

LETTER • **OPEN ACCESS**

Past long-term summer warming over western Europe in new generation climate models: role of large-scale atmospheric circulation

To cite this article: Julien Boé *et al* 2020 *Environ. Res. Lett.* **15** 084038

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

Environmental Research Letters



LETTER

OPEN ACCESS

RECEIVED
3 October 2019

REVISED
15 April 2020

ACCEPTED FOR PUBLICATION
17 April 2020

PUBLISHED
7 August 2020

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.



Past long-term summer warming over western Europe in new generation climate models: role of large-scale atmospheric circulation

Julien Boé^{1,9} , Laurent Terray¹, Marie-Pierre Moine¹, Sophie Valcke¹, Alessio Bellucci², Sybren Drijfhout^{3,4}, Rein Haarsma³, Katja Lohmann⁵, Dian A. Putrasahan⁵, Chris Roberts⁶ , Malcom Roberts⁷, Enrico Scoccimarro² , Jon Seddon⁷ , Retish Senan⁶ and Klaus Wyser⁸

¹ CECI, Université de Toulouse, CERFACS/CNRS, Toulouse, France

² Fondazione Centro Euro-Mediterraneo sui Cambiamenti Climatici (CMCC), Bologna, Italy

³ Royal Netherlands Meteorological Institute (KNMI), PO Box 201, 3730 AE De Bilt, The Netherlands

⁴ University of Southampton, National Oceanography Centre, European Way, Waterfront Campus, Southampton, United Kingdom

⁵ Max-Planck-Institut für Meteorologie, The Ocean in the Earth System Department, Bundesstraße 53, 20146 Hamburg, Germany

⁶ European Centre for Medium Range Weather Forecasting (ECMWF), Reading, United Kingdom

⁷ Met Office, Exeter EX1 3PB, United Kingdom

⁸ Rossby Centre, Swedish Meteorological and Hydrological Institute, Norrköping, Sweden

⁹ Author to whom any correspondence should be addressed.

E-mail: boe@cerfacs.fr

Keywords: climate change, Europe, summer warming, atmospheric circulation, trends, model evaluation

Supplementary material for this article is available [online](#)

Abstract

Past studies have concluded that climate models of previous generations tended to underestimate the large warming trend that has been observed in summer over western Europe in the last few decades. The causes of this systematic error are still not clear. Here, we investigate this issue with a new generation of climate models and systematically explore the role of large-scale circulation in that context.

As an ensemble, climate models in this study warm less over western Europe and warm more over eastern Europe than observed on the 1951–2014 period, but it is difficult to conclude this is directly due to systematic errors given the large potential impact of internal variability. These differences in temperature trends are explained to an important extent by an anti-correlation of sea level pressure trends over the North Atlantic / Europe domain between models and observations. The observed trend tends to warm (cool) western (eastern) Europe but the simulated trends generally have the opposite effect, both in new generation and past generation climate models. The differences between observed and simulated sea level pressure trends are likely the result of systematic model errors, which might also impact future climate projections. Neither a higher resolution nor the realistic representation of the evolution of sea surface temperature and sea ice leads to a better simulation of sea level pressure trends.

1. Introduction

In the last few decades, western Europe has warmed faster than the global average (Bhend and Whetton 2013). Climate models may not capture such regional amplification. Van Oldenborgh *et al* (2009), analysing climate models from the Coupled Model Intercomparison Phase 3 project (CMIP3, Meehl *et al* 2007), have concluded that ‘Western Europe is warming much faster than expected’, especially in summer. Consistently, Bhend and Whetton (2013)

have concluded that the summer temperature trends over western Europe are underestimated in a large number of both CMIP3 and CMIP5 (Taylor *et al* 2012) models. For both Van Oldenborgh *et al* (2009) and Bhend and Whetton (2013) these differences between models and observations are likely the result of systematic model errors rather than internal variability. Obviously, a systematic underestimation of the forced warming signal could have major implications for future climate and impact projections over western Europe.

Note that both in Van Oldenborgh *et al* (2009) and Bhend and Whetton (2013), the regional trends are computed as the linear regression against global temperature, smoothed with a 3-year running mean. Their results therefore do not imply *per se* that simulated temperature trends over western Europe are too small, but only that they are too small relative to global temperature variations. Furthermore, given the narrowness of the window of the running mean, decadal and multi-decadal time-scales also impact the value of regional trends.

Different mechanisms might explain the observed regional amplification of warming over western Europe in summer. The transition from a negative phase of the Atlantic Multidecadal Variability (AMV, Kerr 2000, Enfield *et al* 2001), roughly between 1965–1995, to a positive phase afterwards (Sutton *et al* 2018) may have reinforced the warming trend over Europe. Indeed, observational (Sutton and Dong 2012, Mariotti and Dell'Aquila 2012, O'Reilly *et al* 2017) and model studies (Ruprich-Robert *et al* 2017) have shown that the positive phases of the AMV tend to be associated with warmer temperature over Europe. Dong *et al* (2017) have shown that the rapid summer warming since the mid-1990s over Europe is mainly explained by variations in sea surface temperature (SST, including both internal and forced variations) and sea ice, with also an important direct role of anthropogenic aerosols. This result is consistent with Philipona *et al* (2009) or Nabat *et al* (2014) who have shown that anthropogenic aerosols explain roughly 23% of the summer warming trend on 1980–2012.

Climate models may have difficulties in capturing some of those processes, which could explain the discrepancy between models and observations regarding summer temperature trends. The uncertainties in the simulation of the radiative impact of anthropogenic aerosols (e.g. Boucher *et al* 2013) and in the response of the European climate to anthropogenic aerosols variation (Boé 2016) are large and might influence the simulated temperature trends. Climate models may also have difficulties to capture correctly the AMV properties and the associated teleconnection with summer temperature over Europe (Qasmi *et al* 2017), which could impact simulated temperature trends.

Additionally, Shin and Sardeshmukh (2011) have shown that the atmospheric circulation driven remote impacts of systematic model errors in the pattern of tropical ocean warming are responsible to a large extent for the difficulties of GCMs to capture correctly the annual temperature and precipitation trends over the land masses surrounding the North Atlantic Ocean. Variations in large-scale atmospheric circulation of internal origin may also strongly impact regional temperature trends (Lehner *et al* 2017), making the assessment of the consistency of simulated and observed trends more difficult.

Despite these studies suggesting a potential role of atmospheric circulation on the discrepancy between simulated and observed long-term summer warming over Europe, this role has yet to be systematically explored, which is the main objective of this work.

First, the consistency of simulated and observed summer temperature trends over Europe on the 1951–2014 period is evaluated in both previous generation and new generation climate models from the PRIMAVERA project, at low and high resolution. The role of atmospheric circulation on the differences between observed and simulated warming trends and the realism of simulated large-scale atmospheric circulation trends are then evaluated. The role of potential errors in SSTs on temperature and large-scale circulation trends are characterized thanks to the analysis of both fully coupled climate simulations and corresponding atmospheric simulations forced by observed SSTs.

2. Data and methods

2.1. Data

An ensemble of forced atmospheric and coupled simulations from the PRIMAVERA H2020 European project (e.g. Vannière *et al* 2018) on the 1950–2014 period are studied. The so-called Atmospheric Multimodel Intercomparison Project (AMIP) forced-atmospheric simulations follow the High Resolution Model Intercomparison Project (HighResMIP, Haarsma *et al* 2016) protocol (*highresSST-present* experiments) from the Coupled Model Intercomparison Project phase 6 (Eyring *et al* 2016). The observed daily sea surface temperature (SST) and sea ice concentration at a 0.25° resolution used as forcing are based on the HadISST2 dataset (Kennedy *et al* 2017). AMIP simulations provide an estimate of the evolving atmospheric state as constrained by observed SSTs. The coupled PRIMAVERA simulations follow the hist-1950 HighResMIP protocol (Haarsma *et al* 2016). These historical simulations start in 1950 with atmospheric and oceanic initial states coming from a 1950 control simulation using fixed 1950 forcing.

The six PRIMAVERA models are run (at least) at two resolutions, a higher resolution (HR) and a lower resolution (LR) (table 1), both in the coupled and forced framework. The model tuning, normally performed on the LR version, is as similar as possible in the HR and LR simulations, so that the real impact of resolution can be assessed. Note that, as a result, the tuning of the higher resolution version may not be optimal. It is therefore possible that better performances could have been achieved with higher resolution models had they been specifically tuned. The same number of vertical levels is used in LR and HR simulations. Depending on the model, the lower and higher resolutions can be largely different

Table 1. Climate simulations from the PRIMAVERA project analysed in this study and nominal resolution in km following the CMIP6 definition (see annex 2 in (Taylor *et al* 2018)). The resolution at 50° N in km is also given in italic (from Vannière *et al* 2018). The number of members used in this study are given between brackets. The first figure corresponds to forced atmospheric members and the second figure to coupled members. Note that the 'lower resolution' of PRIMAVERA models may be different from the low or standard resolution of the corresponding CMIP5 models (table 2).

	EC-Earth3P Haarsma <i>et al</i> (submitted to Geosci. Model Dev.)	ECMWF-IFS (Roberts <i>et al</i> 2018)	HadGEM3-GC31 (Roberts <i>et al</i> 2019)	MPI-ESM1-2 (Gutjahr <i>et al</i> 2019)	CMCC-CM2 (Cherchi <i>et al</i> 2019)	CNRM-CM6-1 (Voldoire <i>et al</i> 2019)
Lower resolution	EC-Earth3P 100; 71 (1;4)	ECMWF-IFS-LR 50; 50 (8;8)	HadGEM3- GC31-LM (HadGEM3- GC31-LR for coupled simula- tions) 250; 135 (5;8)	MPI-ESM1- 2-HR 100; 67 (1;1)	CMCC-CM2- HR4 100; 64 (1;1)	CNRM-CM6-1 250; 142 (10;3)
Higher resolution	EC-Earth3P-HR 50; 36 (1;3)	ECMWF-IFS-HR 25; 25 (6;6)	HadGEM3- GC31-HM 50; 25 (3;3)	MPI-ESM1- 2-XR 50; 34 (1;1)	CMCC-CM2- VHR4 25; 18 (1;1)	CNRM-CM6-1- HR 50; 50 (10;3)

Table 2. List of CMIP5 models used in this study.

ACCESS1-0
ACCESS1-3
bcc-csm1-1
BNU-ESM
CanESM2
CCSM4
CESM1-BGC
CNRM-CM5
CSIRO-Mk3-6-0
FGOALS-g2
FIO-ESM
GFDL-CM3
GFDL-ESM2G
GFDL-ESM2M
GISS-E2-H
GISS-E2-R
HadGEM2-ES
inmcm4
IPSL-CM5A-LR
MIROC5
MIROC-ESM
MPI-ESM-LR
MRI-CGCM3
NorESM1-M

(table 1), e.g. the lower resolution of a model may be as high as the higher resolution of other models. The impact of resolution in a multi-model framework therefore cannot be simply assessed as the difference between the higher resolution and lower resolution multi-model ensemble means. A variable number of members is available depending on the model and simulation (table 1).

The six PRIMAVERA models are not independent, which may impact the results of the ensemble (e.g. Boé 2018). Both EC-Earth3 and ECMWF-IFS (Roberts *et al* 2018) are based on the IFS atmospheric model (cycle 36r4 for the former and cycle 43r1 for

the latter). CNRM-CM6-1 (Voldoire *et al* 2019) uses the ARPEGE atmospheric model, whose dynamical core is also derived from IFS (cycle 37t1). The physics packages are however different (Voldoire *et al* 2019). All models except MPI-ESM1-2 use the NEMO oceanic model.

A large ensemble of coupled climate models from CMIP5 (Taylor *et al* 2012) are also analysed. Historical simulations, complemented by the RCP85 scenario after 2005 as well as pre-industrial control simulations (i.e. with constant forcings equal to pre-industrial values) are studied (table 2). One member by model is used (the first available member, i.e. generally the ensemble member named 'r1i1p1'). The models have been selected so that the length of their pre-industrial control simulation (generally at least 500 years) is sufficient to sample low-frequency internal variability.

In order to characterize observational uncertainties, even if it does not allow for a full exploration of the uncertainty space, multiple observational datasets for surface atmospheric temperature and sea level pressure (SLP) are used. For temperature, the Climatic Research Unit (CRU) high-resolution gridded dataset CRU TS4 (Harris *et al* 2014) and the Berkeley Earth System Temperature dataset (BEST, Rohde *et al* 2013) are used. In Supporting information (SI) (stacks.iop.org/ERL/15/084038/mmedia), results with CRUTEM4 (Jones *et al* 2012) and the E-OBS homogenized dataset (E-OBS v19.0eHOM, Squintu *et al* 2019) are also shown. For SLP, the NCAR Sea Level Pressure dataset (Trenberth and Paolino 1980) and the NCEP Reanalysis (Kalnay *et al* 1996) are used. Additionally, SLP trends from the NOAA 20th Century Reanalysis V2c (Compo *et al* 2011) and HadSLP2r (Allan and Ansell 2006) are shown in SI. As shown later, the uncertainties associated with the choice of the dataset are limited for temperature. For

SLP, important differences of trends are noted over continental Europe.

All model and observation datasets are regridded before analysis on a common regular 0.25x0.25-degree resolution grid with a conservative interpolation.

2.2. Methods

In this paper, we focus on the trends on the 1951–2014 period, a longer period compared to previous studies, in order to reduce the impact of internal variability and better capture potential anthropogenic changes. We calculate the linear trends, as a simple indicator of changes on the period.

The significance of observed trends is assessed with the t-statistic. An effective sample size, following Zwiers and von Storch (1995), is used to take into account temporal auto-correlation. The false discovery rate test of Benjamini and Hochberg (1995) is finally applied, with $\alpha_{FDR} = 0.10$ (Wilks 2016).

To compare observed and simulated SLP trends, the weighted centred pattern correlation coefficient (r) is used. Weights (w) to account for spatial variations in grid cell areas are used.

$$r(x, y, w) = \frac{\text{cov}(x, y, w)}{\sqrt{\text{cov}(x, x, w) \cdot \text{cov}(y, y, w)}}$$

with $\text{cov}(x, y, w)$ the weighted spatial covariance:

$$\text{cov}(x, y, w) = \frac{\sum_{i=1}^n w_i \cdot (x_i - \bar{x})(y_i - \bar{y})}{\sum_{i=1}^n w_i}$$

and \bar{x} the weighted spatial average:

$$\bar{x} = \frac{\sum_{i=1}^n w_i \cdot x_i}{\sum_{i=1}^n w_i}$$

i corresponds to the grid points over the domain of interest (green box in figure 3(c)).

The impact of differences in SLP trends between models and observations on temperature trends is assessed by first searching in CMIP5 preindustrial simulations the 64-year periods with a SLP trend very similar either to the observed or simulated 1951–2014 trend. Then, the corresponding surface temperature trends are calculated to estimate the part of the warming trend due to large-scale atmospheric circulation, in the observations or in the models. We first compute all the overlapping 64-year SLP trends in the CMIP5 preindustrial simulations and we compute their centred pattern correlation with the observed 1951–2014 SLP trend (NCAR SLP dataset) on the domain shown in figure 3(c). Then, we select in each CMIP5 model preindustrial simulation the 64-period with the highest pattern correlation with observations and finally compute the corresponding multi-model trend in temperature. A similar analysis is performed searching for the 64-year period

in each preindustrial CMIP5 simulation characterized by the SLP trend with the largest pattern correlation with the PRIMAVERA ensemble mean SLP trend for the forced-atmospheric simulations (shown in figure 3(e)) and then computing the associated mean surface temperature trend.

3. Results

Large warming trends in summer over western Europe have been observed, as high as 2 K over the 1951–2014 period in Spain or the south of France (figures 1(a) and (b)). The temperature trends over the north-east of the domain are much smaller. They barely reach 1 K on the same period and are not always significant. The temperature trends are generally robust among the different observational datasets (figures 1(a) and (b), see also figure S1). Consistently with Van Oldenborgh *et al* (2009) and Bhend and Whetton (2013), the coupled climate models from CMIP5 as an ensemble warm much less over western Europe than observed (figure 1(c)). Conversely, they tend to warm much more than observed over eastern Europe, where the observed trends are small.

The spatial pattern and the magnitude of the differences in temperature trends between observations and PRIMAVERA coupled simulations (including both HR and LR simulations) are very similar to the ones of CMIP5 models (figure 1), except for some very local differences, for example over western France or Iberian Peninsula. There is therefore no major evolution regarding summer temperature trends over Europe in the new generation of climate models.

The difference between observed and simulated temperature trends in PRIMAVERA forced atmospheric simulations is very similar to the one obtained in coupled simulations, except for the Baltic region and United Kingdom of Great Britain and Northern Ireland where the difference tends to be negative in forced atmospheric simulations and positive in coupled simulations. Potential errors in simulated warming trends over the ocean in coupled models are therefore not responsible for the generally smaller (larger) simulated warming trends over western (eastern) Europe compared to observations. Shin and Sardeshmukh (2011) have pointed to an important impact of simulated SST trends on annual temperature trends over land masses surrounding the North Atlantic. The role of SSTs is much smaller regarding specifically summer temperature changes over western Europe on 1951–2014, at least in the PRIMAVERA models.

Note that observed internal oceanic variations are part of the forcing within the forced atmospheric framework. The potential imprint of oceanic internal variability (e.g. of the AMV) on the observed temperature trends is therefore also unlikely to explain the differences between modelled and observed summer

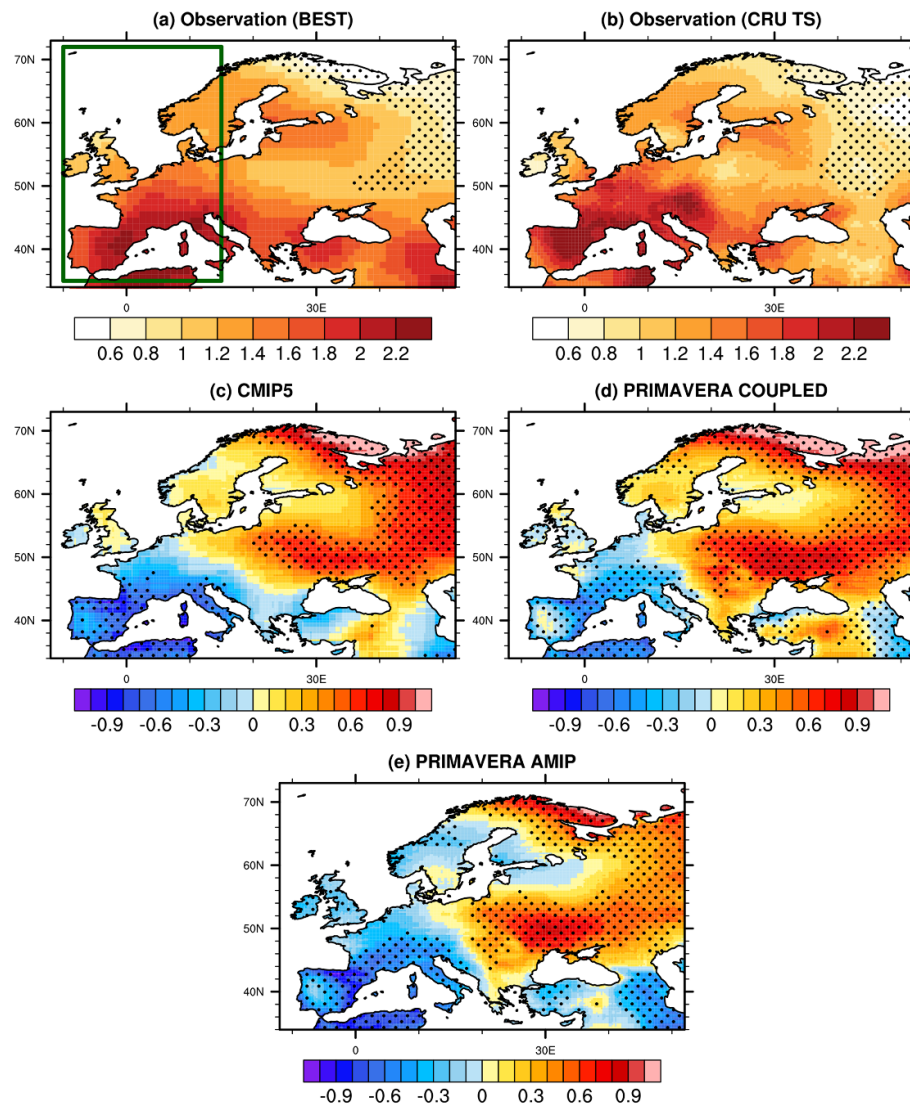


Figure 1. Linear trend in summer (JJA) surface temperature (K over period) over Europe from 1951 to 2014 in (a) BEST observations and (b) CRU TS4 observations. The black dots in (a) and (b) show where the trends are not significant, with $p > 0.05$. See the Data and Methods section for a description of the test. Differences of the linear trends in JJA surface temperature over Europe from 1951 to 2014 (K over period) between (c) the multi-model ensemble mean of CMIP5 models, (d) the multi-model ensemble mean of PRIMAVERA coupled simulations, (e) the multi-model ensemble mean of PRIMAVERA forced atmospheric simulations, and the observations (BEST dataset). For PRIMAVERA, for each model at lower and higher resolution the average over all members is first computed, and then the multi-model ensemble mean is computed. The black dots in (c), (d) and (e) show the grid points where more than 75% of the models agree on the sign of the difference of trend. The green box in (a) shows the domain used to compute the spatial averages over western Europe.

warming trends over Europe, as it should be captured by forced atmospheric simulations.

The strong warming trend observed over western Europe cannot be explained by internal variability alone. Indeed, no 64-year period in preindustrial CMIP5 simulations (i.e. with constant external forcings) shows a warming trend over western Europe (defined with the green box in figure 1(a)) as large as observed on 1951–2014. The differences between 64-year temperature trends simulated in preindustrial simulations and the observed trend are always largely negative (figure 2, ‘CTRL CMIP5’ box). The trends in CMIP5 historical simulations generally come much closer to the observed trend, pointing not surprisingly to the role of external forcings on long-term

warming (figure 2). The potential underestimation of the warming trend over western Europe in CMIP5 models is far from systematic. It is also true for the PRIMAVERA models. For the majority of members, at higher or lower resolution, in the coupled or forced framework, the warming trend is smaller than observed, but all the models have at least one member close to the observations (figure 2). None of the PRIMAVERA models therefore can be said to be inconsistent with observations.

The impact of internal variability on the simulated 1951–2014 trends is strong. In the forced PRIMAVERA simulations, differences of trends as large as 0.8 K between members are noted (for CNRM-CM6-1 at higher resolution, big brown points

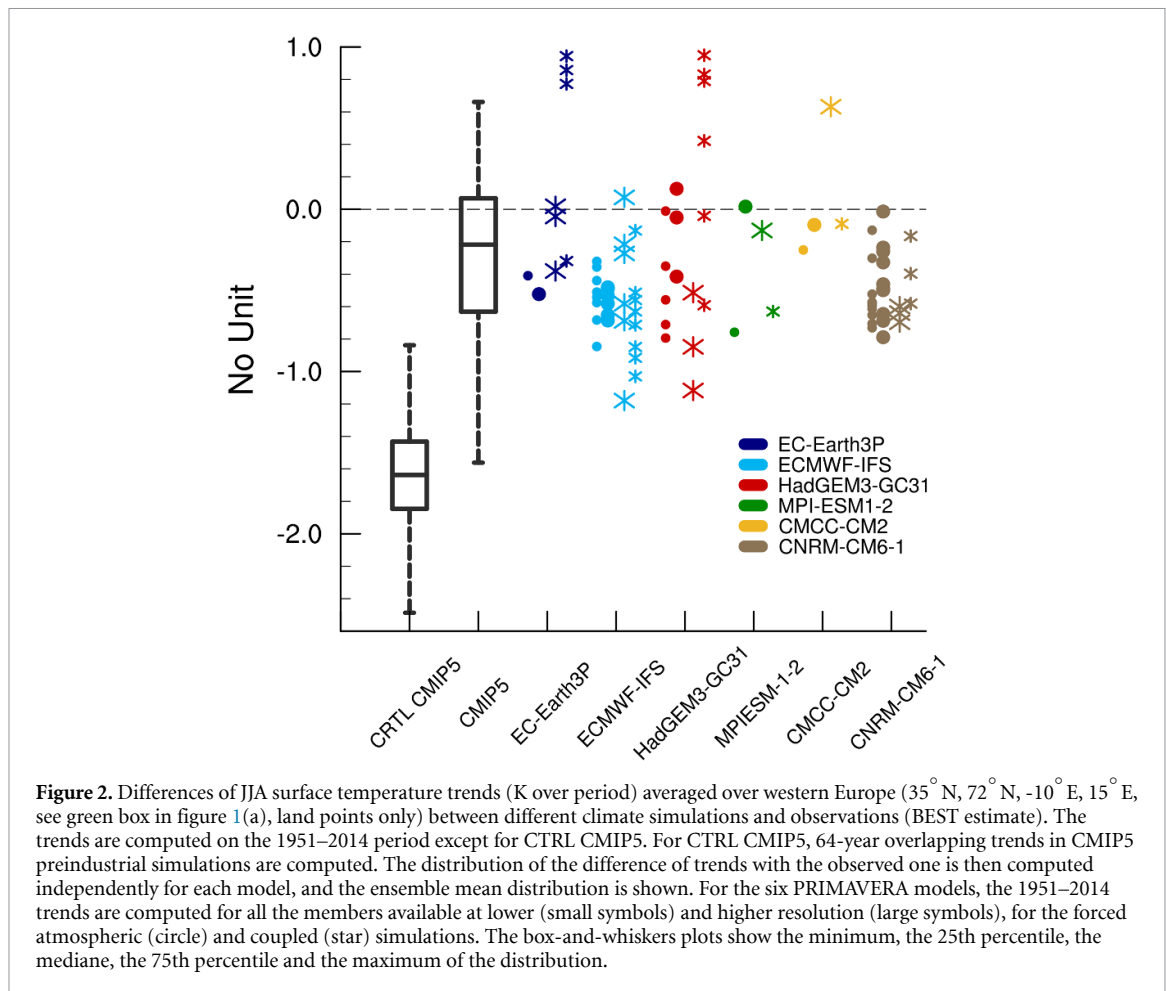


Figure 2. Differences of JJA surface temperature trends (K over period) averaged over western Europe (35° N, 72° N, -10° E, 15° E, see green box in figure 1(a), land points only) between different climate simulations and observations (BEST estimate). The trends are computed on the 1951–2014 period except for CTRL CMIP5. For CTRL CMIP5, 64-year overlapping trends in CMIP5 preindustrial simulations are computed. The distribution of the difference of trends with the observed one is then computed independently for each model, and the ensemble mean distribution is shown. For the six PRIMAVERA models, the 1951–2014 trends are computed for all the members available at lower (small symbols) and higher resolution (large symbols), for the forced atmospheric (circle) and coupled (star) simulations. The box-and-whiskers plots show the minimum, the 25th percentile, the median, the 75th percentile and the maximum of the distribution.

in figure 2). For coupled PRIMAVERA simulations, differences of trends between members as large as 1.6 K are noted (for HadGEM3-GC31 at lower resolution, small red stars in figure 2). Given the large impact of internal variability, even in a forced framework, no clear impact of resolution on summer warming trends emerges. A large number of members would be needed to demonstrate a systematic impact of resolution, if any.

It is possible to obtain observed temperature trends over western Europe over the 1951–2014 period in PRIMAVERA forced atmospheric and coupled simulations, even if the probability is relatively low for some models, especially in the forced framework. The underlying forced signal in simulated summer temperature trends could be realistic and in that case, the observed trend would be seen as an ‘extreme’ reached because of a large additional warming due to internal atmospheric variability. It is also possible that models underestimate the forced signal and that only extreme simulated internal atmospheric variations allow the observed trend to be reached by a few simulations. These results illustrate the overall intrinsic weakness of conclusions that one can reach in model evaluation when dealing with properties strongly impacted by internal variability, such as trends.

Note that the results of the comparison of simulated and observed temperature trends strongly depend on the period used. When the trends are computed on a shorter period, starting in 1980, the observed warming is weak over western Europe and amplified over central Europe. As a result, the models generally tend to warm more than observed over western Europe and less over central Europe (not shown).

In order to assess the impact of atmospheric circulation on the differences in temperature trends between models and observations, the SLP trends are first characterized. A decrease in summer SLP is observed over the North Atlantic off Europe, associated with an increase over Scandinavia (figures 3(a) and (b)). The trend is however significant for a few grid points only. The decrease over the North Atlantic is robust across the different observational datasets, but large observational uncertainties exist regarding the SLP trends over continental Europe south of Scandinavia and over the Mediterranean sea, where even the sign of the trends may differ among the datasets (figures 3(a) and (b) and figure S3).

The ensemble-mean SLP trend in CMIP5 models tends to be anti-correlated with the observed one, with positive trends over the Atlantic off Europe and negative trends over Scandinavia (figure 3(c)).

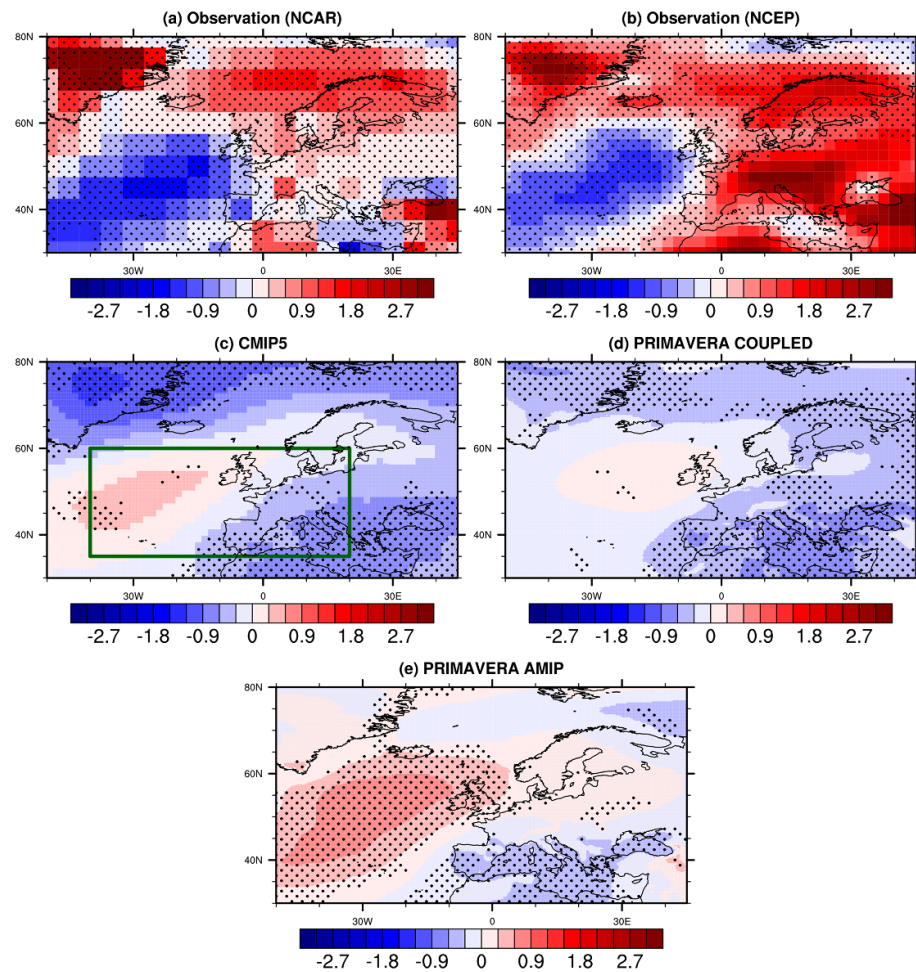


Figure 3. Linear trend in summer (JJA) sea level pressure (hPa over period) over the North Atlantic / Europe domain from 1951 to 2014 in (a) NCAR SLP observations and (b) NCEP reanalysis. The black dots in (a) and (b) show where the trends are not significant with $p > 0.05$. See the Data and Methods section for a description of the test. Linear trend in JJA sea level pressure over the North Atlantic / Europe domain from 1951 to 2014 (hPa over period) in (c) CMIP5 models (d) PRIMAVERA coupled simulations and (e) PRIMAVERA forced atmospheric simulations. For PRIMAVERA, for each model at lower and higher resolution the average over all members is first computed, and then the multi-model ensemble mean is computed. The black dots in (c), (d) and (e) show the grid points where more than 75% of the models agree on the sign of the trend. The green box in (c) shows the domain used to compute the centred pattern correlations of figure 4.

Negative trends are also seen over the Mediterranean sea. Unfortunately, the consistency with observations is hard to assess there, given the large observational uncertainties mentioned earlier.

The SLP trend pattern is generally very similar in the PRIMAVERA and CMIP5 coupled models, with slightly less positive trends over the North-Atlantic in the PRIMAVERA models. The trend pattern is also similar in the forced atmospheric PRIMAVERA simulations, with however a general mean shift on the domain towards more positive values (figure 3(e)). This difference in SLP trends between coupled and forced PRIMAVERA simulations could be due to the realistic representation of SST trends in the forced experiments, which come from the observations, compared to the coupled simulations, in which SSTs are free to evolve. The fact that SLP trends are not closer to observations in forced atmospheric simulations prompts consideration of other possibilities. In particular, it can be contemplated that the lack of

two-way air-sea coupling in forced simulations negatively impact the SLP response.

Note that what matters regarding the impact of SLP trends on western European climate is the evolution of pressure gradients, rather than the absolute changes in SLP. As the evolutions of pressure gradients are generally consistent in forced and coupled simulations, no major associated difference in the impact of SLP trends on temperature over Europe are expected.

To compare SLP trends from individual simulations to observations, we now use as metric the centred pattern correlation (section 2.2) calculated over the green box in figure 3(c). This domain is chosen for its relevance for warming trends over Europe (as demonstrated subsequently) and to avoid the regions of strong observational uncertainties as the Mediterranean sea and central Europe.

The pattern correlations between the observed 1951–2014 SLP trend and the SLP trends over all the

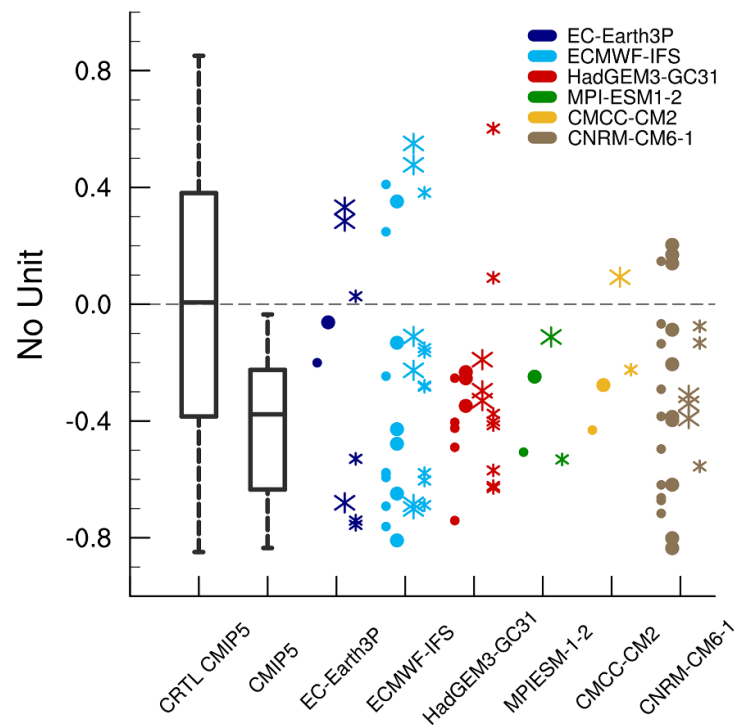


Figure 4. Same as figure 2 except that centred pattern correlations (see section 2.2) over the North Atlantic / Europe domain (35°N , 60°N , -40°E , 20°E , see green box in figure 3(c)) of sea level pressure trends between models and observations replace the differences of surface temperature trends between models and observations. For observations, the NCAR SLP dataset is used.

64-year segments in the preindustrial CMIP5 simulations vary between -0.9 and 0.9 (figure 4). It is therefore possible to obtain because of internal variability alone simulated trends that are very similar to the observed one (in terms of pattern correlation). This is consistent with the fact that the observed trend is significant only for a few grid points (figure 3(a)). Very interestingly, when variations in external forcings are taken into account, i.e. in the historical simulations, none of the CMIP5 models simulates a 1951–2014 SLP trend similar to the observed one. The simulated trend is even anti-correlated with observations for all the CMIP5 models studied. As suggested in figure 3, the external forcings in CMIP5 models lead to a signal that anti-correlates with the observed SLP trends.

The 1951–2014 SLP trend in summer is also anti-correlated with the observed trend in the majority of forced and coupled PRIMAVERA simulations. It is still possible to find some members positively (but weakly) correlated with observations for four coupled and two forced models. Note that the impact of internal variability on the simulated 1951–2014 trends is very large, as some members from the same model (e.g. ECMWF-IFS coupled simulations) may show trends strongly anti-correlated ($r = -0.7$) or weakly positively correlated ($r = 0.5$) with observations. Such a large impact of internal variability is also seen in forced simulations. As for temperature trends, no systematic impact of resolution is discernible given the large internal variability and small ensemble sizes.

Similar results are obtained when using the NCEP reanalysis as reference (figure S4), despite the observational uncertainties noted earlier.

The observed 1951–2014 SLP trend is significant only for a few points and compatible with simulated internal variability in terms of spatial pattern in the absence of external forcings. Even if the positive pattern correlations with the observed trend seen for a few PRIMAVERA simulations suggest that new generation models may have improved, large positive pattern correlations are still never obtained in historical simulations. New generation models, when forced by time-varying external forcings, are still unable to produce a SLP trend really similar to the observed one. It is true even when the models are forced by observed SSTs. Most, if not all, climate models therefore likely suffer from systematic errors in reproducing the long-term SLP trend in summer over the North Atlantic / Europe sector. Either the impact of internal variability on SLP trends is underestimated in climate models, or their response to anthropogenic forcings, generally anti-correlated with the observed trend, is not correct.

This inconsistency has not been discussed before to the best of our knowledge. Gillett *et al* (2013) with a detection and attribution framework have shown the impact of anthropogenic forcings in summer SLP changes, which implies a consistency between simulated and observed evolutions. It is not contradictory with our results as the domain in their study is almost global. The positive simulated SLP trend over

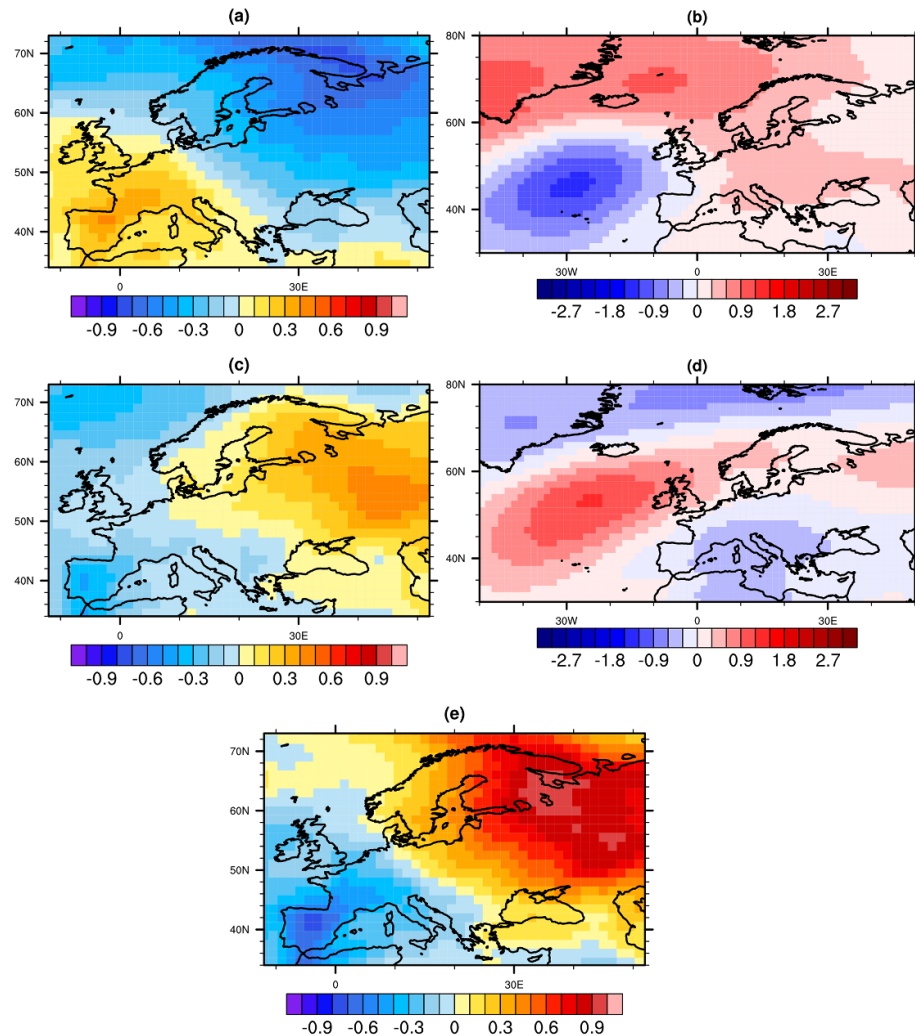


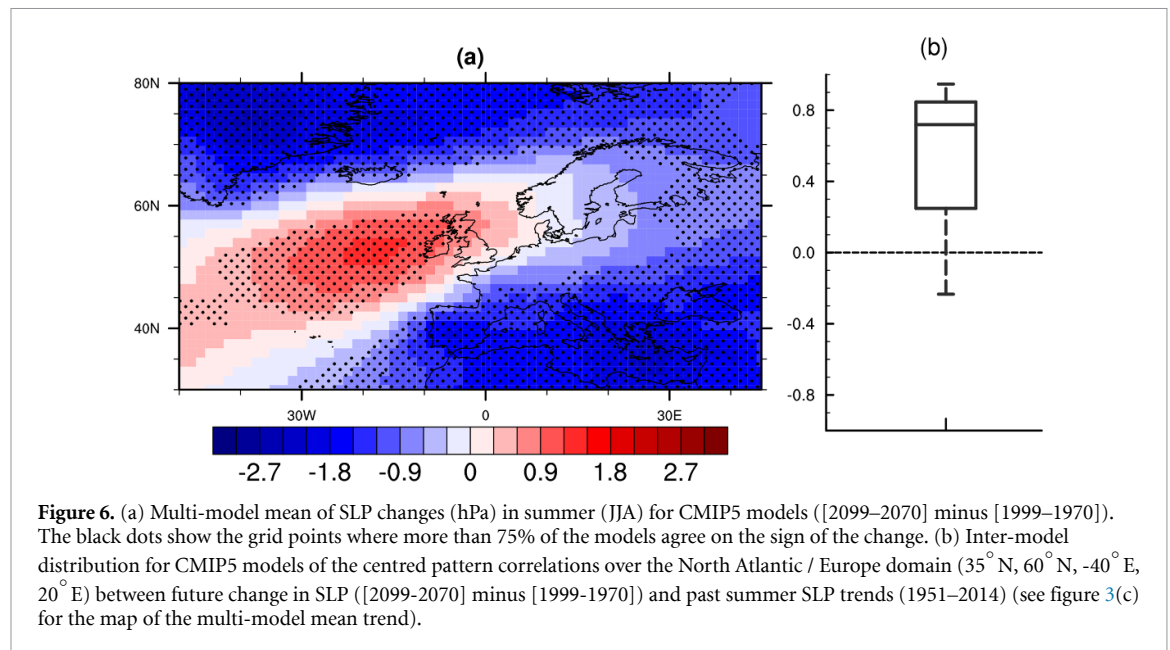
Figure 5. (a) Multi-model ensemble mean trend in surface temperature (K over period) in CMIP5 preindustrial simulations on the 64-year period in each simulation for which the centred pattern correlation of SLP trends over the western North Atlantic / Europe domain (35°N , 60°N , -40°E , 20°E) with the observed 1951–2014 SLP trend (NCAR SLP dataset) is the largest. (b) Same as (a) for SLP (hPa over period). (c) same as (a) except that the 64-year period with the largest centred pattern correlation of SLP trends over the western North Atlantic / Europe domain with the 1951–2014 trend SLP trend simulated in ensemble mean by PRIMAVERA forced atmospheric models is searched in each preindustrial simulation (K over period). (d) Same as (c) for SLP (hPa over period). (e) Difference (c) minus (a), (K over period).

the North-Atlantic opposite to the observed trend can incidentally be seen in their study (based on a few CMIP5 models), and, interestingly, seems to be mainly associated with greenhouse gases and tropospheric ozone.

The differences between observed and simulated temperature trends on the 1951–2014 period in summer (figures 1 and 2) could be related to the apparent general difficulty of climate models to capture the observed SLP trend (figures 3 and 4). Indeed, observed circulation patterns that project on the observed trend, i.e. with a negative pressure anomaly over the North Atlantic off Europe, are associated with anomalous southwesterly advection over western Europe and therefore large positive temperature anomalies there (figure S5(a)). Conversely, circulation patterns that are anti-correlated with the observed SLP trend, as the simulated trends generally are, tend to cool western Europe (figure S5(b)).

To further assess how the dissimilarity of simulated and observed SLP trends impacts temperature trends, we analyse following the methodology described in section 2.2 how purely internally-generated SLP trends, similar either to the observed or simulated SLP 1951–2014 trends, impact summer temperature.

64-year periods with a trend very similar (in terms of pattern correlation) to the observed one on 1951–2014 can be found in all CMIP5 preindustrial simulations. The highest pattern correlation varies between 0.78 and 0.90 depending on the CMIP5 model. Figure 5(b) confirms that the corresponding SLP trend is similar to the observed trend. As physically expected, such a SLP trend tends to warm western Europe and cool eastern Europe (figure 5(a)). 64-year periods with a SLP trend very similar to the PRIMAVERA AMIP ensemble mean are also found in all CMIP5 preindustrial simulations



(the pattern correlation for the selected periods is between 0.84 and 0.95 depending on the CMIP5 model). The associated multi-model mean SLP trend is therefore similar to the 1951–2014 SLP trend in the forced atmospheric simulations, and also CMIP5 and PRIMAVERA coupled simulations (figure 5(d)). Such a SLP pattern tends to cool western Europe and warm eastern Europe (figure 5(c)). In the end, the differences of temperature trends due to the differences of SLP trends reach 0.9 K over period, and are negative over western Europe and positive over eastern Europe (figure 5(e)). This large-scale pattern is very similar to the difference of 1951–2014 temperature trends between climate models and observations shown earlier (figure 1), except for local differences in the models over the Baltic region, almost certainly due to local processes. The magnitudes are also comparable.

Even if the observed and simulated impact of large-scale atmospheric circulation on 1951–2014 warming is relatively modest, as this impact is opposite in the observations and in the models, in the end, large-scale circulation plays a major role in the difference of temperature trends between models and observations.

Boé *et al* (2009) have shown that changes in large-scale circulation play an important role on climate change over Europe in summer and notably on precipitation changes over north-western Europe (e.g. United Kingdom of Great Britain and Northern Ireland, France, Benelux). Interestingly, future SLP changes in CMIP5 GCMs look very similar to the 1951–2014 trends in terms of spatial pattern (figure 6(a)), with a very good model agreement on an increase in SLP over the North-Atlantic off Europe and a decrease over continental Europe and the Mediterranean. Future SLP changes also strongly

project on 1951–2014 trends for the vast majority of individual models (figure 6(b)). These results strongly suggest that both SLP trends over the North Atlantic / Europe domain on the 1951–2014 period and future changes in summer are shaped to an important extent by external forcings. The apparent difficulty of state-of-the-art climate models, even when forced by observed SSTs, to capture past observed SLP trends therefore raises questions about the realism of future projected SLP changes over the North-Atlantic / Europe sector, with implications regarding climate changes over Europe.

4. Conclusion

Since the mid-20th century, a large warming trend has been observed over western Europe in summer. Over eastern Europe, the trend has been much smaller. This spatial pattern is associated with a decrease of sea level pressure over the North Atlantic off Europe and an increase over Scandinavia, which tend to increase warm air advection towards western Europe and cold air advection towards eastern Europe.

Past generation (CMIP5) and new generation (PRIMAVERA) climate models, even when forced by observed SSTs and sea ice, generally tend to warm less than observed over western Europe and more than observed eastern Europe on the 1951–2014 period. However, given the large impact of internal variability on trends, even in a forced-atmospheric framework, new generation climate models are generally compatible with the observed summer warming over western Europe.

The large-scale differences of summer temperature trends on the 1951–2014 period between models

and observations over Europe are explained to an important extent by differences in sea level pressure trends over the North Atlantic / Europe domain. Sea level pressure trends in climate models, even in AMIP simulations, indeed generally tend to be anti-correlated with the observed one, therefore reducing warming over western Europe and amplifying it over eastern Europe, while the opposite is true in the observations. Note that the SST and sea ice forcing fields used in the PRIMAVERA AMIP simulations come from a single dataset, and therefore the impact of observational uncertainties in that context is not assessed.

In addition to SLP trends, other mechanisms suggested by previous studies may still play in the simulated / modelled differences in temperature trends, such as for example the direct response to anthropogenic aerosols for example (e.g. Van Oldenborgh *et al* 2009).

No historical simulation with a sea level pressure trend very close to the observed one is found in past generation and new generation climate models, suggesting either an underestimation of internal variability in the models, or an unrealistic response of SLP to external forcings. The latter would have consequences regarding future climate changes over western Europe, especially since future SLP changes strongly project on past SLP trends.

Our study clearly illustrates the difficulty of evaluating individual models regarding properties strongly impacted by internal variability, such as trends. At best, it can be assessed how likely (or unlikely) is the observed trend within the distribution of simulated trends and large single model initial-condition ensembles are needed to estimate robustly the distribution of trends in a model. Progresses in this context have been made since CMIP5, but most CMIP6 models still do not provide such large ensembles for the moment.

Our study shows that increasing the horizontal model resolution has little impact regarding the discrepancy between observed and simulated SLP trends, for the range of resolutions explored in this study. Our study also shows that potentially unrealistic SST and sea ice cover trends in coupled climate models are very likely not to blame for the discrepancy between simulated and observed summer SLP trends. Very similar results are indeed obtained in coupled simulations and atmospheric simulations forced by observed SSTs and sea ice cover, making this discrepancy all the more puzzling. Future works to better understand the causes of the discrepancy are clearly needed.

Acknowledgments

The PRIMAVERA project is funded by the European Union's Horizon 2020 programme, Grant Agreement

no. 641727. We acknowledge the World Climate Research Program's Working Group on Coupled Modelling, which is responsible for CMIP, and we thank the climate modeling groups for producing and making available their model output. For CMIP the U.S. Department of Energy's Program for Climate Model Diagnosis and Intercomparison provides coordinating support and led development of software infrastructure in partnership with the Global Organization for Earth System Science Portals. Support for the Twentieth Century Reanalysis Project version 2c dataset is provided by the U.S. Department of Energy, Office of Science Biological and Environmental Research (BER), and by the National Oceanic and Atmospheric Administration Climate Program Office. We acknowledge the E-OBS dataset from the EU-FP6 project UERRA (<http://www.uerra.eu>) and the Copernicus Climate Change Service, and the data providers in the ECA&D project (<https://www.ecad.eu>). We acknowledge the developers of the other observational datasets used in this study and described in section 2: CRU TS4, Berkeley Earth System Temperature dataset, CRUTEM4, NCAR SLP, NCEP Reanalysis, HadSLP2r. The figures were produced with the NCAR Command Language Software (Boulder, Colorado: UCAR/NCAR/CISL/TDD. <http://dx.doi.org/10.5065/D6WD3XH5>).

Data availability statements

The data that support the findings of this study are openly available at:

CMCC-CM2-HR4: <https://doi.org/10.22033/ESGF/CMIP6.1359>
 CMCC-CM2-VHR4: <https://doi.org/10.22033/ESGF/CMIP6.1367>
 HadGEM3-GC31-LL: <http://doi.org/10.22033/ESGF/CMIP6.1901>
 HadGEM3-GC31-LM: <http://doi.org/10.22033/ESGF/CMIP6.1321>
 HadGEM3-GC31-HM: <http://doi.org/10.22033/ESGF/CMIP6.446>
 MPI-ESM1-2-HR: <https://doi.org/10.22033/ESGF/CMIP6.762>
 MPI-ESM1-2-XR: <https://doi.org/10.22033/ESGF/CMIP6.10290>
 EC-Earth3P: <https://doi.org/10.22033/ESGF/CMIP6.2322>
 EC-Earth3P-HR: <https://doi.org/10.22033/ESGF/CMIP6.2323>
 CNRM-CM6-1: <https://doi.org/10.22033/ESGF/CMIP6.1925>
 CNRM-CM6-1-HR: <https://doi.org/10.22033/ESGF/CMIP6.1387>
 ECMWF-IFS-LR: <http://doi.org/10.22033/ESGF/CMIP6.2463>
 ECMWF-IFS-HR: <http://doi.org/10.22033/ESGF/CMIP6.2461>

ORCID iDs

Julien Boé  <https://orcid.org/0000-0003-2965-4721>

Chris Roberts  <https://orcid.org/0000-0002-2958-6637>

Enrico Scoccimarro  <https://orcid.org/0000-0001-7987-4744>

Jon Seddon  <https://orcid.org/0000-0003-1302-1049>

Klaus Wyser  <https://orcid.org/0000-0001-9752-3454>

References

- Allan R and Ansell T 2006 A new globally complete monthly historical gridded mean sea level pressure dataset (HadSLP2): 1850–2004 *J. Clim.* **19** 5816–41
- Annan J D and Hargreaves J C 2010 Reliability of the CMIP3 ensemble *Geophys. Res. Lett.* **37** 1–5
- Benjamini Y and Hochberg Y 1995 Controlling the false discovery rate: A practical and powerful approach to multiple testing *J. R. Stat. Soc. Ser. B* **57** 289–300
- Bhend J and Whetton P 2013 Consistency of simulated and observed regional changes in temperature, sea level pressure and precipitation *Clim. Change* **118** 799–810
- Boé J 2016 Modulation of the summer hydrological cycle evolution over western Europe by anthropogenic aerosols and soil-atmosphere interactions *Geophys. Res. Lett.* **43** 7678–85
- Boé J 2018 Interdependency in multimodel climate projections: component replication and result similarity *Geophys. Res. Lett.* **45** 2771–9
- Boé J, Terray L, Cassou C and Najac J 2009 Uncertainties in European summer precipitation changes: role of large-scale circulation *Clim. Dyn.* **33** 265–76
- Boucher O, Randall D, Artaxo P, Bretherton C, Feingold G, Forster P, et al 2013 Clouds and Aerosols *Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change* ed T F Stocker et al (Cambridge: Cambridge University Press) 571–657
- Cherchi A, et al 2019 Global mean climate and main patterns of variability in the CMCC-CM2 coupled model *J. Adv. Model. Earth Syst.* **11** 185–209
- Compo G P, Whitaker J S, Sardeshmukh P D, Matsui N, Allan R J and Yin X et al 2011 The Twentieth Century Reanalysis Project *Q. J. R. Meteorol. Soc.* **137** 1–28
- 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
- Enfield D B, Mestas-Núñez A M and Trimble P J 2001 The Atlantic multidecadal oscillation and its relation to rainfall and river flows in the continental U.S. *Geophys. Res. Lett.* **28** 2077–80
- Eyring V, Bony S, Meehl G A, Senior C A, Stevens B and Stouffer R J et al 2016 Overview of the coupled model intercomparison project phase 6 (CMIP6) experimental design and organization *Geosci. Model Dev.* **9** 1937–58
- Gillett N P, Fyfe J C and Parker D E 2013 Attribution of observed sea level pressure trends to greenhouse gas, aerosol and ozone changes *Geophys. Res. Lett.* **40** 2302–6
- Gutjahr O, Putrasahan D, Lohmann K, Jungclaus J H, von Storch J S, Brüggemann N, Haak H and Stössel A 2019 Max Planck Institute Earth System Model (MPI-ESM1.2) for the High-Resolution Model Intercomparison Project (HighResMIP) *Geosci. Model Dev.* **12** 3241–81
- Haarsma R et al 2016 HighResMIP versions of EC-Earth: EC-Earth3P and EC-Earth3P-HR. Description, model performance, data handling and validation *Geosci. Model Dev. Discuss.* in review <https://doi.org/10.5194/gmd-2019-350>
- Haarsma R J, Roberts M J, Vidale P L, Senior C A, Bellucci A and Bao Q et al 2016 High Resolution Model Intercomparison Project (HighResMIP v1.0) for CMIP6 *Geosci. Model Dev.* **9** 4185–208
- Haarsma R J, Selten F M and Drijfhout S S 2015 Decelerating Atlantic meridional overturning circulation main cause of future west European summer atmospheric circulation changes *Environ. Res. Lett.* **2016** 10 094007
- Harris I, Jones P D, Osborn T J and Lister D H 2014 Updated high-resolution grids of monthly climatic observations - the CRU TS3.10 Dataset. *Int. J. Climatol.* **34** 623–42
- Jones P D, Lister D H, Osborn T J, Harpham C, Salmon M and Morice C P 2012 Hemispheric and large-scale land-surface air temperature variations: An extensive revision and an update to 2010 *J. Geophys. Res. Atmos.* **117**
- Kalnay E, Kanamitsu M, Kistler R, Collins W, Deaven D and Gandin L et al 1996 The NCEP/NCAR 40-Year Reanalysis Project *Bull. Am. Meteorol. Soc.* **77** 437–71
- Kennedy J, Titchner H, Rayner N, and Roberts M 2017 input4MIPs.MOHC.SSTsAndSeaIce.HighResMIP.MOHC-HadISST-2-2-0-0-0. Version 20170505 *Earth System Grid Federation*
- Kerr R A 2000 A North Atlantic Climate Pacemaker for the Centuries *Science* **288** 1984–5
- Lehner F, Deser C and Terray L 2017 Toward a new estimate of “time of emergence” of anthropogenic warming: Insights from dynamical adjustment and a large initial- condition model ensemble *J. Clim.* **30** 7739–56
- Mariotti A and Dell’Aquila A 2012 Decadal climate variability in the Mediterranean region: Roles of large-scale forcings and regional processes *Clim. Dyn.* **38** 1129–45
- Meehl G A, Covey C, Delworth T, Latif M, McAvaney B and Mitchell J F B et al 2007 The WCRP CMIP3 multimodel dataset: A new era in climatic change research *Bull. Am. Meteorol. Soc.* **88** 1383–94
- Nabat P, Somot S, Mallet M, Sanchez-Lorenzo A and Wild M 2014 Contribution of anthropogenic sulfate aerosols to the changing Euro-Mediterranean climate since 1980 *Geophys. Res. Lett.* **41** 5605–11
- 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
- Philipona R, Behrens K and Ruckstuhl C 2009 How declining aerosols and rising greenhouse gases forced rapid warming in Europe since the 1980s *Geophys. Res. Lett.* **36**
- 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
- Roberts C D, Senan R, Molteni F, Boussetta S, Mayer M and Keeley S 2018 Climate model configurations of the ECMWF Integrated Forecast System (ECMWF-42 IFS cycle 43r1) for HighResMIP *Geosci. Model Dev.* **11** 3681–3712
- Roberts M J and Coauthors 2019 Description of the resolution hierarchy of the global coupled HadGEM3-GC3.1 model as used in CMIP6 HighResMIP experiments *Geosci. Model Dev.* **12** 4999–5028
- Rohde R, Muller R, Jacobsen R, Perlmuter S, Rosenfeld A, Wurtele J, Curry J, Wickham C and Mosher S 2013 Berkeley Earth Temperature Averaging Process *Geoinf. Geostatistics An Overv.* **1** 2
- Ruprich-Robert Y, Msadek R, Castruccio F, Yeager S, Delworth T and Danabasoglu G 2017 Assessing the climate impacts of the observed atlantic multidecadal variability using the GFDL CM2.1 and NCAR CESM1 global coupled models *J. Clim.* **30** 2785–2810
- Shin S-I and Sardeshmukh P D 2010 Critical influence of the pattern of Tropical Ocean warming on remote climate trends *Clim. Dyn.* **36** 1577–91
- Squntu A A, van der Schrier G, Brugnara Y and Klein Tank A 2019 Homogenization of daily temperature series in the

- European climate assessment & dataset *Int. J. Climatol* **39** 1243–61
- Sutton R and Dong B 2012 Atlantic Ocean influence on a shift in European climate in the 1990s *Nat. Geosci.* **5** 788–92
- Sutton R T, McCarthy G D, Robson J, Sinha B, Archibald A T and Gray L J 2018 Atlantic multidecadal variability and the U.K. acsis program *Bull. Am. Meteorol. Soc.* **99** 415–25
- Taylor K E et al 2018 CMIP6 Global Attributes, DRS, Filenames, Directory Structure, and CV's 10 September 2018 v 6.2.7 Document short URL: <https://goo.gl/v1drZl>
- Taylor K E, Stouffer R J and Meehl G A 2012 An Overview of CMIP5 and the Experiment Design *Bull. Am. Meteorol. Soc.* **93** 485–98
- Trenberth K E and Paolino D A 1980 The Northern Hemisphere sea-level pressure data set: trends, errors and discontinuities *Mon. Weather Rev* **108** 855–72
- Van Oldenborgh G J, Drijfhout S, Van Ulden A, Haarsma R, Sterl A and Severijns C et al 2009 Western Europe is warming much faster than expected *Clim. Past* **5** 1–12
- Vannière B, Demory M E, Vidale P L, Schiemann R, Roberts M J and Roberts C D et al 2019 Multi-model evaluation of the sensitivity of the global energy budget and hydrological cycle to resolution *Clim. Dyn.* **52** 6817–46
- Voldoire A, SaintMartin D, Sénési S, Decharme B, Alias A and Chevallier M et al 2019 Evaluation of CMIP6 DECK Experiments With CNRMCM61 *J. Adv. Model. Earth Syst.* **11** 2177–213
- Wilks D S 2016 “The stippling shows statistically significant grid points”: how research results are routinely overstated and overinterpreted and what to do about it *Bull. Am. Meteorol. Soc.* **97** 2263–73
- Zwiers F W and von Storch H 1995 Taking serial correlation into account in tests of the mean *J. Clim.* **8** 336–51