Supplementary material for "The strong role of external forcing in seasonal forecasts of European summer temperature"

Matthew Patterson[1], Antje Weisheimer[1,2,3], Daniel Befort[1,3] and Christopher O'Reilly[4]

1. Atmospheric, Oceanic and Planetary Physics, University of Oxford, Parks Road, OX1 3PU, UK

2. European Centre for Medium-Range Weather Forecasts (ECMWF), Shinfield Park, Reading, RG2 9AX, UK

3. National Centre for Atmospheric Science (NCAS), University of Oxford, Parks Road, OX1 3PU, UK

4. Department of Meteorology, University of Reading, Whiteknights Road, Earley Gate, Reading, RG6 6ET, UK

Email: c.h.oreilly@reading.ac.uk

Contents:

Table S1: Further information on C3S seasonal hindcasts (1993-2016)

System	Reference	Start date(s) used in this study	Greenhouse gas forcing (GHG)	Aerosol forcing
CMCC SPS3	Gualdi et al (2020) Technical note: https://www.cmcc.it/wp-content/uploads/2020/09/TN0288-csp-09-2020-1.pdf	1st May (12 members)	Cloud microphysics schemes interact with greenhouse gas concentrations in the model. CMIP5 historical GHG concentrations are used to 2005 and the hindcasts follow RCP8.5 thereafter.	Aerosol forcing taken from fixed climatology based on year 2000.
DWD GCFS2	Fröhlich et al (2021), JAMES https://doi.org/10.1029/2020MS002101	1st May (12 members)	CMIP6 historical GHG concentrations to 2014, constant thereafter as scenario forcing information was not ready when the hindcasts were performed in 2018.	CMIP6 historical aerosol variations to 2014, constant thereafter as scenario forcing information was not ready when the hindcasts were performed in 2018.
ECMWF system 5	Johnson et al (2019), GMD https://doi.org/10.5194/gmd-12-1087-2019	1st May (12 members)	CMIP5 historical GHG is used to 2000 and CMIP5 RCP 3-PD scenario is prescribed thereafter.	Tropospheric sulfate aerosol follows a decadally-varying CMIP5 climatology and stratospheric sulfate aerosol is prescribed using GISS data.
Meteo-France system 7	Batté et al (2021) Technical note: http://www.umr-cnrm.fr/IMG/pdf/system8-technical.pdf	1st May (1 member), last Thursday of previous month (11 members)	CMIP6 historical GHG concentrations to 2014, follows SSP3-7.0 thereafter.	A climatology of aerosol optical depths is calculated as an average over 1995-2014 from runs with the interactive TACTIC_V2 scheme. Volcanic aerosols use a climatology taken from CMIP6 (1850-2014)
UKMO GloSea6	MacLachlan et al (2014), QJRMS https://doi.org/10.1002/qj.2396	1st May (7 members), 25th April (5 members)	CMIP6 historical GHG concentrations to 2015. Concentrations follow SSP1-2.6 thereafter.	All aerosols follow a seasonal climatology (including biogenic aerosols, biomass burning, black-carbon, sea salt, sulphates, dust and organics carbon fossil fuels)

Table S2: A table of CMIP6 models included in this study

		Ensemble members used for each variable			
Model name	**Institute ID**	**tas**	**psl**	**pr**	**zg**
ACCESS-ESM1-5	CSIRO	r1i1p1f1, r2i1p1f1, r3i1p1f1	r1i1p1f1, r2i1p1f1, r3i1p1f1	r1i1p1f1, r2i1p1f1, r3i1p1f1	r1i1p1f1
ACCESS-CM2	CSIRO-ARCCSS	r1i1p1f1	r1i1p1f1	r1i1p1f1	r1i1p1f1
AWI-CM-1-1-MR	AWI	r1i1p1f1		r1i1p1f1	r1i1p1f1
BCC-CSM2-MR	BCC	r1i1p1f1	r1i1p1f1	r1i1p1f1	r1i1p1f1
CAMS-CSM1-0	CAMS	r1i1p1f1	r1i1p1f1, r2i1p1f1	r1i1p1f1, r2i1p1f1	r1i1p1f1
CanESM5	CCCma	r1i1p1f1, r2i1p1f1, r3i1p1f1	r1i1p1f1, r2i1p1f1, r3i1p1f1	r1i1p1f1, r2i1p1f1, r3i1p1f1	r1i1p1f1
CanESM5-CanOE	CCCma	r1i1p2f1, r2i1p2f1, r3i1p2f1	r1i1p2f1	r1i1p2f1	r1i1p2f1
CESM2	NCAR	r1i1p1f1, r2i1p1f1		r1i1p1f1, r2i1p1f1	
CESM2-WACCM	NCAR	r1i1p1f1, r2i1p1f1, r3i1p1f1	r1i1p1f1, r2i1p1f1, r3i1p1f1	r1i1p1f1, r2i1p1f1, r3i1p1f1	r1i1p1f1
CIESM	AS-RCEC	r1i1p1f1		r1i1p1f1	
CMCC-CM2-SR5	CMCC		r1i1p1f1		
CNRM-CM6-1	CNRM	r1i1p1f2, r2i1p1f2, r3i1p1f2	r1i1p1f2	r1i1p1f2	
CNRM-CM6-1-HR	CNRM	r1i1p1f2	r1i1p1f2	r1i1p1f2	
CNRM-ESM2-1	CNRM	r1i1p1f2, r2i1p1f2, r3i1p1f2	r1i1p1f2	r1i1p1f2	
EC-Earth3	EC-Earth-Consortium	r1i1p1f1, r2i1p1f1	r1i1p1f1	r1i1p1f1, r4i1p1f1	
EC-Earth3-Veg	EC-Earth-Consortium	r1i1p1f1, r2i1p1f1, r3i1p1f1	r1i1p1f1	r1i1p1f1, r2i1p1f1, r3i1p1f1	
FGOALS-f3-L	CAS	r1i1p1f1		r1i1p1f1	r1i1p1f1
FGOALS-g3	CAS	r1i1p1f1		r1i1p1f1	r1i1p1f1
GFDL-ESM4	GFDL	r1i1p1f1	r1i1p1f1	r1i1p1f1	
GISS-E2-1-G					
HadGEM3-GC31-LL	MOHC	r1i1p1f3, r2i1p1f3, r3i1p1f3	r1i1p1f3, r2i1p1f3, r3i1p1f3	r1i1p1f3, r2i1p1f3, r3i1p1f3	r1i1p1f3
HadGEM3-GC31-MM	MOHC		r1i1p1f3, r2i1p1f3, r3i1p1f3	r1i1p1f3, r2i1p1f3, r3i1p1f3	r1i1p1f3
IITM-ESM	THu		r1i1p1f1		
INM-CM4-8	INM	r1i1p1f1	r1i1p1f1	r1i1p1f1	
INM-CM5-0	INM	r1i1p1f1	r1i1p1f1	r1i1p1f1	
IPSL-CM6A-LR	IPSL	r1i1p1f1, r2i1p1f1, r3i1p1f1	r1i1p1f1	r1i1p1f1	

KACE-1-0-G	NIMS-KMA	r1i1p1f1, r2i1p1f1, r3i1p1f1		r1i1p1f1, r2i1p1f1, r3i1p1f1	
MCM-UA-1-0	UA	r1i1p1f1			
MIROC-ES2L	MIROC				r1i1p1f2
MIROC6	MIROC	r1i1p1f1, r2i1p1f1	r1i1p1f1, r2i1p1f1, r3i1p1f1	r1i1p1f1, r2i1p1f1, r3i1p1f1	
MPI-ESM1-2-LR	MPI-M	r1i1p1f1, r2i1p1f1, r3i1p1f1	r1i1p1f1, r2i1p1f1, r3i1p1f1	r1i1p1f1, r2i1p1f1, r3i1p1f1	r1i1p1f1
MRI-ESM2-0	MRI	r1i1p1f1	r1i1p1f1	r1i1p1f1	r1i1p1f1
NESM3	NUIST	r1i1p1f1, r2i1p1f1	r1i1p1f1, r2i1p1f1	r1i1p1f1, r2i1p1f1	
NorESM2-LM	NCC	r1i1p1f1			
NorESM2-LM	NCC	r1i1p1f1	r1i1p1f1	r1i1p1f1	
UKESM1-0-LL	MOHC	r1i1p1f2, r2i1p1f2, r3i1p1f1	r1i1p1f2, r2i1p1f2, r3i1p1f2	r1i1p1f2, r2i1p1f2, r3i1p1f2	r1i1p1f2
Total	-	60	44	54	16

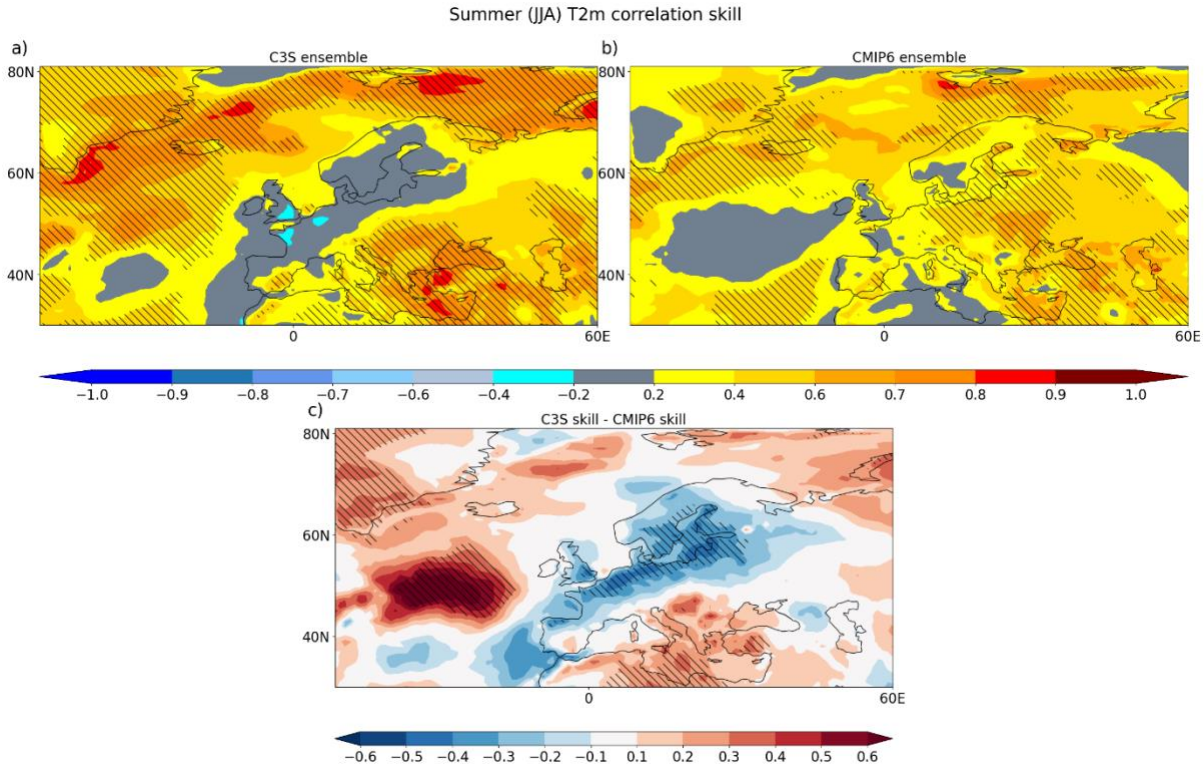

Figure S1: Same as figure 2 of the main manuscript but using 25 ensemble members of the 5 C3S models to create a 125-member ensemble.

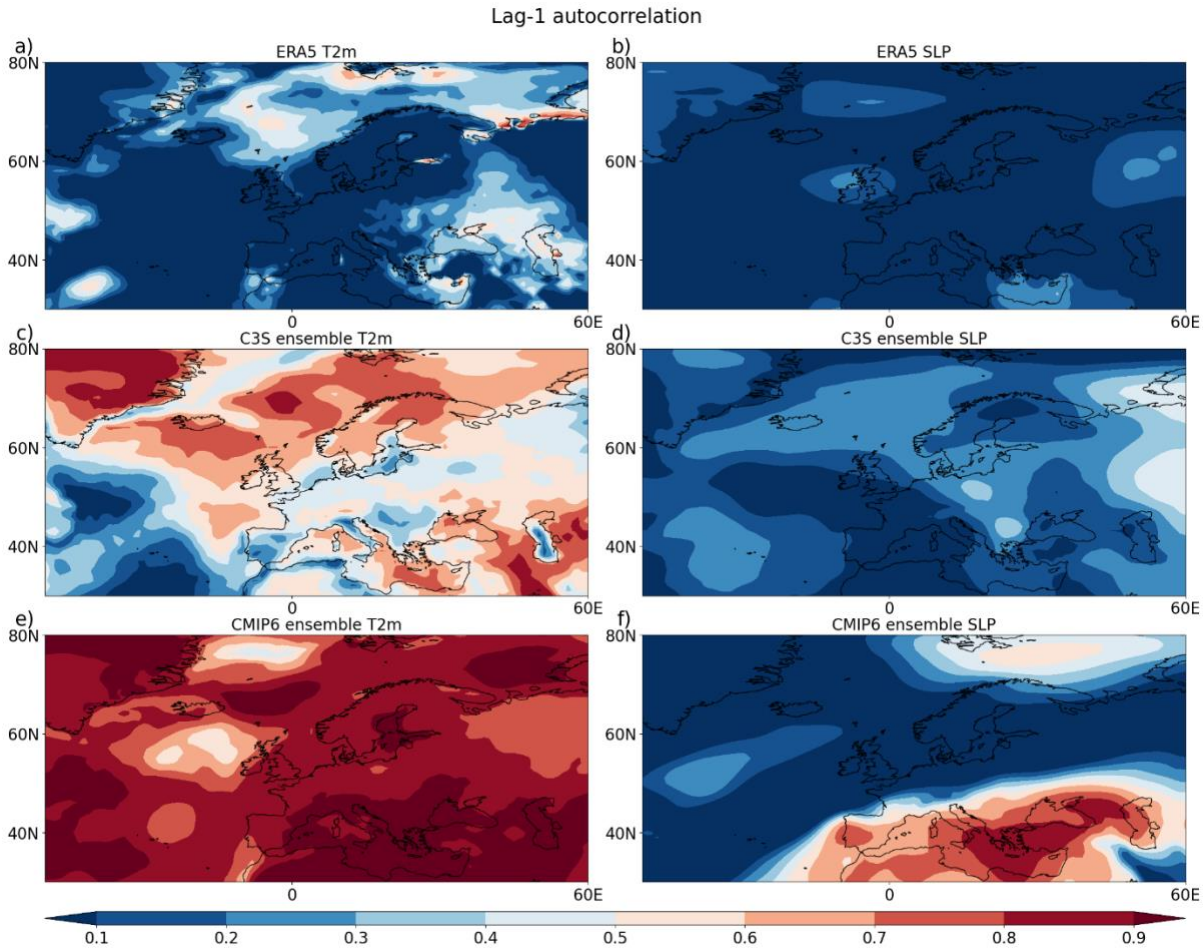

Figure S2: Lag-1 autocorrelation for summer (JJA) 1993-2016 for (left) T2m and (right) SLP using (a,b) ERA5 data (c,d) the C3S ensemble mean and (e,f) the CMIP6 ensemble mean.

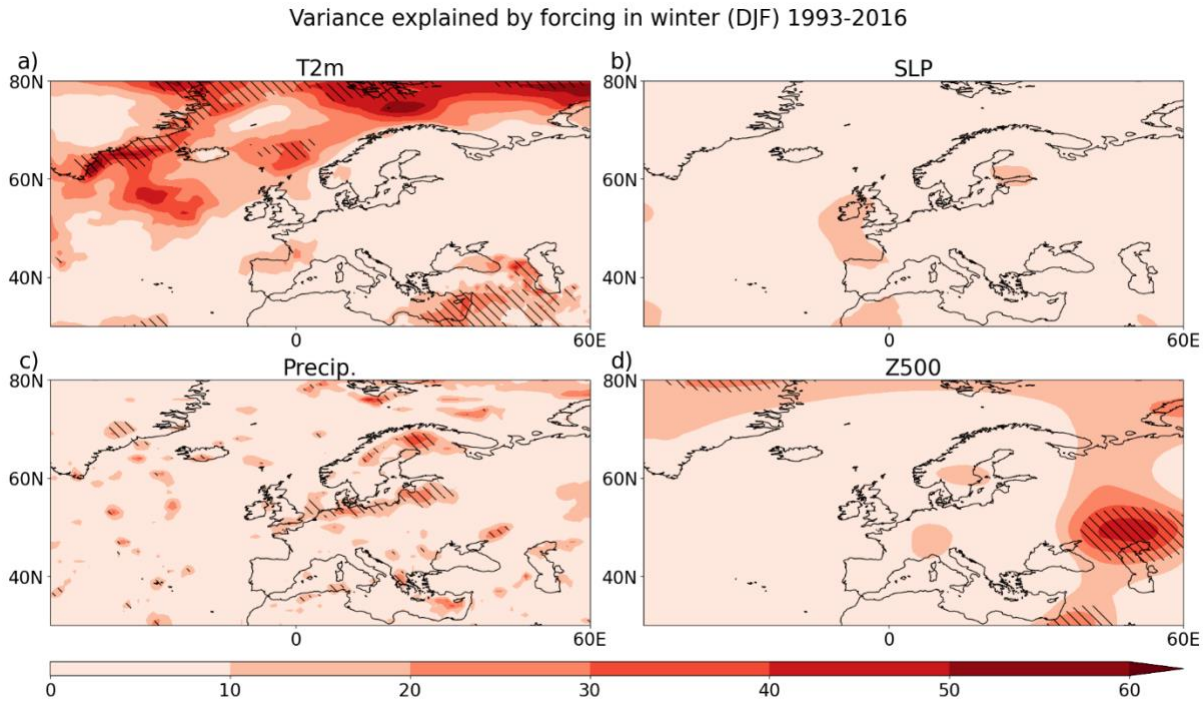

Figure S3: Same as figure 1 of the main manuscript but for boreal winter (December, January, February – DJF).

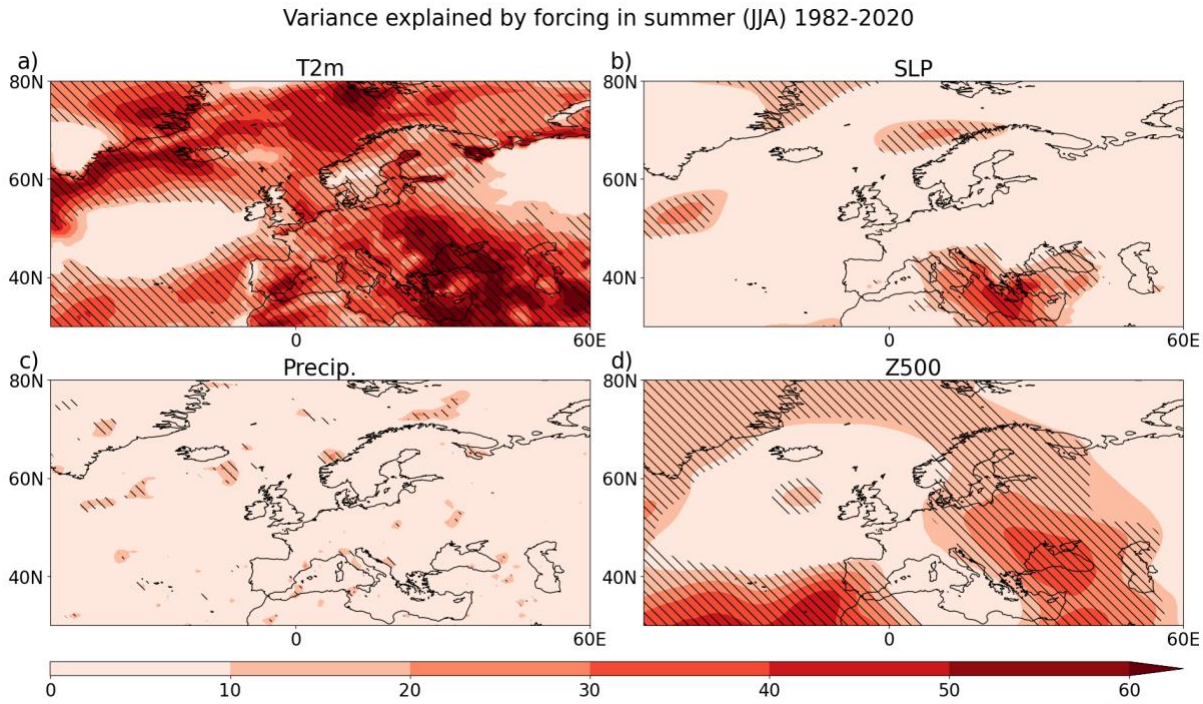

Figure S4: As in figure 1 of the main manuscript, but for the period 1981-2020.

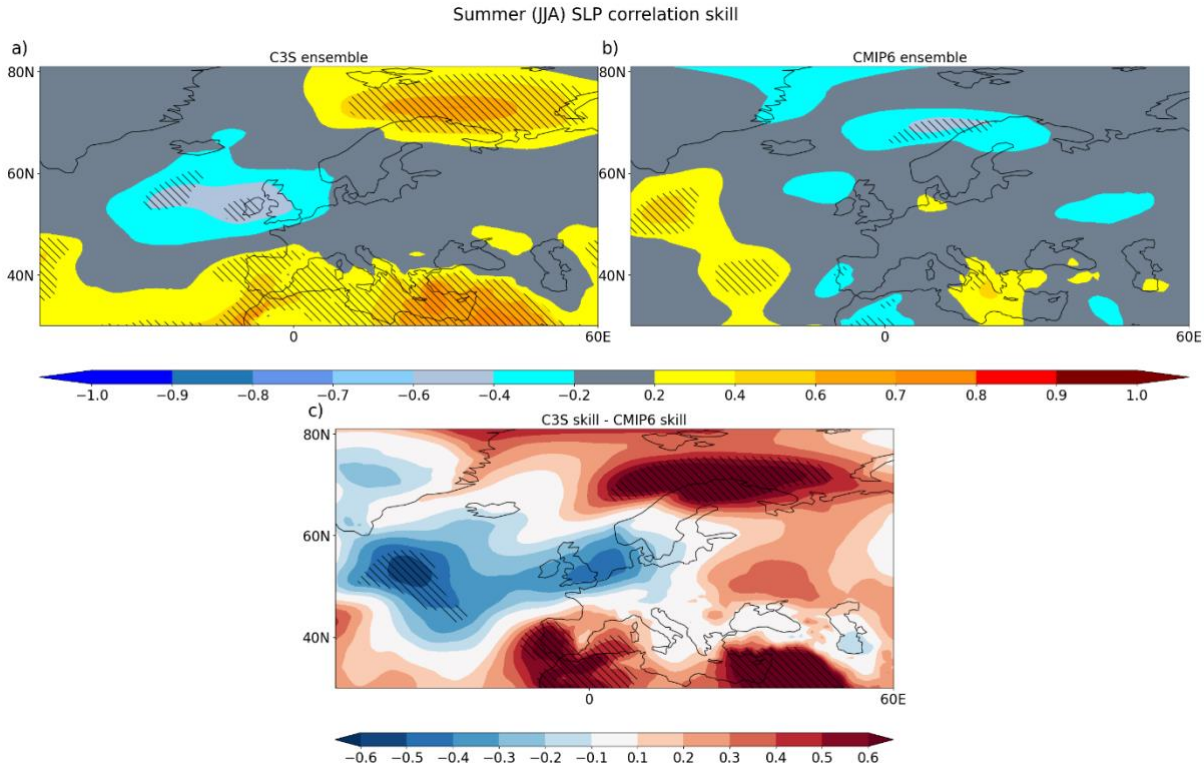

Figure S5: Same as figure 2 of the main manuscript but for sea level pressure.

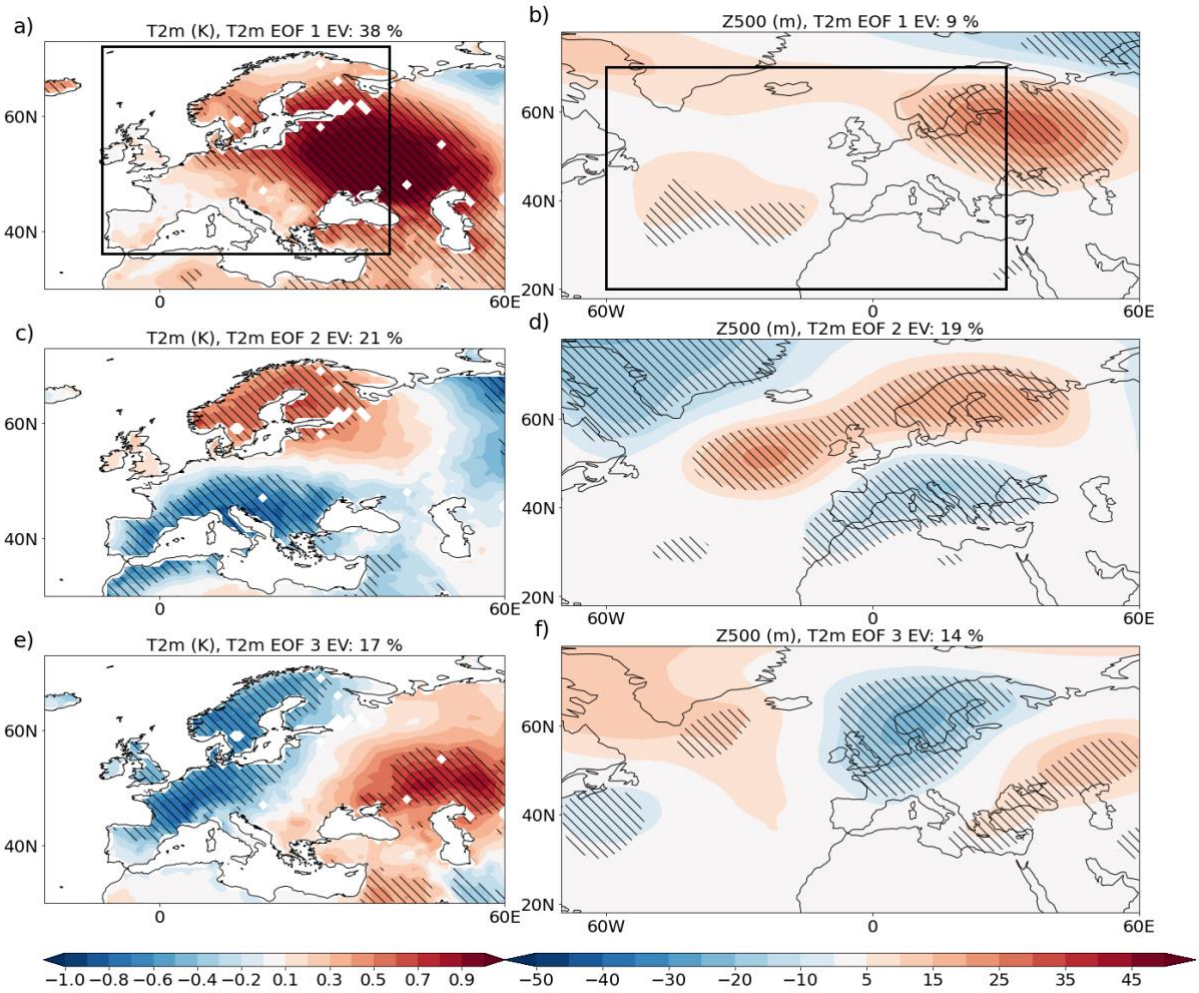

Figure S6: Results of an empirical orthogonal function (EOF) analysis of ERA5 summer T2m (1993-2016) in the European region, indicated by the box in panel a. S6a,c,e) show the T2m regression patterns associated with first three EOFs with the percentage of explained variance (EV) indicated in the title. S6b,d,f) show the 500hPa geopotential height (Z500) regression patterns associated with these EOFs. The title also indicates the proportion of Z500 variance in the Euro-Atlantic region (shown in panel b) is explained by these patterns. Hatching indicates where the regression coefficients are significantly different from zero at the 95% level following a t-test.

Figure S6 indicates that atmospheric circulation is relatively strongly associated with variations in European T2m. The first EOF corresponds to the trend, but the second and third EOFs, which explain 21% and 17% of the variance, respectively, are clearly linked to large scale patterns of atmospheric circulation (figure S6d,f).

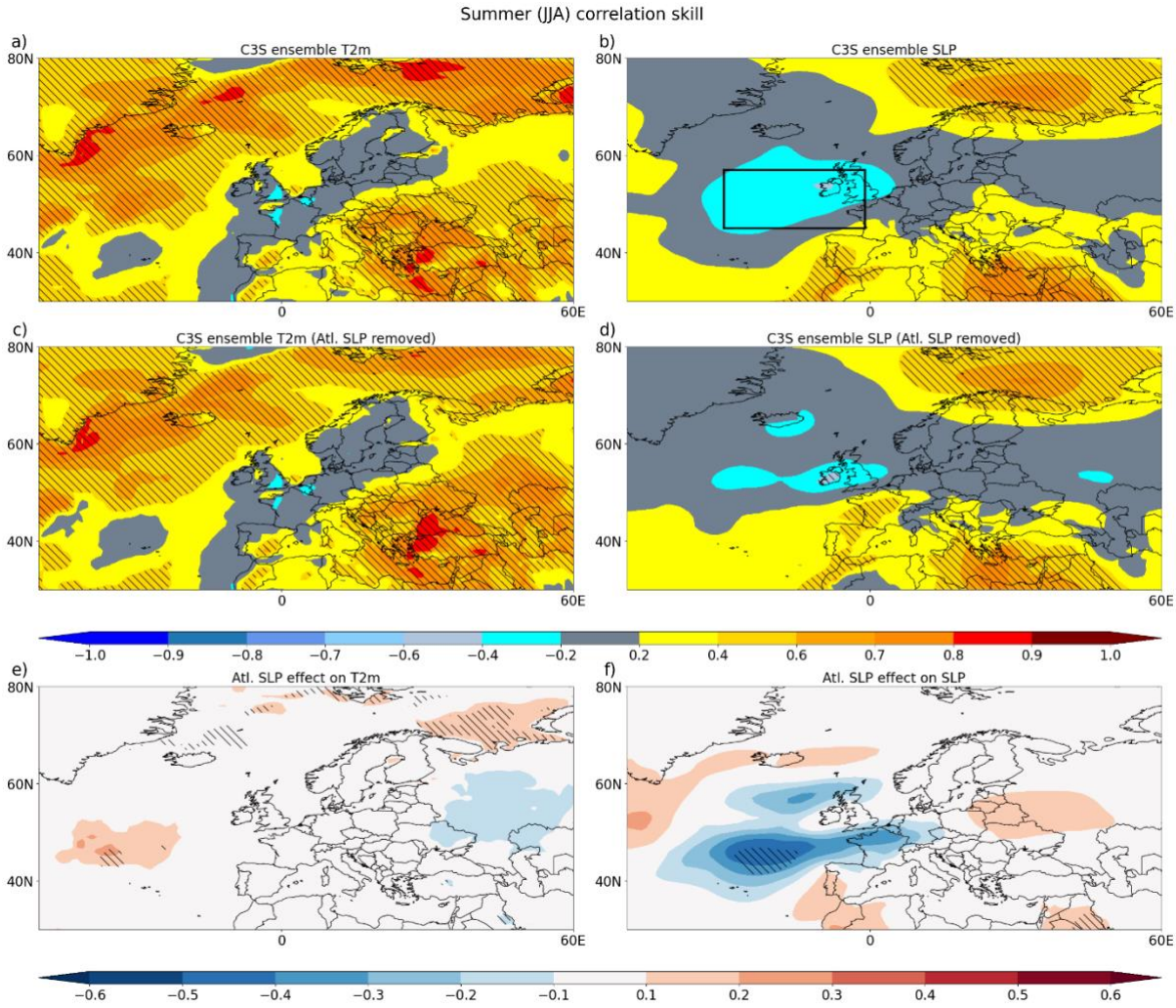

Figure S7: Anomaly correlation coefficient of the C3S ensemble with and without the effect of the region of negative sea level pressure skill to the west of the UK. The top two panels show correlation skill and S7c,d show the correlation skill when the ensemble-mean sea level pressure in the boxed region in S7b (30W-0W,45N-57N) has been linearly removed from the C3S time series' at each grid-point before the correlation was calculated. S7e,f show the differences, i.e. S7c – S7a and S7d– S7b respectively with hatching showing where the p-value is less than 0.05 for the null hypothesis that the correlation skill is the same.

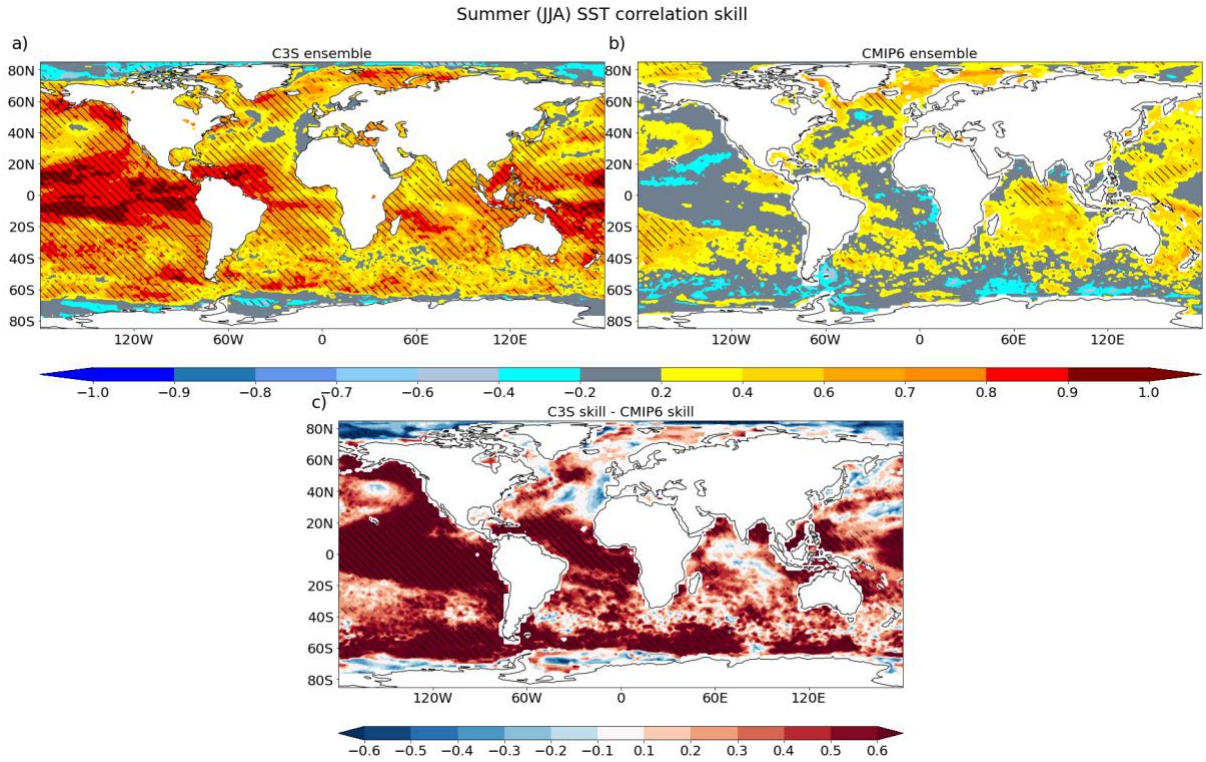

Figure S8: Same as figure 2 of the main manuscript but for sea surface temperature (SST) and showing all longitudes and latitudes to 85N/S.

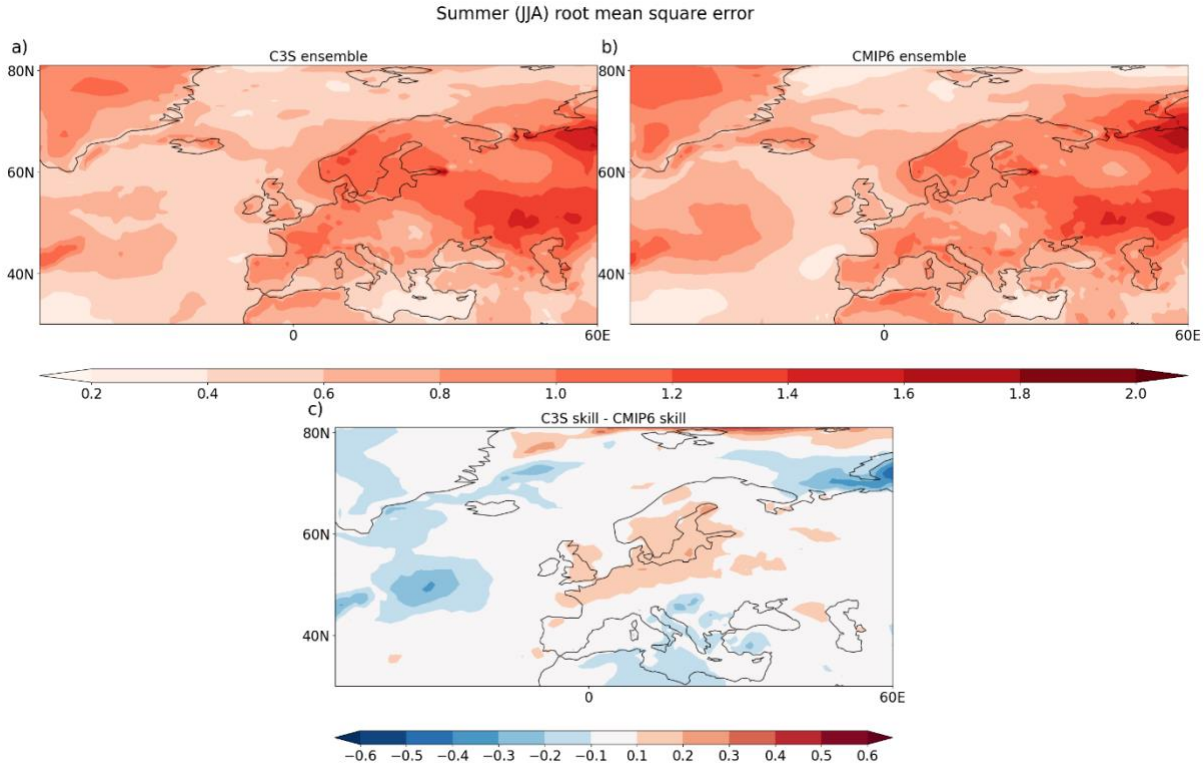

Figure S9: As in figure 2 of the main manuscript but for root mean square error rather than anomaly correlation. Units are Kelvin.

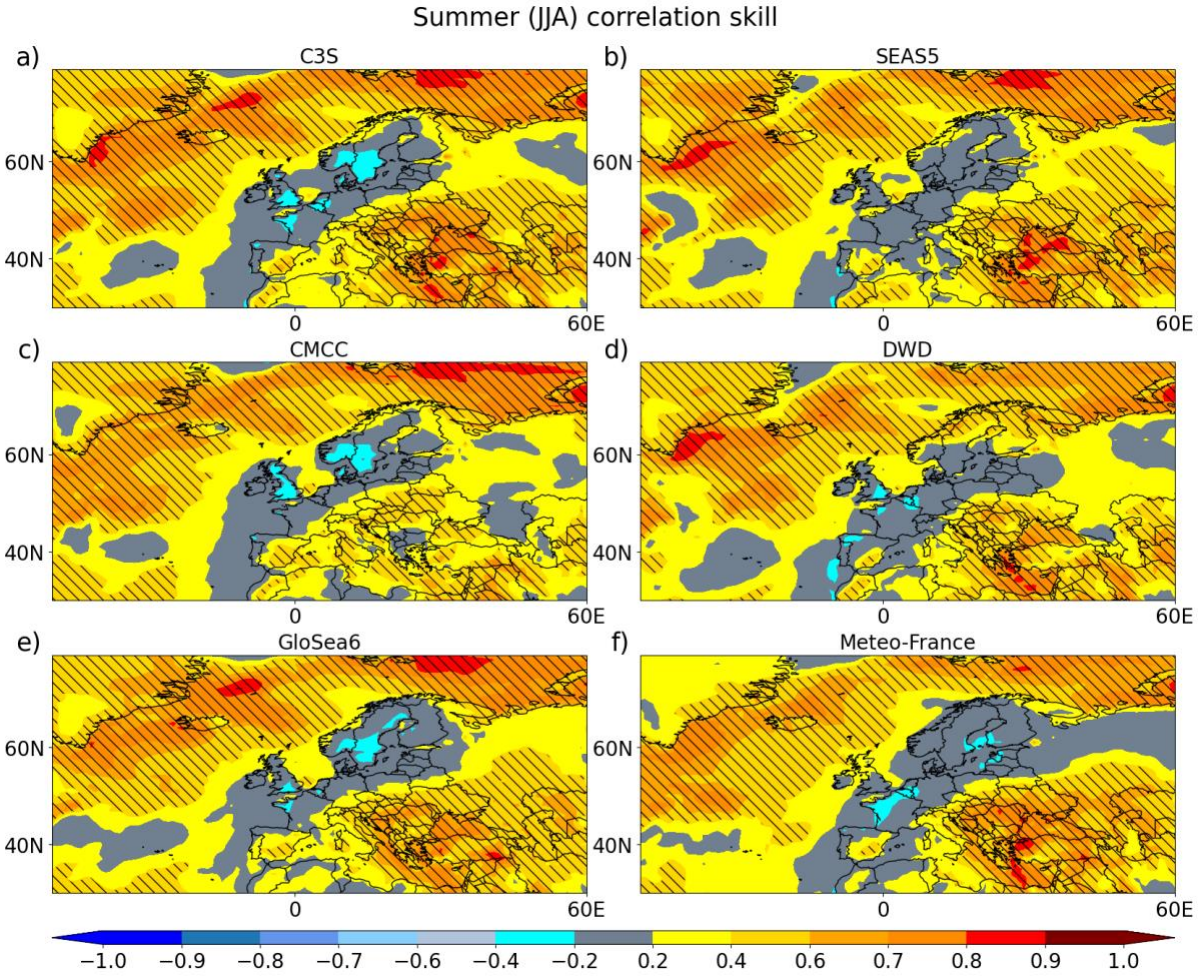

Figure S10: Correlation skill in all C3S models using the same plotting conventions as figure 2a-d of the main paper. S10a is the same as figure 2a of the main paper. For plots with individual models in panels b-f 25 members have been used. Panel a combines all of these members to make a 125 member ensemble.

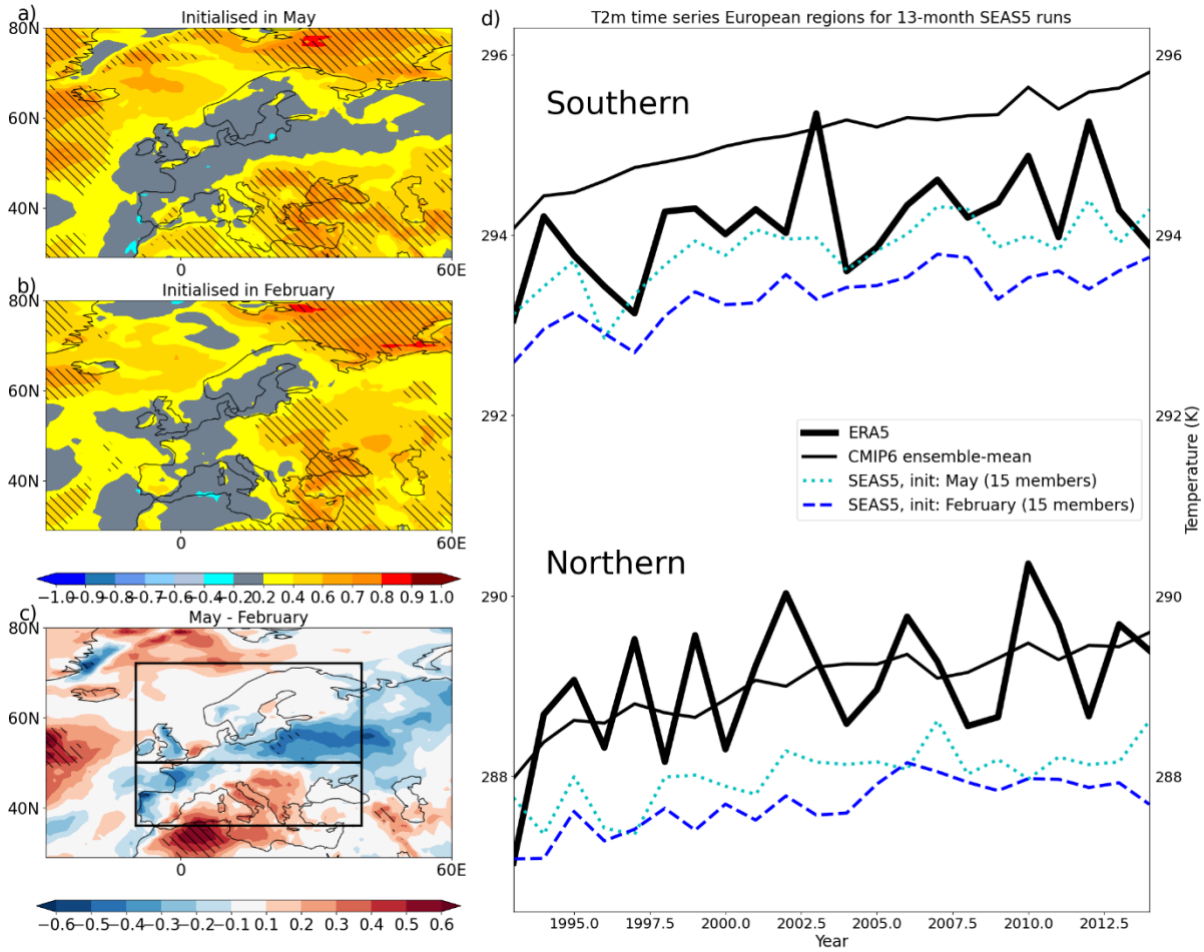

Figure S11: The same as figure 3 from the main manuscript but for hindcasts initialised in February rather than November.

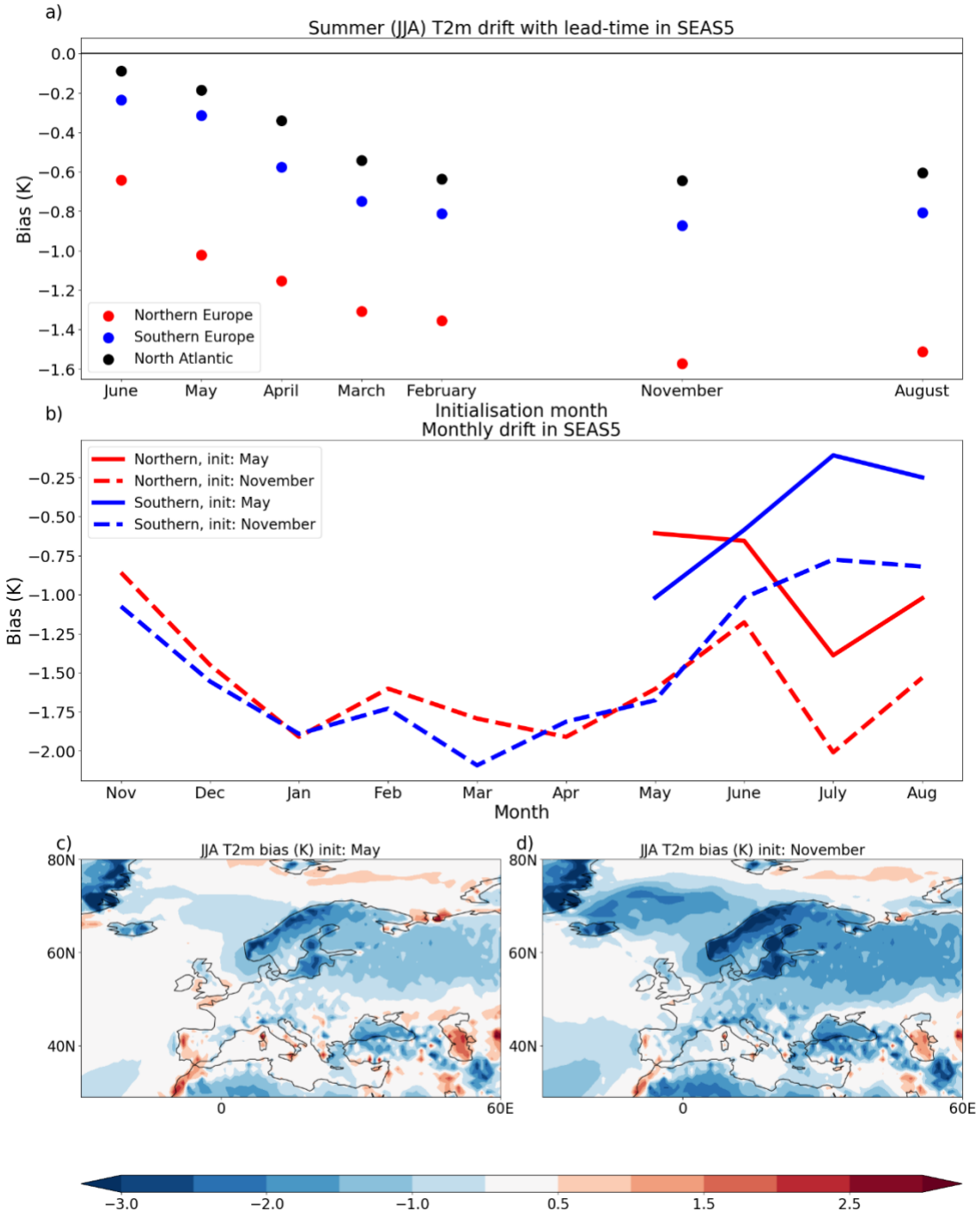

Figure S12 a) Summer (JJA) bias in SEAS5 with respect to ERA5 for different regions as a function of lead-time. The Northern and Southern Europe regions are defined as in the main paper and are averages only over the land and the North Atlantic is defined as 30-50N and 40W-0W and is averaged over open ocean. b) shows the monthly drift in SEAS5 T2m with respect to ERA5 for simulations initialised in November and May. Finally spatial maps of the JJA T2m bias are shown for the hindcasts initialised in c) May and d) November.

Figure S11 shows the bias in T2m ECMWF's system 5 (SEAS5) with respect to ERA5 as a function of lead-time. In general, SEAS5 has a cold bias over Europe which is established within the first month figure S11b) but the model has a drift towards cooler temperatures with lead-time (figure S11a,b). The bias is particularly prevalent over Scandinavia. It is unclear if or to what extent the bias has an impact on seasonal forecast skill.

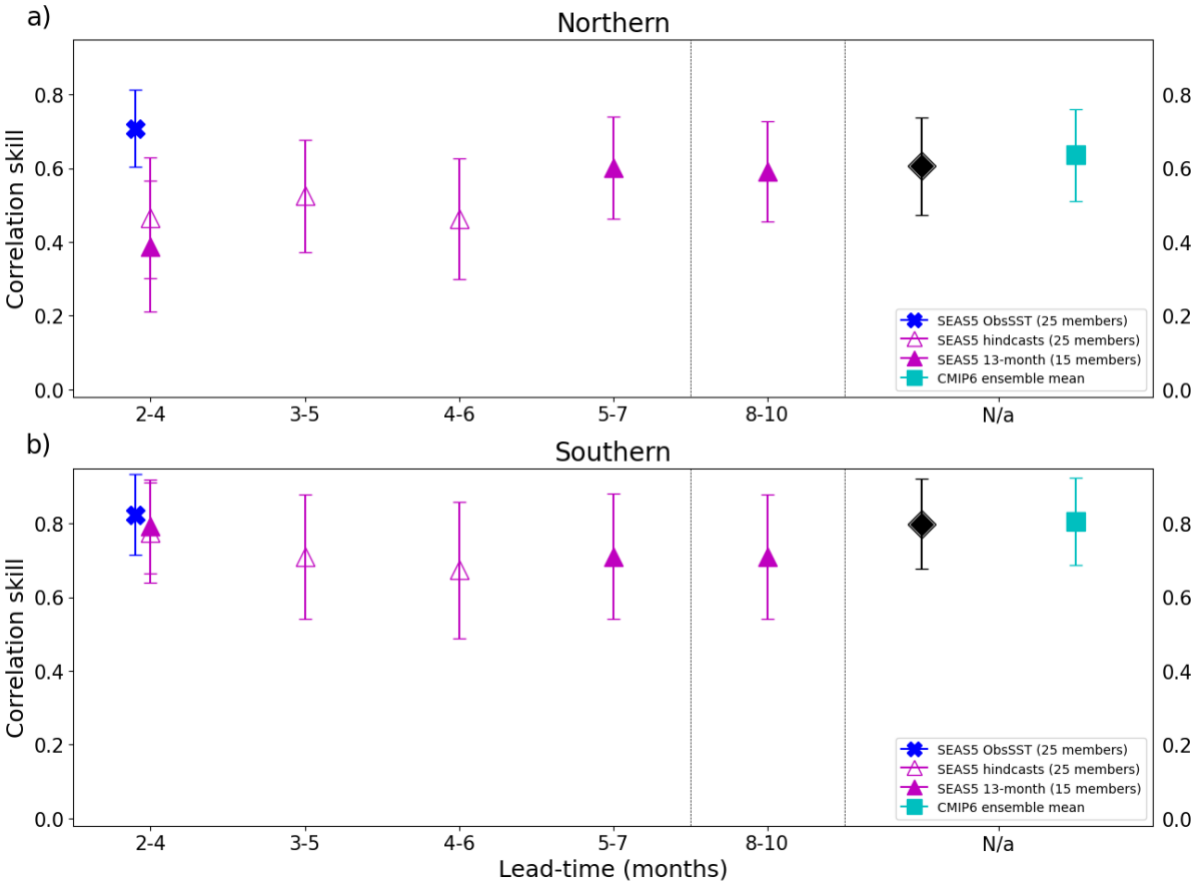

Figure S13: The same as figure 4, but for the period 1982-2016 rather than 1993-2016. Excludes C3S as data is not available for all models before 1993.

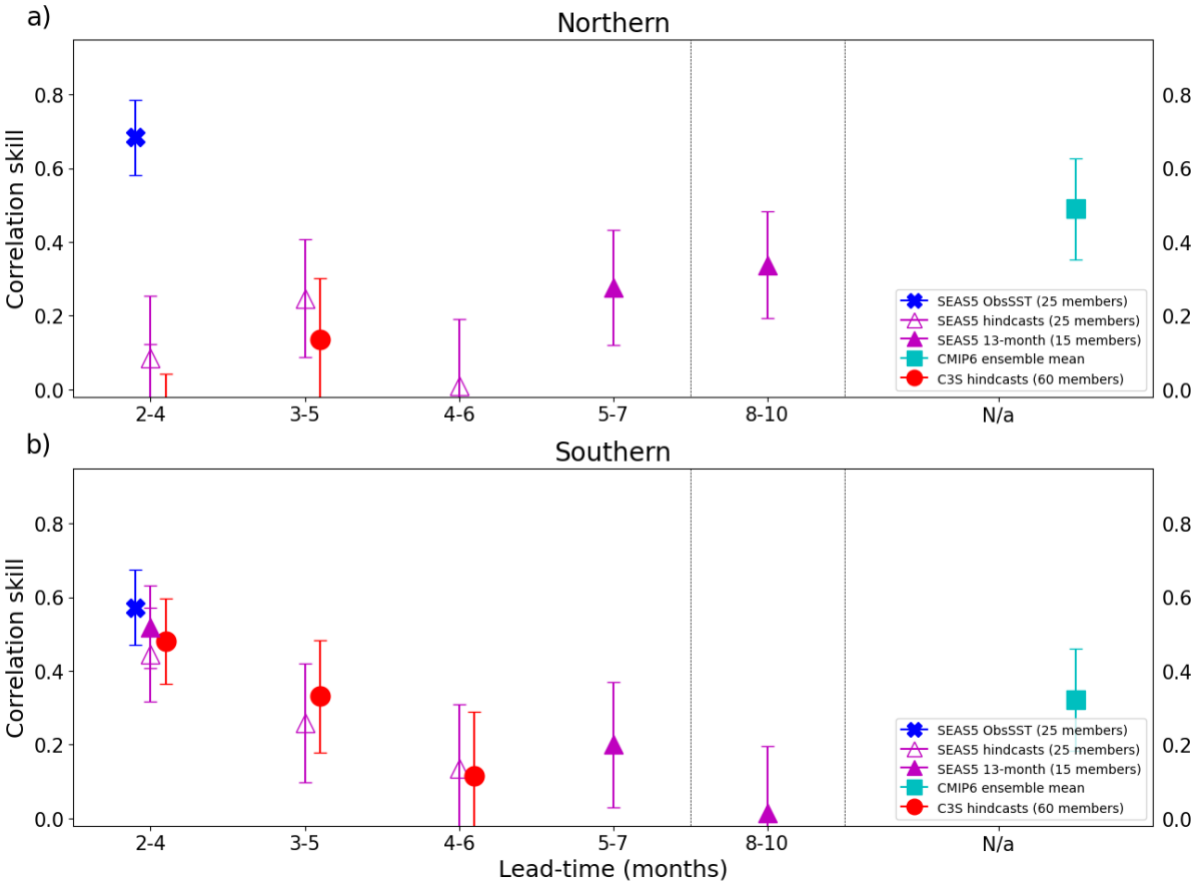

Figure S14: The same as figure 4, but time series' have been detrended before calculating correlations.

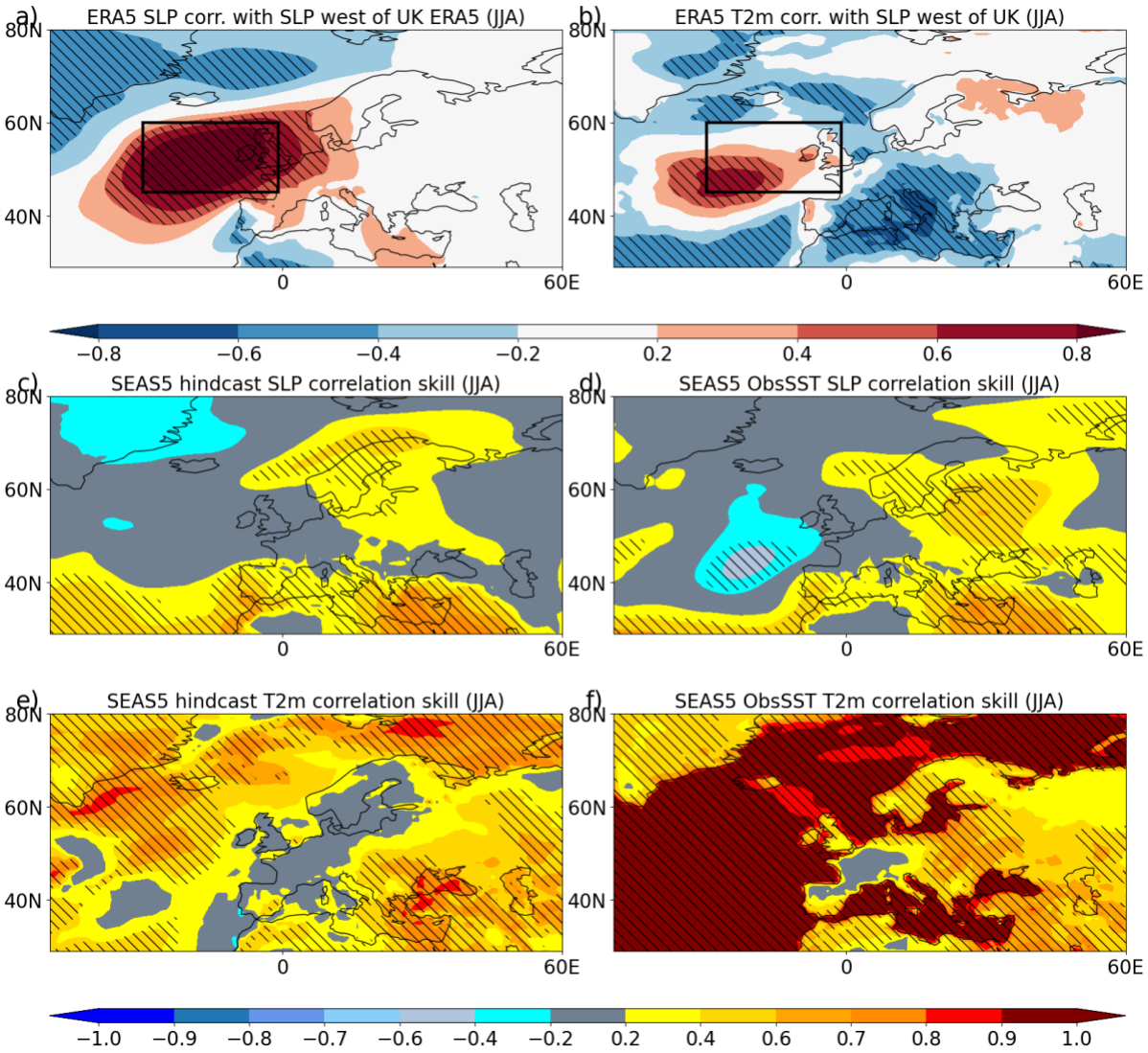

Figure S15: Correlation of the ERA5 mean sea level pressure (SLP) to the west of the UK, calculated over the boxed region, with a) SLP and b) T2m. The following panels show anomaly correlation skill in SEAS5 for the c,e) coupled and d,f) ObsSST (atmosphere-only) hindcast runs for c,d) SLP and e,f) T2m.

Motivated by the fact that there appears to be negative skill for SLP to the west of the UK in the C3S ensemble (figure S5a), figure S14 shows the link between SLP in this region and SLP / T2m over Europe. Figure S14b indicates that high SLP to the west of the UK is not strongly correlated with northern Europe T2m, but is associated with cold anomalies over southern Europe. The following panels explore the fact that T2m skill is higher in the ObsSST runs for northern Europe (figure 4a of the main paper). The atmospheric circulation as indicated by SLP is not improved in the ObsSST runs suggesting that this is not the cause of the higher skill (figure S14c,d). On the other hand, T2m skill over the ocean, particularly close to western European land, is substantially higher in the ObsSST experiments suggesting that advection of warm / cold anomalies by the mean winds might be the cause of the improvements in skill.

Northern Europe Summer 2022 forecast

Motivated by the performance of the CMIP6 ensemble in reproducing observed northern European T2m (using the region defined in the main paper), we produce a forecast for this region for the upcoming summer using the CMIP6 ensemble. In order to make the forecast, we bias-correct each CMIP6 model based on ERA5 (1993-2016) and calibrate the forecast using the method of Doblas-Reyes et al (2005).

Specifically, taking z_{ij} is the bias-corrected CMIP6 projection for northern European T2m at time i for ensemble member j and if z_i is the ensemble-mean at time i, we denote the calibrated ensemble member by w_{ij}.

$$w_{ij} = \alpha z_i + \beta z_{ij}$$

The constants, α and β are constrained such that the standard deviation of the prediction is the same as the standard deviation of the reference time series, s_r (in this case from ERA5) and secondly that the correlation, ρ, is unchanged between the calibrated prediction and reference does not change due to the calibration. Doblas-Reyes et al (2005) use the following solution to these constraints:

$$\alpha = abs(\rho)\frac{s_r}{s_{em}}$$

$$\beta = \sqrt{1-\rho^2}\frac{s_r}{s_e}.$$

Here, s_{em} is the ensemble-mean standard deviation and s_e is the square root of the mean variance of the ensemble.

An examination of the performance of the calibrated SEAS5 predictions and CMIP6 projections for the years 2017-2022 is shown in figure S15. The calibrated CMIP6 projections show slightly better performance over the 2017-2021 period with a root mean square error of 0.37K vs 0.58K, though the period is too short to draw any firm conclusions. Nevertheless, this is consistent with our results from the hindcast period.

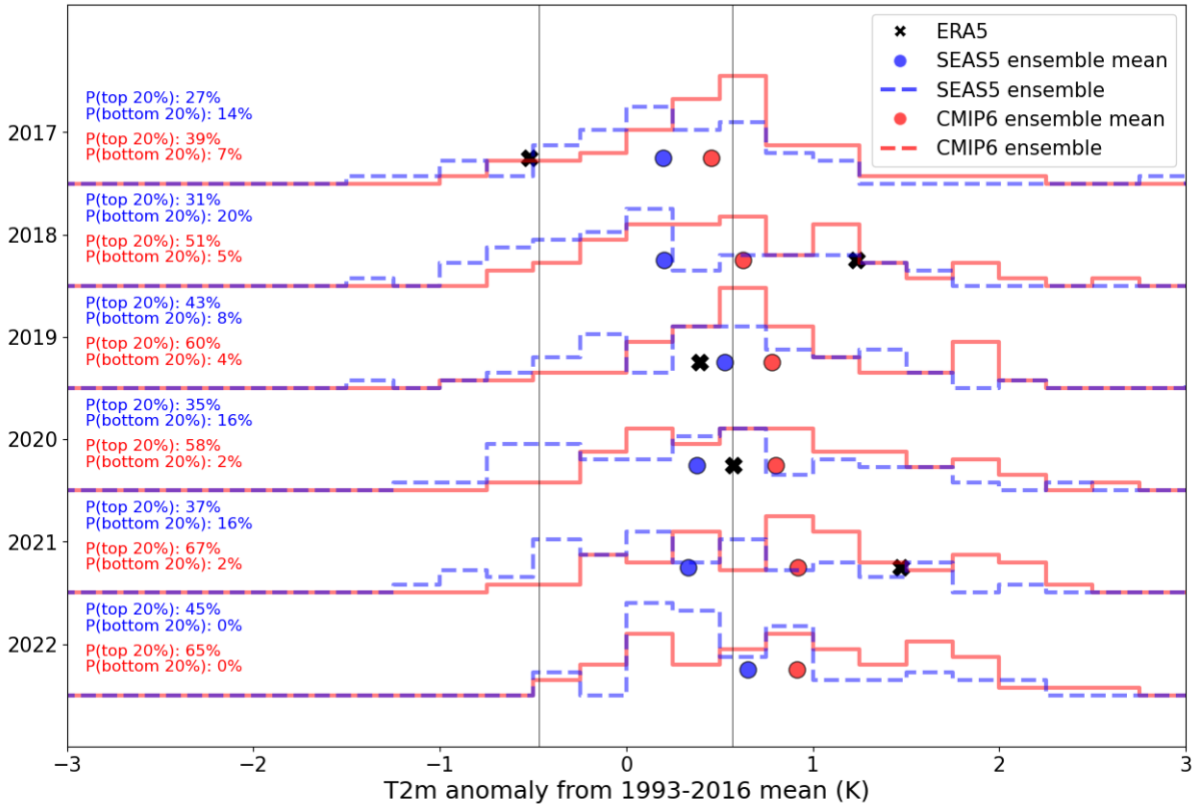

Figure S16: Northern Europe T2m forecasts for the years 2017-2022 based on SEAS5 and CMIP6 ensembles. In each case the ensembles have been bias-corrected with reference to ERA5 (1993-2016) and calibrated following the method of Doblas-Reyes et al (2005). Histograms show the resulting predictions for the ensembles for each year (CMIP6 = red, solid; SEAS5 = blue, dashed). Percentages on the left show indicate the proportions of ensemble members falling below the 20th percentile of ERA5 1993-2016 values and proportions of ensemble members above the 80th percentile. Vertical lines indicate the 20th and 80th percentiles of northern Europe T2m from ERA5 1993-2016.

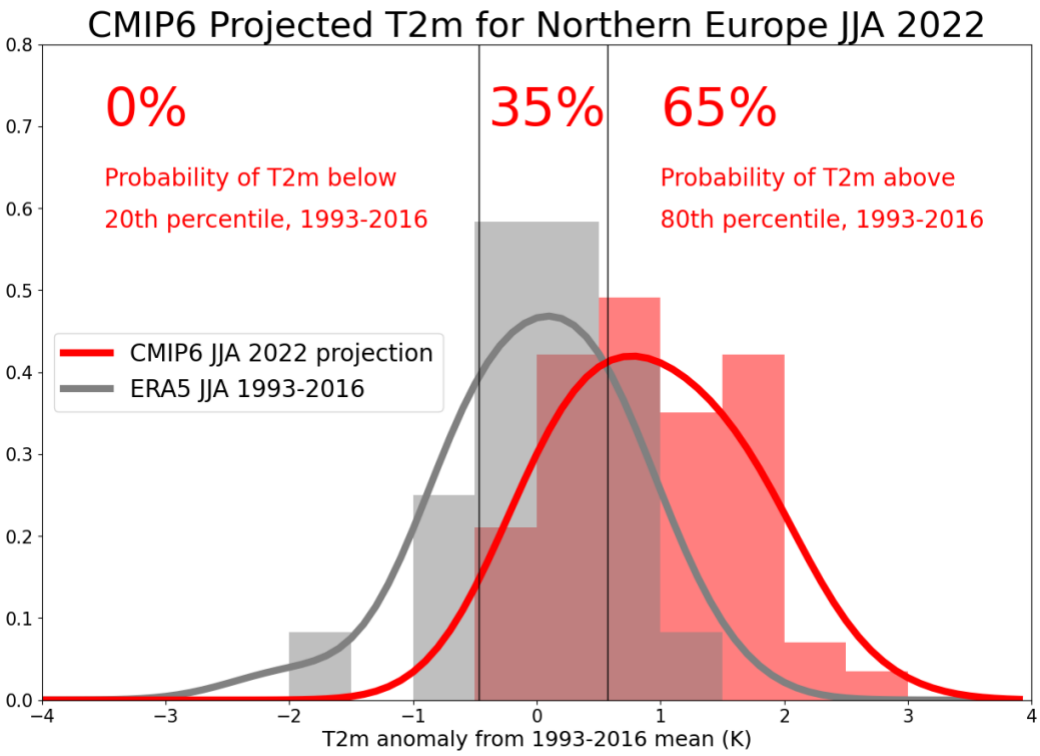

Figure S17: Bias-corrected and calibrated projections for northern European T2m anomalies for JJA 2022 with respect to ERA5 1993-2016. Vertical lines indicate the 20th and 80th percentiles of northern Europe T2m from ERA5 1993-2016.

The outcome of the JJA 2022 prediction in CMIP6 is shown in figure S16 and indicates that northern Europe has a 65% chance of exceeding the 80th percentile of JJA-mean T2m for JJA 2022 with reference to 1993-2016 and 0% chance of falling below the 20th percentile.

References:

Doblas-Reyes, F., Hagedorn, R. & Palmer, T. (2005) "The rationale behind the success of multi-model ensembles in seasonal forecasting – II. Calibration and combination." Tellus A. Volume 57, Issue 3. Pages 234-252