Improved atmospheric circulation over Europe by the new generation of CMIP6 earth system models

Juan A. Fernandez-Granja[1] · Ana Casanueva[1] · Joaquin Bedia[1] · Jesus Fernandez[1]

Abstract

Global Climate Models (GCMs) generally exhibit significant biases in the representation of large-scale atmospheric circulation. Even after a sensible bias adjustment these errors remain and are inherited to some extent by the derived downscaling products, impairing the credibility of future regional projections. In this study we perform a process-based evaluation of state-of-the-art GCMs from CMIP5 and CMIP6, with a focus on the simulation of the synoptic climatological patterns having a most prominent effect on the European climate. To this aim, we use the Lamb Weather Type Classification (LWT, Lamb British isles weather types and a register of the daily sequence 736 of circulation patterns 1861-1971. METEOROL OFF, GEOPHYS MEM; 737 GB; DA 1972; NO 116; PP 1-85; BIBL 2P1/2, 1972), a subjective classification of circulation weather types constructed upon historical simulations of daily mean sea level pressure. Observational uncertainty has been taken into account by considering four different reanalysis products of varying characteristics. Our evaluation unveils an overall improvement of salient atmospheric circulation features consistent across observational references, although this is uneven across models and large frequency biases still remain for the main LWTs. Some CMIP6 models attain similar or even worse results than their CMIP5 counterparts, although in most cases consistent improvements have been found, demonstrating the ability of the new models to better capture key synoptic conditions. In light of the large differences found across models, we advocate for a careful selection of driving GCMs in downscaling experiments with a special focus on large-scale atmospheric circulation aspects.

Keywords Lamb weather type classification · Comparison CMIP5-CMIP6 · Process-based evaluation

1 Introduction

The seasonal variability of large-scale mean sea level pressure patterns exerts a direct influence on the regional European climate. Different mechanisms explain this relationship, such as the influence of the North Atlantic Oscillation pattern (NAO, Hurrell et al. 2003; Folland et al. 2009) or blockings characterized by persistent high pressure systems (Rex 1950; Jury et al. 2019). These are related, for instance, to extreme seasonal temperature events (Buehler et al. 2011; Barriopedro et al. 2011; Favà et al. 2015), precipitation dry/wet spells and extremes (Busuioc et al. 2001; Casanueva et al. 2014; Sousa et al. 2017) or droughts (Bladé et al. 2011)

due to their capability to disturb the predominant cyclonic westerly flow (Sillmann and Croci-Maspoli 2009). As a result, an adequate representation of atmospheric circulation and high/low pressure variability becomes essential for a proper representation of the main regional climate features, although current Global Climate Models (GCMs) exhibit substantial errors in this sense (Vial and Osborn 2012; Dawson et al. 2012; Masato et al. 2013).

Circulation biases affect the centroids location and spatial patterns as well as the frequency and duration of the main Euro-Atlantic wintertime weather regimes (Dawson et al. 2012; Fabiano et al. 2020) and Atlantic and European winter blocking events (Vial and Osborn 2012; Anstey et al. 2013). For instance, the frequency of the latter are systematically underestimated, also by the state-of-the-art simulations of the 5th Coupled Model Intercomparison Project (CMIP5, Taylor et al. 2012). The representation of Northern Hemisphere storm tracks has improved in CMIP5 GCMs with respect to previous model versions (Zappa et al. 2013), but

✉ Juan A. Fernandez-Granja
  juan.fernandez@unican.es

1 Meteorology Group, Department of Applied Mathematics and Computer Science, University of Cantabria, 39005 Santander, Spain

they still underestimate cyclone intensity and present location biases (Chang et al. 2012; Colle et al. 2013). Likewise, CMIP5 GCMs are able to capture eastern Mediterranean weather regimes qualitatively, although they fail in reproducing quantitative features (Dawson et al. 2012; Hochman et al. 2019). The most recent generation of GCMs (CMIP6, Eyring et al. 2016) shows substantial improvements with respect to CMIP5 in the representation of the frequency and persistence of circulation types worldwide (Cannon 2020), although more focused analyses are still needed to adequately assess the implications at a regional scale for downscaling purposes (Addor et al. 2016; Perez et al. 2014; Otero et al. 2018).

GCM uncertainty emerges as an important source of uncertainty in regional future climate projections. To date, coordinated downscaling experiments over Europe have explored uneven combinations of global and regional climate models (RCMs), favouring certain GCMs and being biased towards large ensembles of few RCMs (Fernández et al. 2019). In light of the new EURO-CORDEX activities (Jacob et al. 2014, 2020), decisions must be taken towards the implementation of an optimal experimental design, including the selection of driving GCMs from the CMIP6 ensemble. GCM selection is usually a two-step process (McSweeney et al. 2015), requiring, first, the plausibility of the GCMs climate and, second, that the selected GCMs cover a large fraction of the climate alternatives spanned by the full CMIP ensemble. The selection of GCMs based on their ability to adequately simulate particular surface variables, such as temperatures and/or precipitation, is inadequate and may result in a non-optimal selection of driving GCMs. The idea of evaluating GCM performance by means of atmospheric patterns and weather types started long time ago (Jones et al. 1993; Hulme et al. 1993), although process-based GCM performance assessments have recently come into focus particularly within the downscaling community (Maraun et al. 2017). Such evaluation, which relies on GCM variables used by subsequent downscaling, is preferred (Brands et al. 2013; McSweeney et al. 2015; Addor et al. 2016). A sensible bias correction approach can substantially improve raw model fields from a statistical point of view, and it is advised for specific variables and threshold-dependent climate indices (see e.g. Dosio 2016; Iturbide et al. 2020). However, it is problematic to circumvent fundamental model errors, such as the misrepresentation of large-scale atmospheric circulation (Addor et al. 2016; Maraun et al. 2017). Even though RCMs can add value in this sense by improving the misrepresentation of the driving data defining their lateral boundary conditions (Jones et al. 1995), this improvement is incomplete, particularly when large errors are present in the driving GCM (Diaconescu and Laprise 2013). Moreover, even when bias correction methods improve the applicability of climate simulations, in general it cannot

improve low model credibility, and may even hide the lack of credibility of model outputs when applied inadequately (Maraun et al. 2017) resulting in ill-informed adaptation decisions. As a result, the selection of the driving GCM has a large effect on the skill of RCM simulations (as shown e.g. by Prein et al. 2019, in North America), which also has an impact on the projected signals (Turco et al. 2013), the GCM choice thus being an issue of paramount importance in GCM-RCM intercomparison experiments.

In this study, we categorize the circulation patterns of the new-generation CMIP6 GCMs over Europe according to the Lamb Weather Type (LWT) classification (Lamb 1972). We systematically compare each new model with respect to the previous CMIP5 counterpart. We focus over Europe to address specifically the selection of GCMs for downscaling exercises in the context of EURO-CORDEX. In particular, we aim to (1) assess the potential improvement of CMIP6 over CMIP5 GCMs regarding the representation of the frequency and transition probability between relevant circulation types and (2) provide a quantitative ranking of models, to aid in the plausibility step of model selection over Europe. This work updates earlier work on the ability of GCMs to represent circulation types in this region (Perez et al. 2014; Otero et al. 2018) and introduces transition probabilities as a stringent test on model performance.

2 Methodology and data

2.1 Lamb weather type classification

The LWT classification is a subjective clustering approach where the weather type classification is based on a number of rules relying on meteorological expert knowledge (Lamb 1972). This differs from objective clustering algorithms, which are data driven. Therefore, LWT classification is deterministic and it has a straightforward and well defined physical interpretation. This is an advantage for the aims of this study, since the results obtained can be interpreted in terms of actual meteorological conditions, and there is no source of added uncertainty as in stochastic clustering algorithms, whose results are initialization-dependent.

Following previous studies using the LWT scheme we classify all days in 26 classes that are assigned to a specific circulation type (see e.g.: Trigo and DaCamara 2000; Brands et al. 2014; Ramos et al. 2014; Pereira et al. 2018). In order to produce the LWTs, we follow the formulation developed by Jenkinson and Collison (1977) and Jones et al. (1993) using daily mean sea level pressure (MSLP) over a grid of 16 points, centered in the British Isles (55°N, 5°W) and with a separation of 5° latitude by 10° longitude between each couple of points (Fig. 1). The model grid cells corresponding

to each reference point are located using nearest neighbours (Jones et al. 2013; Pereira et al. 2018).

The formulation of the LWTs uses 6 parameters related with wind-flow characteristics: southerly flow, westerly flow, total flow, southerly shear vorticity, westerly shear vorticity and total shear vorticity. Depending on their values, the daily MSLP is classified in a given weather type. There are 26 LWTs representing pure cyclonic (C) and anticyclonic (A) circulation over the center point, 8 pure directional types (N, NE, E, ..., NW) and hybrid types (mixing A or C with any of the directional types). As an example, Figure 1 shows composite MSLP maps of the 8 most common LWTs over an extended European domain as derived from the ERA-Interim (Dee et al. 2011) reanalysis. These 8 LWTs gather 74% of the days and are consistent with previous studies (Trigo and DaCamara 2000; Brands et al. 2014; Fealy and Mills 2018). In this study, we use the implementation of the LWTs in the R package transformeR (v1.7.3, Iturbide
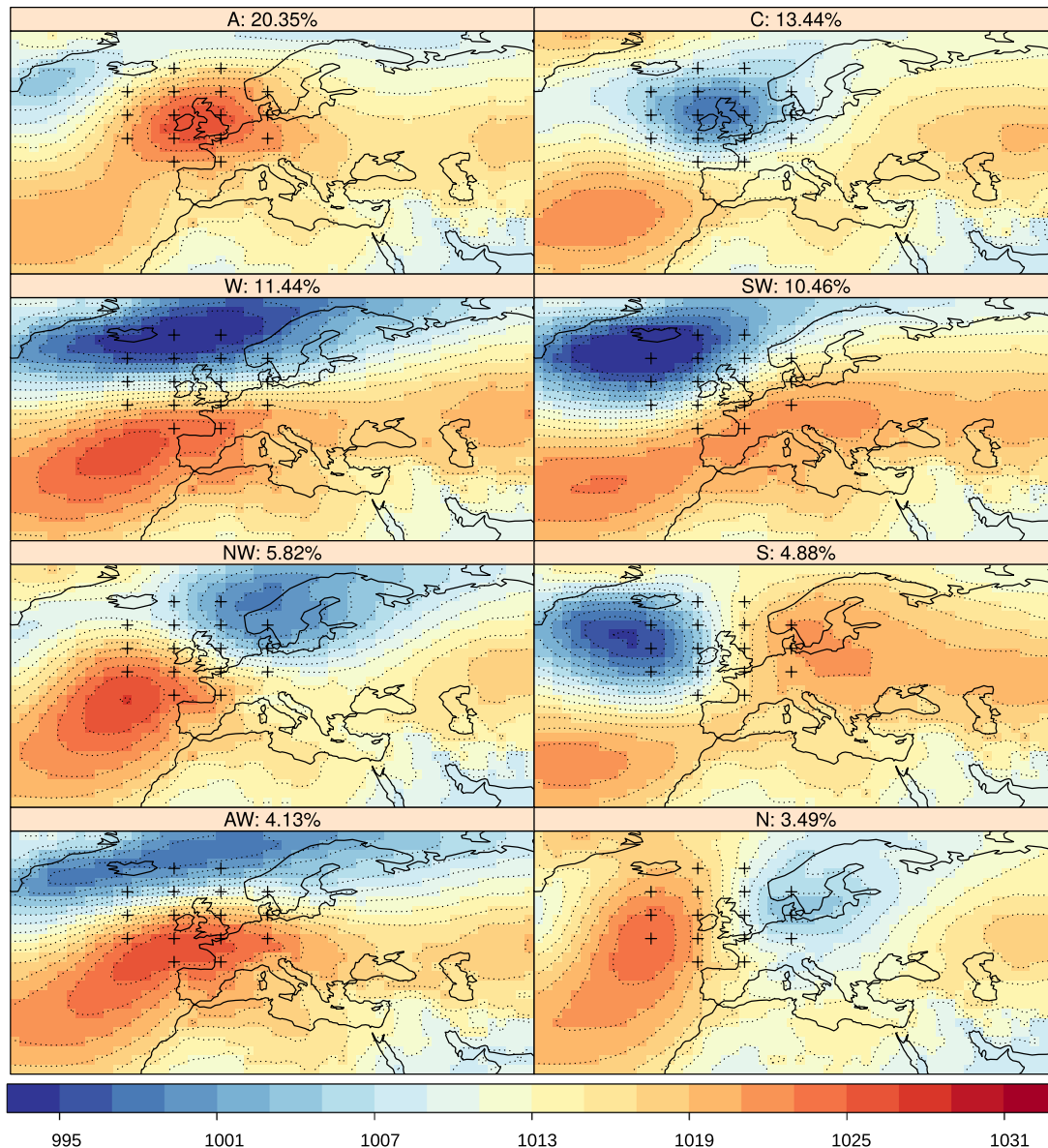

Fig. 1 Composite maps of Lamb Weather Types (LWTs) derived from MSLP (hPa) from ERA-Interim for the period 1981–2010. A subset of the 8 (out of 26) most frequent LWTs annually is displayed. Sub-panels are labelled with their LWT abbreviation (frequency in %) and sorted in decreasing frequency order from top to bottom and from left to right. Colorbar is centered on average sea level atmospheric pressure (reds are highs and blues are lows). Lamb's grid coordinates are also indicated over the British Isles domain. Similar composite maps are calculated for the GCMs and reanalyses gathered in Table 1; their spatial correlations with the ERA-Interim pattern are shown in Fig. A17 in the Electronic Supplementary Material

et al. 2019), illustrated in the companion paper notebook (see "Availability of data and materials").

One salient feature of a weather type is its probability of occurrence, which can be estimated by the relative frequency of occurrence in a sample. Not surprisingly for a mid-latitude region, the A and C types are the most common types in Europe (Fig. 1), followed by all westerly types (W, SW and NW). LWT persistence probabilities (understood here as the probability of staying in the same weather type as the previous day), or more generally, transition probabilities between two different LWTs are also important, since they determine key temporal features such as spell duration, serving as an effective tool for the assessment of the model ability to reproduce atmospheric circulation patterns (Hochman et al. 2019). Let the discrete random variable X_t represent the LWT at time step t, whose values $x_t \in \{1, \ldots, K\}$, with $K = 26$ the total number of LWTs. We consider this variable at two consecutive days, X_{t-1} and X_t, to construct the $K \times K$ transition probability matrix A, where $A_{ij} = p(X_t = j | X_{t-1} = i)$, representing the probability of going from LWT i to LWT j. Hence, each row of the matrix sums to one, $\sum_j A_{ij} = 1$. The transition probability matrix (TPM) thus provides a visual "fingerprint" on how a given model represents the LWT classification, which can be compared to the observational reference through specific evaluation measures (see Sect. 2.3).

2.2 Data

We applied the LWT methodology to classify daily MSLP patterns from GCMs, run under the CMIP historical experiment, and from reanalyses, as quasi-observational reference. In all cases, we considered the 30-year period 1981-2010, which follows the World Meteorological Organization (WMO) guidelines on the calculation of climate normals (WMO 2017) and represents a typical historical period in climate projections assessments. This period leads to a sample of ca. 11000 days per data set.

2.2.1 GCM data

GCM simulations from CMIP5 and CMIP6 historical experiments were used to evaluate different model generations. A set of 9 model pairs (Table 1) was selected to specifically account for model improvement as a factor in our analyses. Each GCM pair was developed in a different modelling center, although this does not guarantee model independence (Boé 2018). As CMIP5 historical experiment ends in 2005, we used the period 2006-2010 from the RCP8.5 scenario run to fill the common 1981-2010 analysis period. This has been done in previous studies (e.g. Casanueva et al. 2020) and there is not an expected impact on the results, since the difference in the forcing across scenarios is very small for the filled period.

Table 1 Set of CMIP5 and CMIP6 models used in the study, their nominal resolution at the equator (in °) and modelling center (top). Reanalysis products used (bottom)

CMIP5	Grid (°)	CMIP6	Grid (°)	Modelling Center (CMIP5) CMIP6
CanESM2	2.81	CanESM5	2.81	Canadian Centre for Climate Modelling and Analysis
CNRM-CM5	1.41	CNRM-CM6-1	1.41	Centre National de Recherches Météorologiques - Centre Européen de Recherche et de Formation Avancée en Calcul Scientifique
EC-EARTH	1.13	EC-EARTH3	0.70	(Irish Centre for High-end Computing) EC-EARTH Consortium
GFDL-ESM2M	2.26	GFDL-ESM4	1.00	NOAA - Geophysical Fluid Dynamics Laboratory
HadGEM2-ES	1.59	UKESM1-0-LL	1.59	Met Office Hadley Centre
IPSL-CM5-LR	2.96	IPSL-CM6A-LR	1.98	Institut Pierre-Simon Laplace
MIROC5	1.41	MIROC6	1.41	(Japan Agency for Marine-Earth Science and Technology, JAMSTEC) JAMSTEC, AORI, NIES and R-CCS
MPI-ESM-LR	1.88	MPI-ESM1-2-LR	1.88	Max Planck Institute for Meteorology
NorESM1-M	2.21	NorESM2-LM	2.21	Norwegian Climate Center - Norwegian Meteorological Institute
		Reanalysis	Grid (°)	Modelling Center
		ERA-Interim	0.75	European Center for Medium Range Weather Forecasts
		JRA-55	0.56	Japanese Meteorological Agency
		NCEP	2.5	National Centers for Environmental Prediction / National Center for Atmospheric Research
		ERA-20C	1.13	European Center for Medium Range Weather Forecasts

2.2.2 Reanalysis data

We used the European Center for Medium Range Weather Forecasts (ECMWF) ERA-Interim reanalysis (Dee et al. 2011) as the main quasi-observational reference to evaluate the model simulations. This state-of-the-art reanalysis is commonly used to evaluate model performance and also provided initial and lateral boundary conditions for CORDEX evaluation simulations. Therefore, it is also natural to use it here to evaluate GCM boundary conditions over Europe.

Moreover we considered three additional reanalysis products (Table 1) to account for observational uncertainty: the Japanese Meteorological Agency 55-year reanalysis (JRA-55; Kobayashi et al. 2015; Harada et al. 2016), the National Centers for Environmental Prediction / National Center for Atmospheric Research (NCEP–NCAR) reanalysis products (hereafter NCEP, Kalnay et al. 1996), and the ECMWF ERA-20C (Poli et al. 2016). The latter assimilated only surface pressure and marine winds, so it is not exactly comparable to the others, which assimilate a wider range of surface, upper-air and satellite observations.

2.3 Evaluation measures

In order to evaluate the accuracy of GCMs from different generations some indices are used such as Kullback–Leibler divergence (KL), Relative Bias, Two-proportions Z-Test and Transition Probability Matrix Score (TPMS). With these metrics, we provide a direct comparison between the GCMs and the ERA-Interim reanalysis and a quantitative value of the degree of similarity/agreement between them.

Kullback–Leibler Divergence This measure (KL; Kullback and Leibler 1951), also known as relative entropy, is used to quantify the degree of disparity between the GCMs and the reanalysis in the representation of the different LWT probabilities. For this purpose, the LWT classifications obtained by the GCMs and reanalysis are handled as discrete Probability Mass Functions (PMFs), whose dissimilarity is measured through KL divergence (see e.g. Jiang et al. 2011; Sharma and Seal 2019). The use of KL divergence in the comparison of two PMFs is more appropriate than using a distance function on a metric space (e.g. Euclidean distance) due to multiple facts: the PMFs may be differently distributed, have different sample sizes, different geometric centers or contain extreme probabilities that may disrupt the comparison negatively (Weijs et al. 2010; Jiang et al. 2011). Therefore, the KL divergence is not symmetric, and it is not affected by any biases derived from the probability of the samples, thus avoiding the more frequent LWTs unduly influencing the evaluation results.

The KL divergence of a discrete probability distribution, $P(x)$, with respect to another, $Q(x)$, both defined on the same probability space X (in our case, spanned by the LWTs) is defined within the Information Theory (Cover and Thomas 2006) as:

$$KL(P \parallel Q) = \sum_{x \in X} P(x) \log \frac{P(x)}{Q(x)} \tag{1}$$

We use it as a measure of the statistical "distance" of the model distribution ($P(x)$) with respect to the reanalysis one ($Q(x)$), which is zero for a perfect match ($P(x) = Q(x) \, \forall x \in X$) and takes positive values with no upper bound for increasingly different distributions. Here, we use the KL divergence implementation of the R package `phylentropy` (v0.4.0, Drost 2018)

Relative Bias From the historical record of observed weather types occurring at discrete time steps $X_1, X_2, \ldots, X_T$, with T days, the frequency of occurrence of the LWT ℓ per season s is denoted as $f(\ell, s)$ and calculated as the number of days falling in type ℓ divided by the total number of days in the season $s \in \{DJF, MAM, JJA, SON\}$. Thus we consider the relative bias ε to compute the deviation of the LWT frequency with respect to a reference data set:

$$\varepsilon_m(\ell, s) = \frac{f_m(\ell, s) - f_o(\ell, s)}{f_o(\ell, s)} \tag{2}$$

where $f_m(\ell, s)$ refers to the frequency in the model m and $f_o(\ell, s)$ is the reference observed frequency (in this case, derived from the ERA-Interim reanalysis). The model (m) can be any of the list of 21 models conformed by the 9 CMIP5 GCMs, the 9 CMIP6 GCMs and the reanalysis products: JRA, ERA-20C and NCEP (Table 1). The relative bias is a non-dimensional measure, which is zero for a perfect agreement of frequencies.

Two-Proportions Z-Test The Two-Proportions Z-Test is used to assess the statistical significance in the differences between models and ERA-Interim. It is used for proportions, which in this case arise from relative frequencies (proportion of days classified in a given LWT) and transition probabilities (proportion of days in LWT i with transition to LWT j). The test statistic takes into account the potentially different sample size in the model and reanalysis data, and the implementation used (`prop.test` function from the R package `stats` (v3.6.3, R Core Team 2020)) includes an exact test for small samples. This test was performed for each combination of LWT ℓ, season s and model m, using a 95% confidence to establish significant probability/relative frequency differences.

Transition probability matrix score In order to summarize the TPM information (Sect. 2.1) we introduce a TPM score (*TPMS*), that allows ranking model performance based on its TPM fingerprint, defined as:

$$TPMS = \sum_{p \in A^*} |p_m - p_o| \qquad (3)$$

where p_m and p_o are the transition probabilities in the model and in the observational reference, respectively, whose (absolute) difference is calculated considering the subset of transition probabilities A^* from the full matrix (A), that are significantly different from the reanalysis, following the two-proportion Z-Test. In order to include the "missing" transitions in the score (i.e. either transitions that exist in the reanalysis but are never simulated by the model, or transitions that are simulated by the model but do not occur in the reanalysis), these are assigned a zero probability (i.e. either $p_m = 0$ or $p_o = 0$) and included in the A^* subset. As a result, the larger the departure from zero (perfect agreement), the larger the dissimilarity of the TPM fingerprints between the GCM and the reanalysis.

3 Results

3.1 Observed LWTs

We first assay the resulting frequencies of the observed LWTs, as represented by the four reanalysis products. In Fig. 2, we show the LWT seasonal frequencies, sorted in decreasing order according to annual ERA-Interim LWTs. In general, small differences in the frequencies are found between the reanalysis for all seasons. The common set of prevailing LWTs has, however, different frequencies among seasons. In winter (DJF), Westerly (W) and Southwesterly
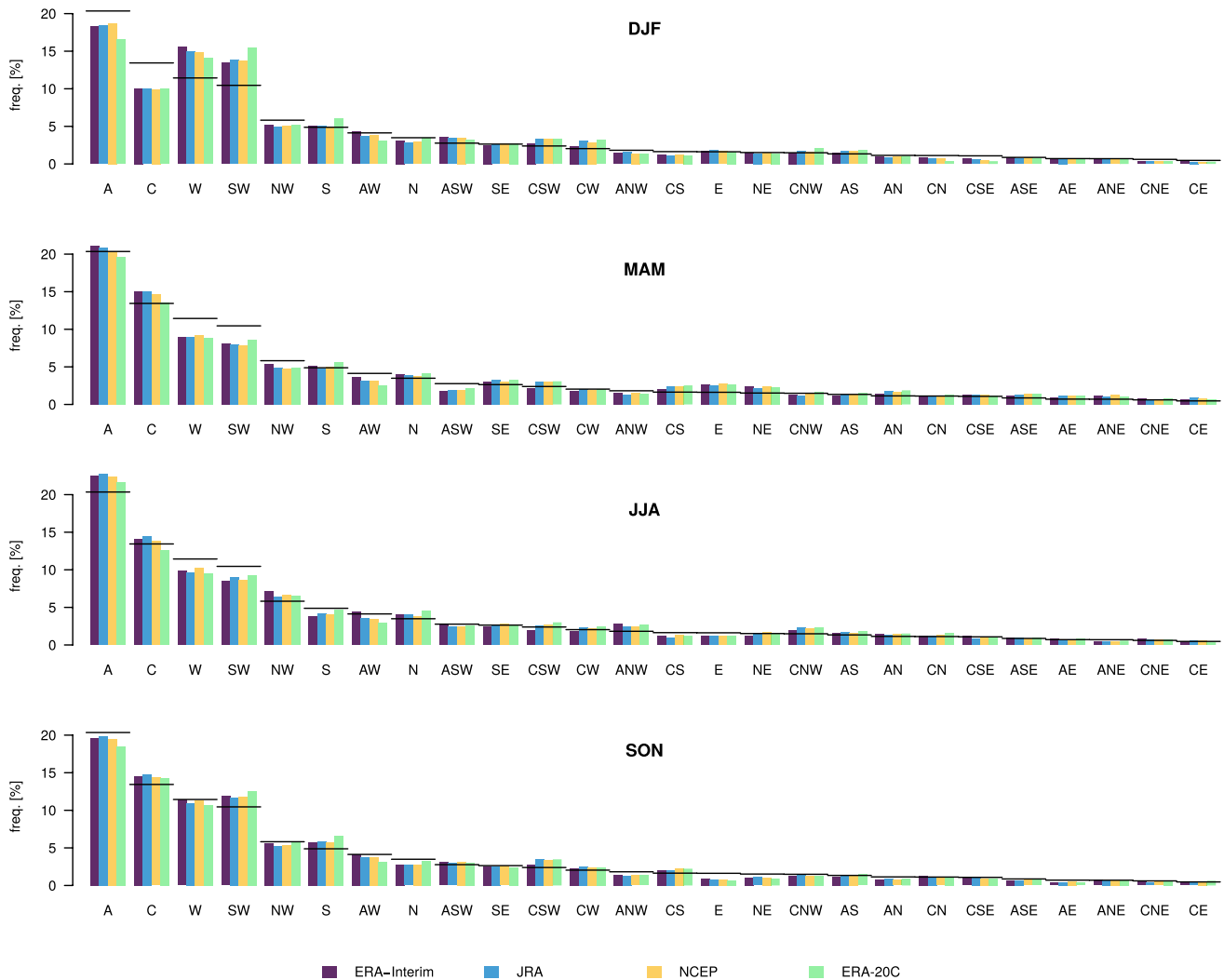

Fig. 2 Comparison of the seasonal relative frequencies of Lamb Weather Types (LWTs) obtained from the four different reanalysis (ERA-Interim, JRA, NCEP and ERA-20C) following the LWTs def-inition of Lamb 1972. The LWTs are sorted in decreasing order of their annual frequencies in ERA-Interim, indicated with horizontal segments as reference

(SW) flow types are more frequent than the Cyclonic (C) type, and they both exceed the annual time-scale reference. Westerly flow decays in spring and summer, and the Anticyclonic (A) type becomes more prevalent in summer. Types A, C , W and SW are the four most frequent LWTs in all seasons. Types S (South), NW (Northwesterly) and AW (Anticyclonic Westerly) are among the 8 most dominant in all seasons. Pure-directional type N (North) is also in the top-8 except in winter, when it is less frequent than type ASW (Anticyclonic Southwesterly). However, N type represents close to 5% of the days in all seasons and also appears among the first 8 LWTs for annual ERA-Interim. In light of these results, we consider the following LWT subset hereinafter for a more detailed analysis of model biases (Sect. 3.2): A, C, W, SW, NW, S, AW and N.

The observational uncertainty in the LWT relative frequencies is small, as their magnitudes are similar among the different reanalysis datasets, with the exception of ERA-20C (Fig. 2). This reanalysis shows lower LWT relative frequencies as compared to ERA-Interim, JRA and NCEP, especially in the two most frequent types (A and C), which is compensated mainly by an increased frequency in the S and SW flow types. This fact could be due to the different data sources of the ERA-20C reanalysis in comparison with the other available reanalysis products, which, in turn, might lead to differences in the LWTs classification. The ERA-20C reanalysis only assimilates sea level pressure data from surface-only observations in order to maintain consistency over time (Poli et al. 2016). In contrast, the rest of the reanalysis products –showing a more consistent LWT frequency distribution– assimilate many surface, upper-air and satellite observations (Fujiwara et al. 2017). Our findings are in line with previous literature, which highlights the poor representation of upper atmospheric processes because data from the free atmosphere are not available in surface-only input reanalyses. For example, lower cyclones in the Northern Hemisphere (Wang et al. 2006), fewer northern high-latitude blocking frequency (Rohrer et al. 2018), and lower occurrence of westerly circulation types (Stryhal and Huth 2017) have been detected for ERA-20C and other surface-only input reanalyses.

Figure 3 (left panel) depicts the similarity between the models (GCMs, ERA-20C, JRA and NCEP) with respect to ERA-Interim by using the KL Divergence. Again, among the reanalyses, ERA-20C shows the largest differences with respect to ERA-Interim ($KL = 0.008$) compared to the other reanalyses (0.003 for both the JRA and NCEP). Given the good agreement in the LWTs classification regardless of the use of ERA-Interim, JRA and NCEP, in the following we use ERA-Interim as reference. Further results considering the other reanalyses as reference are provided in the Electronic Supplementary Material for a more comprehensive picture of the reanalysis uncertainty. Interestingly, JRA and NCEP present a better agreement with ERA-20C than ERA-Interim in terms of KL divergence (Fig. A5 in the Electronic Supplementary Material). This aligns well with Chang and Yau (2016), who found that major storm tracks in the Northern Hemisphere in ERA-20C and JRA are in good agreement with radiosonde observations.

3.2 Modeled LWTs frequency

Model agreement with ERA-Interim reanalysis is analyzed first in terms of the KL divergence (Fig. 3, right panels). However, as annual KL divergence can hide the compensation of large biases, both annual and seasonal timescales are
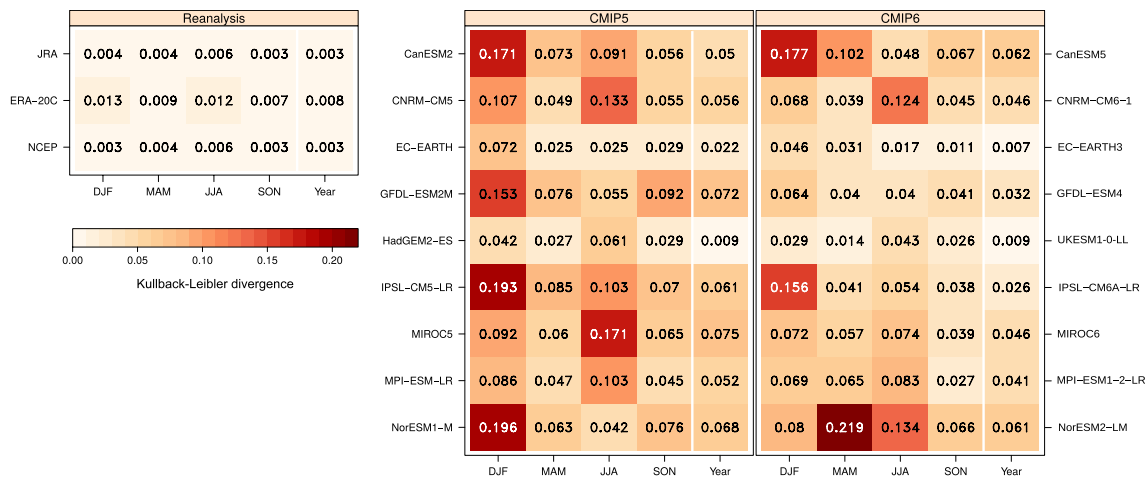

Fig. 3 Kullback–Leibler Divergence (KL) (seasonal and annual values, in columns) for the different reanalyses (left) and GCM experiments (right, CMIP5 and CMIP6). The 26 LWTs are considered in the calculation of KL. The 26 LWTs need to be considered as the KL formulation expects PMFs where the sum of the probabilities of the samples is equal to 1

later considered for the analysis of relative biases. Overall, there is a clear improvement from CMIP5 to CMIP6, although large KL divergences in CMIP5 in specific seasons only slightly diminish or move to another season in CMIP6. Similar conclusions hold when the other three reanalyses are used as reference (see Figs. A1, A3 and A5 in the Electronic Supplementary Material). At annual timescales, CMIP6 EC-EARTH3 exhibits the lowest deviation ($KL = 0.007$), followed by UKESM1-0-LL (0.009), HadGEM2 (0.009), EC-EARTH (0.022) and IPSL-CM6A-LR (0.026). EC-EARTH3 shows also slightly better performance than ERA-20C at the annual scale, which deteriorates in the seasonal analyses (e.g. $KL = 0.046$ in DJF, $KL = 0.031$ in MAM) probably due to biases in the timing along the year and the persistence of the weather types. The largest KL divergences occur in winter for most CMIP5 and CMIP6 models, followed by summer and spring. To explain such differences we next look at the seasonal GCM biases for the main LWTs. The KL divergence of the CMIP5 and CMIP6 models allows to rank them according to their ability to reproduce synoptic conditions with respect to their agreement with ERA-Interim. The general improvement of CMIP6 considering the annual KL divergence (Fig. 3) is also evident in terms of relative biases (Fig. 4). Overall, smaller biases are found for CMIP6, except for IPSL-CM6A-LR in winter, NorESM2-LM and CanESM5 in spring, and NorESM2-LM in summer. All models present the worst performance for the two most frequent LWTs (namely anticyclonic and cyclonic) in winter (in agreement with Fig. 3), with opposite sign biases. Along the four seasons, most models overestimate cyclonic type frequencies whereas they simulate too few anticyclonic conditions. The latter might be associated with the general underestimation of the frequency of the European winter blocking, which is a well-known drawback of CMIP5 models (see e.g. Masato et al. 2013). Overall, CMIP6 GCM

reduce biases in the frequency of the A and C types compared to the CMIP5 counterparts, especially NorESM2-LM and GFDL-ESM4, although statistically significant differences with ERA-Interim still remain.

Results are not conclusive for the other main LWTs, for which different magnitude and sign of biases are found depending on the model. The frequency of W and SW types is overestimated by NorESM2-LM and CanESM5 in spring (also NorESM2-LM in summer), performing worse than their CMIP5 counterparts. AW type is underestimated by most models in spring, regardless of the CMIP experiment. Most GCMs do not exhibit significant differences with respect to ERA-Interim for the least frequent weather types, especially in spring and autumn. GCM evaluation with respect to the three other reanalyses leads to similar conclusions and similar rankings (see Figs. A2, A4, A6 in the Electronic Supplementary Material).

Despite the improvement of CMIP6 models upon their CMIP5 predecessors, some biases still remain, which might be due to the limitations in simulating the most frequent conditions (such as A and C types) and the transitions from one type into another.

3.3 LWT transition probabilities

In order to shed some light on the biases found, we investigate the transition probabilities from one type to another, which might explain the misrepresentation of the synoptic conditions and their frequencies by most GCMs already depicted in Fig. 4. The TPM of ERA-Interim (Fig. 5a) provides the reference *fingerprint* of the transitions among LWTs and the persistence probability of a given LWT (diagonal cells). As expected, the largest probabilities of remaining in the same state are associated with the most frequent LWTs. In particular, more than 60% (50%) of the days with

Fig. 4 Relative Bias of LWT frequencies for the different reanalyses and GCM experiments (in rows: reanalyses in black, CMIP5 GCMs in red, CMIP6 GCMs in blue) for the four seasons (DJF, MAM, JJA and SON in columns). The rows are sorted following the ranking given by the annual KL Divergence in Fig. 3 (the seasonal rankings are given in brackets). Crosses indicate statistically significant values following a Z-test of proportions (Sect. 2.3)
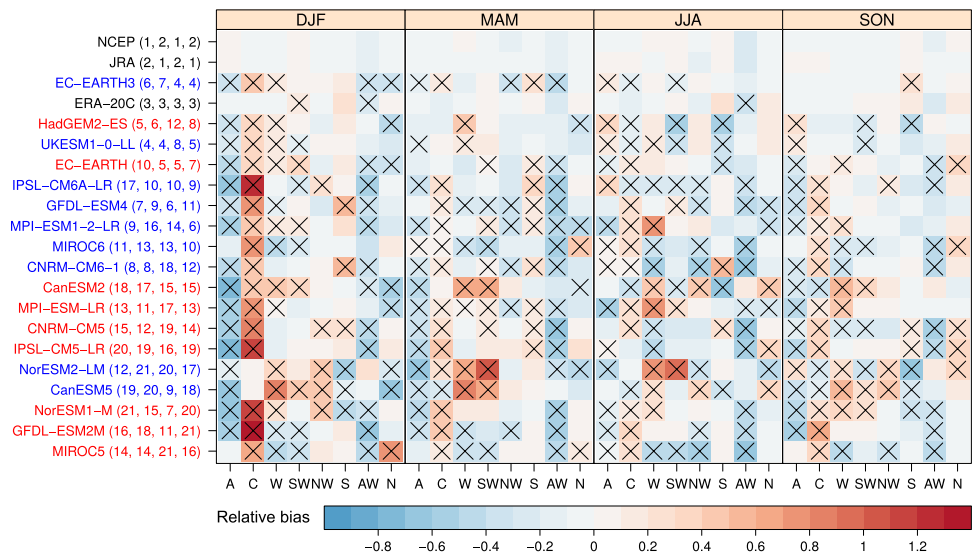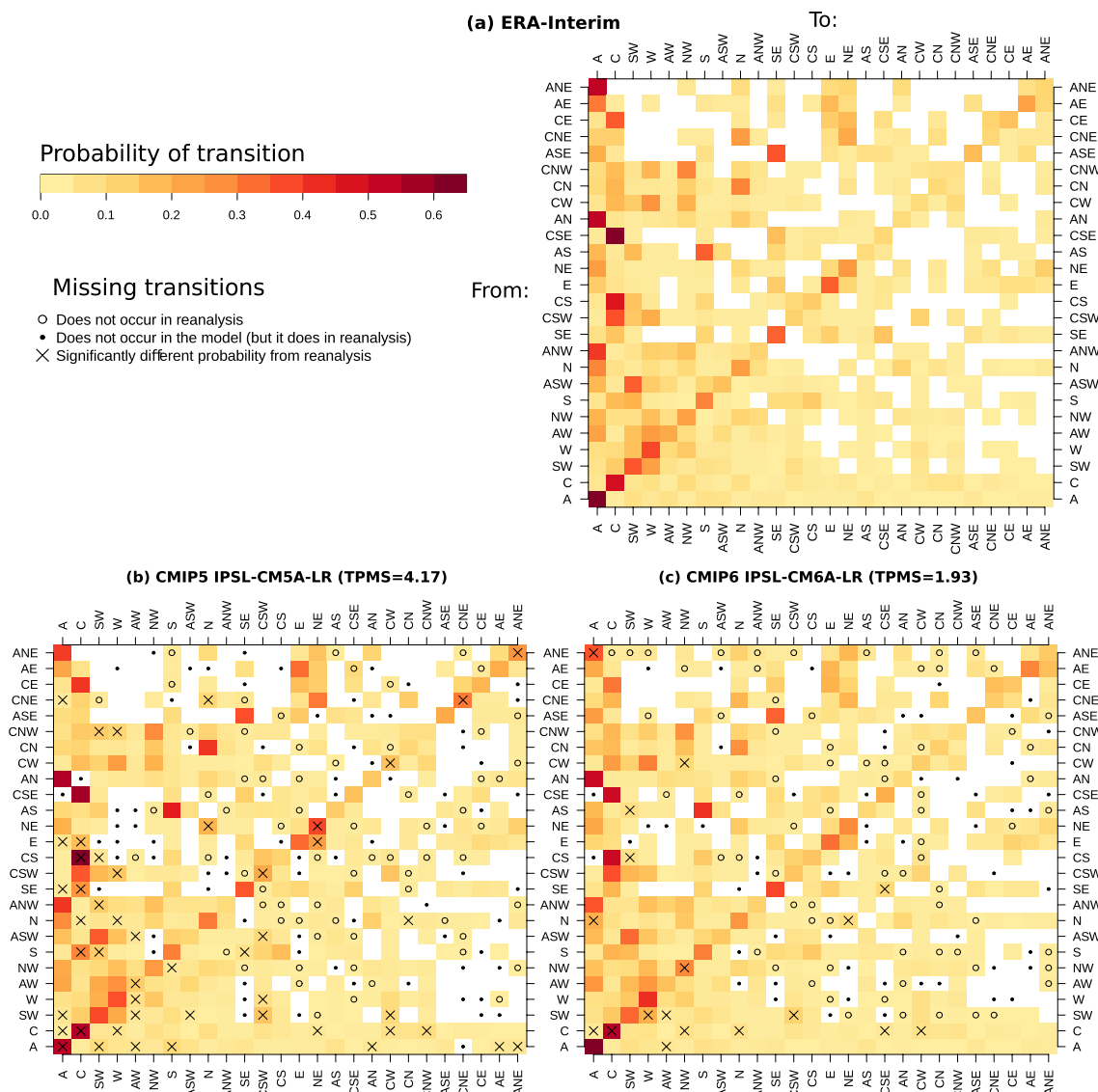
Fig. 5 Example of transition probability matrix (*A*) of Lamb Weather Types for ERA-Interim (**a**), the CMIP5 model IPSL-CM5A-LR (**b**) and its new version CMIP6 IPSL-CM6A-LR (**c**) for the historical period 1981–2010. $A_{ij} = p(X_t = j | X_{t-1} = i)$ represents the probability of going from LWT in row *i* to LWT in column *j*. Therefore, the persistence probability of a LWT can be found by looking at the diagonal of the matrix. Non observed transitions have been blanked to differentiate them from low-probability ones. Transition probabilities significantly different from those observed in ERA-Interim, obtained from a two-proportions Z-Test (Sect. 2.3), are indicated by crosses. In addition, LWT transitions simulated by the model but not observed in ERA-Interim are indicated by empty circles. Likewise, solid circles indicate LWT transitions not simulated by the model, but that occur in ERA-Interim. The corresponding TPMS values attained against ERA-Interim are indicated in parenthesis in the titles of panels (**b**) and (**c**)

type A (C) stay in the same LWT, followed by persistent SW, W, SE and E types (all above 30%). The most frequent transitions to a different state are from ANE, AN and ANW to A type and from CSE, CS, CSW to C type, all with probabilities above 40%. ASE to SE and AS to S type complete the picture of most common transitions. This pattern is in general very similar in the remaining reanalyses used as alternative references, with the largest deviations occurring in ERA-20C (see Figs. A7 and A8 of the Electronic Supplementary Material).

Overall, the ability of the GCMs to reproduce qualitatively the reference TPM regardless of the CMIP generation is remarkable (see example in Figs. 5b-c and also A9-A16 in the Electronic Supplementary Material). All GCMs *fingerprints* are able to capture fairly well the pattern of the reference ERA-Interim, although there are important departures in the magnitude of their probabilities in some cases. As a result, most GCMs fail to achieve the high persistence probabilities of the most frequent cyclonic and anticyclonic LWTs. In particular, attending to statistical significance of

their probabilities, the high persistence probability of the anticyclonic LWT is only adequately reproduced by a few models, namely CMIP5 EC-EARTH and HadGEM2-ES (Figs. A9 and A10 of the Electronic Supplementary Material), and the CMIP6 models IPSL-CM6A-LR (Fig. 5c) and UKESM1-0-LL (Fig. A10 of the Electronic Supplementary Material). The persistence probability of the purely cyclonic LWT (the second most frequent in the historical record) is significantly reproduced by the CMIP5 models EC-EARTH, HadGEM2-ES, MPI-ESM-LR, as well as their CMIP6 counterparts (Figs. A9, A10 and A13, respectively, Electronic Supplementary Material), CMIP5 CanESM2 (Fig. A15, left panel) and CMIP6 NorESM2-LM (Fig. A14, right panel).

The TPM information of each GCM (and reanalysis) is quantitatively summarized with the TPMS in Fig. 6. The improvement in the TPMS of CMIP6 over CMIP5 is especially remarkable for IPSL-CM6A-LR and GFDL-ESM4 models. Both GCMs are able to capture more correctly the transition probabilities between the principal LWTs (such as A, C, SW or W types) than their CMIP5 counterparts, but not yet the persistence probabilities of A and C types (Fig. 5 and A16, Electronic Supplementary Material, respectively). Interestingly, the TPMS spread associated with the observational uncertainty is much reduced in the case of the CMIP6 ensemble, pointing to a better general agreement in their representation of atmospheric circulation, with the exception of two out-lying, poor-performing models, namely NorESM2-LM and CanESM5, which deteriorate in CMIP6 (Fig. 6). Although NorESM2-LM improves on the persistence probability of the cyclonic type, the transitions from CNE to C and from ASW to SW get worse in CMIP6 (Fig. A14, Electronic Supplementary Material), in line with the reduced bias of C type in winter and the large biases

found for SW type in spring and summer (Fig. 4). Similarly, CanESM5 presents too persistent C type and too high transition probabilities from AW and SW to W (Fig. A15, Electronic Supplementary Material), which might be related to the overestimation of the frequencies of W type in winter and spring (Fig. 4).

As for the LWTs frequencies (Fig. 2), very similar TPMs are found for JRA ($TPMS = 0.71$) and NCEP ($TPMS = 0.76$) compared to ERA-Interim and a larger TPMS for ERA-20C ($TPMS = 1.11$, see also Fig. A7 of the Electronic Supplementary Material). The improved performance of CMIP6 with respect to CMIP5 is independent of the reanalysis used as reference (Fig. 6), in line with the results of Cannon (2020). Overall, the differences due to the reference dataset are smaller than model and experiment uncertainties.

4 Summary and conclusions

The present work shows an evaluation of the last two generations of global climate models (CMIP5 and CMIP6) over Europe, in which their ability to represent the atmospheric circulation is assessed by means of the Lamb Weather Type Classification. A set of nine GCM pairs from CMIP5 and CMIP6 are evaluated with respect to four reanalysis products, in order to analyze the sensitivity of the results to the observational dataset. This qualitative, process-based evaluation is intended to help in the design of future downscaling experiments, which are constrained by the boundary conditions provided by GCMs.

A general improvement of CMIP6 over CMIP5 is found in terms of several statistics related to the simulated frequencies of the LWTs and to their temporal sequences
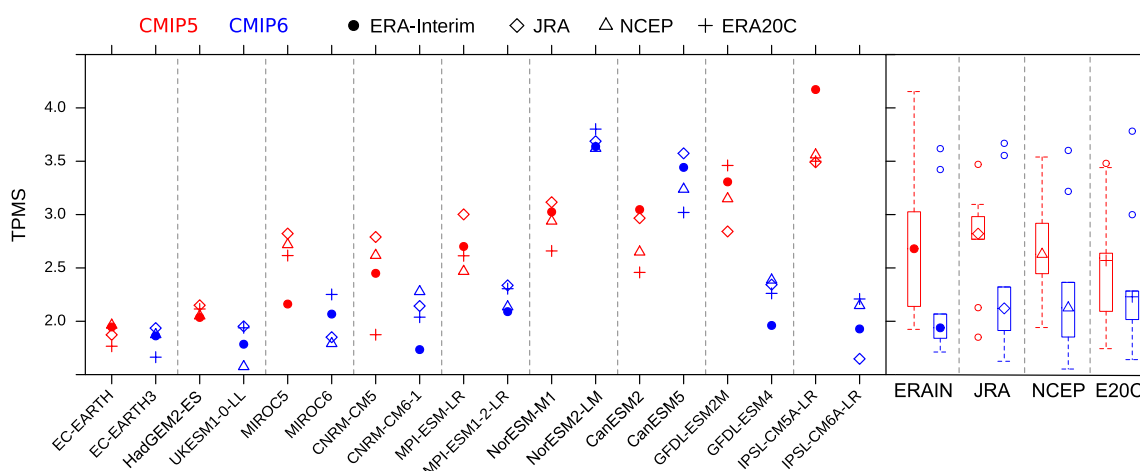

Fig. 6 Transition probability matrix scores (TPMS) attained by the CMIP5/CMIP6 models (red/blue symbols), considering as reference different reanalysis products. The results are presented as CMIP5-CMIP6 GCM pairs, in ascending order of TPMS from left to right,

attained by CMIP5 models and ERA-Interim as reference (solid circles). Boxplots on the right summarize the results for each individual observational reference (see legend symbol indicating the median) and CMIP project (color)

(persistence probability and transition probability from one type to another). Well-performing GCMs in CMIP5 (e.g. EC-EARTH and HadGEM2-ES) also exhibit a good performance in CMIP6. Large improvements are found for IPSL-CM5A-LR and GFDL-ESM4, whereas important biases remain or move along the year in other CMIP6 GCMs (e.g. NorESM2-LM). Such remaining biases relate to their inaccuracies in representing observed transition probabilities, that in general tend to occur for specific seasons. Overall, GCMs show a remarkable ability to represent transition probabilities between LWTs. Despite some significant differences for particular transitions, the GCM TPM *fingerprints* are generally able to faithfully represent the pattern of most likely transitions as represented by the reanalysis, even for the worse performing models. Furthermore, these results are consistent across reference reanalysis products (the extended evaluation results considering alternative reanalysis products are included in the Electronic Supplementary Material).

A general recommendation about the use of specific GCMs is difficult to make, since it depends on the applications of interest, which are usually focused on a given season or might be more sensitive to some weather types (e.g. those leading to extreme events in a particular area). In this sense, based on our results, a user could identify specific seasons and LWTs which particular GCMs fail to reproduce. This application-dependent selection is feasible for statistical downscaling. However, for dynamical downscaling a general performance (all LWTs, all seasons) should be seeked.

While there is a general increase in spatial resolution and an integration of more complex components in CMIP6, these developments take place unevenly for each GCM. For instance, EC-EARTH which is a skillful CMIP5 model improves upon most CMIP6 GCMs, partly due to its rather high resolution (Table 1). A substantial improvement is found for GFDL and IPSL in CMIP6, which have been developed at higher resolutions than their CMIP5 predecessors. Conversely, CanESM and NorESM, which keep a coarse resolution in CMIP6 (the only ones above 2°), deteriorate their TPMS in CMIP6. All the above suggests that the increase of spatial resolution is a factor of improvement in the representation of the atmospheric circulation in the GCMs. Previous studies also find that increasing horizontal resolution of the GCMs leads to a large improvement in the model simulation of the main Euro-Atlantic wintertime weather regimes (Dawson et al. 2012; Strommen et al. 2019) and, particularly, Northern Hemisphere (D'Andrea et al. 1998) and European winter blocking (Matsueda et al. 2009; Berckmans et al. 2013; Davini et al. 2017). The better performance of higher resolution simulations can be attributed to the more realistic orography (Jung et al. 2012) and more realistic representation of Rossby wave breaking processes, which are known to be important in maintaining persistent anomalies (Woollings et al. 2008; Masato et al. 2012). A

recent work based on results of the PRIMAVERA project (Fabiano et al. 2020) shows that the weather regimes tend to be more tightly clustered in the increased resolution simulations, thus resembling more closely the observed ones. However, increased resolution does not improve all aspects in the same way. For instance, Fabiano et al. (2020) find an improvement of the spatial pattern, but limited impact on the frequency of occurrence and persistence of the weather regimes. While resolution stands as a relevant factor, it is not decisive, since some models (here CNRM, HadGEM, MIROC and MPI) improve on TPMS in CMIP6 even though they keep the same resolution. According to Dawson and Palmer (2015) the simulation of spatial and temporal aspects of weather regimes at low resolution can be significantly improved by the introduction of a stochastic physics scheme, highlighting the importance of small-scale processes on large-scale climate variability. Indeed further improvements are needed to remove remaining biases, for instance, better location of the winter blocking is associated with a realistic Gulf Stream sea surface temperature (O'Reilly et al. 2016).

We also show that observational uncertainty stands as a minor source of uncertainty compared to model and experiment uncertainties. With this regard, our results are robust to the selected reference reanalysis and the improvement of CMIP6 over CMIP5 is independent of this choice (in agreement with Cannon 2020).

Note that we did not take into account model internal variability in this study and we use observational uncertainty as reference for substantive changes in the ability of the models to represent the circulation types. Other sources of uncertainty related to the LWT classification remain. For instance, the use of other temporal granularities, 12 UTC (Brands et al. 2014) or 6-hourly (Jones et al. 2013), different from daily-mean data for the LWT classification. This stands as another source of uncertainty and as a very interesting aspect to tackle in future work. Another aspect would be the position of the grid of 16 points considered for the LWT classification, which might shed light on location biases, not addressed in this study. Our results might be sensitive to the circulation classification algorithm used and, therefore, rankings, model performance and even CMIP6 quantitative improvements are particular for the Lamb Weather Types. Cannon (2020) also found an overall improvement in CMIP6 models when using two objective classification algorithms. Thus, a qualitative improvement of CMIP6 is noteworthy regardless of the classification algorithm and evaluation metrics.

Availability of data and materials The data and methods used in this paper are briefly illustrated in the associated paper notebook, available in the public GitHub repository at https://github.com/SantanderMetGroup/notebooks/tree/devel (`2020_Lamb_ClimDyn.*` files). Details on data accessibility and software code reproducing the analyses in this paper is provided.

References

Addor N, Rohrer M, Furrer R, Seibert J (2016) Propagation of biases in climate models from the synoptic to the regional scale: implications for bias adjustment. J Geophys Res Atmos 121(5):2075–2089. https://doi.org/10.1002/2015JD024040

Anstey JA, Davini P, Gray LJ, Woollings TJ, Butchart N, Cagnazzo C, Christiansen B, Hardiman SC, Osprey SM, Yang S (2013) Multi-model analysis of northern hemisphere winter blocking: model biases and the role of resolution. J Geophys Res Atmos 118(10):3956–3971. https://doi.org/10.1002/jgrd.50231

Barriopedro D, Fischer EM, Luterbacher J, Trigo RM, Garcia-Herrera R (2011) The Hot Summer of 2010: redrawing the temperature record map of Europe. Science 332(6026):220–224. https://doi.org/10.1126/science.1201224

Berckmans J, Woollings T, Demory ME, Vidale PL, Roberts M (2013) Atmospheric blocking in a high resolution climate model: influences of mean state, orography and eddy forcing. Atmos Sci Lett 14(1):34–40. https://doi.org/10.1002/asl2.412

Bladé I, Liebmann B, Fortuny D, Oldenborgh GJV (2011) Observed and simulated impacts of the summer NAO in Europe: implications for projected drying in the Mediterranean region. Clim Dyn 39(3–4):709–727. https://doi.org/10.1007/s00382-011-1195-x

Boé J (2018) Interdependency in multimodel climate projections: component replication and result similarity. Geophys Res Lett 45(6):2771–2779. https://doi.org/10.1002/2017GL076829

Brands S, Herrera S, Fernández J, Gutiérrez JM (2013) How well do CMIP5 Earth System Models simulate present climate conditions in Europe and Africa? Clim Dyn 41(3):803–817. https://doi.org/10.1007/s00382-013-1742-8

Brands S, Herrera S, Gutiérrez J (2014) Is Eurasian snow cover in October a reliable statistical predictor for the wintertime climate on the Iberian Peninsula?: Is the wintertime climate in Iberia driven by Eurasian snow cover? Int J Climatol 34:1615–1627. https://doi.org/10.1002/joc.3788

Buehler T, Raible CC, Stocker TF (2011) The relationship of winter season North Atlantic blocking frequencies to extreme cold or dry spells in the ERA-40. Tellus A Dyn Meteorol Oceanogr 63(2):174–187. https://doi.org/10.1111/j.1600-0870.2010.00492.x

Busuioc A, Chen D, Hellström C (2001) Performance of statistical downscaling models in GCM validation and regional climate change estimates: application for Swedish precipitation: STATISTICAL DOWNSCALING FOR SWEDISH PRECIPITATION. Int J Climatol 21(5):557–578. https://doi.org/10.1002/joc.624

Cannon A (2020) Reductions in daily continental-scale atmospheric circulation biases between generations of Global Climate Models: CMIP5 to CMIP6. Environ Res Lett. https://doi.org/10.1088/1748-9326/ab7e4f

Casanueva A, Rodríguez-Puebla C, Frías MD, González-Reviriego N (2014) Variability of extreme precipitation over Europe and its relationships with teleconnection patterns. Hydrol Earth Syst Sci 18(2):709–725. https://doi.org/10.5194/hess-18-709-2014

Casanueva A, Kotlarski S, Fischer A, Flouris A, Kjellstrom T, Lemke B, Nybo L, Schwierz C, Liniger M (2020) Escalating environmental summer heat exposure—a future threat for the European workforce. Regional Environ Change 20(2), https://doi.org/10.1007/s10113-020-01625-6

Chang EKM, Guo Y, Xia X (2012) Cmip5 multimodel ensemble projection of storm track change under global warming. J Geophys Res Atmos 117(D23), https://doi.org/10.1029/2012JD018578

Chang E, Yau A (2016) Northern hemisphere winter storm track trends since 1959 derived from multiple reanalysis datasets. Clim Dyn 47. https://doi.org/10.1007/s00382-015-2911-8

Colle BA, Zhang Z, Lombardo KA, Chang E, Liu P, Zhang M (2013) Historical Evaluation and Future Prediction of Eastern North American and Western Atlantic Extratropical Cyclones in the CMIP5 Models during the Cool Season. J Clim 26(18):6882–6903. https://doi.org/10.1175/JCLI-D-12-00498.1

Cover TM, Thomas JA (2006) Elements of Information Theory. John Wiley & Sons, Hoboken

D'Andrea F, Tibaldi S, Blackburn M, Boer G, Déqué M, Dix MR, Dugas B, Ferranti L, Iwasaki T, Kitoh A, Pope V, Randall D, Roeckner E, Strauss D, Stern W, Van den Dool H, Williamson D (1998) Northern Hemisphere atmospheric blocking as simulated by 15 atmospheric general circulation models in the period 1979–1988. Clim Dyn 14(6):385–407. https://doi.org/10.1007/s003820050230

Davini P, Corti S, D'Andrea F, Rivière G, von Hardenberg J (2017) Improved winter European atmospheric blocking frequencies in high-resolution global climate simulations. J Adv Model Earth Syst 9(7):2615–2634. https://doi.org/10.1002/2017MS001082

Dawson A, Palmer TN (2015) Simulating weather regimes: impact of model resolution and stochastic parameterization. Clim Dyn 44(7):2177–2193. https://doi.org/10.1007/s00382-014-2238-x

Dawson A, Palmer TN, Corti S (2012) Simulating regime structures in weather and climate prediction models: REGIMES IN WEATHER AND CLIMATE MODELS. Geophys Res Lett 39(21):n/a–n/a, https://doi.org/10.1029/2012GL053284

Dee DP, Uppala SM, Simmons AJ, Berrisford P, Poli P, Kobayashi S, Andrae U, Balmaseda MA, Balsamo G, Bauer P, Bechtold P, Beljaars ACM, van de Berg L, Bidlot J, Bormann N, Delsol C, Dragani R, Fuentes M, Geer AJ, Haimberger L, Healy SB, Hersbach H, Hólm EV, Isaksen L, Kaallberg P, Kohler M, Matricardi M, McNally AP, Monge-Sanz BM, Morcrette J, Park B, Peubey C, de Rosnay P, Tavolato C, Thépaut JN, Vitart F (2011) The ERA-Interim reanalysis: configuration and performance of the data assimilation system. Q J R Meteorol Soc 137:553–597. https://doi.org/10.1002/qj.828

Diaconescu EP, Laprise R (2013) Can added value be expected in RCM-simulated large scales? Clim Dyn 41(7–8):1769–1800. https://doi.org/10.1007/s00382-012-1649-9

Dosio A (2016) Projections of climate change indices of temperature and precipitation from an ensemble of bias-adjusted high-resolution EURO-CORDEX regional climate models: BIAS-ADJUSTED CLIMATE CHANGE INDICES. J Geophys Res Atmos 121(10):5488–5511. https://doi.org/10.1002/2015JD024411

Drost HG (2018) Philentropy: information theory and distance quantification with R. J Open Source Softw 3(26):765. https://doi.org/10.21105/joss.00765

Eyring V, Bony S, Meehl GA, Senior CA, Stevens B, Stouffer RJ, Taylor KE (2016) Overview of the coupled model intercomparison project phase 6 (cmip6) experimental design and organization. Geosci Model Dev 9(5):1937–1958. https://doi.org/10.5194/gmd-9-1937-2016

Fabiano F, Christensen HM, Strommen K, Athanasiadis P, Baker A, Schiemann R, Corti S (2020) Euro-Atlantic weather regimes in the primavera coupled climate simulations: impact of resolution and mean state biases on model performance. Clim Dyn 54(11):5031–5048. https://doi.org/10.1007/s00382-020-05271-w

Favà V, Curto JJ, Llasat MC (2015) Relationship between the summer NAO and maximum temperatures for the Iberian Peninsula. Theor Appl Climatol pp 1–15, https://doi.org/10.1007/s00704-015-1547-2

Fealy R, Mills G (2018) Deriving Lamb weather types suited to regional climate studies: A case study on the synoptic origins of precipitation over Ireland. Int J Climatol 38(8):3439–3448. https://doi.org/10.1002/joc.5495, 00000

Fernández J, Frías MD, Cabos WD, Cofiño AS, Domínguez M, Fita L, Gaertner MA, García-Díez M, Gutiérrez JM, Jiménez-Guerrero P, Liguori G, Montávez JP, Romera R, Sánchez E (2019) Consistency of climate change projections from multiple global and regional model intercomparison projects. Clim Dyn 52(1):1139–1156. https://doi.org/10.1007/s00382-018-4181-8

Folland CK, Knight J, Linderholm HW, Fereday D, Ineson S, Hurrell JW (2009) The Summer North Atlantic Oscillation: past, present, and future. J Clim 22(5):1082–1103. https://doi.org/10.1175/2008JCLI2459.1

Fujiwara M, Wright JS, Manney GL, Gray LJ, Anstey J, Birner T, Davis S, Gerber EP, Harvey VL, Hegglin MI, Homeyer CR, Knox JA, Krüger K, Lambert A, Long CS, Martineau P, Molod A, Monge-Sanz BM, Santee ML, Tegtmeier S, Chabrillat S, Tan DGH, Jackson DR, Polavarapu S, Compo GP, Dragani R, Ebisuzaki W, Harada Y, Kobayashi C, McCarty W, Onogi K, Pawson S, Simmons A, Wargan K, Whitaker JS, Zou CZ (2017) Introduction to the sparc reanalysis intercomparison project (s-rip) and overview of the reanalysis systems. Atmos Chem Phys 17(2):1417–1452. https://doi.org/10.5194/acp-17-1417-2017

Harada Y, Kamahori H, Kobayashi C, Endo H, Kobayashi S, Ota Y, Onoda H, Onogi K, Miyaoka K, Takahashi K (2016) The jra-55 reanalysis: representation of atmospheric circulation and climate variability. J Meteorol Soc Jpn Ser II 94(3):269–302. https://doi.org/10.2151/jmsj.2016-015

Hochman A, Alpert P, Harpaz T, Saaroni H, Messori G (2019) A new dynamical systems perspective on atmospheric predictability: Eastern mediterranean weather regimes as a case study. Sci Adv 5(6):eaau0936. https://doi.org/10.1126/sciadv.aau0936

Hulme M, Briffa K, Jones P, Senior C (1993) Validation of gcm control simulations using indices of daily airflow types over the British isles. Clim Dyn 9(2):95–105. https://doi.org/10.1007/BF00210012

Hurrell JW, Kushnir Y, Ottersen G, Visbeck M (2003) An overview of the North Atlantic Oscillation. In: Hurrell JW, Kushnir Y, Ottersen G, Visbeck M (eds) Geophysical Monograph Series, vol 134. American Geophysical Union. Washington, D. C., pp 1–35

Iturbide M, Bedia J, Herrera S, Baño-Medina J, Fernández J, Frías M, Manzanas R, San-Martín D, Cimadevilla E, Cofiño A, Gutiérrez J (2019) The R-based climate4R open framework for reproducible climate data access and post-processing. Environ Model Softw 111:42–54. https://doi.org/10.1016/j.envsoft.2018.09.009

Iturbide M, Casanueva A, Bedia J, Herrera S, Milovac J, Gutiérrez J (2020) On the need of bias adjustment for more plausible climate change projections of extreme heat. Atmos Sci Lett Submitted

Jacob D, Petersen J, Eggert B, Alias A, Christensen OB, Bouwer LM, Braun A, Colette A, Déqué M, Georgievski G, Georgopoulou E, Gobiet A, Menut L, Nikulin G, Haensler A, Hempelmann N, Jones C, Keuler K, Kovats S, Kröner N, Kotlarski S, Kriegsmann A, Martin E, Ev Meijgaard, Moseley C, Pfeifer S, Preuschmann S, Radermacher C, Radtke K, Rechid D, Rounsevell M, Samuelsson P, Somot S, Soussana JF, Teichmann C, Valentini R, Vautard R, Weber B, Yiou P (2014) EURO-CORDEX: new high-resolution climate change projections for European impact research. Regional Environ Change 14(2):563–578. https://doi.org/10.1007/s10113-013-0499-2

Jacob D, Teichmann C, Sobolowski S, Katragkou E, Anders I, Belda M, Benestad R, Boberg F, Buonomo E, Cardoso RM, Casanueva A, Christensen OB, Christensen JH, Coppola E, De Cruz L, Davin EL, Dobler A, Domínguez M, Fealy R, Fernandez J, Gaertner MA, García-Díez M, Giorgi F, Gobiet A, Goergen K, Gómez-Navarro JJ, Alemán JJG, Gutiérrez C, Gutiérrez JM, Güttler I, Haensler A, Halenka T, Jerez S, Jiménez-Guerrero P, Jones RG, Keuler K, Kjellström E, Knist S, Kotlarski S, Maraun D, van Meijgaard E, Mercogliano P, Montávez JP, Navarra A, Nikulin G, de Noblet-Ducoudré N, Panitz HJ, Pfeifer S, Piazza M, Pichelli E, Pietikäinen JP, Prein AF, Preuschmann S, Rechid D, Rockel B, Romera R, Sánchez E, Sieck K, Soares PMM, Somot S, Srnec L, Sørland SL, Termonia P, Truhetz H, Vautard R, Warrach-Sagi K, Wulfmeyer V (2020) Regional climate downscaling over europe: perspectives from the EURO-CORDEX community. Regional Environ Change 20(2):51. https://doi.org/10.1007/s10113-020-01606-9

Jenkinson A, Collison F (1977) An initial climatology of gales over the north sea. synoptic climatology branch memorandum. Meteorol Off pp 1–62

Jiang B, Pei J, Tao Y, Lin X (2011) Clustering uncertain data based on probability distribution similarity. IEEE Trans Knowl Data Eng 25(4):751–763

Jones PD, Hulme M, Briffa KR (1993) A comparison of Lamb circulation types with an objective classification scheme. Int J Climatol 13(6):655–663. https://doi.org/10.1002/joc.3370130606

Jones RG, Murphy JM, Noguer M (1995) Simulation of climate change over europe using a nested regional-climate model. I: Assessment of control climate, including sensitivity to location of lateral boundaries. Q J R Meteorol Soc 121(526):1413–1449. https://doi.org/10.1002/qj.49712152610

Jones PD, Harpham C, Briffa KR (2013) Lamb weather types derived from reanalysis products. Int J Climatol 33(5):1129–1139. https://doi.org/10.1002/joc.3498

Jung T, Miller MJ, Palmer TN, Towers P, Wedi N, Achuthavarier D, Adams JM, Altshuler EL, Cash BA, Kinter IJL, Marx L, Stan C, Hodges KI (2012) High-Resolution Global Climate Simulations with the ECMWF Model in Project Athena: Experimental Design, Model Climate, and Seasonal Forecast Skill. J Clim 25(9):3155–3172. https://doi.org/10.1175/JCLI-D-11-00265.1

Jury MW, Herrera S, Gutiérrez JM, Barriopedro D (2019) Blocking representation in the ERA-Interim driven EURO-CORDEX RCMs. Clim Dyn 52(5–6):3291–3306. https://doi.org/10.1007/s00382-018-4335-8

Kalnay E, Kanamitsu M, Kistler R, Collins W, Deaven D, Gandin L, Iredell M, Saha S, White G, Woollen J, Zhu Y, Chelliah M, Ebisuzaki W, Higgins W, Janowiak J, Mo KC, Ropelewski C, Wang J, Leetmaa A, Reynolds R, Jenne R, Joseph D (1996) The NCEP/NCAR 40-Year Reanalysis Project. Bulletin of the American Meteorological Society 77(3):437–472. DOI https://doi.org/10.1175/1520-0477(1996)077%3c0437:TNYRP%3e2.0.CO;2, publisher: American Meteorological Society

Kobayashi S, Ota Y, Harada Y, Ebita A, Moriya M, Onoda H, Onogi K, Kamahori H, Kobayashi C, Endo H, Miyaoka K, Takahashi K (2015) The jra-55 reanalysis: general specifications and basic

characteristics. J Meteorol Soc Jpn Ser II 93(1):5–48. https://doi.org/10.2151/jmsj.2015-001

Kullback S, Leibler RA (1951) On information and sufficiency. Ann Math Stat 22(1):79–86

Lamb H (1972) British isles weather types and a register of the daily sequence of circulation patterns 1861-1971. METEOROL OFF, GEOPHYS MEM; GB; DA 1972; NO 116; PP 1-85; BIBL 2P1/2

Maraun D, Shepherd TG, Widmann M, Zappa G, Walton D, Gutiérrez JM, Hagemann S, Richter I, Soares PMM, Hall A, Mearns LO (2017) Towards process-informed bias correction of climate change simulations. Nat Clim Change 7(11):664–773. https://doi.org/10.1038/nclimate3418

Masato G, Hoskins BJ, Woollings TJ (2012) Wave-breaking characteristics of midlatitude blocking. Q J R Meteorol Soc 138(666):1285–1296. https://doi.org/10.1002/qj.990

Masato G, Hoskins BJ, Woollings T (2013) Winter and Summer Northern Hemisphere Blocking in CMIP5 Models. J Clim 26(18):7044–7059. https://doi.org/10.1175/JCLI-D-12-00466.1

Matsueda M, Mizuta R, Kusunoki S (2009) Future change in wintertime atmospheric blocking simulated using a 20-km-mesh atmospheric global circulation model. J Geophys Res Atmos 114(D12), https://doi.org/10.1029/2009JD011919

McSweeney C, Jones R, Lee RW, Rowell D (2015) Selecting CMIP5 GCMs for downscaling over multiple regions. Clim Dyn 44(11):3237–3260. https://doi.org/10.1007/s00382-014-2418-8

O'Reilly CH, Minobe S, Kuwano-Yoshida A (2016) The influence of the gulf stream on wintertime European blocking. Clim Dyn 47(5):1545–1567. https://doi.org/10.1007/s00382-015-2919-0

Otero N, Sillmann J, Butler T (2018) Assessment of an extended version of the Jenkinson-Collison classification on CMIP5 models over Europe. Clim Dyn 50(5):1559–1579. https://doi.org/10.1007/s00382-017-3705-y

Pereira S, Ramos A, Rebelo L, Trigo R, Zêzere J (2018) A centennial catalogue of hydro-geomorphological events and their atmospheric forcing. Adv Water Resour 122:98–112. https://doi.org/10.1016/j.advwatres.2018.10.001

Perez J, Menendez M, Mendez FJ, Losada IJ (2014) Evaluating the performance of CMIP3 and CMIP5 global climate models over the north–east Atlantic region. Clim Dyn 43(9):2663–2680. https://doi.org/10.1007/s00382-014-2078-8

Poli P, Hersbach H, Dee DP, Berrisford P, Simmons AJ, Vitart F, Laloyaux P, Tan DG, Peubey C, Thépaut JN, Trémolet Y, Hólm EV, Bonavita M, Isaksen L, Fisher M (2016) ERA-20C: an atmospheric reanalysis of the twentieth century. J Clim 29(11):4083–4097. https://doi.org/10.1175/JCLI-D-15-0556.1

Prein AF, Bukovsky MS, Mearns LO, Bruyère CL, Done JM (2019) Simulating North American weather types with regional climate models. Front Environ Sci 7:36. https://doi.org/10.3389/fenvs.2019.00036

R Core Team (2020) R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria, https://www.R-project.org/

Ramos AM, Cortesi N, Trigo RM (2014) Circulation weather types and spatial variability of daily precipitation in the Iberian Peninsula. Front Earth Sci 2. https://doi.org/10.3389/feart.2014.00025, 00037

Rex DF (1950) Blocking action in the middle troposphere and its effect upon regional climate. Tellus 2(3):196–211. https://doi.org/10.1111/j.2153-3490.1950.tb00331.x

Rohrer M, Brönnimann S, Martius O, Raible CC, Wild M, Compo GP (2018) Representation of Extratropical Cyclones, Blocking Anticyclones, and Alpine Circulation Types in Multiple Reanalyses and Model Simulations. J Clim 31(8):3009–3031. https://doi.org/10.1175/JCLI-D-17-0350.1

Sharma KK, Seal A (2019) Modeling uncertain data using monte Carlo integration method for clustering. Expert Syst Appl 137:100–116. https://doi.org/10.1016/j.eswa.2019.06.050

Sillmann J, Croci-Maspoli M (2009) Present and future atmospheric blocking and its impact on European mean and extreme climate. Geophys Res Lett 36(10):L10,702. https://doi.org/10.1029/2009GL038259

Sousa PM, Trigo RM, Barriopedro D, Soares PMM, Ramos AM, Liberato MLR (2017) Responses of European precipitation distributions and regimes to different blocking locations. Clim Dyn 48(3–4):1141–1160. https://doi.org/10.1007/s00382-016-3132-5

Strommen K, Mavilia I, Corti S, Matsueda M, Davini P, von Hardenberg J, Vidale PL, Mizuta R (2019) The sensitivity of euro-Atlantic regimes to model horizontal resolution. Geophys Res Lett 46(13):7810–7818. https://doi.org/10.1029/2019GL082843

Stryhal J, Huth R (2017) Classifications of Winter Euro-Atlantic Circulation Patterns: An Intercomparison of Five Atmospheric Reanalyses. J Clim 30(19):7847–7861. https://doi.org/10.1175/JCLI-D-17-0059.1

Taylor KE, Stouffer RJ, Meehl GA (2012) An overview of cmip5 and the experiment design. Bull Am Meteorol Soc 93(4):485–498. https://doi.org/10.1175/BAMS-D-11-00094.1

Trigo RM, DaCamara CC (2000) Circulation weather types and their influence on the precipitation regime in Portugal. Int J Climatol p 23, DOI 10.1002/1097-0088(20001115)20:13%3C1559::AID-JOC555%3E3.0.CO;2-5

Turco M, Sanna A, Herrera S, Llasat MC, Gutiérrez JM (2013) Large biases and inconsistent climate change signals in ENSEMBLES regional projections. Clim Change 120(4):859–869. https://doi.org/10.1007/s10584-013-0844-y

Vial J, Osborn T (2012) Assessment of atmosphere-ocean general circulation model simulations of winter northern hemisphere atmospheric blocking. Clim Dyn 39:95–112. https://doi.org/10.1007/s00382-011-1177-z

Wang XL, Swail VR, Zwiers FW (2006) Climatology and Changes of Extratropical Cyclone Activity: Comparison of ERA-40 with NCEP-NCAR Reanalysis for 1958–2001. J Clim 19(13):3145–3166. https://doi.org/10.1175/JCLI3781.1

Weijs SV, van Nooijen R, van de Giesen N (2010) Kullback–Leibler divergence as a forecast skill score with classic reliability-resolution-uncertainty decomposition. Mon Weather Rev 138(9):3387–3399. https://doi.org/10.1175/2010MWR3229.1

WMO (2017) WMO Guidelines on the Calculation of Climate Normals. Tech. Rep. WMO No. 1203, World Meteorological Organization, ISBN:978-92-63-11203-3

Woollings T, Hoskins B, Blackburn M, Berrisford P (2008) A New Rossby Wave-Breaking Interpretation of the North Atlantic Oscillation. J Atmos Sci 65(2):609–626. https://doi.org/10.1175/2007JAS2347.1

Zappa G, Shaffrey LC, Hodges KI (2013) The Ability of CMIP5 Models to Simulate North Atlantic Extratropical Cyclones*. J Clim 26(15):5379–5396. https://doi.org/10.1175/JCLI-D-12-00501.1