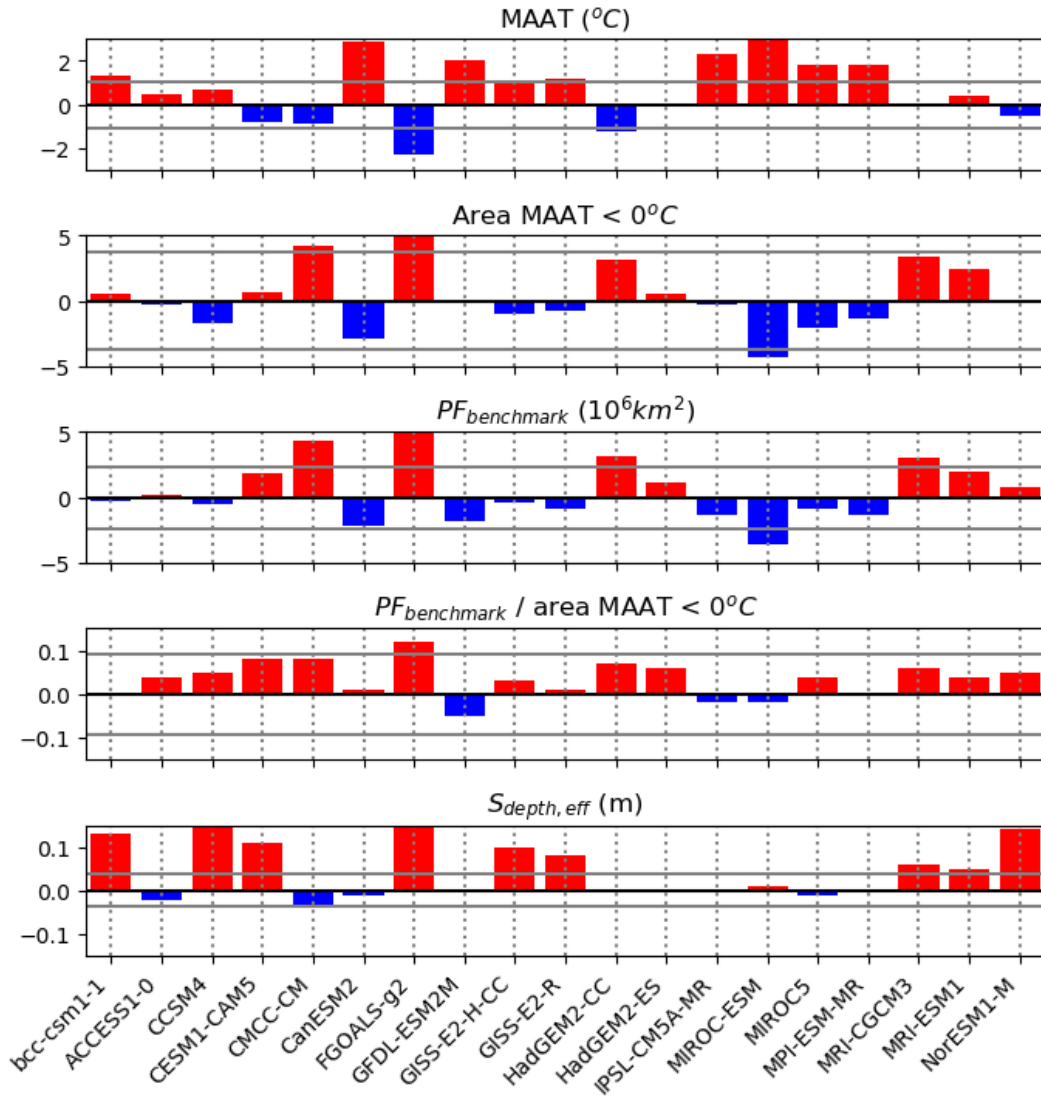
Figure S2.1. The climate characteristics of the CMIP5 multi-model ensemble compared with the observations for the period 1986-2005. The red bars are where the model value is greater than the observations and the blue bars are where the model value is less than the observations. The $MAGT$ is a single value representing the period 2000-2016 and the $S_{depth,eff}$ is for the period 1998-2016. $S_{depth,eff}$ is not available for every ensemble member.
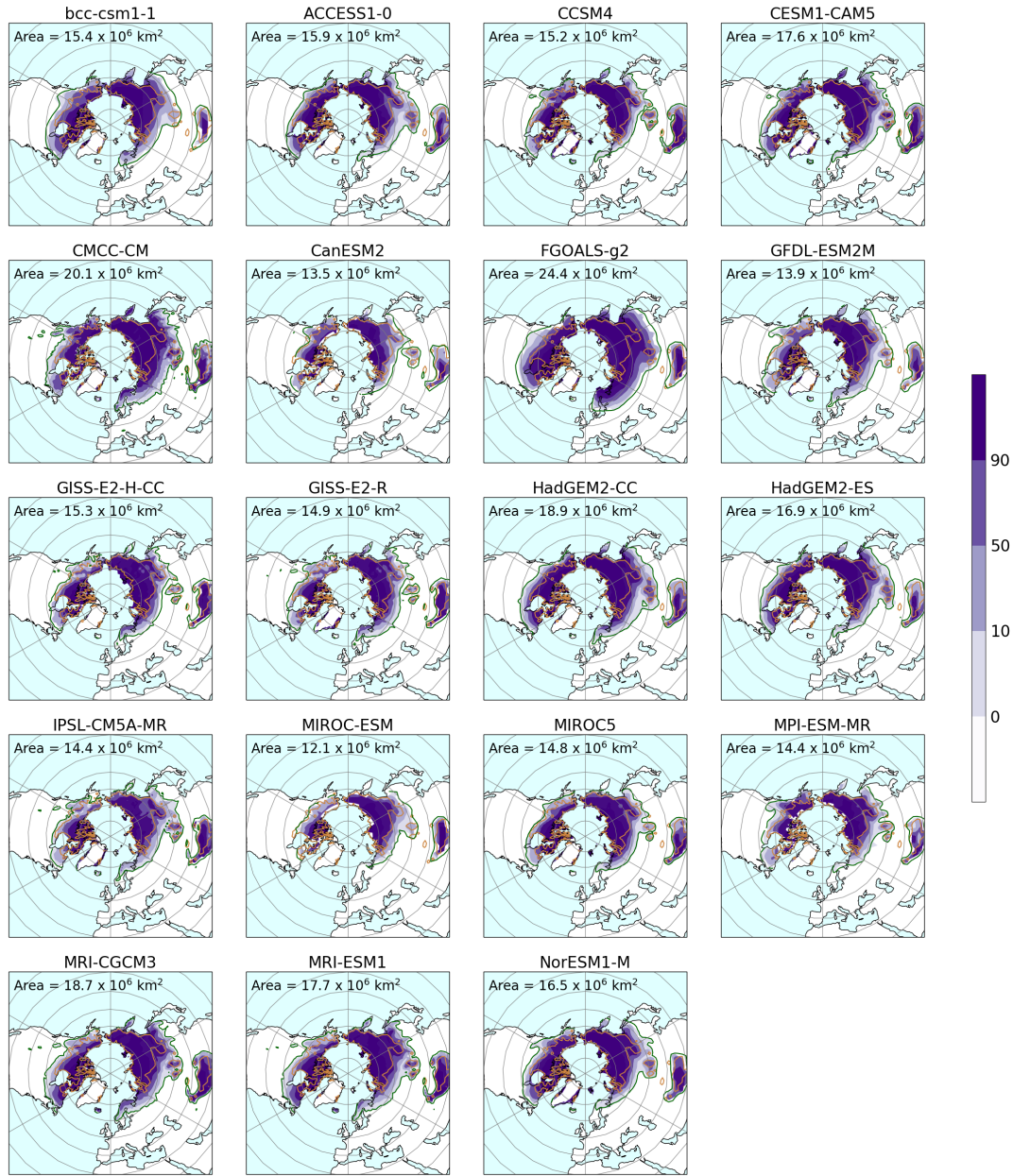
Figure S2.2. The probability of permafrost for the individual CMIP5 models derived using the mean $MAAT$ from each model for the period 1986-2005 and the Chadburn et al. (2017) relationship. This is denoted $PF_{benchmark}$ in the text. Any region where there is permafrost using this definition is shaded in purple with the continuous permafrost shaded dark purple. The orange line shows the threshold for the 50% probability of permafrost derived from the CCI-PF observations re-gridded to the same resolution as UKESM1-0-LL (2.5° latitude x 3.75° longitude). The green line show the threshold for $MAAT < 0°C$.

Figure S2.3. The climatology of the effective snow depth ($S_{depth,eff}$ in m) for the period 1986-2005 for each individual CMIP5 model. All grid cells with $S_{depth,eff}$ less than 2 cm are masked.
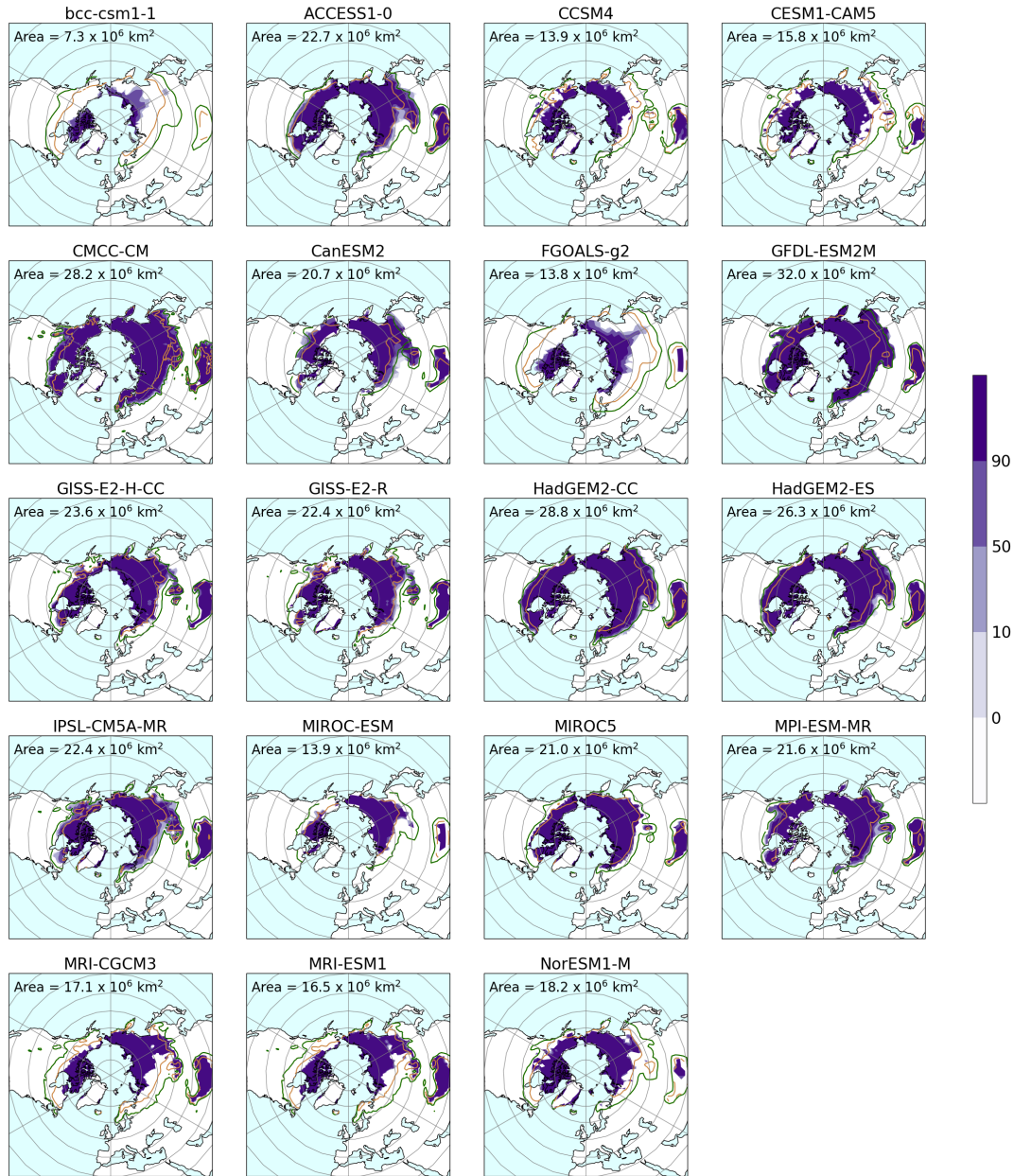
Figure S2.4. Permafrost extent (PF_{ex}) derived using the temperature at D_{zaa} or the lowest model level if the soil profile is too shallow. This is for the CMIP5 models for the period 1986-2005. Each model grid cell has either a 0% or 100% chance of finding permafrost in any particular year and the figure shows the mean for 20 years. Any region where there is permafrost is shaded in purple. Superimposed as an orange line on each plot is the 50% chance of finding permafrost using the model specific $PF_{benchmark}$ derived from the Chadburn et al. (2017) observed relationship. The green lines show the threshold where $MAAT < 0°C$
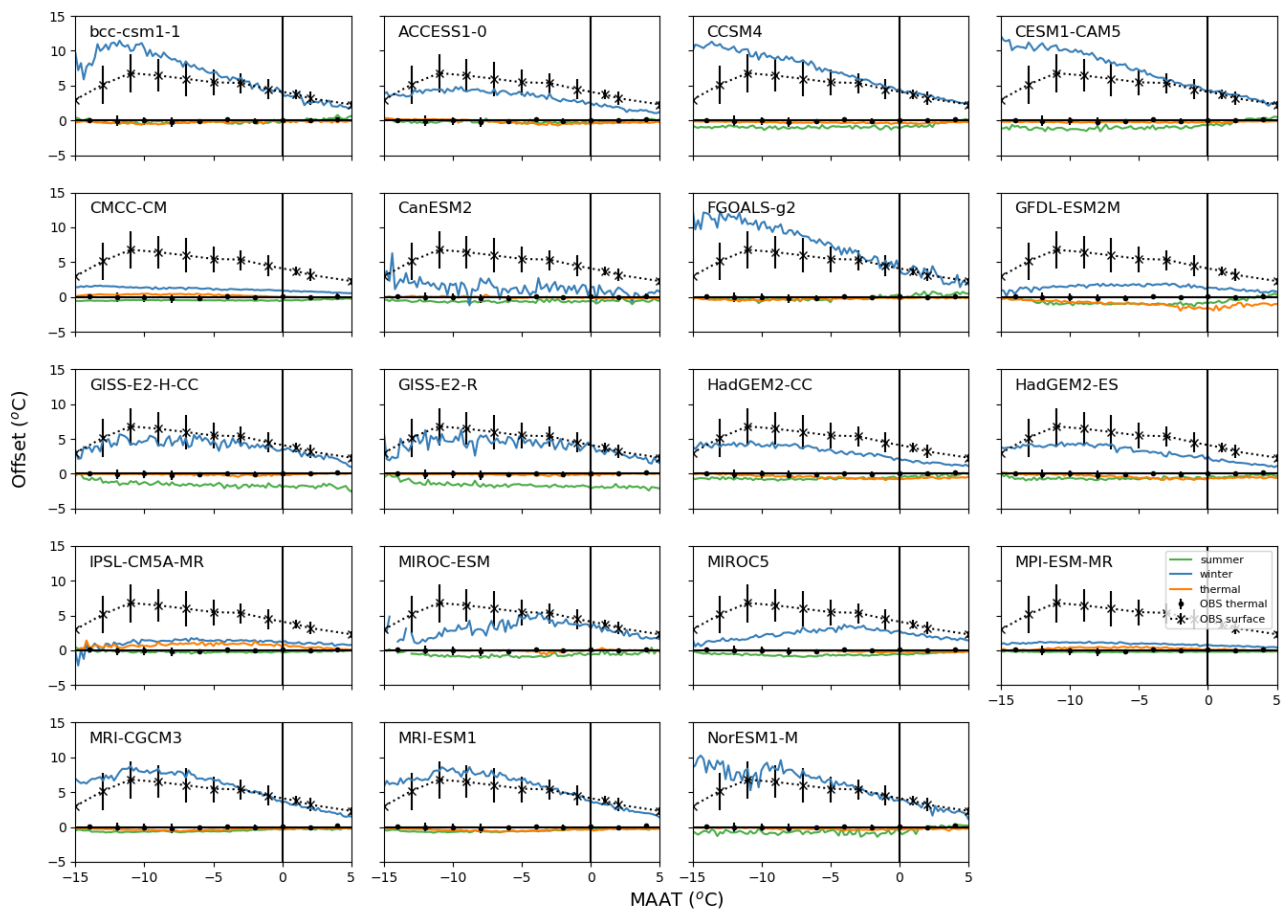
Figure S2.5. The winter, summer and thermal offsets for the CMIP5 models as a function of $MAAT$ for the climatology period of 1986-2005. These offset data are binned into $0.5°C$ bins and the median value of each offset taken. The observed surface and thermal offsets are added in black.
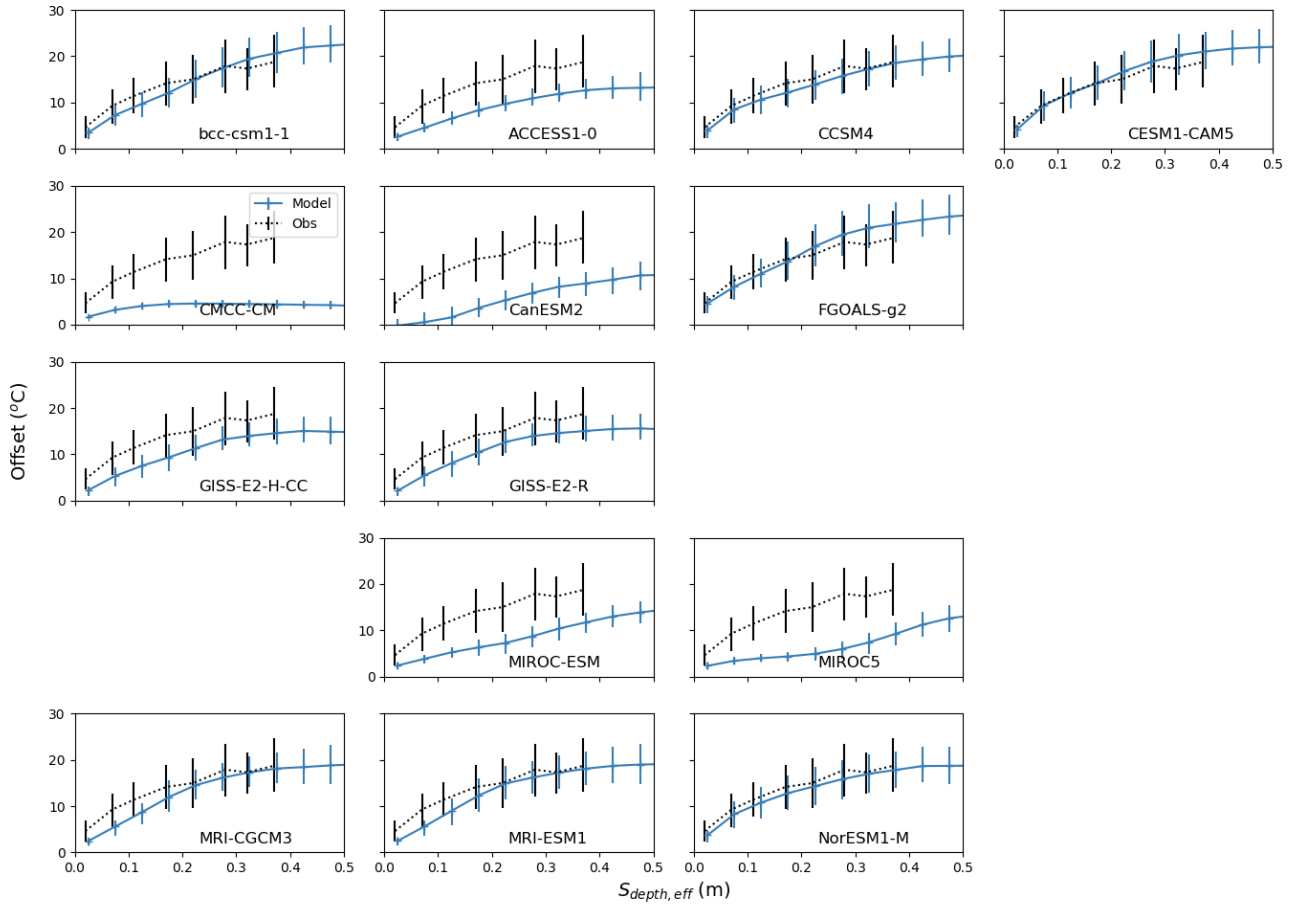
Figure S2.6. Differences between the $MAAT$ and the annual mean soil temperature at 0.2 m for the winter as a function of $S_{depth,eff}$. The climatological period of 1995–2014 is shown for the CMIP6 models. The blue points with the error bars are the model data and the dotted black lines and errorbars are the observations derived using the data from Zhang et al. (2018).
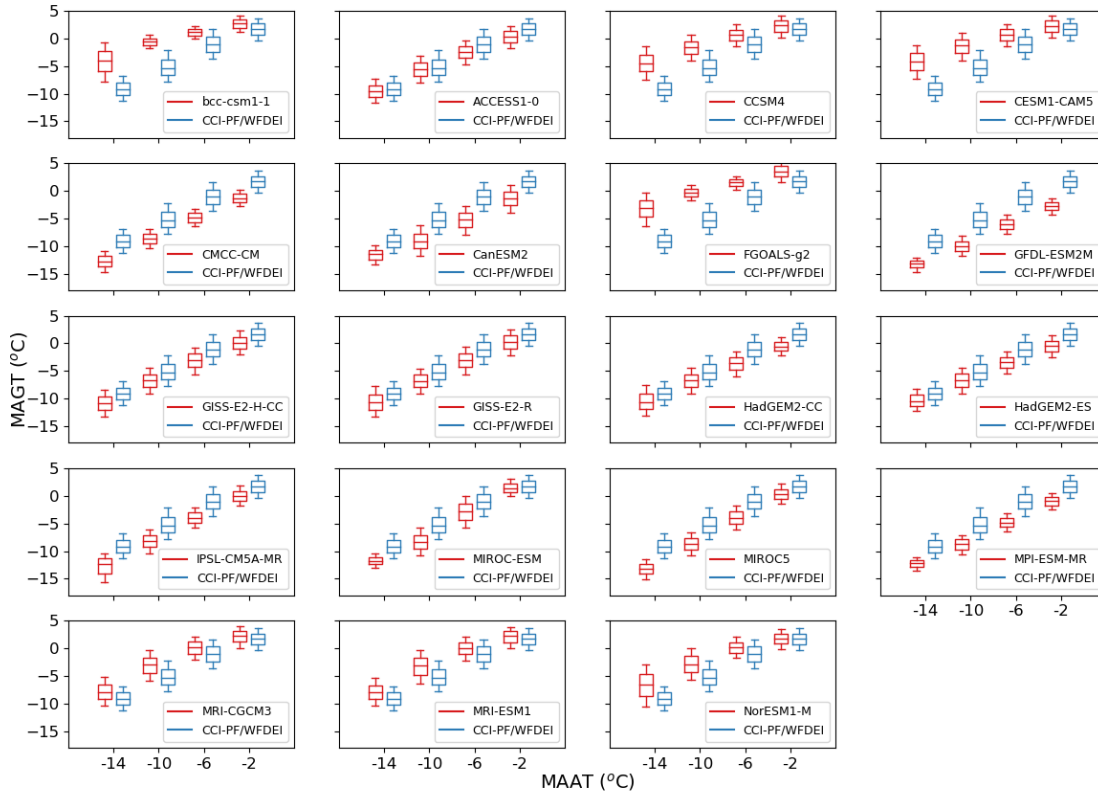
Figure S2.7. $MAGT$ as a function of local $MAAT$ for the CMIP5 models and the climatological period 1986-2005. The $MAGT$ observations were taken from the CCI-PF data set and the $MAAT$ from the WFDEI data.
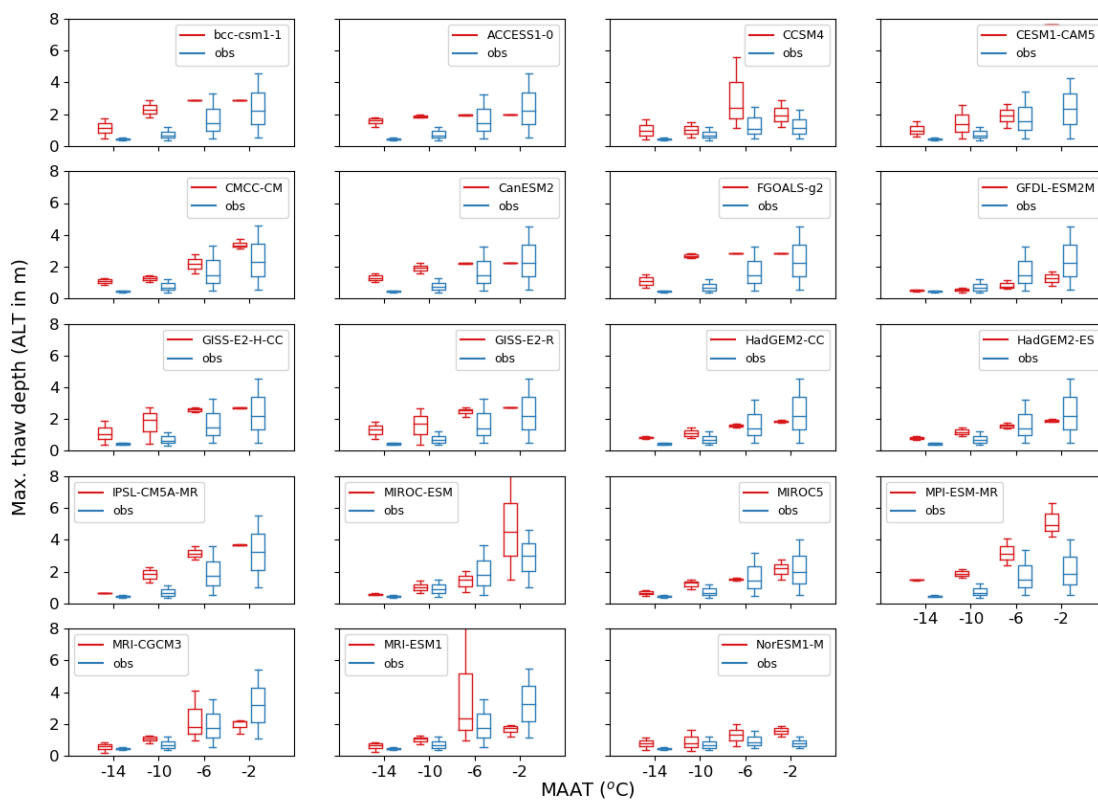
Figure S2.8. Active layer thickness (ALT) as a function of local $MAAT$ for the CMIP5 models and the climatological period 1986-2005. Observations of active layer are from the CALM sites and the air temperatures are from the large scale WFDEI data set. Only sites where the specified model simulates permafrost are included in this analysis.
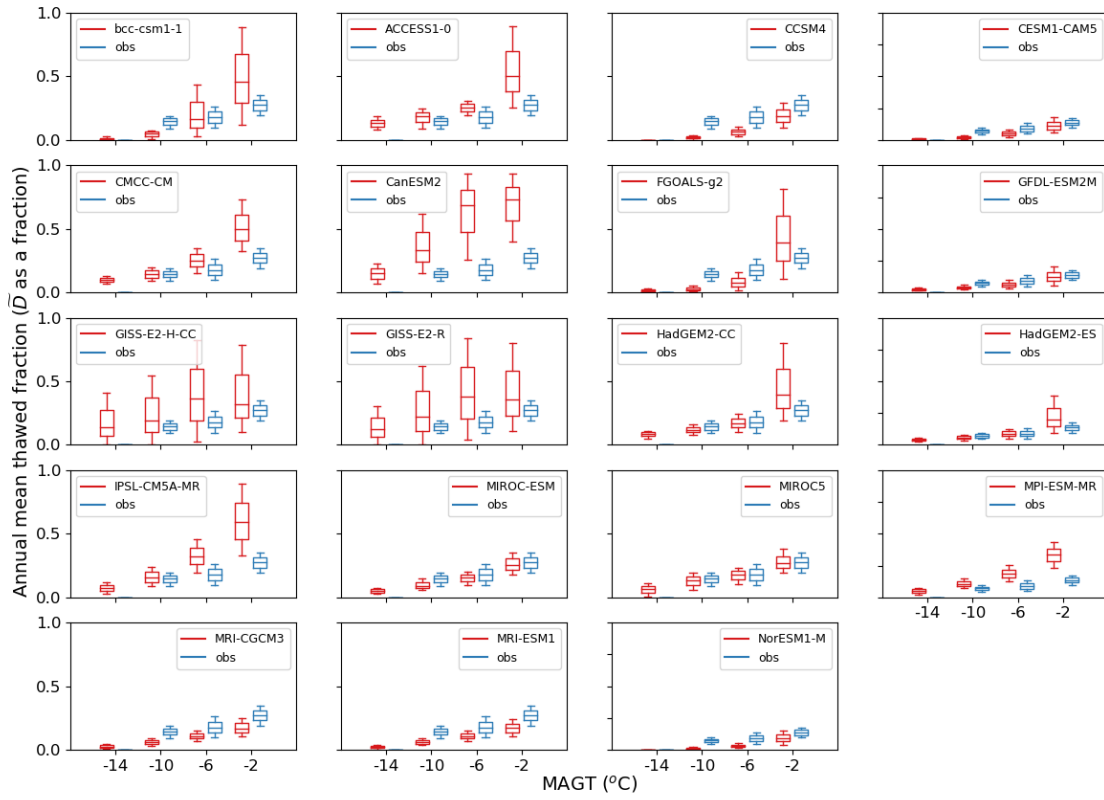
Figure S2.9. Relationship between the annual mean thawed fraction ($\widetilde{D}$) and the $MAGT$ from the site observations and the models in CMIP5 for the climatological period 1986-2005.
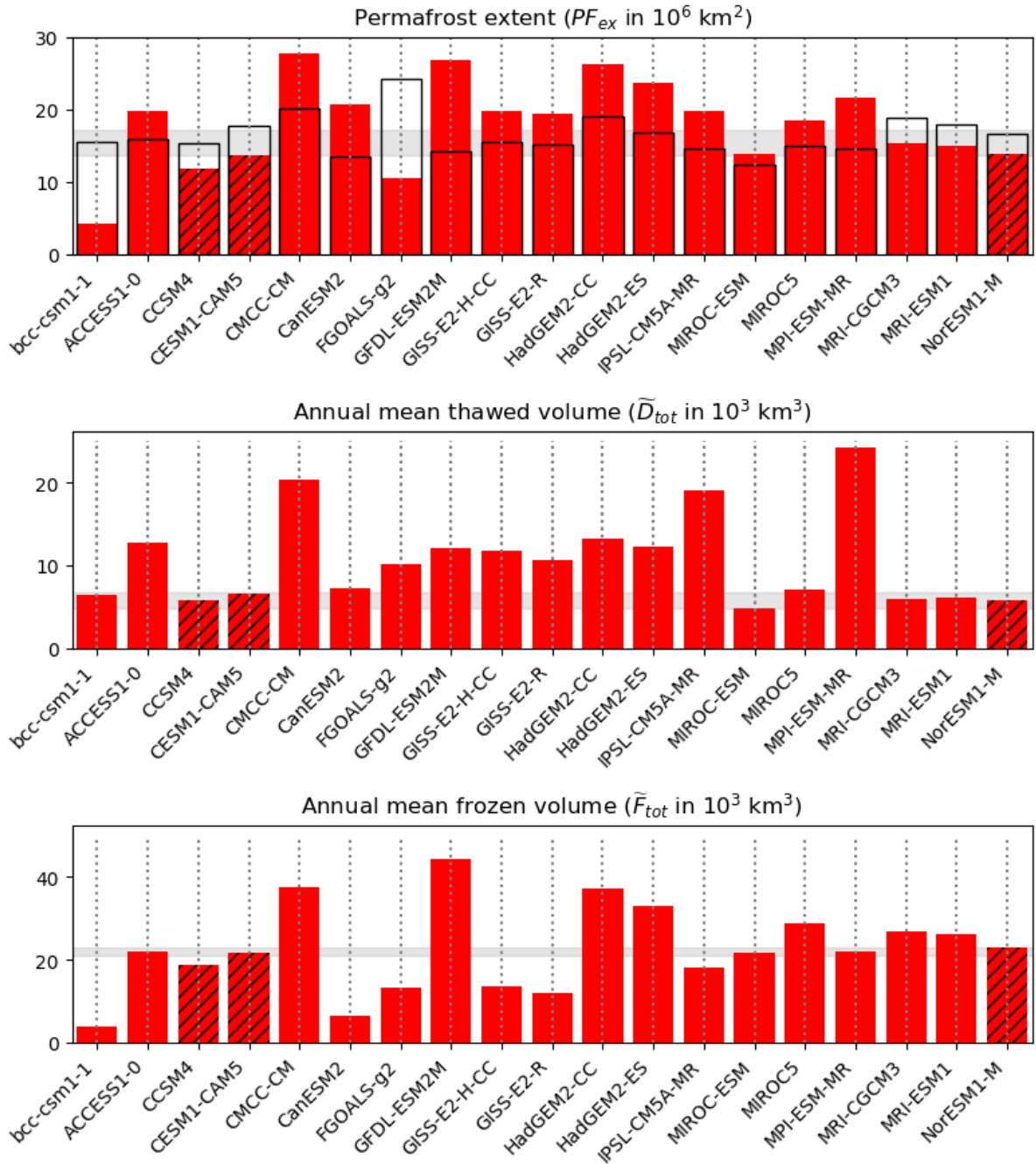
Figure S2.10. (a) Permafrost extent derived using the mean temperature at D_{zaa} (or the mean temperature at the bottom of the soil layer if the soil profile is too shallow) is shaded in red. The models where the permafrost is defined using D_{zaa} have black hatching. In (a) the empty bars with black outlines are the $PF_{benchmark}$ derived from the Chadburn et al. (2017) relationship. The grey shaded area is the range expected from observations. All the results are for the CMIP5 models for the climatological period of 1986-2005. (b) shows the annual thawed volume ($\widetilde{D}_{tot}$) of the top 2 m for the model-specified permafrost region and (c) shows the annual frozen volume of the top 2 m ($\widetilde{F}_{tot}$) for the model-specified permafrost region. These two should add to the permafrost extent shown in (a) multiplied by 2m.

Model	PF_{ex}/area $MAAT<0°C$	$\tilde{D}_{tot}$/area $MAAT<0°C$ (m)	snow off. (°C)	veg. off. (°C)	surface off. (°C)	thermal off. (°C)	$MAGT/MAAT$ (R²)	ALT (m; -12 < MAAT <-10°C)	ALT (m; -6 < MAAT <-4°C)
bcc-csm1-1	0.04	0.15	8.1	-0.2	7.9	-0.3	0.52 (68)	2.6	2.9
ACCESS1-0	**0.72**	0.54	3.8	-0.2	3.7	-0.2	0.85 (79)	1.9	2.0
CCSM4	0.50	0.18	8.3	-0.9	7.4	-0.3	0.53 (58)	1.3	**1.7**
CESM1-CAM5	0.53	**0.20**	8.2	-1.0	7.3	-0.2	0.51 (51)	1.7	2.1
CMCC-CM	0.88	0.48	1.3	-0.4	0.8	0.4	**0.95 (90)**	1.3	2.8
CanESM2	0.16	0.19	1.3	-0.6	0.7	**0.0**	**0.90 (68)**	2.0	2.2
FGOALS-g2	0.12	0.12	8.9	-0.3	8.6	-0.3	0.49 (57)	2.8	2.9
GFDL-ESM2M	1.06	0.37	1.7	-0.9	0.8	-0.9	**0.95 (94)**	**0.6**	**0.9**
GISS-E2-H-CC	0.23	0.27	4.4	-1.4	2.9	**-0.1**	**0.92 (82)**	2.1	2.7
GISS-E2-R	0.20	**0.25**	4.5	-1.5	3.0	**-0.1**	**0.92 (80)**	2.2	2.7
HadGEM2-CC	0.91	0.46	3.6	-0.7	2.9	-0.4	0.80 (76)	1.1	**1.8**
HadGEM2-ES	0.92	0.49	3.6	-0.7	3.0	-0.4	0.83 (79)	1.1	**1.8**
IPSL-CM5A-MR	0.45	0.41	1.4	-0.2	1.2	0.8	1.09 (91)	1.7	3.5
MIROC-ESM	**0.63**	**0.21**	3.4	-0.8	2.6	**-0.1**	1.26 (92)	0.8	**1.5**
MIROC5	0.80	0.31	2.6	-0.7	1.9	**-0.1**	1.17 (94)	1.3	**1.5**
MPI-ESM-MR	0.81	0.62	1.0	-0.2	0.8	0.3	1.00 (98)	1.8	4.2
MRI-CGCM3	0.54	0.18	7.0	-0.6	**6.3**	-0.4	0.85 (74)	1.0	8.5
MRI-ESM1	0.54	**0.20**	6.9	-0.6	**6.3**	-0.5	**0.86 (75)**	1.0	8.5
NorESM1-M	**0.55**	0.18	7.4	-0.8	**6.6**	-0.3	0.69 (57)	0.8	**1.8**
Mean observations	0.62	0.23	-	-	5.75	0.03	0.91	0.42	1.15
Min. observations	0.55	0.20	-	-	4.2	-0.15	0.86	0.49	0.64
Max. observations	0.77	0.25	-	-	7.1	0.15	0.95	0.65	1.98

Table S2.3. Evaluation metrics for the CMIP5 land surface modules. All of the offsets are calculated for the $MAAT$ range between -14°C and -2°C. The values within the range of the observations are highlighted in bold.
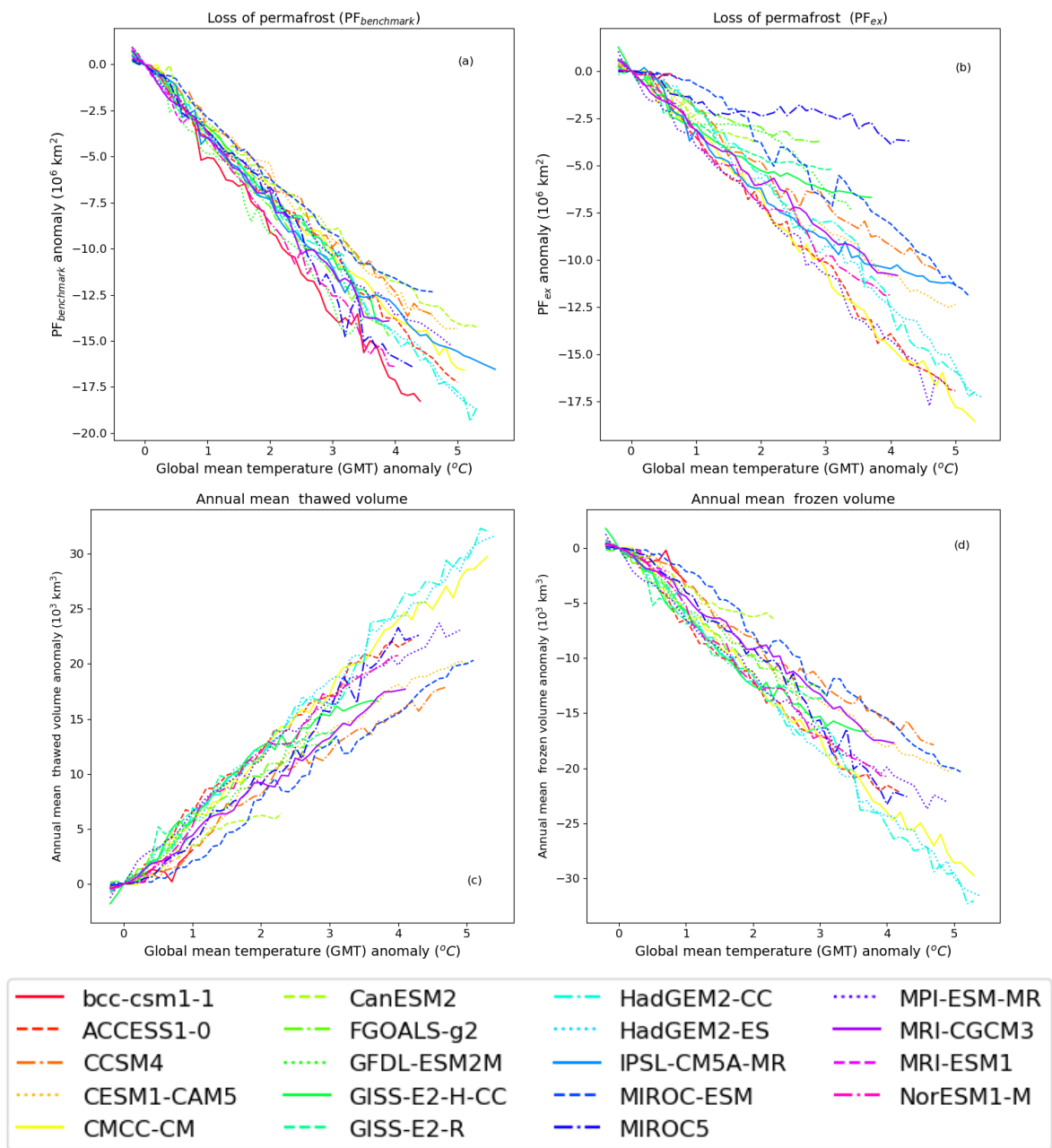
Figure S2.11. Projections of (a) loss of permafrost extent defined as $PF_{benchmark}$ derived from the $MAAT$; (b) loss of permafrost extent defined as PF_{ex} derived from the soil temperatures; (c) increase in annual mean thawed volume; and (d) loss of annual mean frozen volume as a function of global mean temperature change for the CMIP5 models. All the available scenarios are superimposed on one figure and the results binned into $0.1°$C global mean temperature change (GMT) bins.

5 References

Chadburn, S. E., Burke, E., Cox, P., Friedlingstein, P., Hugelius, G., and Westermann, S.: An observation-based constraint on permafrost loss as a function of global warming, Nature Climate Change, 7, 340, 2017.

Zhang, Y., Sherstiukov, A. B., Qian, B., Kokelj, S. V., and Lantz, T. C.: Impacts of snow on soil temperature observed across the circumpolar north, Environmental Research Letters, 13, 044 012, 2018.