

LETTER • OPEN ACCESS

## The strong role of external forcing in seasonal forecasts of European summer temperature

To cite this article: Matthew Patterson *et al* 2022 *Environ. Res. Lett.* **17** 104033

View the [article online](#) for updates and enhancements.

### You may also like

- [Internal variability in European summer temperatures at 1.5 °C and 2 °C of global warming](#)  
Laura Suarez-Gutierrez, Chao Li, Wolfgang A Müller et al.
- [Ocean and atmosphere influence on the 2015 European heatwave](#)  
J V Mecking, S S Drijfhout, J J-M Hirschi et al.
- [Seasonal forecasts of Eurasian summer heat wave frequency](#)  
Rui Zhi Zhang, Xiao Jing Jia and Qi Feng Qian

ENVIRONMENTAL RESEARCH  
LETTERS

## LETTER

## OPEN ACCESS

RECEIVED  
28 January 2022REVISED  
12 August 2022ACCEPTED FOR PUBLICATION  
15 September 2022PUBLISHED  
3 October 2022

Original Content from  
this work may be used  
under the terms of the  
[Creative Commons  
Attribution 4.0 licence](#).

Any further distribution  
of this work must  
maintain attribution to  
the author(s) and the title  
of the work, journal  
citation and DOI.

The strong role of external forcing in seasonal forecasts  
of European summer temperatureMatthew Patterson<sup>1,\*</sup> , Antje Weisheimer<sup>1,2,3</sup> , Daniel J Befort<sup>1</sup> and Christopher H O'Reilly<sup>4</sup> <sup>1</sup> Atmospheric, Oceanic and Planetary Physics, University of Oxford, Parks Road, OX1 3PU Oxford, United Kingdom<sup>2</sup> European Centre for Medium-Range Weather Forecasts (ECMWF), Shinfield Park, RG2 9AX Reading, United Kingdom<sup>3</sup> National Centre for Atmospheric Science (NCAS), University of Oxford, Parks Road, OX1 3PU Oxford, United Kingdom<sup>4</sup> Department of Meteorology, University of Reading, Whiteknights Road, Earley Gate, RG6 6ET Reading, United Kingdom

\* Author to whom any correspondence should be addressed.

E-mail: [matthew.patterson@physics.ox.ac.uk](mailto:matthew.patterson@physics.ox.ac.uk)**Keywords:** seasonal forecasts, climate change, EuropeSupplementary material for this article is available [online](#)

## Abstract

Since the 1980s, external forcings from increasing greenhouse gases and declining aerosols have had a large effect on European summer temperatures. These forcings may therefore provide an important source of forecast skill, even for timescales as short as a season ahead. However, the relative importance of external forcings for seasonal forecasts has thus far received little attention, particularly on a regional scale. In this study, we investigate forcing-induced skill by comparing the near-surface temperature skill of a multi-model ensemble of seasonal predictions from the Copernicus Climate Change Service archive to that of an uninitialised ensemble of Coupled Model Intercomparison Project phase 6 projections for European summers (June–July–August) spanning the years 1993–2016. As expected, predictive skill over southern Europe is larger for initialised seasonal predictions compared to uninitialised climate projections. However, for northern Europe, we find that predictive skill is generally small in current seasonal models and surprisingly even smaller compared to uninitialised climate projections. These results imply that further research is necessary to understand the role of external forcing on seasonal temperature variations over Europe.

## 1. Introduction

Accurate seasonal forecasts of European summers have the potential to provide useful information to a wide range of stakeholders including the water (Samaniego *et al* 2019), energy (Bett *et al* 2022) and agricultural (Calanca *et al* 2011) sectors. However, whilst current forecast systems have shown relatively high skill in simulating European winter variability (e.g. Scaife *et al* 2014, Athanasiadis *et al* 2017), accurate predictions of summer circulation have proved to be more challenging (e.g. Weisheimer *et al* 2011, Maclachlan *et al* 2015, Johnson *et al* 2019). This dichotomous behaviour may result from the weak El Niño Southern Oscillation signal in boreal summer (O'Reilly *et al* 2018, 2019) and the relative inactivity of the stratosphere, both of which are significant sources of winter circulation predictability (Domeisen *et al* 2015).

In spite of this, some studies have shown potential sources of skill for summer seasonal forecasts such as soil moisture feedbacks (Seneviratne *et al* 2010, Quesada *et al* 2012, Miralles *et al* 2014, Prodhomme *et al* 2016, 2022, Ardilouze *et al* 2017), Caribbean and eastern tropical Pacific rainfall (Cassou *et al* 2005, Wulff *et al* 2017, O'Reilly *et al* 2018, Rieke *et al* 2021) and North Atlantic sea surface temperatures (SSTs) (Ossó *et al* 2018, Dunstone *et al* 2019). Dunstone *et al* (2018) demonstrated that there is weak to moderate skill ( $R \approx 0.5$ ) in predicting European summer rainfall at 2–4 month lead-time when using a large ensemble of simulations from the Met Office HadGEM3-GC2 model. It also appears that extremely hot summers are more predictable than others due to land–atmosphere feedbacks and persistent circulation patterns (Wulff and Domeisen 2019).

Additionally, Doblas-Reyes *et al* (2006) and Liniger *et al* (2007) showed that seasonal hindcasts

forced with observed greenhouse gas forcing perform substantially better than a set of hindcasts with constant greenhouse gases. One typically thinks of seasonal forecasts as obtaining skill via initialisation with observed climate conditions. However, given the large trends in boundary conditions such as greenhouse gases and aerosols, these likely also represent a significant source of predictability, particularly when the year-to-year variability is low, as is often the case in summer. The role of this external forcing in predictability is frequently referenced in the decadal prediction community (Smith *et al* 2019, Borchert *et al* 2021), but rarely is this discussed in the context of seasonal forecasts.

The aim of this work is to investigate the extent to which external forcing provides a source of near-surface temperature skill in the European summer season. This paper begins with a brief description of the seasonal hindcast experiments in section 2 followed by an examination of the skill in these hindcasts in comparison to uninitialised Coupled Model Intercomparison Project phase 6 (CMIP6) models in section 3, an examination of the skill in different European regions in section 4 and finishing with discussion and conclusions in section 5.

## 2. Data and methods

### 2.1. Seasonal hindcasts

In this study we make use of monthly hindcast data from five models in the Copernicus Climate Change Service (C3S) archive spanning the common period 1993–2016. The models are the ECMWF System 5 seasonal forecast model (SEAS5) (Johnson *et al* 2019), CMCC Operational Seasonal Prediction System SPS3 (Gualdi *et al* 2020), DWD German Climate Forecasting System Version 2 (Fröhlich *et al* 2021), UK Met Office's GloSea6 (Williams *et al* 2018) and Meteo-France System 7 (Batte and Deque 2016). All members from the CMCC, DWD and ECMWF systems are initialised on 1 May, whereas the Met Office and Meteo-France systems include members initialised on 1 May and at lagged start dates. For all models we analyse months 2–4, i.e. June–July–August (JJA) and take the first 12 members from each model (for models with lagged start dates we take the most recent 12), combining them to form a 60 member ensemble. We specifically choose an ensemble of 60 members for the seasonal hindcasts to compare with the same number of CMIP6 model runs (detailed below), however the results of this study are unchanged if 25 members are chosen from each model to form a 125 member ensemble (supplementary figure S1). Note that for each year, the seasonal hindcast models are forced with fixed greenhouse gas concentrations, but the concentrations increase with each subsequent year. Consequently, the models all contain some representation of the greenhouse gas-forced warming trend, although the specific scenario of greenhouse

gas forcing varies between models. On the other hand, sulphate aerosols are represented by a seasonally varying climatology in all models apart from the DWD and ECMWF models which include the tropospheric sulphate trend. For further details on the forcing see supplementary table S1.

In addition to the multi-model seasonal forecast ensemble, we utilise several experiments using SEAS5 including a 25 member, atmosphere-only hindcast forced with observed SSTs (known as ObsSSTs) and an ensemble of 15 member, 13 month long hindcasts, created by extending the first 15 members of the main coupled hindcast.

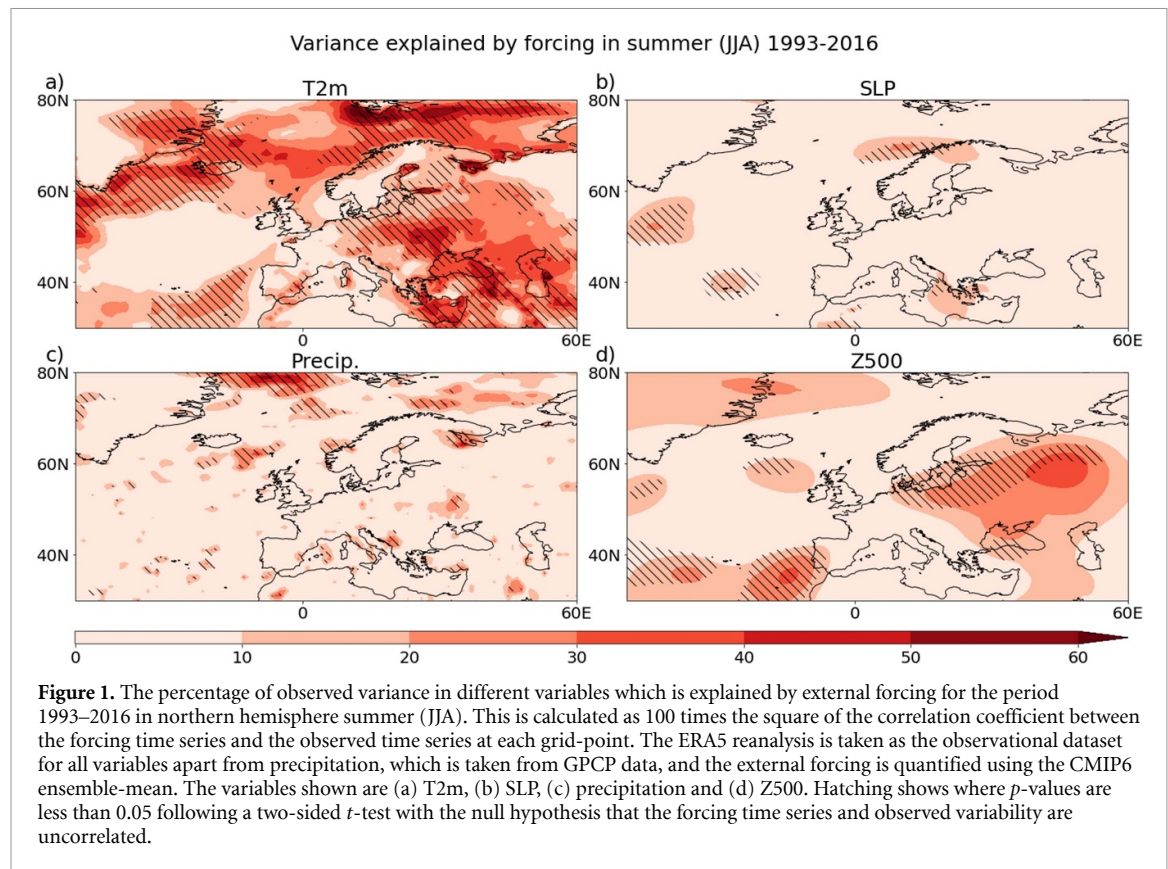
### 2.2. Other datasets

We compare the seasonal hindcasts to an ensemble of 60 uninitialised CMIP6 simulations (Eyring *et al* 2016) to quantify the forced response to external forcings over the C3S model hindcast period. Given that historical experiments in CMIP6 end in 2014, we combine these with the first two years of the scenario run, ssp585, to create the full period of 1993–2016. A full list of models used in this study is provided in supplementary table S2.

Finally, we compare the model data with ERA5 reanalysis (Hersbach *et al* 2020) and the Global Precipitation Climatology Project (GPCP) (Adler *et al* 2018) which we take as observational reference datasets.

### 2.3. Statistical significance

Many standard statistical tests assume that data points are independent, an assumption which is often violated by atmospheric variables due to serial correlation in time, particularly on daily timescales. This study largely concerns interannual atmospheric variability of 2m temperature (T2m) which exhibits some serial correlation in reanalysis on this timescale. For the European and wider North Atlantic system, the lag-1 autocorrelation in T2m is less than 0.1 over much of the North Atlantic with the exception of the sub-polar North Atlantic and eastern Europe (supplementary figure S2(a)). On the other hand, the seasonal hindcast and CMIP6 model ensembles are substantially serially correlated for T2m (supplementary figures S2(c) and (e)). In this study, we account for the reduced sample size associated with serial correlation by introducing an effective sample size for all statistical tests,  $N_{\text{eff}} = N(1 - \rho)/(1 + \rho)$ , following Wilks (2019). Here  $N$  is the number of years in the time series and  $\rho$  is the observed lag-1 autocorrelation in the verifying observed ERA5 data, though strictly for a correlation, the serial correlation of both the observed and forecast should be considered together (Bretherton *et al* 1999). To test whether correlation skill in different simulations is significantly different we use the method of Siegert *et al* (2017). However, we note that this method is not strictly valid due to the presence of some serial correlation.



### 3. Results

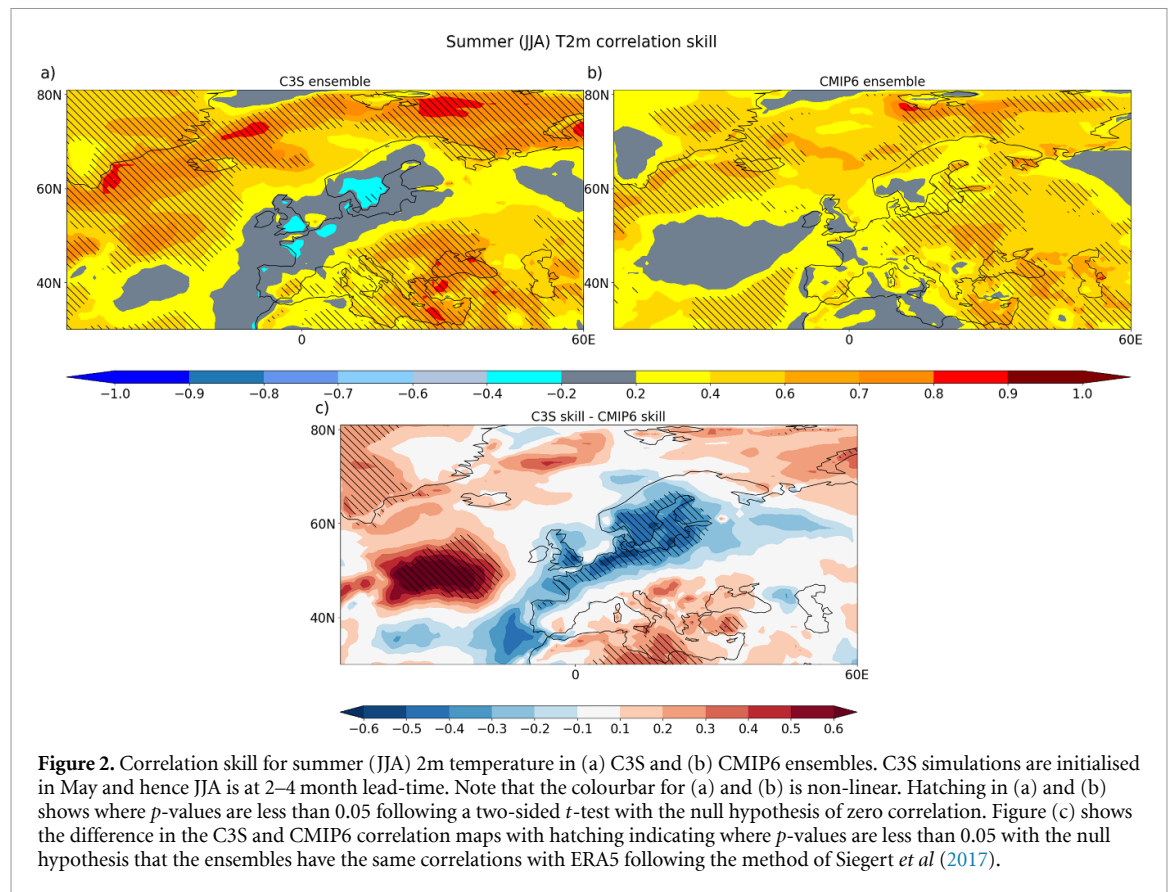
To what extent does external forcing explain recent temperature variability and does it contribute substantially to circulation variability? To investigate this we show the proportion of observed summertime variance explained by an ensemble of CMIP6 models over the period 1993–2016 for T2m, precipitation and summary circulation variables, sea level pressure (SLP) and 500 hPa geopotential height (Z500) in figure 1. In spite of the relative shortness of the period, external forcing explains between 20% and 40% of the variance in T2m at individual grid-points over most of central and eastern Europe and the sub-polar to polar North Atlantic (figure 1(a)). This suggests that even uninitialised simulations could provide skill in summer season T2m over these regions. For comparison, in winter, external forcing explains less than 10% of the variance at individual grid-points over almost the whole of Europe (supplementary figure S3(a)). External forcing explains even more of the summer T2m variance when considering the longer period of 1981–2020, with the forcing explaining 30%–60% of the T2m variability at individual grid-points over eastern Europe (supplementary figure S4(a)). This seems reasonable given that the longer period allows the trend to emerge more clearly from interannual variability.

On the other hand, neither SLP nor precipitation is dominated by external forcing (figures 1(b) and (c)) which is unsurprising given the large interannual

variability in extratropical atmospheric circulation. Observations within this period do show a large, positive trend in Greenland blocking, however this is not captured by current climate models (Hanna *et al* 2018, Davini and D’Andrea 2020) and hence does not feature in figure 1(b). Interestingly, external forcing does explain over 20% of Z500 variance over Russia and eastern Europe (figure 1(d)). Surface temperature warming trends influence Z500 trends and the region of strongest warming being over eastern Europe might explain the localisation of this Z500 anomaly, though there may also be a dynamical component. Moreover, for the longer period of 1981–2020, the region in which external forcing of Z500 explains the most variance is shifted further south, co-located with the strongest warming signal (supplementary figure S4(d)).

We now consider how much T2m skill is linked to external forcing alone by calculating the anomaly correlation coefficient of the CMIP6 ensemble with ERA5 data, shown in figure 2. We then compare this to the skill of the ensemble of seasonal hindcasts in the C3S archive for simulations initialised in May for the period 1993–2016. The CMIP6 ensemble should provide a baseline level of skill given that the C3S and CMIP6 ensembles have the same number of ensemble members and contain largely the same external forcing such as greenhouse gas variations as the CMIP6 models, but the C3S ensemble should derive additional skill from initialisation of the land, atmosphere and oceans. The C3S models have significant skill in





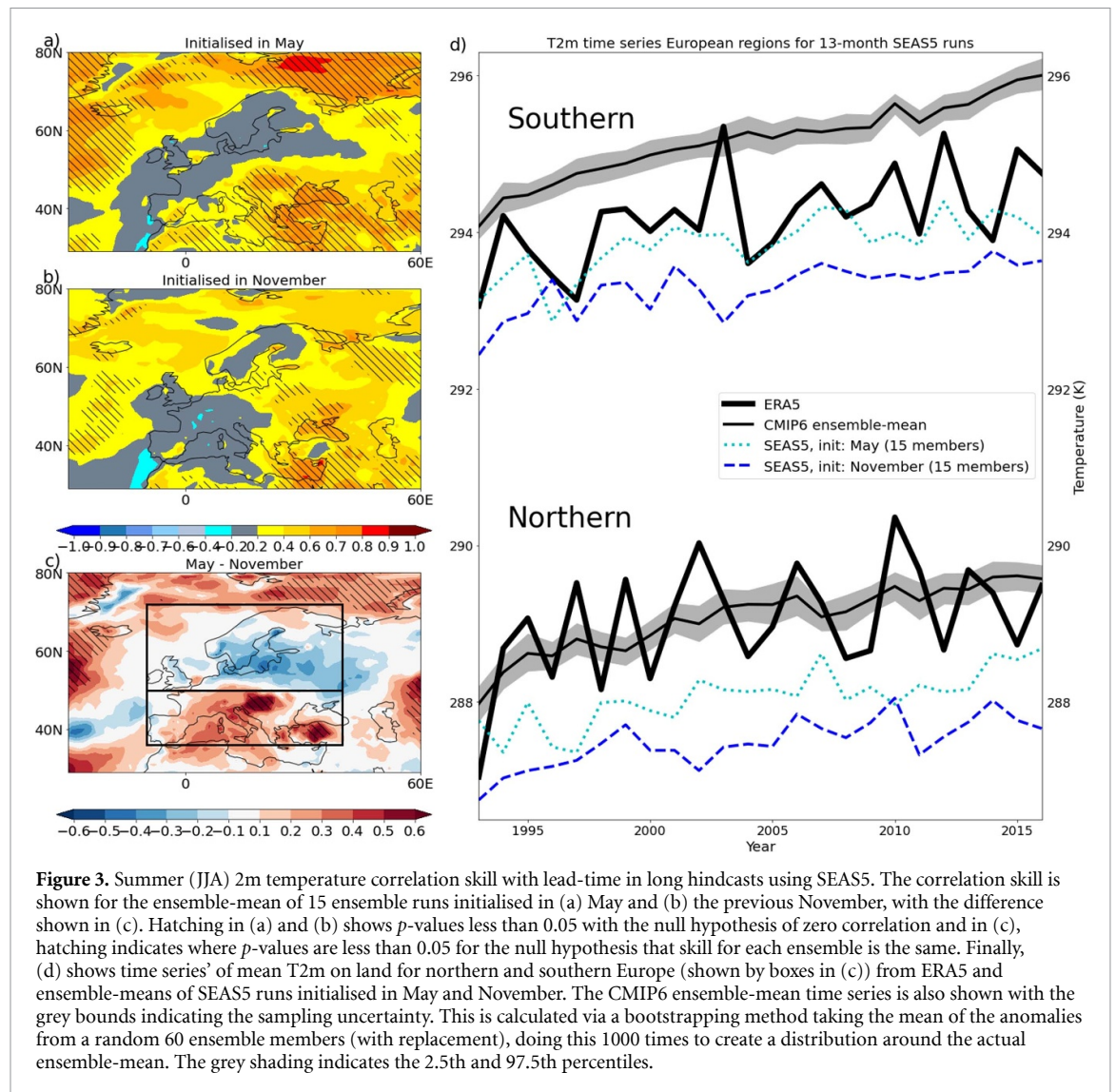
predicting T2m over central, eastern and southern parts of Europe (figure 2(a)), which is to be expected given the strong trends in these regions under external forcing (figure 1(a)). There is, however, no significant skill over northern and western Europe, including the United Kingdom (UK) and Scandinavia (figure 2(a)). This is consistent with a lack of skill in SLP in mid-latitudes (supplementary figure S5) as atmospheric circulation, while less dominant than in winter, still plays a strong role in northern and western European summer T2m variability (supplementary figure S6) (Folland *et al* 2009). The C3S ensemble actually shows negative SLP skill over and to the west of the UK (supplementary figure S5(a)), though linearly regressing out SLP in this region does not improve the T2m skill in C3S (supplementary figure S7). On the other hand, in subtropical regions there is significant SLP skill (supplementary figure S5(a)) which may contribute to T2m skill in the Mediterranean (figure 2(a)).

Comparing skill from the CMIP6 ensemble with that of the C3S ensemble reveals that not only does the CMIP6 ensemble exhibit significant skill in T2m over central and eastern Europe, where external forcing plays a strong role, but it also appears to show greater skill over northern and western Europe (figure 2(c)). It should be noted that much of this T2m skill is not statistically significant over the UK and Scandinavia, however, it is significantly larger than C3S (hatching in figure 2(c)) suggesting that it genuinely performs better. This is particularly surprising given that the

C3S ensemble SSTs are more skilful over much of the North Atlantic (supplementary figure S8) as a result of initialisation and hence should provide some additional skill for T2m through advection or dynamical processes (figures 2(a) and (c)). T2m over the Mediterranean and North Atlantic has a higher correlation skill in the C3S ensemble than in CMIP6 potentially highlighting the beneficial role of initialisation in these regions. Note that all of these results also hold if the forecast quality is measured using the root mean square error (supplementary figure S9) and that the individual C3S models all show very similar patterns of T2m anomaly correlation skill to the full ensemble (supplementary figure S10).

#### 4. Spatial and lead-time dependent variations in European T2m skill

The result that T2m skill is higher in the CMIP6 ensemble is surprising given the extra information provided by initialisation. In this section we seek to investigate this result further by considering a number of additional experiments using SEAS5 as this model behaves reasonably similarly to the C3S ensemble-mean (supplementary figure S10(b)). One might anticipate that in the limit of long lead-times, information from the initial conditions will become less important and hence the ensemble-mean of initialised predictions will tend towards that of uninitialised projections. In this case, the SEAS5 forecasts of

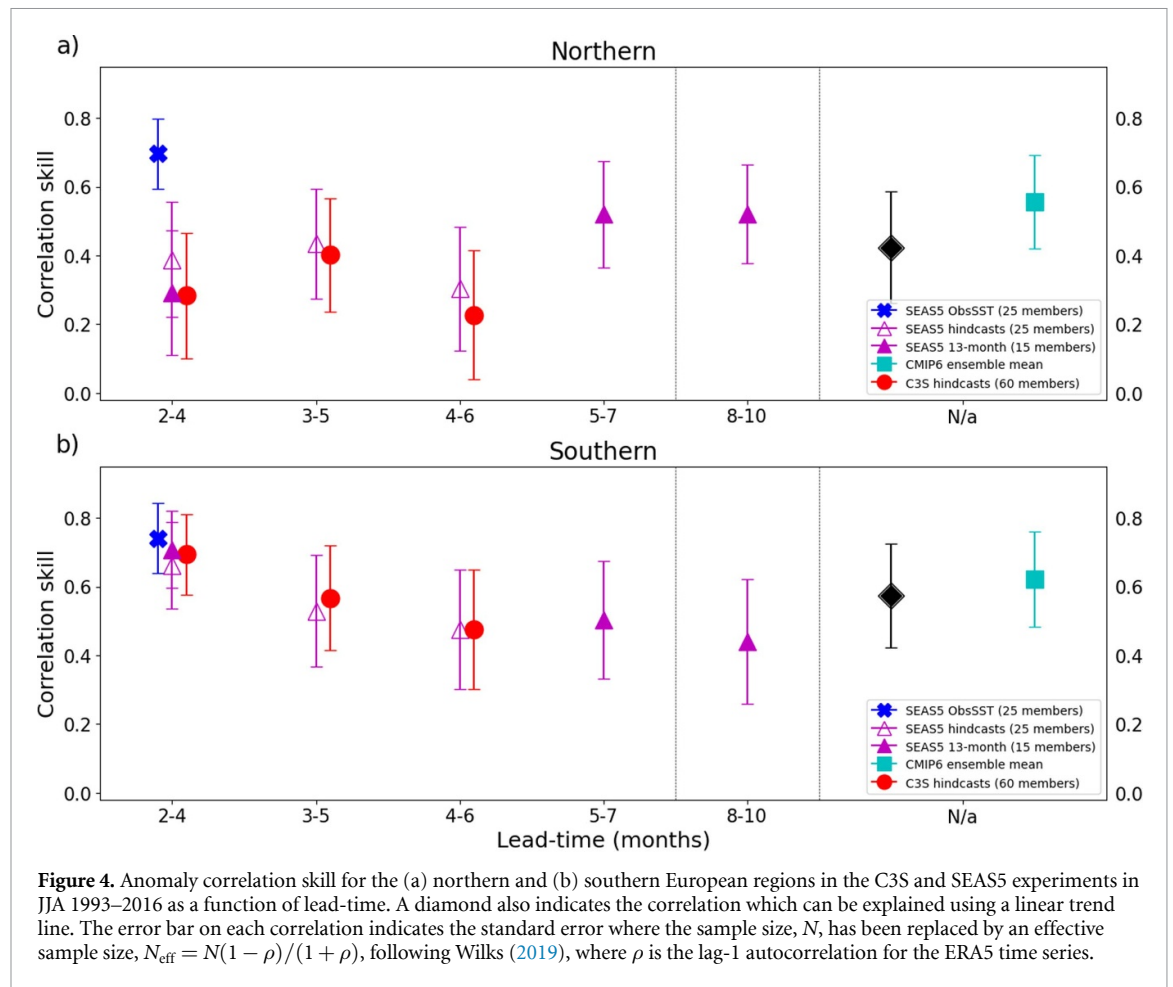


T2m would actually improve at longer lead-times in those regions where the CMIP6 ensemble has more skill than C3S. This is indeed what we see in 13 month hindcast runs. Figure 3 shows the T2m skill in JJA for runs initialised in May (i.e. 2–4 month lead-time) and November (8–10 month lead-time) using 15 ensemble members. The difference pattern (figure 3(c)) is relatively similar to the difference between CMIP6 and C3S ensemble skill (figure 2(e)) in terms of the fact that skill is lower over northern Europe with a shorter lead-time and at the same time improved over southern Europe. A similar pattern is seen when comparing hindcasts initialised in May and February (supplementary figure S11).

The slight improvement with lead-time for northern Europe could come about for a number of different reasons. Firstly, the model may be responding incorrectly to initial conditions like SSTs or soil moisture and hence this may degrade the forecast at short lead-times. For instance, the predecessor of SEAS5, System 4 (Molteni *et al* 2011), fails to correctly represent the observed atmospheric teleconnection

from the tropical Pacific to the Atlantic in summer (O'Reilly *et al* 2018). A second possibility is that some initial conditions may be incorrect. For example, decadal SST variability in the winter season over the sub-polar North Atlantic appears to be degraded in SEAS5 with respect to System 4 as a result of new ocean initial conditions (Johnson *et al* 2019, Tietsche *et al* 2020). Thirdly, the seasonal forecast may suffer from an initialisation shock (e.g. Balmaseda and Anderson 2009, Hudson *et al* 2011, Magnusson *et al* 2013, Hermanson *et al* 2018) as the model climatology differs from that of the real earth system (supplementary figure S12). The exact reason for this discrepancy is beyond the scope of this article but its existence suggests a potential area of improvement in summer seasonal forecasts.

Time series' of T2m averaged over land in northern and southern European regions are also instructive in terms of how variability differs between the different parts of Europe and with lead-time (figure 3(d)). These regions are both defined by the longitudes 10° W–40° E and by 36° N–50° N and



50° N–72° N for southern and northern Europe, respectively. The boundaries of these regions are also shown in figure 3(c). ERA5 T2m in the southern region shows a relatively strong warming trend, suggesting that the trend can provide a large amount of predictability. The SEAS5 prediction initialised in May manages to capture both this upward trend and some of the observed interannual variation. The CMIP6 ensemble-mean time series is also plotted for comparison and appears to be steeper in the 1990s before warming more slowly in the 2000s, coinciding with the so-called global warming ‘hiatus’ period (e.g. Medhaug *et al* 2017), subsequently becoming steeper once again.

Northern Europe in ERA5 also shows a positive trend and particularly large interannual variability, little of which appears to be well captured in the SEAS5 ensemble-means for either May or November initialisations. Note that the year 1993 appears to have been particularly cold over northern Europe, accentuating the trend. The large year-to-year variability likely comes as a result of atmospheric circulation, which plays a substantial role in northern European T2m variability (supplementary figure S6).

Figure 4 provides a summary of the lead-time dependence of correlation skill in the northern and southern European regions. For northern Europe, JJA

skill increases for longer lead-times for the 13 months SEAS5 runs from  $\approx 0.3$  at months 2–4 to  $\approx 0.5$  at months 5–7 and stays at a similar level for months 8–10. However, this same pattern of increasing skill with lead-time is not seen for the 25 member SEAS5 hindcast between 2–4 and 4–6 months. The full C3S ensemble also shows similar correlations for lead-times of 2–4 and 4–6 months to the 25 member SEAS5 hindcast. Furthermore, the short time series means that the uncertainty on each correlation is considerable with some overlap of the error bars on the 2–4 month and 5–7 month for the 13 month SEAS5 runs. Moreover, there is some uncertainty associated with ensemble size as demonstrated by the slight difference between 15 and 25 member SEAS5 ensembles for northern Europe, initialised in May. Note that removing the particularly cold year, 1993, from the time series causes the SEAS5 2–4 month, 5–7 and 8–10 month correlations to drop to 0.23, 0.35 and 0.34 respectively. Nevertheless, when a longer period of 1982–2016 is used instead (supplementary figure S13), the results are similar to figure 4 with northern Europe T2m skill increasing noticeably with lead-time for the 13 month runs, suggesting that this result is indeed robust. What is clear is that the northern European skill in the hindcasts is unexpectedly low at short lead-times when compared to both skill at

longer lead-times and skill of the forcing-only CMIP6 ensemble.

Notably, the ObsSST SEAS5 hindcast in which the model is forced with ObsSSTs, shows relatively high correlation skill at 2–4 month lead-time, with a correlation of about 0.7 over northern Europe. This improved skill in the ObsSST runs may occur for thermodynamic reasons, e.g. due to advection of heat anomalies over land by the mean winds (O'Reilly *et al* 2017) as dynamical skill is not improved in the ObsSST relative to the coupled hindcasts (supplementary figures S15(c) and (d)). One interpretation of this is that improvements to skill in SSTs might improve northern European T2m skill in the fully coupled runs. On the other hand, it is worth noting that some SST variability may be inherently unpredictable, hence using atmosphere-only experiments may overestimate the true potential for improved T2m skill.

The European temperature forcing is a combination of the direct effect of greenhouse gas forcing, forcing from warming SSTs and forcing from the decline in aerosol emissions (Dong *et al* 2017). Consequently, the forcing is likely to vary both spatially and non-linearly with time and hence provide more information than a simple linear trend (figure 3(d)). However, is the CMIP6 ensemble more skilful than a simple linear trend? In figure 4(a) it appears that the CMIP6 ensemble-mean correlation skill is slightly higher than that of a linear trend (the slope of the trend has no bearing on the correlation), though the short time series and substantial error bars again prevent us from drawing firm conclusions.

T2m correlation skill in the southern European region is high for 2–4 month lead-times at around 0.7 for both the 25 member SEAS5 hindcasts and 60 member C3S ensemble and decays with lead-time (figure 4(b)), but is higher than the skill from the CMIP6 ensemble or a linear trend. This suggests that skill comes not just from the external forcing, but that the model derives significant skill from the initial conditions, unlike for northern Europe. This is consistent with the skill in atmospheric circulation seen in the Mediterranean in contrast to the lack of skill further north (supplementary figure S5(a)). The ObsSST hindcasts perform only slightly better than the coupled hindcasts at 2–4 month lead-time implying that the relevant SST variability to the T2m in the southern Europe region is generally well captured in the coupled hindcasts.

Finally, motivated by the considerable societal interest in seasonal predictions of summer T2m and by our results, it is of interest to consider the forecasts made for summers since the hindcast period and how these compare to the CMIP6 ensemble. Specifically, we apply a simple bias-correction and then calibrate the T2m SEAS5 predictions and CMIP6 projections for the northern Europe region for the 2017–2022 summers using the method of Doblas-Reyes *et al*

(2005), based on the models' 1993–2016 climatologies and using ERA5 as a reference. This allows us to produce probabilistic forecasts for T2m in northern Europe based on these ensembles. Interestingly, the ensemble-means of the forecasts using the CMIP6 ensemble have a slightly lower root mean square error than the calibrated SEAS5 predictions over the period 2017–2021 (0.37 K vs 0.58 K, supplementary figure S16), consistent with our results. The JJA 2022 CMIP6 prediction yields an ensemble with 65% of members exceeding the 80th percentile of the JJA-mean northern European T2m in the 1993–2016 period (supplementary figure S17). The increase in occurrence probability of what previously might have been considered an anomalously warm summer again underlines our argument that external forcing can play a significant role in summer season T2m predictions.

## 5. Discussion and conclusions

Anthropogenically-forced trends in T2m represent a substantial component of European summer T2m variability (figure 1(a)) even on seasonal timescales. Forcings hence can provide a significant amount of skill. We have shown this by comparing a multi-model ensemble of initialised seasonal hindcasts to the uninitialised projections from an ensemble of CMIP6 models for the seasonal hindcast model period of 1993–2016.

Our findings indicate that the relative roles of initial conditions and external forcing vary regionally. Near-surface temperature variability over northern Europe is largely governed by atmospheric circulation and hence low dynamical skill in the seasonal hindcasts means that there is little skill in T2m for this region. It is not clear whether the lack of dynamical skill in this region is due to a poor representation of remote teleconnections in models or whether the potential for predictability is low in this season. Perplexingly, the CMIP6 models appear to show a higher level of T2m skill over northern Europe than the C3S models (figures 2(a) and (c)) and T2m skill in this region actually increases slightly with lead-time (figures 3(a)–(c)). The exact reason for this discrepancy is beyond the scope of this article but could be caused by an incorrect representation of teleconnections, inaccurate initial conditions, or initialisation shocks. Furthermore, this discrepancy suggests that there is potential for improving summer seasonal forecasts if its cause can be identified and addressed. Interestingly, skill for northern European T2m does improve considerably if ObsSSTs are prescribed in an atmosphere-only setup (figure 4(a)), though this improvement does not appear to occur through improvements to circulation. Instead the skill increase may be a result of air, warmed or cooled by North Atlantic SSTs, being advected over land.

Southern Europe, on the other hand, does show both a strong trend and dynamical skill in coupled



hindcasts and T2m is therefore well predicted at short lead-times (figure 4(c)). Moreover, the correlation skill decays with lead-time as one would anticipate, plateauing at around 0.5, likely due to the externally forced trend.

From this study we draw out a number of implications and make some suggestions for further work:

- (a) The magnitude of the warming trend on European summer temperatures means that the uninitialised CMIP6 ensemble shows skill on timescales of seasonal predictions. Depending on their specific requirements, the CMIP6 ensemble may therefore provide useful information concerning summer T2m for some users. On the other hand, the lack of skill for northern Europe in operational seasonal forecast models means that current predictions of T2m for this region should be treated with caution.
- (b) The representation of the externally forced component in seasonal forecast models and the response to this forcing may be important for making accurate forecasts, though confirming this would require additional targeted seasonal forecast experiments. For example, if interannual aerosol variations are not represented in a seasonal forecast model or if the model has an overly high transient climate sensitivity, one might expect that the pattern or magnitude of warming over land will be incorrect.
- (c) This work has shown that mid-latitude atmospheric circulation is poorly captured over the North Atlantic in current seasonal forecast models. It is highly likely that improving skill in circulation in this region will improve summer northern European T2m prediction skill, but improving circulation skill will require further research.

## Data availability statement

No new data were created or analysed in this study.

## Acknowledgments

This work was funded by the EUCP project (Horizon 2020; Grant Agreement 776613). C O R was also supported by a Royal Society University Research Fellowship. We thank the modelling centres within C3S for providing their seasonal hindcast data. We also acknowledge the World Climate Research Programme, which, through its Working Group on Coupled Modelling, coordinated and promoted CMIP6. We thank the climate modelling groups for producing and making available their model output, the Earth System Grid Federation (ESGF) for archiving the data and providing access, and the multiple funding agencies who support CMIP6 and ESGF.

Finally, we thank the three anonymous reviewers for their well-considered comments which helped to substantially improve the manuscript.

## ORCID iDs

Matthew Patterson  <https://orcid.org/0000-0002-9484-8410>

Antje Weisheimer  <https://orcid.org/0000-0002-7231-6974>

Daniel J Befort  <https://orcid.org/0000-0002-2851-0470>

Christopher H O'Reilly  <https://orcid.org/0000-0002-8630-1650>

## References

- Adler R F *et al* 2018 The global precipitation climatology project (GPCP) monthly analysis (new version 2.3) and a review of 2017 global precipitation *Atmosphere* **9** 138
- Ardilouze C *et al* 2017 Multi-model assessment of the impact of soil moisture initialization on mid-latitude summer predictability *Clim. Dyn.* **49** 3959–74
- Athanasiadis P J, Bellucci A, Scaife A A, Hermanson L, Materia S, Sanna A, Borrelli A, MacLachlan C and Gualdi S 2017 A multisystem view of wintertime NAO seasonal predictions *J. Clim.* **30** 1461–75
- Balmaseda M and Anderson D 2009 Impact of initialization strategies and observations on seasonal forecast skill *Geophys. Res. Lett.* **36** L01701
- Batte L and Deque M 2016 Randomly correcting model errors in the ARPEGE-Climate v6.1 component of CNRM-CM: applications for seasonal forecasts *Geosci. Model Dev.* **9** 2055–76
- Bett P E, Thornton H E, Troccoli A, Felice M D, Suckling E, Dubus L, Saint-Drenan Y-M and Brayshaw D J 2022 A simplified seasonal forecasting strategy, applied to wind and solar power in Europe *EarthArXiv Preprint* (<https://doi.org/10.31223/osf.io/kzwqx>) (Accessed 1 December 2021)
- Borchert L F, Menary M B, Swingedouw D, Sgubin G, Hermanson L and Mignot J 2021 Improved decadal predictions of North Atlantic subpolar gyre SST in CMIP6 *Geophys. Res. Lett.* **48** e2020GL091307
- Bretherton C S, Widmann M, Dymnikov V P, Wallace J M and Bladé I 1999 The effective number of spatial degrees of freedom of a time-varying field *J. Clim.* **12** 1990–2009
- Calanca P, Bolius D, Weigel A P and Liniger M A 2011 Application of long-range weather forecasts to agricultural decision problems in Europe *J. Agric. Sci.* **149** 15–22
- Cassou C, Terray L and Phillips A S 2005 Tropical Atlantic influence on European heat waves *J. Clim.* **18** 2805–11
- Davini P and D'Andrea F 2020 From CMIP3 to CMIP6: Northern Hemisphere atmospheric blocking simulation in present and future climate *J. Clim.* **33** 10021–38
- Doblas-Reyes F J, Hagedorn R and Palmer T N 2005 The rationale behind the success of multi-model ensembles in seasonal forecasting—II. Calibration and combination *Tellus A* **57** 234–52
- Doblas-Reyes F J, Hagedorn R, Palmer T N and Morcrette J-J 2006 Impact of increasing greenhouse gas concentrations in seasonal ensemble forecasts *Geophys. Res. Lett.* **33** L07708
- Domeisen D I V, Butler A H, Fröhlich K, Bittner M, Müller W A and Baehr J 2015 Seasonal predictability over Europe arising from El Niño and stratospheric variability in the MPI-ESM seasonal prediction system *J. Clim.* **28** 256–71
- Dong B, Sutton R T and Shaffrey L 2017 Understanding the rapid summer warming and changes in temperature extremes since the mid-1990s over Western Europe *Clim. Dyn.* **48** 1537–54

- Dunstone N *et al* 2018 Skilful seasonal predictions of summer European rainfall *Geophys. Res. Lett.* **45** 3246–54
- Dunstone N, Smith D, Hardiman S, Eade R, Gordon M, Hermanson L, Kay G and Scaife A 2019 Skilful real-time seasonal forecasts of the dry Northern European summer 2018 *Geophys. Res. Lett.* **46** 12368–76
- Eyring V *et al* 2016 Overview of the Coupled Model Intercomparison Project Phase 6 (CMIP6) experimental design and organization *Geosci. Model Dev.* **9** 1937–58
- Folland C K, Knight J, Linderholm H W, Fereday D, Ineson S and Hurrell J W 2009 The summer North Atlantic Oscillation: past, present, and future *J. Clim.* **22** 1082–103
- Fröhlich K *et al* 2021 The German climate forecast system: GCFs *J. Adv. Model. Earth Syst.* **13** e2020MS002101
- Gualdi S *et al* 2020 The new CMCC Operational Seasonal Prediction System *CMCC Technical Notes* TN0288
- Hanna E, Fettweis X and Hall R J 2018 Brief communication: recent changes in summer Greenland blocking captured by none of the CMIP5 models *Cryosphere* **12** 3287–92
- Hermanson L *et al* 2018 Different types of drifts in two seasonal forecast systems and their dependence on ENSO *Clim. Dyn.* **51** 1411–26
- Hersbach H *et al* 2020 The ERA5 global reanalysis *Q. J. R. Meteorol. Soc.* **146** 1999–2049
- Hudson D, Alves O, Hendon H H and Wang G 2011 The impact of atmospheric initialisation on seasonal prediction of tropical Pacific SST *Clim. Dyn.* **36** 1155–71
- Johnson S J *et al* 2019 SEAS5: the new ECMWF seasonal forecast system *Geosci. Model Dev.* **12** 1087–117
- Liniger M A, Mathis H, Appenzeller C and Doblas-Reyes F J 2007 Realistic greenhouse gas forcing and seasonal forecasts *Geophys. Res. Lett.* **34** L04705
- MacLachlan C *et al* 2015 Global Seasonal forecast system version 5 (GloSea5): a high-resolution seasonal forecast system *Q. J. R. Meteorol. Soc.* **141** 1072–84
- Magnusson L, Alonso-Balmaseda M, Corti S, Molteni F and Stockdale T 2013 Evaluation of forecast strategies for seasonal and decadal forecasts in presence of systematic model errors *Clim. Dyn.* **41** 2393–409
- Medhaug I, Stolpe M B, Fischer E M and Knutti R 2017 Reconciling controversies about the ‘global warming hiatus’ *Nature* **545** 41–47
- Miralles D G, Teuling A J, van Heerwaarden C C and De Arellano J V G 2014 Mega-heatwave temperatures due to combined soil desiccation and atmospheric heat accumulation *Nat. Geosci.* **7** 345–9
- Molteni F *et al* 2011 The new ECMWF seasonal forecast system (System 4) (ECMWF) (available at: [www.ecmwf.int/sites/default/files/elibrary/2011/11209-new-ecmwf-seasonal-forecast-system-system-4.pdf](http://www.ecmwf.int/sites/default/files/elibrary/2011/11209-new-ecmwf-seasonal-forecast-system-system-4.pdf))
- O'Reilly C H, Woollings T and Zanna L 2017 The dynamical influence of the Atlantic multidecadal oscillation on continental climate *J. Clim.* **30** 7213–30
- O'Reilly C H, Woollings T, Zanna L and Weisheimer A 2018 The impact of tropical precipitation on summertime Euro-Atlantic circulation via a circumglobal wave train *J. Clim.* **31** 6481–504
- O'Reilly C H, Woollings T, Zanna L and Weisheimer A 2019 An interdecadal shift of the extratropical teleconnection from the tropical Pacific during boreal summer *Geophys. Res. Lett.* **46** 13379–88
- Ossó A, Sutton R, Shaffrey L and Dong B 2018 Observational evidence of European summer weather patterns predictable from spring *Proc. Natl Acad. Sci. USA* **115** 59–63
- Prodhomme C *et al* 2021 Seasonal prediction of European summer heatwaves *Clim. Dyn.* **58** 2149–66
- Prodhomme C, Doblas-Reyes F, Bellprat O and Dutra E 2016 Impact of land-surface initialization on sub-seasonal to seasonal forecasts over Europe *Clim. Dyn.* **47** 919–35
- Quesada B, Vautard R, Yiou P, Hirschi M and Seneviratne S I 2012 Asymmetric European summer heat predictability from wet and dry southern winters and springs *Nat. Clim. Change* **2** 736–41
- Rieke O, Greatbatch R J and Gollan G 2021 Nonstationarity of the link between the Tropics and the summer East Atlantic pattern *Atmos. Sci. Lett.* **22** e1026
- Samaniego L *et al* 2019 Hydrological forecasts and projections for improved decision-making in the water sector in Europe *Bull. Am. Meteorol. Soc.* **100** 2451–72
- Scaife A A *et al* 2014 Skillful long-range prediction of European and North American winters *Geophys. Res. Lett.* **41** 2514–9
- Seneviratne S I, Corti T, Davin E L, Hirschi M, Jaeger E B, Lehner I, Orlowsky B and Teuling A J 2010 Investigating soil moisture–climate interactions in a changing climate: a review *Earth-Sci. Rev.* **99** 125–61
- Siebert S, Bellprat O, Ménégoz M, Stephenson D B and Doblas-Reyes F J 2017 Detecting improvements in forecast correlation skill: statistical testing and power analysis *Mon. Weather Rev.* **145** 437–50
- Smith D M *et al* 2019 Robust skill of decadal climate predictions *npj Clim. Atmos. Sci.* **2** 1–10
- Tietsche S, Balmaseda M, Zuo H, Roberts C, Mayer M and Ferranti L 2020 The importance of North Atlantic Ocean transports for seasonal forecasts *Clim. Dyn.* **55** 1995–2011
- Weisheimer A, Doblas-Reyes F J, Jung T and Palmer T N 2011 On the predictability of the extreme summer 2003 over Europe *Geophys. Res. Lett.* **38** L05704
- Wilks D S 2019 *Statistical Methods in the Atmospheric Sciences* (Amsterdam: Elsevier) pp 617–68
- Williams K D *et al* 2018 The Met Office global coupled model 3.0 and 3.1 (GC3.0 and GC3.1) configurations *J. Adv. Model. Earth Syst.* **10** 357–80
- Wulff C O and Domeisen D I V 2019 Higher subseasonal predictability of extreme hot European summer temperatures as compared to average summers *Geophys. Res. Lett.* **46** 11520–9
- Wulff C O, Greatbatch R J, Domeisen D I V, Gollan G and Hansen F 2017 Tropical forcing of the Summer East Atlantic pattern *Geophys. Res. Lett.* **44** 11166–73