

LETTER • OPEN ACCESS

## Skillful decadal prediction of unforced southern European summer temperature variations

To cite this article: L F Borchert *et al* 2021 *Environ. Res. Lett.* **16** 104017

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.
- [A seven-fold rise in the probability of exceeding the observed hottest summer in India in a 2 °C warmer world](#)  
Nanditha J S, Karin van der Wiel, Udit Bhatia et al.
- [European summer temperatures since Roman times](#)  
J Luterbacher, J P Werner, J E Smerdon et al.

ENVIRONMENTAL RESEARCH  
LETTERS

## LETTER

## OPEN ACCESS

RECEIVED  
22 June 2021REVISED  
16 August 2021ACCEPTED FOR PUBLICATION  
25 August 2021PUBLISHED  
23 September 2021

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.

Skillful decadal prediction of unforced southern European  
summer temperature variationsL F Borchert<sup>1,2,\*</sup> , V Koul<sup>3,4</sup>, M B Menary<sup>1,2</sup> , D J Befort<sup>5</sup> , D Swingedouw<sup>6</sup>, G Sgubin<sup>6</sup> and J Mignot<sup>1</sup><sup>1</sup> Sorbonne Universités (SU/CNRS/IRD/MNHN), IPSL Laboratory, Institut Pierre Simon Laplace (IPSL), Paris, France<sup>2</sup> Laboratoire de Météorologie Dynamique (LMD), École Normale Supérieure (ENS), Paris, France<sup>3</sup> Institute of Coastal Systems Analysis and Modeling, Helmholtz-Zentrum Hereon, Geesthacht, Germany<sup>4</sup> Institute for Oceanography, CEN, Universität Hamburg, Hamburg, Germany<sup>5</sup> Atmospheric, Oceanic and Planetary Physics, Department of Physics, University of Oxford, Oxford, United Kingdom<sup>6</sup> EPOC, Université de Bordeaux, Pessac, France

\* Author to whom any correspondence should be addressed.

E-mail: [leonard.borchert@lmd.ipsl.fr](mailto:leonard.borchert@lmd.ipsl.fr)**Keywords:** climate prediction, CMIP6, European summer temperatureSupplementary material for this article is available [online](#)**Abstract**

We assess the capability of decadal prediction simulations from the Coupled Model Intercomparison Project phase 6 (CMIP6) archive to predict European summer temperature during the period 1970–2014. Using a multi-model ensemble average, we show that Southern European (SEU) summer temperatures are highly predictable for up to ten years in CMIP6. Much of this predictive skill, is related to the externally forced response: historical simulations explain about 90% of observed SEU summer temperature variance. Prediction skill for the unforced signal of SEU summer temperature is low: initialized model simulations explain less than 10% of observed variance after removing the externally forced response. An observed link between unforced SEU summer temperature and preceding spring Eastern North Atlantic—Mediterranean sea surface temperature (SST) motivates the application of a dynamical-statistical model to overcome the low summer temperature skill over Europe. This dynamical-statistical model uses dynamical spring SST predictions to predict European summer temperature, and significantly increases decadal prediction skill of unforced European summer temperature variations, showing significant prediction skill for unforced Southern European summer temperature 2–9 years ahead. As a result, dynamical-statistical models can benefit the decadal prediction of variables with initially limited skill beyond the forcing, such as summer temperature over Europe.

**1. Introduction**

Prediction of climate up to ten years into the future, so-called *decadal* climate prediction, has received considerable scientific attention due to its potentially high impact on society (Yeager and Robson 2017, Solaraju-Murali *et al* 2019, Merryfield *et al* 2020, Meehl *et al* 2021). On this time scale, the influence of chaotic internal climate variability and the response to external forcing strongly overlap (Lehner *et al* 2020). While the response to external forcing from greenhouse gas concentration is slowly varying and therefore well-predictable, internal variability poses a challenge to predictions (Doblas-Reyes *et al* 2013,

Marotzke *et al* 2016). However, due to its relatively slow characteristic time scales, internal variations of the ocean can be predicted in some places like the North Atlantic (Yeager and Robson 2017, Borchert *et al* 2021). Such predictable internal ocean variations could yield skillful prediction of unforced European summer temperature. Here, we examine whether ocean variations influence the skill of decadal-scale summer temperature predictions over Europe and, if so, whether this influence can be invoked to improve such predictions.

Most previous attempts to predict European temperature analyzed models from the 5th phase of the Coupled Model Intercomparison Project (CMIP).

Meanwhile, the decadal prediction simulations from the CMIP6 archive's Decadal Climate Prediction Project (DCPP) (Boer *et al* 2016) have been shown to have improved in predicting North Atlantic SST (Borchert *et al* 2021), both due to an improved response to volcanic forcing and improved capability to simulate internal climate variations. DCP simulations have also contributed to studies showing accurate reproduction of observed extremely warm summer frequencies over Europe (Borchert *et al* 2019), winter North Atlantic Oscillation (Smith *et al* 2020), jet stream position (Ruggieri *et al* 2021), and atmospheric blocking (Athanasiadis *et al* 2020). These findings inspire hope that CMIP6 decadal prediction simulations also show skillful prediction of European summer temperature variability.

However, previous studies generally found low-to-insignificant decadal temperature prediction skill over Europe beyond the forced response (Hanlon *et al* 2013, Liu *et al* 2019, Wu *et al* 2019). At the same time, some studies did demonstrate skillful temperature predictions beyond the linear trend in parts of Europe (Müller *et al* 2012, Arthun *et al* 2017, Feldmann *et al* 2019, Sgubin *et al* 2021). European summer temperature was shown to be relatively well-predictable compared to other seasons (Müller *et al* 2012), notwithstanding relatively low skill beyond the linear trend for this metric (Wu *et al* 2019). Any skill over Europe is commonly attributed to variability of climatic modes in the climate system (Doblas-Reyes *et al* 2013), many of which are in turn related to the relatively slowly varying ocean (Yeager and Robson 2017). As a result, numerous studies demonstrated skillful decadal predictions of ocean temperature (Robson *et al* 2012, Yeager *et al* 2012, Mignot *et al* 2016), which is commonly highest in the North Atlantic region (Matei *et al* 2012, Reintges *et al* 2020), due to ocean circulation (Robson *et al* 2012, Yeager *et al* 2012, Swingedouw *et al* 2013). However, models have been shown to be deficient in simulating observed atmospheric teleconnection mechanisms that transport that information from oceans to land (Qasmi *et al* 2017, Borchert *et al* 2018), which limit prediction skill over Europe. Any accounts of skillful prediction over Europe are often limited to relatively small regions and/or time scales, and draw on the memory of the ocean (Arthun *et al* 2017).

A few studies presented evidence that coupling a statistical model to dynamical and skillful SST predictions can yield skillful decadal prediction of European summer air temperature based on average North Atlantic SST (Wu *et al* 2019), European late summer precipitation (Simpson *et al* 2019), and the Indian summer monsoon (Sahastrabudhe and Ghosh 2021). Such hybrid dynamical-statistical (henceforth *dyn-stat*) predictions rely on observed links between SST (the predictor) and European climate (the predictand) to translate predicted SST into predictions of continental climate, thus circumventing

the deficiencies of climate prediction models concerning potential teleconnections. Crucially, such models require that predictions of the predictor are skillful, and that the connection between predictor and predictand is robust (Simpson *et al* 2018).

Here, we provide a first stocktake of the capability of CMIP6 decadal prediction systems to predict European summer temperature. We also develop and apply a simple dynamical-statistical model based on North Atlantic SST prediction to investigate its potential to predict the unforced component of European summer climate. By comparing *dyn-stat* model and the decadal predictions, we examine the added value of a *dyn-stat* model for predicting unforced modulations of European summer temperature using CMIP6, which identifies the potential that exists for decadal temperature predictions over Europe in these simulations.

## 2. Models and methods

### 2.1. CMIP6 models

Our analysis is based on an ensemble of eight initialized model systems, with ten ensemble members each, from the DCP (table S1 available online at [stacks.iop.org/ERL/16/104017/mmedia](https://stacks.iop.org/ERL/16/104017/mmedia)). The simulations provided using these systems are called re-forecast or hindcasts (*HC*). DCP simulations are initialized with observations every year after 1960 and run out for ten years (or *lead years*). We here analyze the period 1970–2015, which is the longest possible period that contains simulations from all ten lead years, since the last starting year for some models is 2014. For example, year 1970 is represented in lead year 10 of hindcasts started in 1960, lead year 9 of the 1961 hindcast, lead year 8 of the 1962 hindcast, and so on. Additionally, ensembles of historical simulations (HIST) with 28 CMIP6 models are analyzed (see table S1 for details). In total, we analyze a large multi-model ensemble (MME) of 192 historical members, subject to common forcing with observed atmospheric greenhouse gas concentrations, anthropogenic aerosols, solar variations and volcanic eruptions (Eyring *et al* 2016). These historical simulations are otherwise run freely to simulate diverse internal climate variability around the common forced response. Forming an ensemble mean across these members thus approximates climate variations that arise in response to external forcing (Maher *et al* 2020). In contrast, the 80 hindcast members have external forcing as well as the initialised observed climatic state at the start of their integration in common. This improves the ability of these systems to predict decadal internal climate variability (Keenlyside *et al* 2008, Doblas-Reyes *et al* 2013, Marotzke *et al* 2016). Forming an ensemble mean across these members thus aims at approximating the full climate evolution of a given period.

All model output is remapped to a latitude-longitude  $1 \times 1^\circ$  grid prior to analysis. We analyze the MME mean as the mean of the single-model ensemble means (i.e. a one-model-one-vote hierarchy). In decadal hindcasts, we analyze averages across lead year ranges 2–5, 6–9 and 2–9. We subtract the individual ensemble member mean for each lead year individually over the period 1970–2014 from the respective member prior to all analysis. As this is done for individual lead years, equating to a lead-time dependent mean bias correction (Boer *et al* 2016).

## 2.2. Observations

For observational reference, we consider the gridded observational temperature data sets from the Hadley Center, namely HadCRUTv5 (Morice *et al* 2021) for surface air temperature (SAT), and interpolated HadISST (Rayner *et al* 2003) for SST. Sea level pressure (SLP) is obtained from the ERA5 reanalysis (Hersbach *et al* 2020). Our results are robust to the use of ERA5 instead of HadCRU/ISST for temperature over both Europe and the North Atlantic (not shown). In all cases, we remap the information to a regular  $1 \times 1^\circ$  grid, and analyze the period 1970–2015.

## 2.3. Methods

We analyze spring (March–April–May, hereafter MAM) mean SST as well as summer (June–July–August, hereafter JJA) mean SAT. We evaluate the skill of HC and HIST against observations using the anomaly correlation coefficient (ACC), and the explained observed variance for different types of model simulations as  $ACC^2$  (Goddard *et al* 2013). Significance is tested at the 5% level using a Monte Carlo bootstrapping method, where the underlying data is resampled 1000 times with replacement (Smith *et al* 2020). In addition to the bootstrapping, we benchmark the predictions against persistence forecast, which is calculated as the correlation of the observational time series with itself at different time lags. Finally, we evaluate the sensitivity of residual skill estimates to the exclusion of individual years in the analysis, using a leave-one-out cross validation analysis.

We test two different approaches of extracting the forced response. As a first approximation of the forcing trend, we extract the long-term linear trend from simulations and observations. This technique has, however, been shown to potentially introduce spurious variability (Trenberth and Shea 2006, Mann *et al* 2014, Tandon and Kushner 2015). We also extract the forced component using the *residuals* approach (Smith *et al* 2019). This technique uses the historical MME mean to estimate the forced component, which is linearly regressed on HC or observed climate variability, and then subtracted from each. We use residuals to assess unforced (i.e. internally generated)

decadal prediction skill as well as study unforced physical interactions in the climate system.

## 3. Decadal predictions of European summer temperature in CMIP6

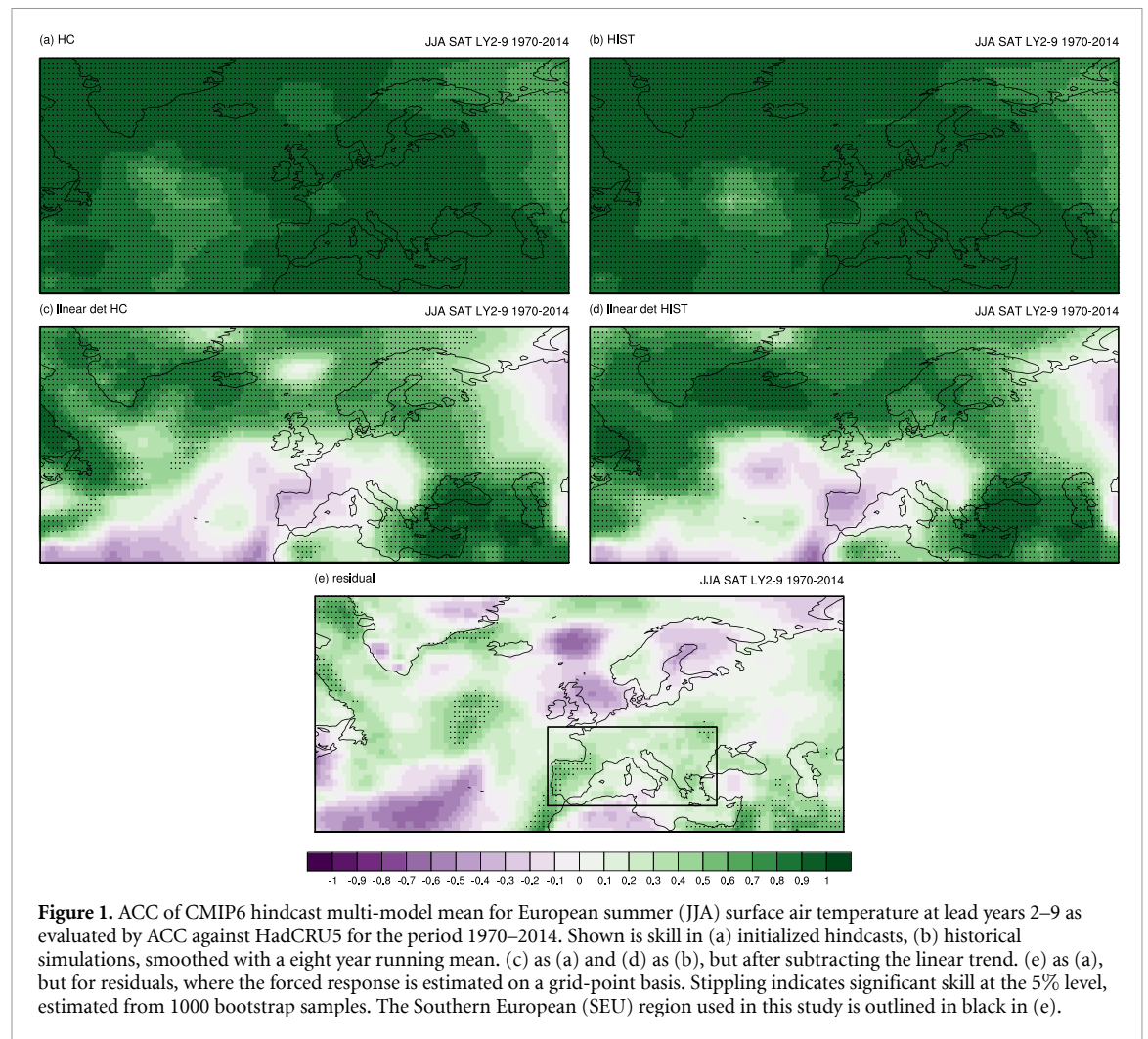
Decadal prediction skill in terms of ACC for summer temperature throughout Europe is highly significant at lead years 2–9 in the HC MME (figure 1(a)). Skill inferred from HIST simulations that represent the forced response is similarly high (figure 1(b)), indicating that much of the signal in decadal European summer temperature variations that can be predicted in initialized prediction systems arises from externally forced variations. The high skill in HIST simulations is also found when using the same eight models as used in HC instead of the full 28 model ensemble (not shown). We use the full ensemble in this study for robustness.

An attempt to characterize unforced skill is shown in figures 1(c) and (d), using linear detrending. This approach, however, appears ill-fit to extract the forced response: as HIST approximates the forced response of the system, skill should be close to 0 after linear detrending if linear detrending was appropriate to extract the forced response (Branstator and Teng 2012, Gastineau and Frankignoul 2015, Wu *et al* 2019). Skill of HC and HIST after linearly detrending is non-zero and significant in large parts of Europe, particularly in central-to-eastern as well as northern and southeastern Europe. A more elaborate technique thus needs to be used when aiming to assess unforced variability and predictability of climate (Tandon and Kushner 2015).

One such approach of estimating the unforced variability based on historical simulations is the *residual* (Smith *et al* 2019, see Methods). Applying this methodology on a grid point basis in our skill calculation (figure 1(e)) shows that there is little summer temperature ACC skill in the HC simulations beyond the forced response: residual ACC is insignificant across much of Europe, with a small band of significant residual ACC across western Iberia, and a tendency of higher residual ACC towards southern Europe.

## 4. Using decadal predictions of seasonal SST to predict unforced European summer temperature

Given the overall low skill for unforced summer temperature variability over Europe (cf figure 1), we now develop a simple dynamical-statistical model to overcome this issue. Here, a connection between SST and SAT over land is identified, explored and exploited in such a way where skillful dynamical decadal predictions of residual SST are invoked to predict the residual of JJA SAT. Because of the—modestly—high residual ACC for summer temperature in southern



**Figure 1.** ACC of CMIP6 hindcast multi-model mean for European summer (JJA) surface air temperature at lead years 2–9 as evaluated by ACC against HadCRU5 for the period 1970–2014. Shown is skill in (a) initialized hindcasts, (b) historical simulations, smoothed with a eight year running mean. (c) as (a) and (d) as (b), but after subtracting the linear trend. (e) as (a), but for residuals, where the forced response is estimated on a grid-point basis. Stippling indicates significant skill at the 5% level, estimated from 1000 bootstrap samples. The Southern European (SEU) region used in this study is outlined in black in (e).

Europe (cf figure 1(e)), we construct a dyn-stat model aimed at predictions of SEU ( $10^{\circ}$  W– $30^{\circ}$  E,  $35^{\circ}$ – $50^{\circ}$  N; black outlines) summer temperature over land, masking out ocean regions.

#### 4.1. Observed SST impact on European summer temperature

To avoid overlap between the training and validation period of our dyn-stat model, we here explore the impact of unforced spring SST variations on summer SAT over southern Europe for the period 1900–1969, and MAM SST prediction skill during 1970–2014 (figure 2). Unforced (i.e. residual) spring SST shows high correlation to SEU summer SAT in the subpolar and eastern North Atlantic, as well as the Mediterranean in the observations (figure 2(a)). Much of this pattern is significant. The season lag between the two signals suggests that spring SST influences summer SAT.

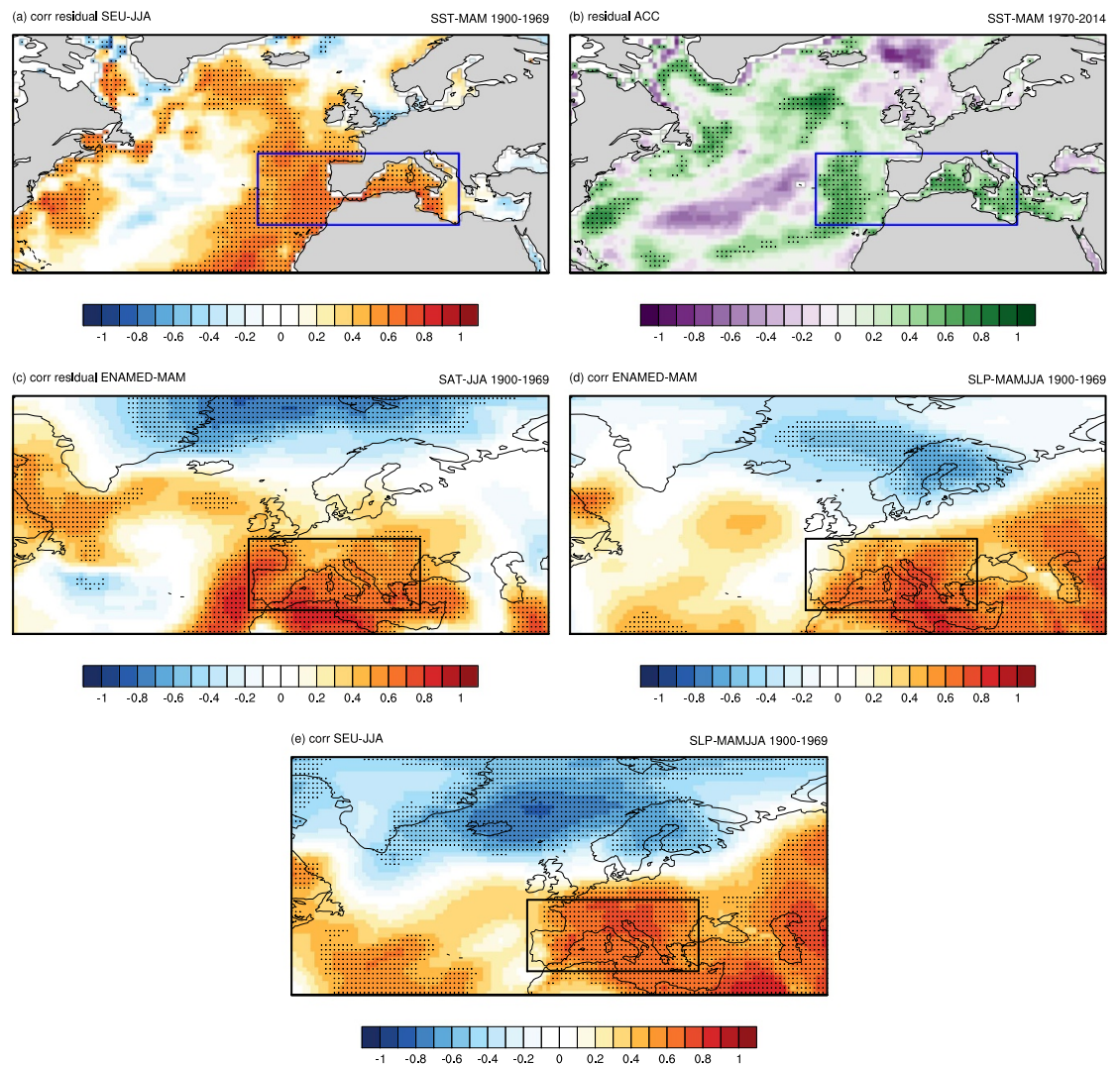
The predictor needs to be predictable to be a useful predictor for SEU summer temperature in the dyn-stat framework. Significant residual skill for spring SST in the HC ensemble during 1970–2014 (figure 2(b)) is confined to the northeast subpolar as well as the eastern North Atlantic and the

Mediterranean Sea. The eastern part of the North Atlantic off the coast of Portugal and the western Mediterranean (the ENAMED region  $25^{\circ}$  W– $15^{\circ}$  E,  $35^{\circ}$ – $45^{\circ}$  N; blue lines in figures 2(a) and (b)) appears particularly suitable to construct the dyn-stat model, where unforced spring SST is both connected to summer SEU SAT and significantly predictable.

Spring SST in the ENAMED region is correlated with summer SAT over much of southern Europe, confirming the selected SST region to have an impact there (figure 2(c)). Spring-summer sea level pressure patterns suggest that this connection is accomplished by a combination of dynamic transport of atmospheric heat of oceanic origin around a high pressure system centered over the Mediterranean (figures 2(d) and (e)) and the persistence of heat in the Mediterranean region itself.

Findings presented here are robust to the use of different seasonal lead times between SST and summer SAT (e.g. using SST in preceding winter or simultaneous summer; not shown), illustrating the broad influence of ENAMED SST on SEU summer SAT. Heat persistence in the ocean is therefore the likely carrier of the spring SST signal to summer. Spring does, however, show larger residual SST prediction



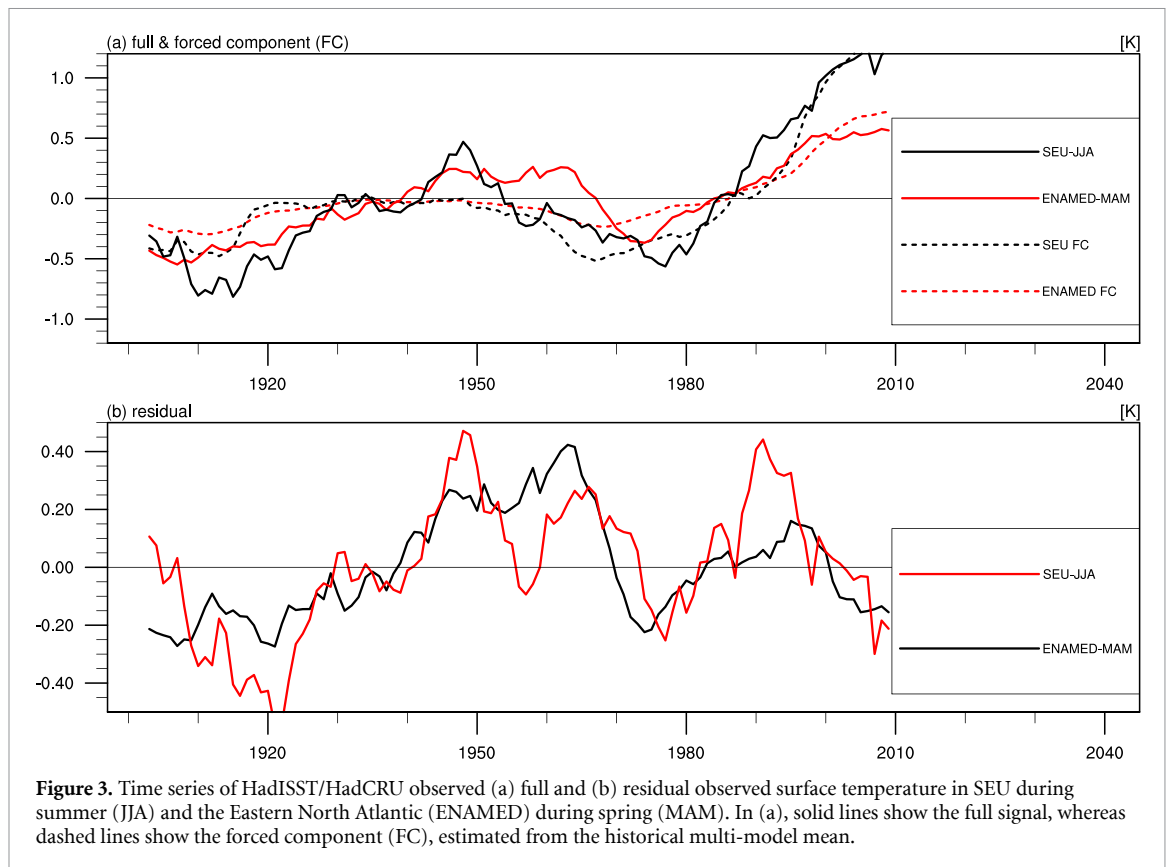


**Figure 2.** (a) observed (HadISST/HadCRU) correlation of residual spring (MAM) SST to residual summer (JJA) SAT in the SEU region (outlined in black in (c)–(e)), applying a eight year running mean on the period 1900–1969. (b) Residual ACC for spring SST at lead years 2–9, estimated from the MME for 1970–2014. Also shown is the correlation of spring Eastern North Atlantic—Mediterranean (ENAMED) SST (outlined in blue in (a) and (b) to observed (c) residual summer SAT and (d) residual spring-summer (MAMJJA) sea level pressure for 1900–1969. (e) as (d), but using SEU SAT instead of ENAMED SST. Stippling indicates significant skill at the 5% level, estimated from 1000 bootstrap samples.

skill compared to the other seasons (summer SST is particularly unpredictable; not shown), which is why we here concentrate on spring SST as predictor of European summer temperature.

The local SST and SAT correlation patterns shown for the period 1900–1969 in figures 2(a) and (c) are also found during the period 1970–2014 (figure S1). A tripole structure observed in the recent period in the residual spring SST correlation pattern to SEU summer SAT, characterized by negative correlations in the subpolar and tropical North Atlantic and a band of positive correlations in between, inhibits the use of the high subpolar ACC for spring SST (figure 2(b)). Nonetheless, these findings highlight that the ENAMED-based statistical model trained on 1900–1969 is also applicable during the 1970–2014 period.

The correlation of the ENAMED-MAM and the SEU-JJA indices is also exemplified in their time series (figure 3). During the period 1900–1969, the forced response is characterised by a cold-warm-cold pattern (figure 3(a)). Both residuals show pronounced multidecadal variation, ending at a warm state at the end of the 1960s (figure 3(b)). After 1970, the forced response causes both indices to rise over time, and both residuals show a cold-warm-cold pattern in time (figure 3(b)). Summer SEU SAT thus follows the low-frequency behavior of spring ENAMED SST reasonably well and robustly throughout time. As the residual potentially explains about half of the forced response—standard deviations for observed JJA SEU SAT are: full signal 0.52; forced component 0.47; residual 0.23—predicting both forced response and unforced variability accurately is important to



**Figure 3.** Time series of HadISST/HadCRU observed (a) full and (b) residual observed surface temperature in SEU during summer (JJA) and the Eastern North Atlantic (ENAMED) during spring (MAM). In (a), solid lines show the full signal, whereas dashed lines show the forced component (FC), estimated from the historical multi-model mean.

achieve decadal prediction of the full summer temperature signal in southern Europe.

#### 4.2. Dynamical-statistical predictions of European summer temperature

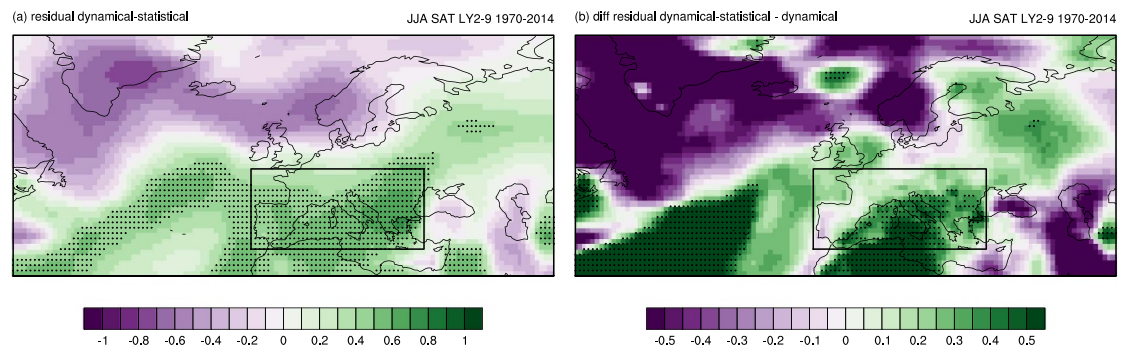
The dyn-stat model developed here is inspired by more complex models (Simpson *et al* 2019, Wu *et al* 2019). The underlying idea is that potential deficiencies in the atmospheric component of climate models that hamper prediction skill over land (Meehl *et al* 2021) can be overcome with statistical models, after a physical link between a predictor and a predictand has been identified. Our dynamical-statistical model of JJA SEU SAT rescales the predicted MAM ENAMED SST with the ratio of standard deviations between the two indices observed during 1900–1969. The dyn-stat model thus requires skillful prediction of MAM SST and as such highlights the importance of initialization to achieve skillful decadal climate prediction.

There is reasonable area overlap between the area of observed positive correlation between ENAMED SST and SEU SAT (figure 2(c)) and skill in the dyn-stat model (figure 4(a)), indicating that the skill of the dyn-stat predictions arises from the observed link between ENAMED SST and SEU SAT. Dyn-stat residual ACC is significant over most of the SEU region south of 45° N, extending from the Iberian Peninsula and southern France to the coasts of the Black Sea (figure 4(a)). However, significant skill improvements compared to dynamical predictions are significant only over Italy and part of the Balkans, as well as

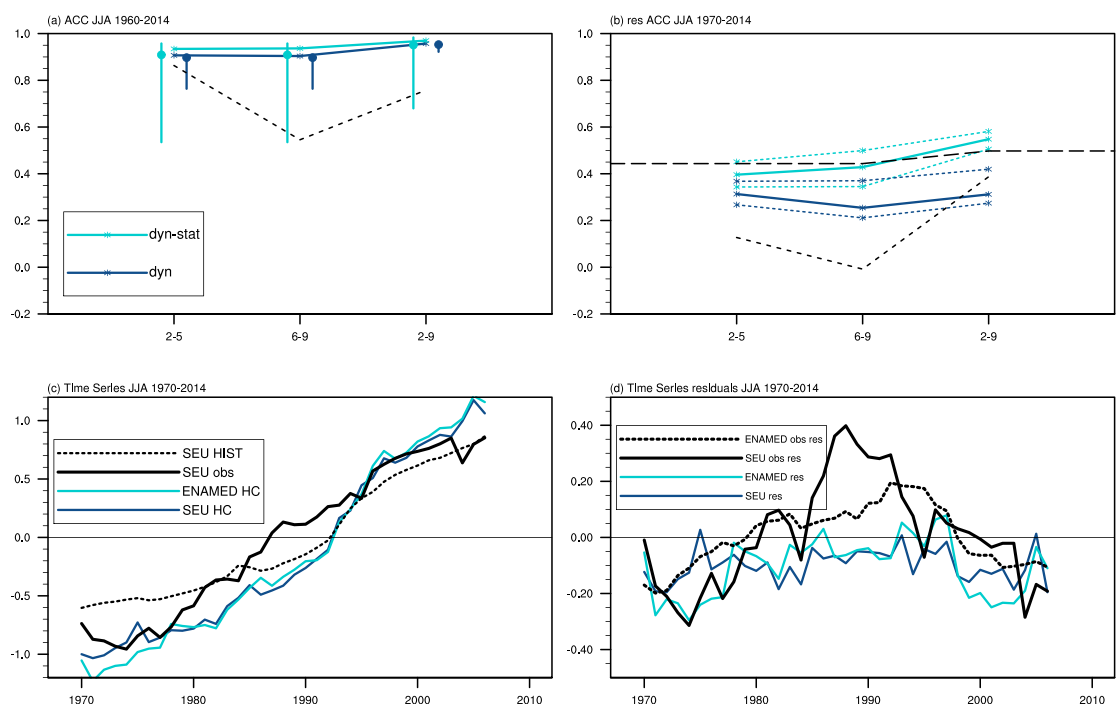
the Mediterranean itself (figure 4(b)). Beyond these regions, we find positive but generally insignificant improvement of residual ACC for European summer temperature in the dyn-stat model compared to the fully dynamical prediction.

The value of the dyn-stat model over purely dynamical predictions can further be illustrated in area mean skill for SEU summer temperature over land, explicitly masking out skill improvements over the Mediterranean Sea (figure 5). For the full SEU summer SAT signal, both dyn-stat and the HC ensemble are highly skillful across all examined lead time ranges, beating persistence forecast (figure 5(a)). The dynamical hindcast MMEs explain between 82% and 92% (dyn), and between 85% and 92% (dyn-stat), of observed SEU SAT variance, depending on lead time, the historical MMEs explain between 85% and 90%. This highlights the importance of the forced response for prediction skill, and is related to strong trends in both indices (figure 5(c)).

We find insignificant residual skill in both the dyn-stat and dynamical model for individual lead years (not shown) as well as four year lead time averages (figure 5(b)). Only at eight year lead time averages does significant residual ACC emerge in the dyn-stat model. Such relatively high skill at long temporal averages relates to filtering of noise through smoothing of the time series (Hegerl *et al* 2021). Dyn-stat skill at lead years 2–9 is also the instance where we find significant residual skill that is robust to a leave-one year-out cross validation. This analysis highlights the



**Figure 4.** (a) Residual ACC for summer (JJA) temperature, based on dynamical-statistical predictions using spring (MAM) ENAMED SST. (b) Difference of this dyn-stat residual ACC to the dynamical residual ACC shown in figure 1(e). Stippling indicates significance at the 5% level, based on 1000 bootstraps. Outlines indicate the SEU region.



**Figure 5.** ACC (y-axis) for (a) the full and (b) the residual summer (JJA) SEU temperature signal across lead time averages (x-axis). The dynamical model is shown in blue and the dyn-stat model based on spring (MAM) ENAMED SST predictions in cyan. Skill of the historical models (vertical line) and the historical MME mean (dots) is shown. The dashed line in (a) shows persistence skill as a benchmark, as do the short dashes in (b). Long dashes in (b) indicate the 5% significance threshold, estimated based on bootstrapping a dummy time series with similar characteristics as the observations 1000 times with replacement. Colored dashes in (b) denote maximum and minimum skill of a leave-one-out cross validation skill analysis, where each individual year was left out of the skill calculation once to highlight the importance of individual years for the diagnosed skill. (c) and (d) show the time series corresponding to lead years 2–9 in (a) and (b), respectively. The solid black line shows residual SEU observations, the dashed line in (d) shows the observed residual ENAMED index, and the dashed line in (c) indicates the forced component of SEU summer temperature based on the multi-model mean of the historical simulations.

value of the dyn-stat model developed here as well as low-pass filtering to achieve robustly significant skill 2–9 years into the future.

The underlying time series of these analyses (figures 5(c) and (d)) illustrate the nuances that yield the skill improvement described above. On the full time series, the dyn-stat and dynamically predicted SEU summer temperature variations are very similar, explaining their similar full ACC (figure 5(c)). A comparable impression arises for the predicted residuals (figure 5(d)), due to the relationship between the two

indices in their unforced variability described above. The dyn-stat model reproduces the observed low-frequency cold-warm-cold variations of unforced SEU summer SAT better than the dynamical prediction does, explaining the elevated skill in dyn-stat. This confirms our hypothesis from the observed time series (cf figure 3) that the ENAMED SST index mainly captures low-frequency changes in unforced summer SEU SAT while the teleconnections are not well reproduced in the dynamical hindcasts.



## 5. Discussion and conclusions

Analyzing a MME of DCPD simulations, we have shown that these simulations show high skill for predicting European summer temperature. Much of this can be attributed to skillful historical simulations representing the forced response. As a result, DCPD simulations show low skill at predicting decadal unforced (residual) European summer temperature variations. In other words, the role of forcing for predictability of decadal variations of European summer temperature appears to be slightly larger than previously thought (Smith *et al* 2019). Based on observed teleconnections between spring ENAMED SST and summer SEU SAT, we present a dynamical-statistical approach to overcome the issue of limited residual skill over Europe. The dyn-stat model beats the HC ensemble as well as persistence at predicting decadal variations of SEU summer temperature.

In CMIP6, the skill of both initialised (HC) and non-initialised (HIST) decadal predictions of summer SAT are comparable, and improved compared to CMIP5 (e.g. Smith *et al* 2019). However, the strength of the external forcing appears to mitigate the potential for additional skill to arise from initialisation more strongly in CMIP6 than it used to be the case in CMIP5 (Smith *et al* 2019). This is probably related to improved model response to forcing in CMIP6 compared to CMIP5 (Borchert *et al* 2021). Based on our dyn-stat model results, limited residual summer temperature ACC may be caused by underestimated teleconnection between ENAMED SST and SEU SAT in models, either due to structural deficiencies in models or due to shock from the initialization procedure. The problem of underestimation of atmospheric signals is exemplified by the amount of summer SAT variance that is explained by spring SST: at lead years 2–9, the unforced ENAMED index explains 28% of variance of the SEU index (based on  $ACC^2$ ) in HC simulations, while the same calculation in observations yields a value 50%. This, as well as the low amplitude in predicted indices presented in this study, is likely related to the general underestimation of decadal climatic signals in climate models, known as the signal-to-noise problem (Smith *et al* 2020). As such, improving the response of atmospheric models to oceanic variability is crucial to improved skill in dynamical predictions at least to the level demonstrated by our dyn-stat model.

In this study we have used the residuals approach (Smith *et al* 2019) to remove the forced signal from observed and predicted time series. Recently, an alternative method was also presented (Sospedra-Alfonso and Boer 2020). Both of these methods utilize the historical MME to estimate the forced response and, for large enough ensembles, should yield similar mean-skill values. While the residuals (Smith *et al* 2019) are relatively

straight-forward to calculate and thus interpret, the alternative method (Sospedra-Alfonso and Boer 2020) provides additional estimates of unpredictable noise based on the ensemble spread. Here, given our focus on multi-model means, both methods should give similar results, so we use the more established residual method for its simplicity complementing it with additional bootstrap analyses to estimate uncertainties.

In previous studies, it was demonstrated that SST impact on European summer temperature was accomplished via a Rossby wave train (Borchert *et al* 2019, Wu *et al* 2019). This mechanism is partly different from what we describe here. In our case, ENAMED SST during spring and summer influences summer SEU SAT through a combination of advective transport of local heat anomalies and dynamic transport of heat from the North Atlantic around a high-pressure system over the Mediterranean. This difference in physical mechanisms might originate from either the consideration of the SEU region specifically, or from the use of residuals instead of the linearly detrended indices considered in the other studies. When using linear detrending instead of residuals, we found the Rossby wave train atmospheric response (not shown). As we show linear detrending to be deficient to extract the forced response, the Rossby wave train discussed in earlier work results at least partly from a response to forcing.

The timeseries of residual SEU summer temperature after 1970 (cf figure 3(b)) is reminiscent of the winter NAO low-frequency time series (Smith *et al* 2020, their figure 2), which might be related to the SST pattern that characterizes the Atlantic Multidecadal Variability (AMV) (Sutton and Hodson 2005, Klavans *et al* 2019). There is evidence that these two indices are not independent, and indeed that a AMV-like SST pattern is forced by fluctuations of the winter NAO (Klavans *et al* 2019, Oelsmann *et al* 2020), so there might be a potential role of the NAO in driving recent unforced summer temperature variations in the SEU region (Sutton and Hodson 2005). The ENAMED SST index representing the low-frequency part of SEU variations (cf figure 3(b)) fits this narrative. It is thus possible that NAO plays a role in modulating the ENAMED SST that is related to SEU summer temperature as described here, probably through an impact on SST that persists into later seasons.

The dyn-stat model used here is kept very simple for the sake of easy interpretability. There are more complex options, such as EOF-based models (Wu *et al* 2019) or multivariate regression models (Simpson *et al* 2019). Considering more complex processes or SST dipoles thus holds the potential to increase the skill found in the dyn-stat model here even further; the skill estimates shown in this study are a lower bound on expectable skill in that sense.

A general shortcoming of the residual approach is that it includes the impact of volcanic eruptions in the forced response, although eruptions would probably benefit the skill of initialized predictions in a forecast setting. The HIST simulations used in residual calculation include the effect of volcanic forcing, which would not be known in forecasts of the future. Volcanoes, however, appear to be particularly important for decadal climate predictions in the North Atlantic region (Hermanson *et al* 2020, Borchert *et al* 2021). For a more nuanced assessment of the value of initialization, only greenhouse gas and anthropogenic aerosol forcings—as used in climate projections—should therefore be used to calculate residuals in hindcast studies. We do not do this here because the current protocol for hindcast initialization is also unfit for a hindcast assessment fully consistent with forecasts, in that the information of volcanic eruptions is included in hindcast forcing terms at the exact time of the eruption and not (as would be the case for a forecast) at the next initialization time after the eruption.

Our work highlights the importance of external forcing for the capability of decadal prediction systems to predict European summer temperature. Invoking links to predictable unforced SST, however, the resulting limited skill for internal summer temperature variations can be overcome to some extent, emphasizing the importance of skillful initialization of SST prediction to predict climate over land. Our results thus provide hope for skillful predictions of European summer temperature variations in future modeling systems when the representation of teleconnections is improved.

### Data availability statement

The data that support the findings of this study are openly available at the following URL/DOI: <https://esgf-node.ipsl.upmc.fr/>.

### Acknowledgments


The authors acknowledge the World Climate Research Programme, which, through its Working Group on Coupled Modelling, coordinated and promoted CMIP6. The authors thank the climate modeling groups (listed in table S1) 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. L F B, D J B, D S, G S and J M were supported by the EUCP project funded by the European Union's Horizon2020 programme, Grant Agreement Number 776613. J M and D S were also supported by the H2020 Blue-Action project, Grant Agreement Number 727852. M B M was supported by the H2020 EPICE project, Grant Agreement Number 789445. L F B and

M B M received additional support from the ANR-TREMPLEIN ERC Project HARMONY, Grant Agreement Number ANR-20-ERC9-0001. L F B especially thanks Johanna Baehr for the warm welcome to her group in the difficult times of COVID-19. HadISST and HadCRU data were obtained from [www.metoffice.gov.uk/hadobs](http://www.metoffice.gov.uk/hadobs) and are © British Crown Copyright, Met Office, 2021, provided under a Non-Commercial Government Licence [www.nationalarchives.gov.uk/doc/non-commercial-government-licence/version/2/](http://www.nationalarchives.gov.uk/doc/non-commercial-government-licence/version/2/).

The authors thank anonymous reviewers for their valuable help in improving the manuscript.

### ORCID iDs

L F Borchert  <https://orcid.org/0000-0002-3232-7409>

M B Menary  <https://orcid.org/0000-0002-9627-2056>

D J Bafort  <https://orcid.org/0000-0002-2851-0470>

### References

- Årthun M, Eldevik T, Viste E, Drange H, Furevik T, Johnson H L and Keenlyside N S 2017 Skillful prediction of northern climate provided by the ocean *Nat. Commun.* **8** 15875
- Athanasiadis P J, Yeager S, Kwon Y O, Bellucci A, Smith D W and Tibaldi S 2020 Decadal predictability of North Atlantic blocking and the NAO *npj Clim. Atmos. Sci.* **3** 1–10
- Boer G J *et al* 2016 The decadal climate prediction project (DCPP) contribution to CMIP6 *Geosci. Model Dev.* **9** 3751–77
- 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
- Borchert L F, Müller W A and Baehr J 2018 Atlantic ocean heat transport influences interannual-to-decadal surface temperature predictability in the North Atlantic region *J. Clim.* **31** 6763–82
- Borchert L F, Pohlmann H, Baehr J, Neddermann N C, Suarez-Gutierrez L and Müller W A 2019 Decadal predictions of the probability of occurrence for warm summer temperature extremes *Geophys. Res. Lett.* **46** 14042–51
- Branstator G and Teng H 2012 Potential impact of initialization on decadal predictions as assessed for CMIP5 models *Geophys. Res. Lett.* **39** L12703
- Doblas-Reyes F J, Andreu-Burillo I, Chikamoto Y, García-Serrano J, Guemas V, Kimoto M, Mochizuki T, Rodrigues L R L and Oldenborgh G J V 2013 Initialized near-term regional climate change prediction *Nat. Commun.* **4** 1715
- Eyring V, Bony S, Meehl G A, Senior C A, Stevens B, Stouffer R J and Taylor K E 2016 Overview of the coupled model intercomparison project phase 6 (CMIP6) experimental design and organization *Geosci. Model Dev.* **9** 1937–58
- Feldmann H, Pinto J G, Laube N, Uhlig M, Moemken J, Pasternack A, Früh B, Pohlmann H and Kottmeier C 2019 Skill and added value of the MiKlip regional decadal prediction system for temperature over Europe *Tellus A* **71** 1618678
- Gastineau G and Frankignoul C 2015 Influence of the North Atlantic SST variability on the atmospheric circulation during the twentieth century *J. Clim.* **28** 1396–416
- Goddard L *et al* 2013 A verification framework for interannual-to-decadal predictions experiments *Clim. Dyn.* **40** 245–72

- Hanlon H M, Hegerl G C, Tett S F B and Smith D M 2013 Can a decadal forecasting system predict temperature extreme indices? *J. Clim.* **26** 3728–44
- Hegerl G C et al 2021 Towards consistent observational constraints in climate predictions and projections *Front. Clim.* **3** 4343
- Hermanson L et al 2020 Robust multiyear climate impacts of volcanic eruptions in decadal prediction systems *J. Geophys. Res.: Atmos.* **125** e2019JD031739
- Hersbach H et al 2020 The ERA5 global reanalysis *Q. J. R. Meteorol. Soc.* **146** 1999–2049
- Keenlyside N S, Latif M, Jungclauss J, Kornblueh L and Roeckner E 2008 Advancing decadal-scale climate prediction in the North Atlantic sector *Nature* **453** 84–88
- Klavans J M, Clement A C and Cane M A 2019 Variable external forcing obscures the weak relationship between the NAO and North Atlantic multidecadal SST variability *J. Clim.* **32** 3847–64
- Lehner F, Deser C, Maher N, Marotzke J, Fischer E M, Brunner L, Knutti R and Hawkins E 2020 Partitioning climate projection uncertainty with multiple large ensembles and CMIP5/6 *Earth Syst. Dyn.* **11** 491–508
- Liu Y, Donat M G, Rust H W, Alexander L V and England M H 2019 Decadal predictability of temperature and precipitation means and extremes in a perfect-model experiment *Clim. Dyn.* **53** 3711–29
- Maher N, Lehner F and Marotzke J 2020 Quantifying the role of internal variability in the temperature we expect to observe in the coming decades *Environ. Res. Lett.* **15** 054014
- Mann M E, Steinman B A and Miller S K 2014 On forced temperature changes, internal variability and the AMO *Geophys. Res. Lett.* **41** 3211–19
- Marotzke J et al 2016 MiKlip: a national research project on decadal climate prediction *Bull. Am. Meteorol. Soc.* **97** 2379–94
- Matei D, Pohlmann H, Jungclauss J, Müller W, Haak H and Marotzke J 2012 Two tales of initializing decadal climate prediction experiments with the ECHAM5/MPI-OM model *J. Clim.* **25** 8502–23
- Meehl G A et al 2021 Initialized earth system prediction from subseasonal to decadal timescales *Nat. Rev. Earth Environ.* **2** 1–18
- Merryfield W J et al 2020 Current and emerging developments in subseasonal to decadal prediction *Bull. Am. Meteorol. Soc.* **101** E869–96
- Mignot J, García-Serrano J, Swingedouw D, Germe A, Nguyen S, Ortega P, Guilyardi E and Ray S 2016 Decadal prediction skill in the ocean with surface nudging in the IPSL-CM5A-LR climate model *Clim. Dyn.* **47** 1225–46
- Morice C P et al 2021 An updated assessment of near-surface temperature change from 1850: the HadCRUT5 data set *J. Geophys. Res.: Atmos.* **126** e2019JD032361
- Müller W A et al 2012 Forecast skill of multi-year seasonal means in the decadal prediction system of the Max Planck Institute for Meteorology *Geophys. Res. Lett.* **39** L22707
- Oelsmann J, Borchert L, Hand R, Baehr J and Jungclauss J H 2020 Linking ocean forcing and atmospheric interactions to atlantic multidecadal variability in MPI-ESM1.2 *Geophys. Res. Lett.* **47** e2020GL087259
- Qasmi S, Cassou C and Boé J 2017 Teleconnection between atlantic multidecadal variability and european temperature: diversity and evaluation of the coupled model intercomparison project phase 5 models *Geophys. Res. Lett.* **44** 11140–9
- Rayner N A, Parker D E, Horton E B, Folland C K, Alexander L V, Rowell D P, Kent E C and Kaplan A 2003 Global analyses of sea surface temperature, sea ice and night marine air temperature since the late nineteenth century *J. Geophys. Res.: Atmos.* **108** 4407
- Reintges A, Latif M, Bordbar M H and Park W 2020 Wind stress-induced multiyear predictability of annual extratropical North Atlantic sea surface temperature anomalies *Geophys. Res. Lett.* **47** e2020GL087031
- Robson J I, Sutton R T and Smith D M 2012 Initialized decadal predictions of the rapid warming of the North Atlantic Ocean in the mid 1990s *Geophys. Res. Lett.* **39** L19713
- Ruggieri P et al 2021 Atlantic multidecadal variability and North Atlantic jet: a multimodel view from the decadal climate prediction project *J. Clim.* **34** 347–60
- Sahastrabudhe R and Ghosh S 2021 Does statistical model perform at par with computationally expensive general circulation model for decadal prediction? *Environ. Res. Lett.* **16** 064028
- Sgubin G, Swingedouw D, Borchert L F, Menary M B, Noel T, Loukos H and Mignot J 2021 Systematic investigation of the skill opportunities in decadal temperature prediction over Europe and potential implications for integrated applications *Clim. Dyn.* (<https://doi.org/10.1007/s00382-021-05863-0>)
- Simpson I R, Deser C, McKinnon K A and Barnes E A 2018 Modeled and observed multidecadal variability in the North Atlantic jet stream and its connection to sea surface temperatures *J. Clim.* **31** 8313–38
- Simpson I R, Yeager S G, McKinnon K A and Deser C 2019 Decadal predictability of late winter precipitation in western Europe through an ocean–jet stream connection *Nat. Geosci.* **12** 613–19
- Smith D M et al 2019 Robust skill of decadal climate predictions *npj Clim. Atmos. Sci.* **2** 1–10
- Smith D M et al 2020 North Atlantic climate far more predictable than models imply *Nature* **583** 796–800
- Solaraju-Murali B, Caron L P, Gonzalez-Reviriego N and Doblas-Reyes F J 2019 Multi-year prediction of European summer drought conditions for the agricultural sector *Environ. Res. Lett.* **14** 124014
- Sospedra-Alfonso R and Boer G J 2020 Assessing the impact of initialization on decadal prediction skill *Geophys. Res. Lett.* **47** e2019GL086361
- Sutton R T and Hodson D L R 2005 Atlantic Ocean forcing of North American and European summer climate *Science* **309** 115–18
- Swingedouw D, Rodehacke C B, Behrens E, Menary M, Olsen S M, Gao Y, Mikolajewicz U, Mignot J and Biastoch A 2013 Decadal fingerprints of freshwater discharge around Greenland in a multi-model ensemble *Clim. Dyn.* **41** 695–720
- Tandon N F and Kushner P J 2015 Does external forcing interfere with the AMOC's influence on North Atlantic sea surface temperature? *J. Clim.* **28** 6309–23
- Trenberth K E and Shea D J 2006 Atlantic hurricanes and natural variability in 2005 *Geophys. Res. Lett.* **33** L12704
- Wu B, Zhou T, Li C, Müller W A and Lin J 2019 Improved decadal prediction of northern-hemisphere summer land temperature *Clim. Dyn.* **53** 1357–69
- Yeager S G and Robson J I 2017 Recent progress in understanding and predicting atlantic decadal climate variability *Curr. Clim. Change Rep.* **3** 112–27
- Yeager S, Karspeck A, Danabasoglu G, Tribbia J and Teng H 2012 A decadal prediction case study: late twentieth-century North Atlantic Ocean heat content *J. Clim.* **25** 5173–89