Supplementary material

Table S1: List of PAMIP models that were used for the analysis in Figs. S1–S3. Listed is also the available ensemble size for each experimental setup (in number of years)

Institute	Model	pdSST/pdSIC	futArcSIC	futBKSIC
Max Planck Institute for Meteorology, Hamburg	ECHAM6	100	100	100
National Center for Atmospheric Research Canadian	CESM2	200	200	-
Norwegian Meteorological Institute	NorESM2-LM	200	200	-
University of Tokyo/National Institute for Environmental Studies/Japan Agency for Marine-Earth	MIROC6	100	100	100
Canadian Centre for Climate Modelling and Analysis	CanESM5	300	300	-
US Department of Energy/University of California Irvine	E3SMv1	185	190	100
Met Office UK	HadGEM3-GC31-MM	300	300	200
National Center of Atmospheric Research/ University of California Irvine	CESM1-WACCM-SC	300	100	100
Institute of Atmospheric Physics, Beijing	FGOALS-f3-L	100	100	-
Centre Européen de Recherche et de Formation Avancée en Calcul Scientifique US	CNRM-CM6-1	300	300	-
Institute Pierre Simon Laplace University	IPSL-CM6A-LR	200	200	100

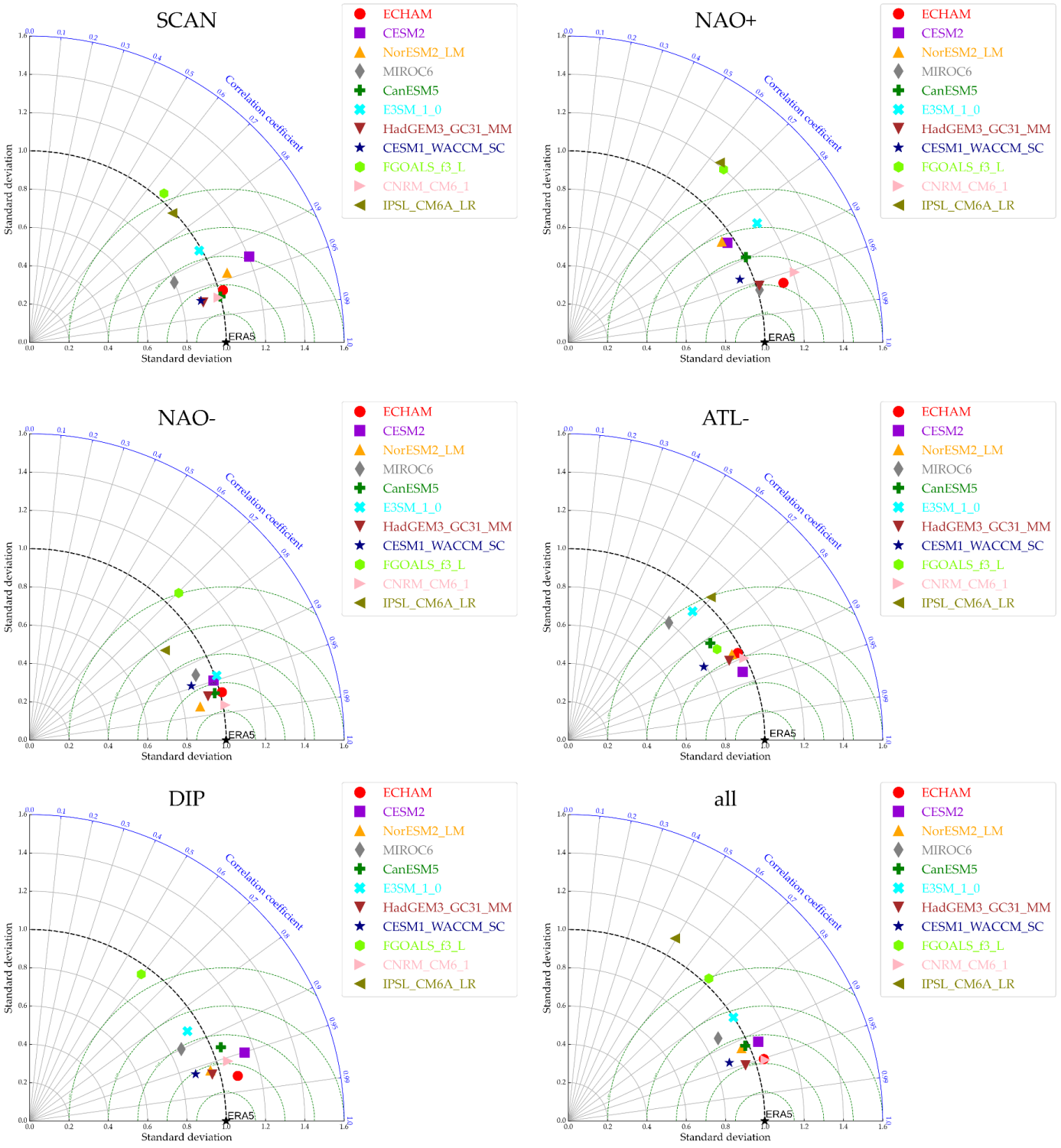

Figure S1: Taylor diagrams that summarize different statistics in order to compare computed model patterns from different PAMIP models with regime patterns obtained from ERA5 (see e.g. Fig. A2). Different symbols/colors indicate different models. Regimes were computed by merging the pdSST/pdSIC with the futArcSIC data set. Individual Taylor diagrams compare individual model regimes with the respective ERA5 regimes. The Taylor diagram in the lower right ("all") compares the statistics averaged over all regimes.

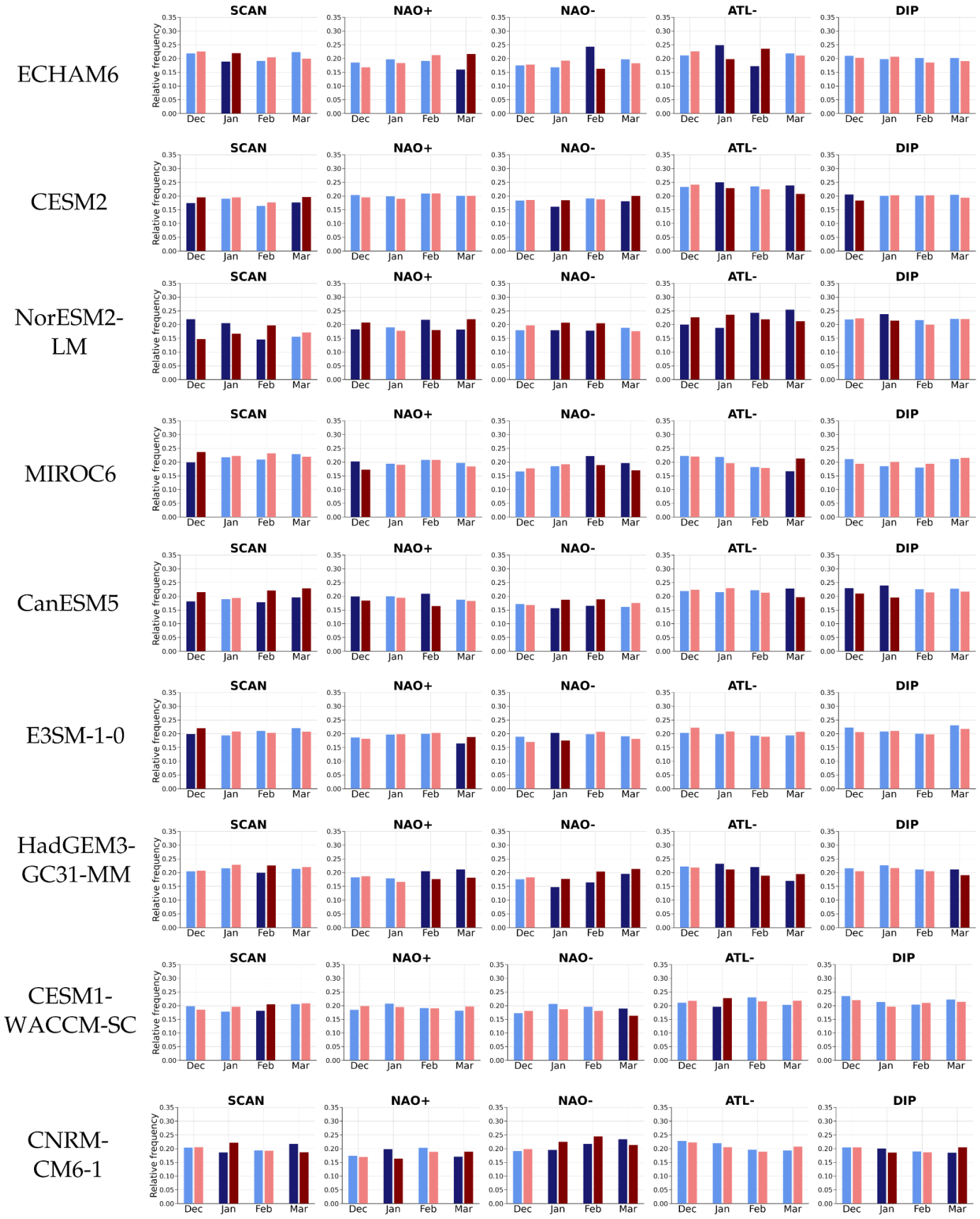

Figure S2: Relative regime occurrence frequencies for different winter months as in Figs. 2a–e, but for different PAMIP models (ERA5 not shown). Compared are the pdSST/pdSIC reference simulation (blueish bars) and the futArcSIC (redish bars). Only the nine models that according to Fig. S1 are able to realistically reproduce most of the ERA5 regime pattern structures were considered here. Dark-colored bars indicate significant differences.

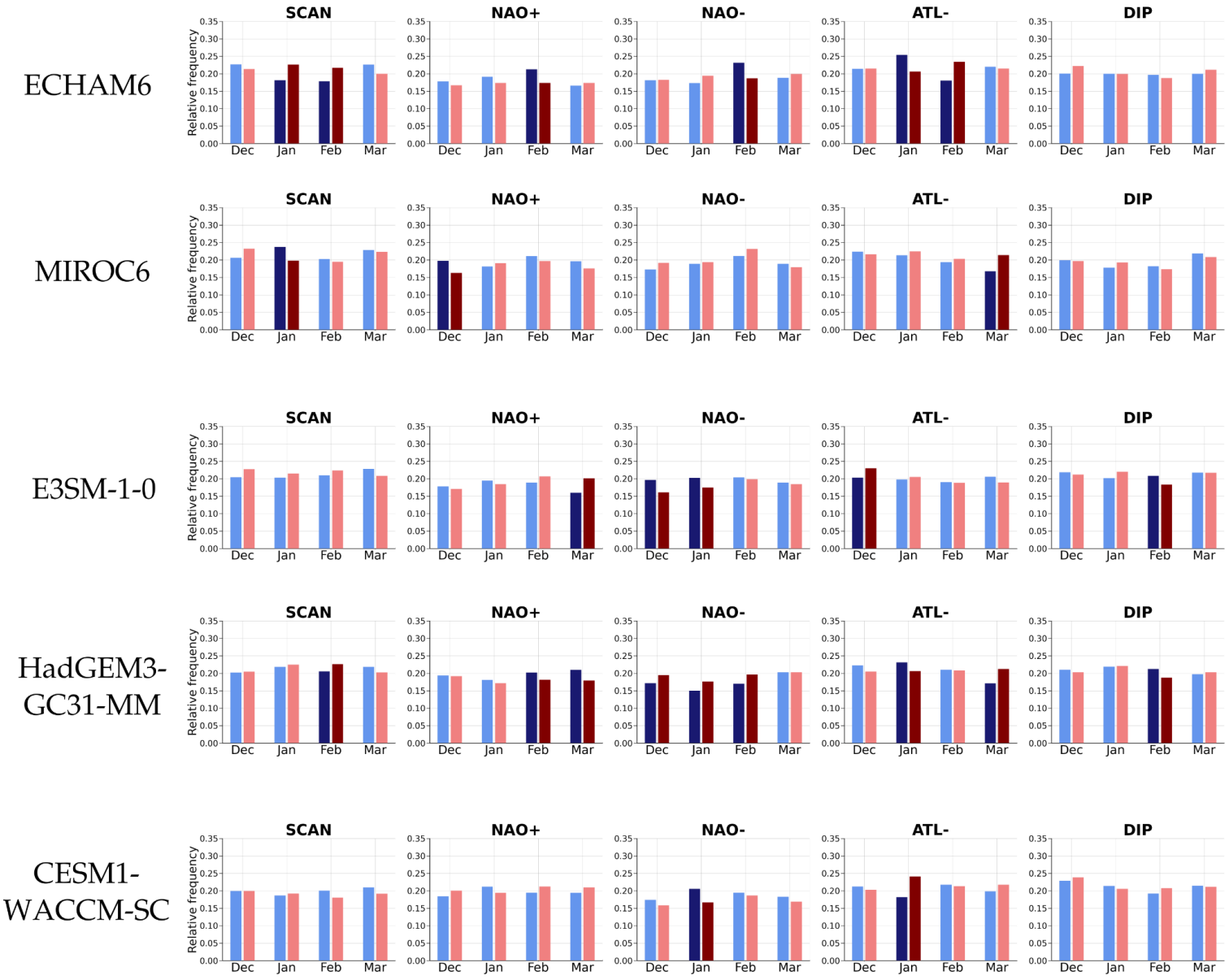

Figure S3: Relative regime occurrence frequencies for different winter months similar to Figs. 2f–j, but for different PAMIP models. Compared are the pdSST/pdSIC reference simulation (blueish bars) and the futBKSIC (redish bars). Only the five available models that according to Fig. S1 are able to realistically reproduce the regime pattern structure were considered here. Dark-colored bars indicate significant differences.

Table S2: Absolute number of regime days for each winter month that were used for the computation of the ERA5 regime frequencies in Fig. 2. Blue numbers indicate the number of regime days for high Arctic sea ice conditions, red numbers for low sea ice conditions.

	SCAN	NAO+	NAO-	ATL-	DIP
December	112/168	115/124	94/119	156/102	143/107
January	81/181	151/100	94/110	163/100	131/129
February	100/183	142/82	94/94	96/119	128/82
March	98/147	142/137	114/127	107/116	159/93.

Table S3: Absolute number of regime days for each winter month that were used for the computation of the relative frequencies in Figs. 2a–e. Blue numbers indicate the day count for the (100 year-long) ECHAM6 pdSST/pdSIC experiment, red numbers for the (100 year-long) ECHAM6 futArcSIC experiment. Note, that the sum over all regimes for a specific month and an experiment sums up to 100 times the number of days within the respective month (e.g. 100*31 days =3100 days for December).

	SCAN	NAO+	NAO-	ATL-	DIP
December	676/699	576/521	540/552	657/702	651/626
January	585/682	610/569	522/596	770/611	613/642
February	534/572	536/594	682/455	484/660	564/519
March	693/619	494/671	612/566	674/654	627/590

Table S4: Same as Table S3 but for pdSST/pdSIC and futBKSIC. Blue numbers indicate the day count for the (100 year-long) ECHAM6 pdSST/pdSIC experiment, red numbers indicate the counts for the (100 year-long) ECHAM6 futBKSIC experiment that were used for the computation of the relative frequencies in Figs. 2f–j .

	SCAN	NAO+	NAO-	ATL-	DIP
December	706/ 666	546/ 521	564/ 567	662/ 654	622/ 692
January	567/ 703	585/ 537	534/ 605	792/ 634	622/ 621
February	498/ 612	598/ 482	646/ 524	505/ 652	553/ 530
March	700/ 618	518/ 542	585/ 619	679/ 666	618/ 655

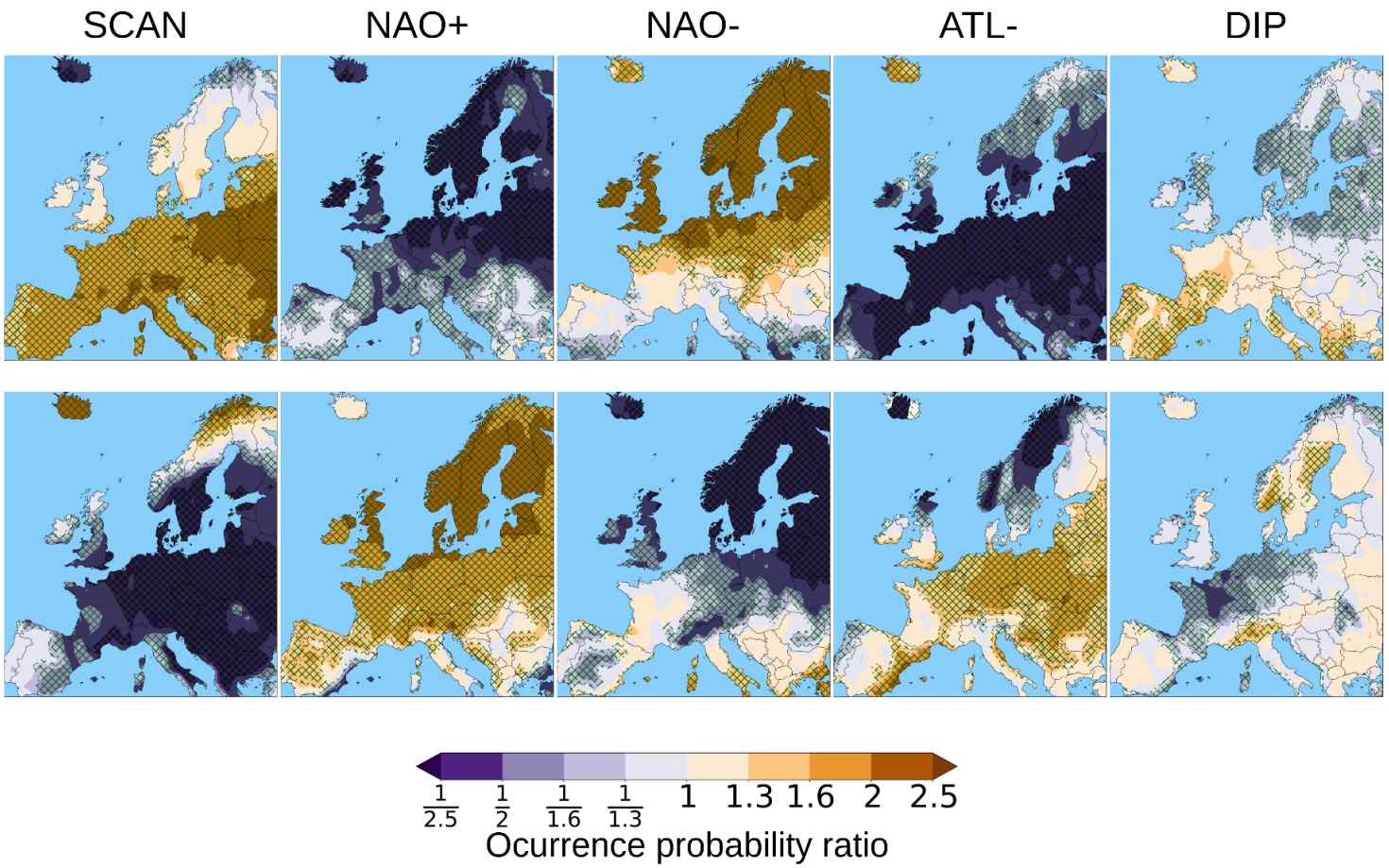

Figure S4: Same analysis as in Fig. 3, but for ERA5 over the period 1979–2018. Thus, regime patterns computed from ERA5 were used for computing these plots. ERA5 T2max/T2min times series at each grid point were linearly detrended beforehand.