

Direct and component-wise bias correction of multi-variate climate indices: The percentile adjustment function diagnostic tool

A. Casanueva · J. Bedia · S. Herrera ·
J. Fernández · J.M. Gutiérrez

Abstract The use and development of bias correction (BC) methods has grown fast in recent years, due to the increased demand of unbiased projections by many sectoral climate change impact applications. Case studies are frequently based on multi-variate climate indices (CIs) combining two or more essential climate variables, that are frequently individually corrected prior to CI calculation. This poses the question of whether the BC method modifies the inter-variable dependencies and eventually the climate change signal. The direct bias correction of the multi-variate CI stands as a usual alternative, since it preserves the physical and temporal coherence among the primary variables as represented in the dynamical model output, at the expense of incorporating the individual biases on the CI with an effect difficult to foresee, particularly in the case of complex CIs bearing in their formulation non-linear relationships between components. Such is the case of the Fire Weather Index (FWI), a meteorological fire danger indicator frequently used in forest fire prevention and research. In the present work we test the suitability of the direct BC approach on FWI as a representative multi-variate CI, assessing its performance in present climate conditions and its effect on the climate change signal when applied to future projections. Moreover, the results are compared with the

A. Casanueva (✉)

Meteorology Group, Dept. Applied Mathematics and Computer Sciences, University of Cantabria, 39005 Santander, Spain

Currently at: Federal Office of Meteorology and Climatology MeteoSwiss, 8058 Zurich, Switzerland

Tel.: +41 58 460 92 55

E-mail: ana.casanueva@meteoswiss.ch

S. Herrera · J. Fernández

Meteorology Group, Dept. Applied Mathematics and Computer Sciences, University of Cantabria, 39005 Santander, Spain

J. Bedia · J.M. Gutiérrez

Meteorology Group, Institute of Physics of Cantabria, CSIC-University of Cantabria, 39005 Santander, Spain

common approach of correcting the input variables separately. To this aim, we apply the widely used empirical quantile mapping method (QM), adjusting the 99 empirical percentiles. The analysis of the Percentile Adjustment Function (PAF) provides insight into the effect of the QM on the climate change signal. Although both approaches present similar results in present climate, the direct correction introduces a greater modification of the original change signal. These results warn against the blind use of QM, even in the case of essential climate variables or uni-variate CIs.

Keywords Bias Correction · Bias adjustment · Fire Weather Index · Climate Change · Quantile Mapping · Regional Climate Models

1 Introduction

The assessment of climate change impacts on the different human activities and ecosystems has become a major challenge in the last decades, especially for those regions that are particularly vulnerable to climate change (IPCC, 2014). Regional Climate Models (RCMs) are the primary tools to provide regional climate information (particularly climate change projections) needed for this task. However, in spite of the continuous increase of spatial resolution and physical processes solved by the successive generations of RCMs —e.g. PRUDENCE (Christensen et al, 2007), ENSEMBLES (van der Linden and Mitchell, 2009) and EURO-CORDEX (Kotlarski et al, 2014) in Europe,— their output cannot be directly used for impact studies due to their systematic biases (Christensen et al, 2008; Hagemann et al, 2011). Therefore, some sort of bias correction (BC) or adjustment of the model output is needed. These two terms are used interchangeably in the literature, with a recent preference for the latter, due to the misleading expectation of “correct” data after BC. In the present work we explore the adjustments applied by these techniques in a multi-variate context and suggest a diagnostic tool to ease the detection of statistical artifacts.

In the climate change context, bias adjustment is typically done using distributional model output statistics to bring modeled probability distributions closer to those of the observations (see e.g. Déqué, 2007; Piani et al, 2010a). These methods are simple to apply and provide, by construction, strong bias reductions under current climate. However, despite these advantages, their limitations need to be considered. For instance, the physical causes of model errors are not taken into account (Teutschbein and Seibert, 2012), i.e. biases are not related to the (mis)representation of physical processes in the models, such as synoptic circulation (Addor et al, 2016), and changes in trends or spatio-temporal structures may appear due to the scale mismatch between RCM and observations (Maraun, 2013). As for all statistical downscaling methods, BC techniques are applied under the stationarity (time invariance) assumption in climate change conditions (Ehret et al, 2012). Furthermore, they constitute an additional source of uncertainty in climate change projections and can modify the climate change signal of the raw model output (Teng et al, 2015). Although

these changes may be defensible in the case of stationary, intensity-dependent biases (Gobiet et al, 2015), they should be carefully analyzed from a physical perspective case by case — considering also time-invariance,— in order to avoid statistical artifacts.

In a multi-variate context, the dynamical nature of the relationships between variables simulated by a climate model is subject to an ongoing debate (see Piani et al, 2010b, and references therein). On the one hand, RCMs present biases in inter-variable correlation fields (Wilcke et al, 2013), thus requiring specific multi-variate bias adjustment techniques (e.g. Li et al, 2014; Vrac and Friederichs, 2015; Cannon, 2016) able to reconstruct the inter-variable dependences of the observations. On the other hand, the ability of RCMs to respond in a physically consistent way to external forcings is one of their basic foundations (Wilby et al, 2000). Due to the purely statistical nature of bias adjustment, the potential modification of the inter-variable dependencies caused by univariate techniques is an inherent caveat of the methodology (Piani et al, 2010b; Ehret et al, 2012; Muerth et al, 2013). Whereas Wilcke et al (2013) show that univariate bias adjustment is able to retain the quality of the temporal structure and the inter-variable dependencies of the uncorrected data, Rocheta et al (2014) show that bias correcting variables independently results in increased errors in the potential vorticity field, which are improved when the latter is directly corrected.

Forest management is an important socio-economic sector highly sensitive to climate change, since forest fire activity is greatly controlled by climate variability (see e.g. Pechony and Shindell, 2010). Thus, it is important to identify and understand the relationships between fires and weather, and their implications in a changing climate (Bedia et al, 2015). Several fire danger indices, based on the compound effect of key meteorological variables (temperature, precipitation, relative humidity, wind, etc.) are widely used to quantify fire danger (see e.g. Stocks et al, 1998; Williams et al, 2001), such as the Canadian Fire Weather Index (FWI van Wagner, 1987). FWI is here taken as an illustrative example of a sector-specific CI requiring a multi-variable bias adjustment. The most common approach is to adjust independently the component variables prior to the index calculation (i.e. *the component-wise approach*), as Yang et al (2015). In this multi-variate context, the BC method should be carefully chosen to retain the physical consistency between variables, otherwise inconsistencies in spatio-temporal fields may appear and the inter-variable dependencies may be modified (Ehret et al, 2012; Vrac and Friederichs, 2015). The alternative of directly adjusting the index (i.e. *direct approach*) could circumvent this constraint and has not been widely tested in a BC context (see Casanueva et al (2014) for an example in statistical downscaling under the perfect prognosis approach and Rocheta et al (2014) for an example of direct BC with a scaling method).

The present work aims at exploring the application of the direct BC approach as an alternative to the component approach for climate projections, illustrated for the FWI in Spain. In this way, a single parameter (FWI) is adjusted at the expense of including the individual biases of the input variables

in the index calculation. In particular, the different nature of the FWI component variables (near-surface temperature, relative humidity, wind and 24-h accumulated precipitation) and their physical interdependence (e.g. relative humidity depends on 2-meter temperature), make this index a good candidate to test the direct BC approach.

Among the different BC methods available in the literature, we use the empirical Quantile Mapping (QM, Panofsky and Brier, 1968) since it can be successfully applied to either temperature, precipitation, winds or relative humidity (Wilcke et al, 2013) —unlike scaling or parametric methods that depend on the nature of each specific variable— and is one of the most popular BC methods.

The objectives of this work are (1) to assess the applicability of the direct BC of a multi-variate index (FWI), and (2) to analyze the effect of the QM (direct and component) on the climate change signal. For the latter, the percentile adjustments applied by the QM are explored.

2 Data and Methods

2.1 The Fire Weather Index

The Canadian forest Fire Weather Index (FWI, van Wagner, 1987; Wotton, 2009), is one of the most popular fire danger indices worldwide. In particular, its suitability for different Mediterranean ecosystems has been highlighted by several authors (see, e.g. Viegas et al, 1999; Dimitrakopoulos et al, 2011) and is the reference fire danger indicator used by the European Commission to assess current and future fire danger in Europe, in the framework of the European Forest Fire Information System (EFFIS, San-Miguel-Ayán et al, 2013, <http://forest.jrc.ec.europa.eu/effis/>). The FWI is a dimensionless daily indicator of potential fire conditions based on four weather variables i.e. instantaneous values of temperature, relative humidity and wind velocity at noon local standard time, and accumulated precipitation in the previous 24 hours. It is calculated through a complex system based on three indices tracking the moisture content in different fuel layers (depending on precipitation, temperature and relative humidity) and two additional indices characterizing fire behavior (rate of spread and fuel consumption), through non-linear combinations of the moisture indices and wind speed. The final index (FWI) is the result of combining all the previous ones, representing fire intensity as energy output per unit length of fire front. The reader is referred to van Wagner (1987) for an overview and van Wagner and Pickett (1985) for more details in the calculation (including FORTRAN code routines). In the present work, the FWI is computed using the R package *fireDanger* (Bedia et al, 2017), that implements the FWI System in the R language, linking it to the climate data model implemented in the *climate4R* bundle (Cofiño et al, 2017).

The FWI has been previously applied to the estimation of future regional fire danger scenarios in Europe by several authors, considering different down-

scaling approaches (Moriondo et al, 2006; Bedia et al, 2013, 2014, 2017). In general, due to the lack of instantaneous public data, daily mean values from the primary variables are usually considered in the FWI calculation leading to a clear underestimation of the FWI (Bedia et al, 2014) which yields to great differences in climate projections (Herrera et al, 2013). Therefore, instantaneous data is preferred for the FWI calculation although this requirement is a limiting factor for the broad use of this index.

Here, we focus on the seasonal mean FWI and its 90th percentile (FWI90), the latter used as an indicator for extreme fire danger conditions, considering the standard fire season in the Mediterranean region (June to September, JJAS). All FWI-derived indices are computed from the daily series.

2.2 Observational data

Impact studies usually focus on local scale information to drive impact models in order to produce projections at very high spatial resolution (see e.g. Quintana Seguí et al, 2010). To also meet the instantaneous data requirement of FWI, we used observational data from meteorological stations over Spain provided by the Spanish Meteorological Agency (AEMET). These historical data were recorded using standard automatic stations, which provide subdaily (6-hourly) data. The period covered by these automatic stations is 1979-2003. These data have passed an automated data quality control and those with a missing value percentage above 20% in the period of study were discarded (Bedia et al, 2013). After this process, 45 high-quality stations remained to be used in the study (see Fig. 1a-b).

2.3 Regional Climate Models

Standard RCM databases rarely provide instantaneous data outputs, as required for the FWI calculation. Two RCMs from the ENSEMBLES project (van der Linden and Mitchell, 2009), driven by the same GCM were used in this study (RACMO2 and RegCM3, see Table 1), after contacting all the institutions producing the simulations to get the data.

In the present work, we used hourly or 3-hourly climate change projections nested into a single GCM (ECHAM5 r3), forced by a single emissions scenario (SRES A1B). Thus, the uncertainty due to the global forcing and scenario was not considered. A region covering the Iberian Peninsula was chosen from the European-wide simulation domain. Both models used a similar horizontal resolution (0.22° regular on a rotated latitude-longitude grid and 25 Km on a Lambert conic conformal grid). RCM data were paired to point station observations by selecting the RCM gridcells closest to each station location.

The selection of these particular RCMs was motivated by the availability of high temporal resolution data from the producing centers. In any case, the RACMO2 ranks among the best in previous evaluation studies in the Iberian

Table 1 Summary of the ENSEMBLES RCM simulations used in this work.

RCM	Temporal Res.	Horiz. Res.	Institution	Reference
RACMO2	Hourly	0.22°	KNMI, Netherlands	van Meijgaard et al (2008)
RegCM3	3-Hourly	25 Km	ICTP, Italy	Pal et al (2007)

Peninsula (Herrera et al, 2010), which is consistent with more recent analysis based on EURO-CORDEX (Casanueva et al, 2016), and both RCMs were free from coupling artifacts according to Turco et al (2013). We also checked that the projected delta changes of these models are not outliers for precipitation or temperature, especially RACMO2, which ranks near the ENSEMBLES multi-model median for both variables (not shown).

2.4 Time consistency of input data

As mentioned, FWI requires instantaneous data at noon local time. In this case (summer in Spain), local time is Central European Summer Time (CEST), which is 2 hours ahead of UTC (used by the RCMs). The AEMET data set (Section 2.2) provides instantaneous values of temperature, relative humidity and wind speed at 15 CEST, and precipitation accumulated in 24 hours (from 9 CEST of day D to 9 CEST of day $D + 1$), recorded as the daily precipitation value at 07 UTC of day D .

As Bedia et al (2013) and Casanueva et al (2014), precipitation data were shifted one day ahead the whole series, since precipitation at noon accumulated in the last 24 hours is required for FWI. By doing this, the precipitation accumulated from 9 CEST of day D to 9 CEST of day $D + 1$ are assigned to day $D + 1$. As a result, it must be noted that there is a 5-hour lag between the RCM precipitation data, accumulated up to 12 UTC (14 CEST), and AEMET observations. This lag is just 1 hour for the rest of variables.

2.5 Bias adjustment

In order to adjust the bias of the FWI projections, we considered a single method based on the empirical Quantile Mapping (hereafter denoted as QM, Panofsky and Brier, 1968), defined by Eq. 1,

$$q_m^* = F_o^{-1}(F_m(q_m)) \quad (1)$$

where q_m and q_m^* are the model original and adjusted quantiles, respectively, and F_m and F_o denote the empirical cumulative distribution function of the modelled and observed variable, as obtained from a calibration historical period.

The procedure used here is the same as Déqué (2007), where the 99 empirical percentiles are adjusted to match those observed and linear interpolation is used for the values between two percentiles. Constant extrapolation is applied for values outside the calibration range, i.e. the adjustment applied to the last (first) percentile is applied to all the values above (below) it. The function indicating the adjustment factors $q_o - q_m$ as a function of the percentile probabilities provides overall information of the variability of the local adjustments along the range of values of the variable. We will refer to this function as the Percentile Adjustment Function (PAF). The effective adjustments carried out by the QM would depend on the distribution of the RCM data in the test period with respect to the distribution in the calibration. This is denoted as effective Percentile Adjustment Function (PAFeff) and are obtained following: $q_m^* - q_m$.

For precipitation, QM adjusts the excess of light precipitation frequency ('drizzle effect') when a model overestimates the wet-day frequency. The frequency adaptation proposed by (Thiemeßl et al, 2012; Wilcke et al, 2013) was applied in this work to overcome the opposite situation, when the model underestimates wet-day frequency. The frequency adaptation consists in randomly sampling the observational distribution into the simulated first bin (0-1 mm) in order to generate dry days, and has been used to adjust future precipitation scenarios (Thiemeßl et al, 2012).

The QM was applied to JJAS station and RCM daily series following two approaches: 1) individually to all component variables (i.e. the component-wise approach, QMc) and 2) directly to the uncorrected FWI (i.e. the direct approach, QMd). In order to maximize the length of the training period, we used the full observational data set extent (1979-2003) to train the QM. Therefore, in the RCMs, we extended the 20C3M experiment (1979-2000) with a few initial years (2001-2003) of the A1B experiment. First, for the validation of the direct and component QM in present climate (Section 3.1) a k-fold cross-validation approach was considered by splitting the data into k equal-size sub-samples ($k = 5$). We applied the same k-fold cross validation approach used in Bedia et al (2013) and Casanueva et al (2014), considering a stratified sampling (i.e. first test sample formed by the years 1979, 1984, 1989, 1994, 1999, the second by the years 1980, 1985, 1990, 1995 and 2000, etc). Secondly, to adjust the biases in the future projections (Section 3.2), we used the whole period 1979-2003 to train the QM and applied the PAF obtained to three future periods (2011-2040, 2041-2070 and 2071-2100) from the A1B scenario and to a reference period (1971-2000) from the 20C3M experiment. Note that the 30-year reference period shares 22 years with the training period. The (adjusted) climate change signal is defined as the difference (delta) between the future and reference bias-adjusted data.

3 Results

3.1 Cross-validation of direct- and component-wise QM

The modelled and observed JJAS series of the four meteorological drivers (temperature, precipitation, wind and relative humidity) and the FWI are used to cross-validate the QM under a 5-fold stratified cross-validation approach (Section 2.5) following the two BC approaches (QM_c and QM_d).

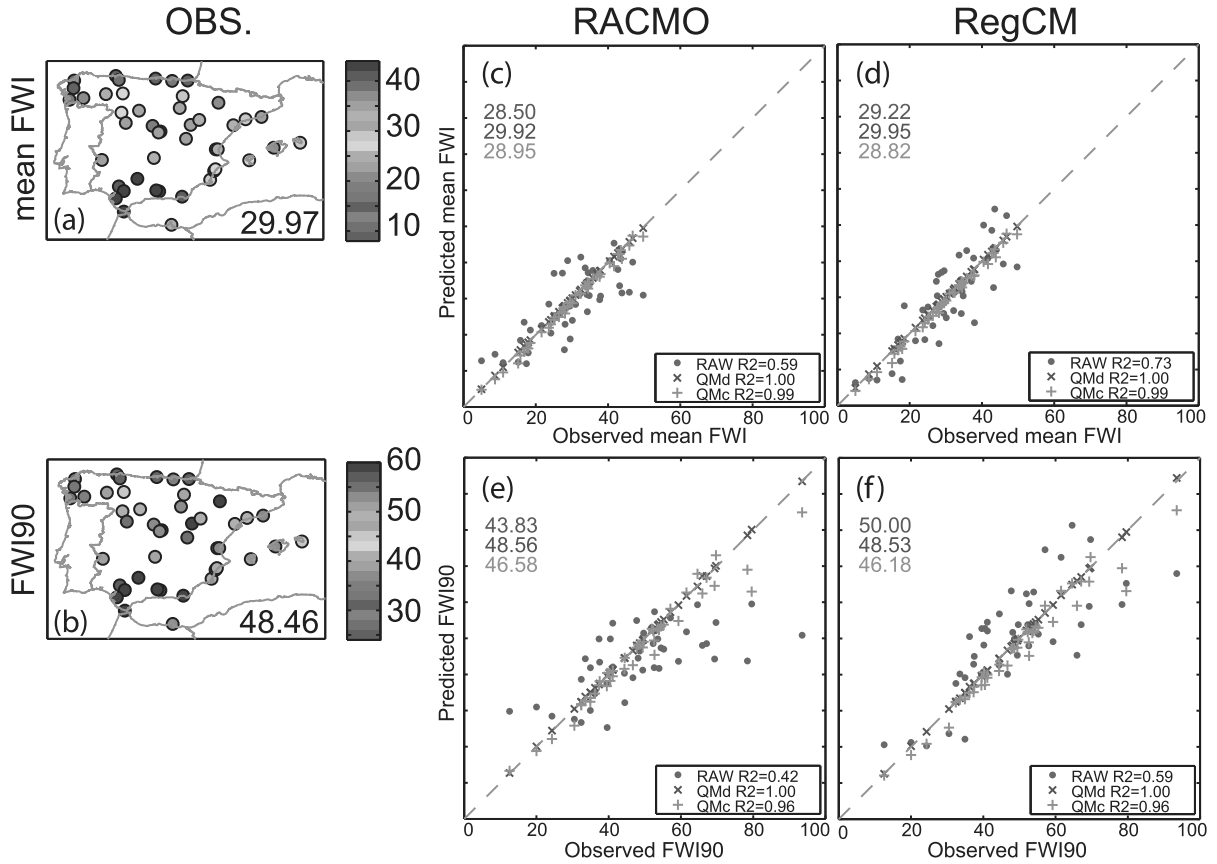


Fig. 1 (a-b) Observed climatology for the period 1979-2003. The numbers indicate the spatial average over the 45 stations. Scatter plots for the point-based observed versus simulated mean FWI (c-d) and FWI90 (e-f), for the two RCMs (columns) in the cross-validation experiment. Within each scatter plot, red points represent the respective values for the raw RCM in the closest grid box, blue markers for the QM_d and green markers for the QM_c. The coloured numbers represent the spatial average over the 45 stations. Goodness-of-fit estimates for the linear regression (R^2) are provided in each case, using a robust fit with a weighting function (Andrews, 1974), the same as used by Brands et al (2014), to avoid the effect of outliers.

As expected, the application of QM (Fig. 1, c-f) drastically reduces biases from the raw RCM output for mean FWI and FWI90. In all cases, QM_d agrees better than QM_c with the observations, although the latter also presents high R^2 values (above 0.95 in both indices and RCMs). These results are consistent with the maps of mean biases from the raw and adjusted data (Supplementary material, Fig. S1).

Note that QMc is based on the non-linear combination of variables that are independently adjusted which seem to retain their physical relationships (e.g. between temperature and relative humidity) to provide a very reasonable estimate of the indicators ($R^2 > 0.95$). Given that QMd and QMc, with their pros and cons, perform equally well under current climate conditions, any errors in the FWI introduced by the individual biases through QMd are implicitly equivalent to potential inter-variable inconsistencies introduced through QMc. This relative equivalence has also been found for seasonal forecasts (Bedia et al, 2017) but may not be preserved under future climate projections, when greater differences are expected between the distributions in the test (future) and the training periods (Sect. 3.2).

3.2 Future FWI projections

In this Section the QM is trained with the observations in the period 1979-2003 and the PAFs are then applied to the reference historical simulations (1971-2000) and to three 30-year future periods from the A1B scenario (near: 2011-2040, mid: 2041-2070 and long-term: 2071-2100).

Relative delta changes (Fig. 2) for mid- and long-term future periods show an increase of mean FWI in most stations (panels a-d), ranging on (spatial) average between 7-17% (mid-term) and 20-37% (long-term), with larger spatial variability in the far future period. Overall, for the near and mid-term future periods (green and purple markers), the climate change signals after QM are consistent with their raw counterparts. For the long-term future period (red circles), QM modifies the climate change signal to a greater extent and differently for each RCM. For RACMO, QMd modifies the raw climate change signal of mean FWI by largely increasing it, whereas QMc preserves better the raw climate change signal. For RegCM, both QM approaches reduce the climate change signal in the mean FWI.

Climate change signals for FWI90 (Fig. 2, e-h) are more sensitive to the QM choice, especially in the long-term future period. For both RCMs, QMd largely increases the climate change signal with respect to the original one, whereas QMc preserves better the RCM signal. Possible reasons of the large modifications of the QMd in the climate change signal are investigated in Sect. 3.3.

3.3 A diagnostic tool to assess the effect of bias adjustment on the climate change signal

As shown in the previous section and displayed spatially in Fig. 3, QMd modifies the climate change signal to a larger extent than QMc in both RCMs. This is particularly noticeable in RACMO for FWI90, in which the difference between the QMd-adjusted and the raw climate change signal is larger than 10 in most of the stations. Three stations (squared in Fig. 3 and labeled in Fig. 2)

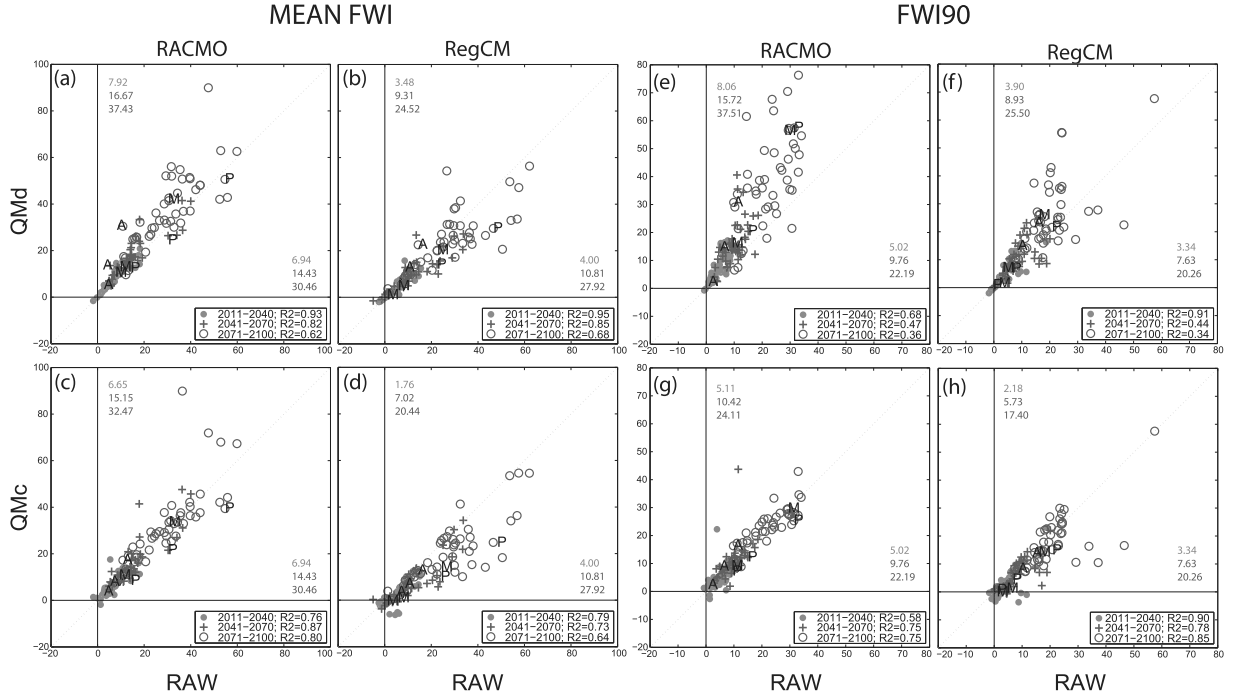


Fig. 2 Point-based relative delta changes for mean FWI (a-d) and FWI90 (e-h), for the raw versus QMd (first row) and raw versus QMc (second row). Results for the three future periods are shown with different markers; goodness-of-fit estimates for the robust linear regression are provided in each case. Three stations (Almería -A-, Madrid -M- and Ponferrada -P) are labeled for further analysis. The coloured numbers represent the geometric mean across the stations for each period for the respective values represented in the X and Y axis. The geometric mean over the deltas is equivalent to the delta of the spatial geometric mean.

have been selected as illustrative examples for further analysis: Ponferrada (in the northwest), Madrid-Retiro (center) and Almería-Aeropuerto (southeast).

To better understand the larger contribution of the QMd approach to the signal, we analyze the percentile adjustment function (PAF) calibrated in the training phase (period 1979-2003) and the way it is applied (PAFeff) in the reference period and long-term projections (Section 2.5). Consistently with the stationarity assumption of the QM, the PAF is assumed to be invariant in time. During the application of the QM, the model data are mapped into the percentiles of the training data, and each value is adjusted according to the corresponding PAF value. Ideally, a smooth PAF is preferred, meaning that the model is systematically over- or underestimating the variable or, at least, adjustments vary slowly across percentiles (see for example the comparison against the constant PAF from an additive scaling method, depicted in magenta in Fig. 4a-f). Otherwise, a largely-varying PAF may produce an effect on the climate change signal difficult to foresee. Discerning which changes in the climate change signal are accepted and which ones lack physical foundations is not straightforward, requiring deep process-understanding (Maraun et al, 2017). The analysis of the PAF can facilitate the detection of potential statistical artifacts, serving as a convenient tool for a first diagnosis.

In the examples shown (Fig. 4a-f, blue lines), the PAFs calibrated through QMd present some variability across percentiles and are considerably large for the upper tail of the FWI distribution. For instance, in Almería the PAFs

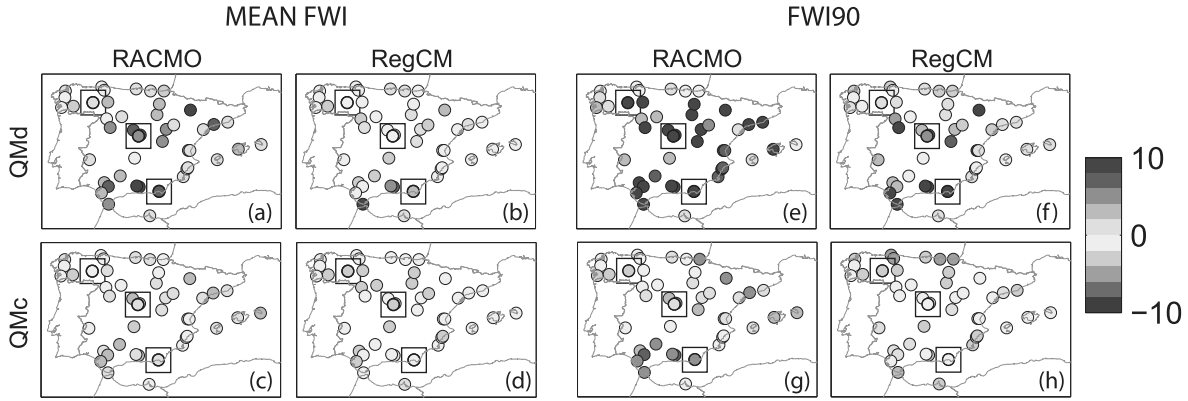


Fig. 3 Differences between the bias-adjusted and the raw climate change signal for the long-term future period, for the mean FWI (a-d) and the FWI90 (e-h). Results for the QMd (QMc) are presented in the first (second) row. The stations in Ponferrada (north-west), Madrid-Retiro (center) and Almeria-Aeropuerto (south-east) have been highlighted for further analysis.

of both RCMs start being positive (FWI is underestimated and requires an increase to fit the observations), then turn negative in the central part of the distribution and turn again into large positive values in the upper tail of the distribution (up to +50 and +35, respectively). For Madrid with RegCM, despite of the changing PAFs, positive and negative adjustments compensate in the climate change signal of mean FWI (see Figs. 2b and 3b), leading to a small effect of QMd on mean FWI. However, the effect becomes apparent for FWI90 (Figs. 2f and 3f). Conversely, RACMO in Madrid mainly presents a large negative bias in the upper tail (thus positive PAF) that, together with the application of constant extrapolation (Section 2.5), leads to larger climate change signals in the mean (Figs. 2a and 3a) and especially, in FWI90 (Figs. 2e and 3e).

In addition to the PAFs, it is essential to know how the target values are distributed with respect to the model percentiles in the training phase, which will determine the shape of the PAFeff. The relative frequency of values of the reference period mapped into the training data (Fig. 4a-f, gray lines) are very close to the frequencies in the training period (i.e. 0.01 by definition), since the reference and the training period largely overlap. This means that, as long as the distribution of values in the test period does not differ much from that in the training phase, the frequency of the adjustments applied will be similar along the distribution (PAFeff resembles PAF, see red dashed lines in Figure 4g-l). This is also the case for the near and mid-term future periods (not shown), where QMd does not modify the raw climate change signal to a great extent. Conversely, the values from the long-term future period mapped into the percentiles in the training phase are not regularly distributed, being more frequent those corresponding to the highest FWI percentiles (Fig. 4a-f, black lines). This effect, together with the usually higher values of the PAF in the upper tail (blue lines) leads to a PAFeff with high frequency of large positive adjustment values (Figure 4g-l, red solid lines) and, thus, to larger climate

change signals under QMd (as discussed by Themefl et al, 2012; Hagemann et al, 2011, for uni-variate BC).

As mentioned before, the effect of QM on the mean FWI could be compensated depending on the PAF (Fig. 3b), but it becomes more evident in FWI90. The vertical dashed lines in Figure 4a-f depict the percentile probabilities (in the training data) associated to the FWI90 in the reference (gray) and long-term future period (black). Whereas the gray lines correspond to the 90th or 91st percentiles, the FWI90 in the long-term future varies between 96-99th. Thus, their intersections with the blue lines (PAFs) give the adjustments applied in the reference and future periods, whose difference is the direct contribution of the QMd to the raw climate change signal (as in Fig. 3, right panel for FWI90).

Although the change in the frequency distribution may also be found in the component variables (higher temperature and lower relative humidity episodes become more frequent in a warmer climate), the adjustments applied to the individual variables have converse effects on the adjusted FWI projections, as shown by the PAFeff of the QMc (green lines in Figure 4g-l). The QMc approach also allows to track down biases in the multi-variate index by analyzing the individual corrections (Fig.S2). Even in the cases of the largest increase of the signal due to QMd (e.g. Almería with the two RCMs), the PAFeff for the individual variables for the end of the century (green lines in Fig.S2) are relatively small and stable in the part of the distribution which contributes to higher fire risk (i.e. upper part for temperature and lower part for relative humidity). Thus, the amplification of the signal by the direct QM cannot be attributed to the biases of the individual variables.

4 Discussion and Conclusions

This study assesses climate projections of an impact-relevant, multi-variable CI that characterizes fire danger (the Fire Weather Index, FWI) over Spain, using empirical Quantile Mapping (QM) as bias-adjustment method. QM is a popular method which is used often to this aim, as it can be applied to variables with different statistical properties. Due to the biases of the simulated component variables forming the FWI, it is common practice to adjust the biases in the individual variables prior to the index calculation (see e.g. Yang et al, 2015). Here, this approach (component-wise) is compared to the direct adjustment of FWI. By doing this, the physical and temporal coherence among the primary variables given by the RCM is preserved, at the expense of including the individual biases in the FWI calculation.

We illustrated the effect of bias adjustment methods in two different setups (direct and component-wise) on a single climate impact index and using two RCMs nested into a single GCM. Any conclusion about future fire risk in Spain should rely on an ensemble of simulations encompassing all known uncertainty sources (i.e. including several GCMs, RCMs and emission scenarios), which are out of the scope of this work.

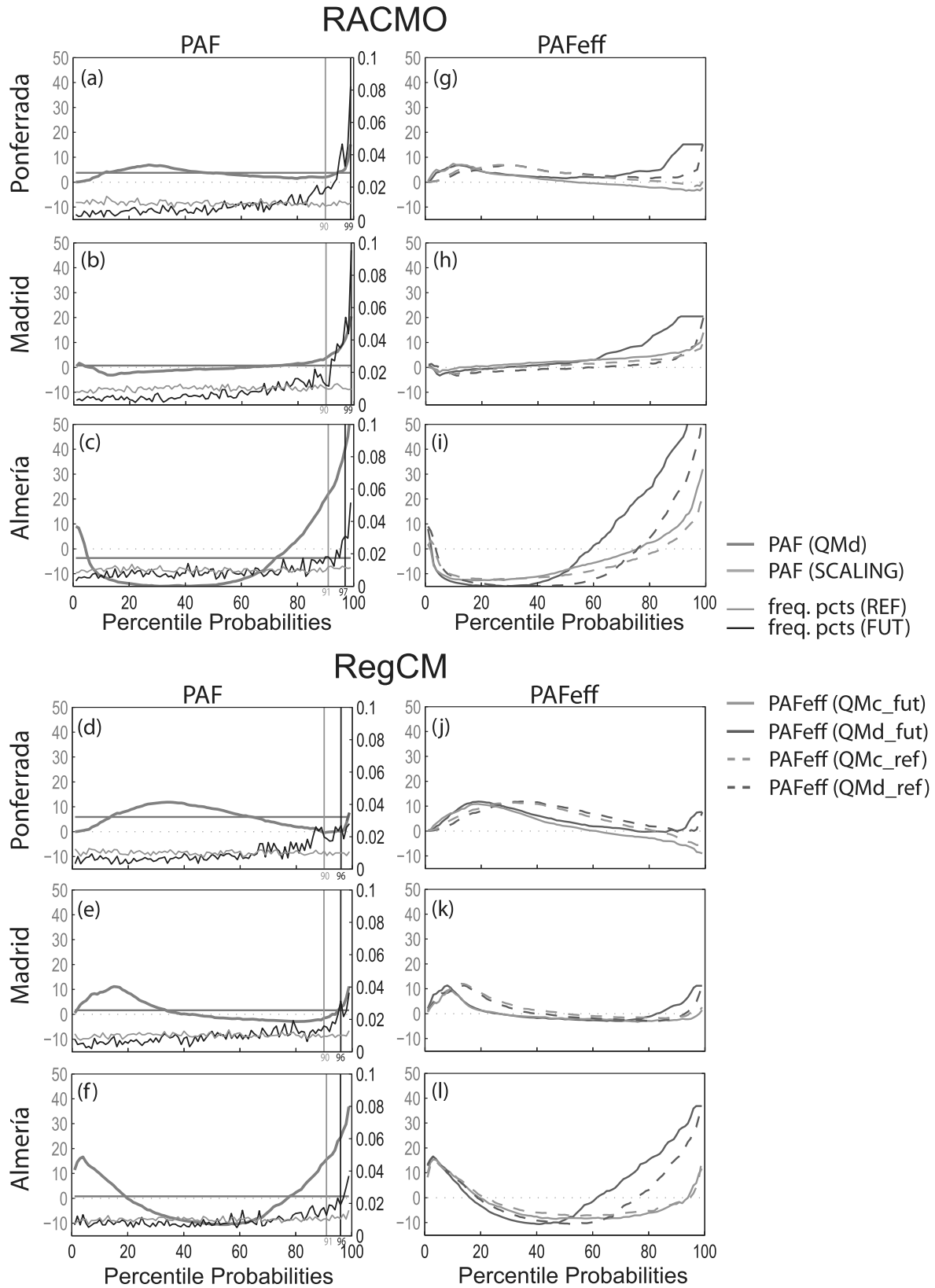


Fig. 4 [a-f] Percentile adjustment functions (PAF) calibrated by the QMd of the FWI during the training phase for the 99 percentiles (blue lines, referred to Y-left axis). Frequency of RCM values in the reference (gray) and future (black) periods mapped into the percentiles of the model data in the training period (referred to Y-right axis). The reference and future periods are 1971-2000 and 2071-2100, respectively. The vertical lines represent the percentile corresponding to the FWI90 in the reference (gray) and in the future (black) referred to the RCM training data. The PAF obtained from an additive scaling bias adjustment method is depicted in magenta. [g-l] Effective percentile adjustment functions (PAFeff) calculated from the QMd (red) and QMc (green) in the reference (dashed) and future (solid lines) periods.

From our results, both the direct and component-wise bias adjustment approaches exhibit a similar performance in current climate under a cross-validation framework, with a slightly better performance for the direct approach. This is expected from the experimental design, since under the direct approach the FWI is cross-validated as a single variable, regardless of the biases in the input variables and in their interdependences. The relative agreement of the two approaches in current climate suggests that QM performs equally well despite the presence of biases in the input variables (direct approach) or the potential inconsistencies among the corrected variables (component-wise approach). The main advantage of the direct approach consists on adjusting a single variable instead of several variables with different statistical properties. With this promising results, we examine the suitability of the direct bias adjustment in climate projections.

Our results for the FWI projections suggest that the component-wise approach is preferable for extreme percentiles, such as the FWI90. This is due to the smoother effective adjustments (PAFeff) at the upper part of the FWI distribution, as a result of compensating non-linear effects of more stable individual corrections. Therefore the inter-variable dependencies and the climate change signal from the uncorrected data are retained. The component approach can be considered more tractable, since examining the individual biases can provide some hints for the multi-variate index. Conversely, the direct approach may be artificially inflating the FWI values due to the larger magnitude of the local adjustments in the upper percentiles, which become more frequent under the warmer climate conditions projected for the long-term future.

QM allows for the correction of intensity-dependent biases, thus to some extent changes in the climate change signal are acknowledged. However, it is debatable if QM improves or deteriorates the climate change signal (Gobiet et al, 2015; Maraun et al, 2017) and, moreover, either the direct or component-wise adjustment of a multi-variate index produce a more feasible signal. It is well known that climate models may simulate implausible changes on large-scale phenomena due to unrealistically represented processes (see several examples in Maraun et al, 2017). Such fundamental climate model errors cannot be improved by QM, which is a mere statistical post-processing method. Trend-preserving corrections are a sensible choice when a climate model simulates a credible climate change signal (Maraun, 2016), otherwise neither the raw nor the bias-corrected signal may be physically justified, since QM will inherit the implausible projected changes anyway. The modification of the climate change signal by QM comes from the rescaling of modelled day-to-day variability. Therefore, non-trend preserving methods may sensibly be used if the transfer function calibrated on short timescales can be reasonably applied to correct biases on long timescales, and the modifications in the signal are physically justified (Maraun et al, 2017). Analyzing relevant climatic processes and their model representation is essential to assess if the modifications of the signal are plausible. Other modifications of the climate change signal than those based on physical arguments would probably be statistical artifacts. The use of the percentile adjustment function (PAF) for any uni- or multi-variate future

experiment constitutes an effective tool for the analysis of the modification of the signal due to bias adjustment.

This work is focused on a multi-variate index, however the main conclusions are equally valid for any uni-variate QM transformation. We show that a proper cross-validation experiment is not useful to detect problems that may arise in a climate change context due to the changes in the frequency distribution in a future climate. Also, the effect of the QM on the climate change signal depends on the parameter of the distribution under consideration, since some effects with opposite sign may compensate (e.g. in the mean). The modifications of the climate change signal go hand in hand with the percentile adjustment function. Although QM is a powerful technique, it should not be indistinctly applied before analyzing the PAF, which depends on the variable, model, location, etc. Therefore, the appropriateness of the direct or component-wise bias adjustment for other climate indices, RCMs and locations should be tested on case-specific analyses.

Acknowledgements All the statistical downscaling experiments have been computed using the MeteoLab software (<http://www.meteo.unican.es/software/meteolab>), an open-source Matlab toolbox for statistical downscaling. The authors are grateful to the Spanish Meteorological Agency (AEMET) for providing the observational data and Erika Coppola from the International Center of Theoretical Physics (ICTP) and Erik van Meijgaard from the Royal Netherlands Meteorological Institute (KNMI) for making available the ENSEMBLES RegCM3 and RACMO2 regional climate models, respectively. A.C. thanks to the Spanish Ministry of Economy and Competitiveness for the funding provided within the FPI programme (BES-2011-047612). J.F. acknowledges support from the INSIGNIA project, co-funded by the Spanish R&D programme (CGL2016-79210-R) and the European Regional Development Fund. This work was partially supported by the project MULTI-SDM (CGL2015-66583-R, MINECO/FEDER). We also thank two anonymous referees for their useful comments that helped to improve the original manuscript.

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. *Journal of Geophysical Research: Atmospheres* 121(5):2075–2089, DOI 10.1002/2015JD024040, URL <http://dx.doi.org/10.1002/2015JD024040>, 2015JD024040
- Andrews D (1974) A robust method for multiple linear-regression. *Technometrics* 16(4):523–531, DOI 10.2307/1267603
- Bedia J, Herrera S, San-Martín D, Koutsias N, Gutiérrez JM (2013) Robust projections of Fire Weather Index in the Mediterranean using statistical downscaling. *Climatic Change* 120(1-2):229–247, DOI 10.1007/s10584-013-0787-3
- Bedia J, Herrera S, Camia A, Moreno JM, Gutiérrez JM (2014) Forest fire danger projections in the Mediterranean using ENSEMBLES regional climate change scenarios. *Climatic Change* 122(1-2):185–199, DOI 10.1007/s10584-013-1005-z

- Bedia J, Herrera S, Gutiérrez JM, Benali A, Brands S, Mota B, Moreno JM (2015) Global patterns in the sensitivity of burned area to fire-weather: Implications for climate change. *Agricultural and Forest Meteorology* 214:369 – 379, DOI <http://dx.doi.org/10.1016/j.agrformet.2015.09.002>
- Bedia J, Golding N, Casanueva A, Iturbide M, Buontempo C, Gutierrez JM (2017) Seasonal predictions of Fire Weather Index: Paving the way for their operational applicability in Mediterranean Europe. *Climate Services* DOI 10.1016/j.cliser.2017.04.001
- 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? *Int J Climatol* 34(5):1615–1627, DOI 10.1002/joc.3788
- Cannon AJ (2016) Multivariate bias correction of climate model output: Matching marginal distributions and intervariable dependence structure. *Journal of Climate* 29(19):7045–7064, DOI 10.1175/JCLI-D-15-0679.1
- Casanueva A, Frías MD, Herrera S, San-Martín D, Zaninovic K, Gutiérrez JM (2014) Statistical downscaling of climate impact indices: testing the direct approach. *Climatic Change* 127(3-4):547–560, DOI 10.1007/s10584-014-1270-5
- Casanueva A, Kotlarski S, Herrera S, Fernández J, Gutiérrez J, Boberg B, Colette A, Christensen OB, Goergen K, Jacob D, Keuler K, Nikulin G, Teichmann C, Vautard R (2016) Daily precipitation statistics in a EURO-CORDEX RCM ensemble: added value of raw and bias-corrected high-resolution simulations. *Climate Dynamics* 47:719–737, DOI 10.1007/s00382-015-2865-x
- Christensen JH, Carter TR, Rummukainen M, Amanatidis G (2007) Evaluating the performance and utility of regional climate models: the PRUDENCE project. *Climatic Change* 81(1):1–6, DOI 10.1007/s10584-006-9211-6
- Christensen JH, Boberg F, Christensen OB, Lucas-Picher P (2008) On the need for bias correction of regional climate change projections of temperature and precipitation. *Geophys Res Lett* 35(20):L20,709, DOI 10.1029/2008GL035694
- Cofiño A, Bedia J, Iturbide M, Vega M, Herrera S, Fernández J, Frías M, Manzanas R, Gutiérrez J (2017) The ecoms user data gateway: Towards seasonal forecast data provision and research reproducibility in the era of climate services. *Climate Services* DOI <https://doi.org/10.1016/j.cliser.2017.07.001>, URL <http://www.sciencedirect.com/science/article/pii/S2405880717300079>
- Déqué M (2007) Frequency of precipitation and temperature extremes over France in an anthropogenic scenario: Model results and statistical correction according to observed values. *Global and Planetary Change* 57(1?2):16–26, DOI 10.1016/j.gloplacha.2006.11.030
- Dimitrakopoulos A, Bemmerzouk A, Mitsopoulos I (2011) Evaluation of the Canadian fire weather index system in an eastern Mediterranean environment. *Meteorol Appl* 18:83–93
- Ehret U, Zehe E, Wulfmeyer V, Warrach-Sagi K, Liebert J (2012) HESS Opinions “Should we apply bias correction to global and regional cli-

- mate model data?". *Hydrol Earth Syst Sci* 16(9):3391–3404, DOI 10.5194/hess-16-3391-2012
- Gobiet A, Suklitsch M, Heinrich G (2015) The effect of empirical-statistical correction of intensity-dependent model errors on the temperature climate change signal. *Hydrology and Earth System Sciences* 19(10):4055–4066, DOI 10.5194/hess-19-4055-2015
- Hagemann S, Chen C, Haerter JO, Heinke J, Gerten D, Piani C (2011) Impact of a Statistical Bias Correction on the Projected Hydrological Changes Obtained from Three GCMs and Two Hydrology Models. *J Hydrometeor* 12(4):556–578, DOI 10.1175/2011JHM1336.1
- Herrera S, Fita L, Fernández J, Gutiérrez JM (2010) Evaluation of the mean and extreme precipitation regimes from the ENSEMBLES regional climate multimodel simulations over Spain. *J Geophys Res* 115(D21):D21,117, DOI 10.1029/2010JD013936
- Herrera S, Bedia J, Gutiérrez J, Fernández J, Moreno J (2013) On the projection of future fire danger conditions with various instantaneous/mean-daily data sources 118:827–840
- IPCC (2014) *Climate Change 2014: Impacts, Adaptation, and Vulnerability. Part A: Global and Sectoral Aspects. Contribution of Working Group II to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*. In: [Field, C.B., V.R. Barros, D.J. Dokken, K.J. Mach, M.D. Mastrandrea, T.E. Bilir, M. Chatterjee, K.L. Ebi, Y.O. Estrada, R.C. Genova, B. Girma, E.S. Kissel, A.N. Levy, S. MacCracken, P.R. Mastrandrea, and L.L. White (eds.)], Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, p 1132 pp
- Kotlarski S, Keuler K, Christensen OB, Colette A, Déqué M, Gobiet A, Gorgen K, Jacob D, Lüthi D, van Meijgaard E, Nikulin G, Schär C, Teichmann C, Vautard R, Warrach-Sagi K, Wulfmeyer V (2014) Regional climate modeling on European scales: a joint standard evaluation of the EURO-CORDEX RCM ensemble. *Geosci Model Dev* 7(4):1297–1333, DOI 10.5194/gmd-7-1297-2014
- Li C, Sinha E, Horton DE, Diffenbaugh NS, Michalak AM (2014) Joint bias correction of temperature and precipitation in climate model simulations. *Journal of Geophysical Research: Atmospheres* 119(23):13,153–13,162, DOI 10.1002/2014JD022514, URL <http://dx.doi.org/10.1002/2014JD022514>, 2014JD022514
- van der Linden P, Mitchell J (eds) (2009) *ENSEMBLES: Climate Change and its Impacts: Summary of research and results from the ENSEMBLES project*. Met Office Hadley Centre, FitzRoy Road, Exeter EX1 3PB, UK
- Maraun D (2013) Bias Correction, Quantile Mapping, and Downscaling: Revisiting the Inflation Issue. *J Climate* 26(6):2137–2143, DOI 10.1175/JCLI-D-12-00821.1
- Maraun D (2016) Bias correcting climate change simulations - a critical review. *Current Climate Change Reports* 2(4):211–220, DOI 10.1007/s40641-016-0050-x, URL <https://doi.org/10.1007/s40641-016-0050-x>

- Maraun D, Shepherd TG, Widmann M, Zappa G, Walton D, Gutiérrez J, Hagemann S, Richter I, Soares PMM, Hall A, Mearns LO (2017) Towards process-informed bias correction of climate change simulations. *Nature Climate Change* 7:764–773, DOI 10.1038/nclimate3418
- van Meijgaard E, van Ulft L, van de Berg W, Bosveld B, van der Hurk B, Lenderik A Gand Siebesma (2008) The KNMI regional atmospheric climate model RACMO version 2.1. Tech Rep 302
- Moriondo M, Good P, Durao R, Bindi M, Giannakopoulos C, Corte-Real J (2006) Potential impact of climate change on fire risk in the Mediterranean area 31:85–95
- Muerth MJ, Gauvin St-Denis B, Ricard S, Velázquez JA, Schmid J, Minville M, Caya D, Chaumont D, Ludwig R, Turcotte R (2013) On the need for bias correction in regional climate scenarios to assess climate change impacts on river runoff. *Hydrology and Earth System Sciences* 17(3):1189–1204, DOI 10.5194/hess-17-1189-2013, URL <https://www.hydrol-earth-syst-sci.net/17/1189/2013/>
- Pal JS, Giorgi F, Bi X, Elguindi N, Solmon F, Rauscher SA, Gao X, Francisco R, Zakey A, Winter J, Ashfaq M, Syed FS, Sloan LC, Bell JL, Diffenbaugh NS, Karmacharya J, Konar? A, Martinez D, da Rocha RP, Steiner AL (2007) Regional Climate Modeling for the Developing World: The ICTP RegCM3 and RegCNET. *Bull Amer Meteor Soc* 88(9):1395–1409, DOI 10.1175/BAMS-88-9-1395
- Panofsky HA, Brier GW (1968) Some applications of statistics to meteorology. University Park : Penn. State University, College of Earth and Mineral Sciences
- Pechony O, Shindell DT (2010) Driving forces of global wildfires over the past millennium and the forthcoming century. *P Natl Acad Sci USA* 107(45):19,167–19,170, DOI 10.1073/pnas.1003669107
- Piani C, Haerter JO, Coppola E (2010a) Statistical bias correction for daily precipitation in regional climate models over Europe. *Theoretical and Applied Climatology* 99(1-2):187–192, DOI 10.1007/s00704-009-0134-9
- Piani C, Weedon G, Best M, Gomes S, Viterbo P, Hagemann S, Haerter J (2010b) Statistical bias correction of global simulated daily precipitation and temperature for the application of hydrological models. *Journal of Hydrology* 395(3):199 – 215, DOI <https://doi.org/10.1016/j.jhydrol.2010.10.024>, URL <http://www.sciencedirect.com/science/article/pii/S0022169410006475>
- Quintana Seguí P, Ribes A, Martin E, Habets F, Boé J (2010) Comparison of three downscaling methods in simulating the impact of climate change on the hydrology of Mediterranean basins. *Journal of Hydrology* 383(1?2):111–124, DOI 10.1016/j.jhydrol.2009.09.050
- Rocheta E, Evans JP, Sharma A (2014) Assessing atmospheric bias correction for dynamical consistency using potential vorticity. *Environmental Research Letters* 9(12):124,010, URL <http://stacks.iop.org/1748-9326/9/i=12/a=124010>

- San-Miguel-Ayanz J, Schulte E, Schmuck G, Camia A (2013) The European Forest Fire Information System in the context of environmental policies of the European Union. *Forest Policy and Economics* 29(SI):19–25, DOI 10.1016/j.forpol.2011.08.012
- Stocks B, Fosberg M, Lynham T, Mearns L, Wotton B, Yang Q, Jin J, Lawrence K, Hartley G, Mason J, McKenney D (1998) Climate change and forest fire potential in russian and canadian boreal forests. *Climatic Change* 38:1 – 13
- Teng J, Potter NJ, Chiew FHS, Zhang L, Wang B, Vaze J, Evans JP (2015) How does bias correction of regional climate model precipitation affect modelled runoff? *Hydrol Earth Syst Sci* 19(2):711–728, DOI 10.5194/hess-19-711-2015
- Teutschbein C, Seibert J (2012) Bias correction of regional climate model simulations for hydrological climate-change impact studies: Review and evaluation of different methods. *Journal of Hydrology* 456?457:12–29
- Thiemeßl MJ, Gobiet A, Heinrich G (2012) Empirical-statistical downscaling and error correction of regional climate models and its impact on the climate change signal. *Climatic Change* 112(2):449–468, DOI 10.1007/s10584-011-0224-4
- Turco M, Sanna A, Herrera S, Llasat MC, Gutiérrez JM (2013) Large biases and inconsistent climate change signals in ensembles regional projections. *Climatic Change* 120(4):859–869, DOI 10.1007/s10584-013-0844-y, URL <https://doi.org/10.1007/s10584-013-0844-y>
- Viegas D, Bovio G, Ferreira A, Nosenzo A, Sol B (1999) Comparative study of various methods of fire danger evaluation in southern Europe. *Int J Wildland Fire* 9:235–246
- Vrac M, Friederichs P (2015) Multivariate-intervariable, spatial, and temporal-bias correction. *J Climate* 28(1):218–237, DOI 10.1175/JCLI-D-14-00059.1
- van Wagner C, Pickett T (1985) Equations and FORTRAN program for the Canadian forest fire weather index system. *Forestry Tech. Rep. 33*, Canadian Forestry Service, Ottawa, Canada
- van Wagner CE (1987) Development and structure of the Canadian Forest Fire Weather Index. *Forestry Tech. Rep. 35*, Canadian Forestry Service, Ottawa, Canada
- Wilby RL, Hay LE, Gutowski WJ, Arritt RW, Takle ES, Pan Z, Leavesley GH, Clark MP (2000) Hydrological responses to dynamically and statistically downscaled climate model output. *Geophysical Research Letters* 27(8):1199–1202, DOI 10.1029/1999GL006078, URL <http://dx.doi.org/10.1029/1999GL006078>
- Wilke RAI, Mendlik T, Gobiet A (2013) Multi-variable error correction of regional climate models. *Climatic Change* 120(4):871–887, DOI 10.1007/s10584-013-0845-x
- Williams A, Karoly D, Tapper N (2001) The sensitivity of australian fire danger to climate change. *Climatic Change* 49(1-2):171 – 191

- Wotton BM (2009) Interpreting and using outputs from the Canadian Forest Fire Danger Rating System in research applications. *Environ Ecol Stat* 16:107–131, DOI 10.1007/s10651-007-0084-2
- Yang W, Gardelin M, Olsson J, Bosshard T (2015) Multi-variable bias correction: application of forest fire risk in present and future climate in Sweden. *Nat Hazards Earth Syst Sci* 15(9):2037–2057, DOI 10.5194/nhess-15-2037-2015

SUPPLEMENTARY MATERIAL

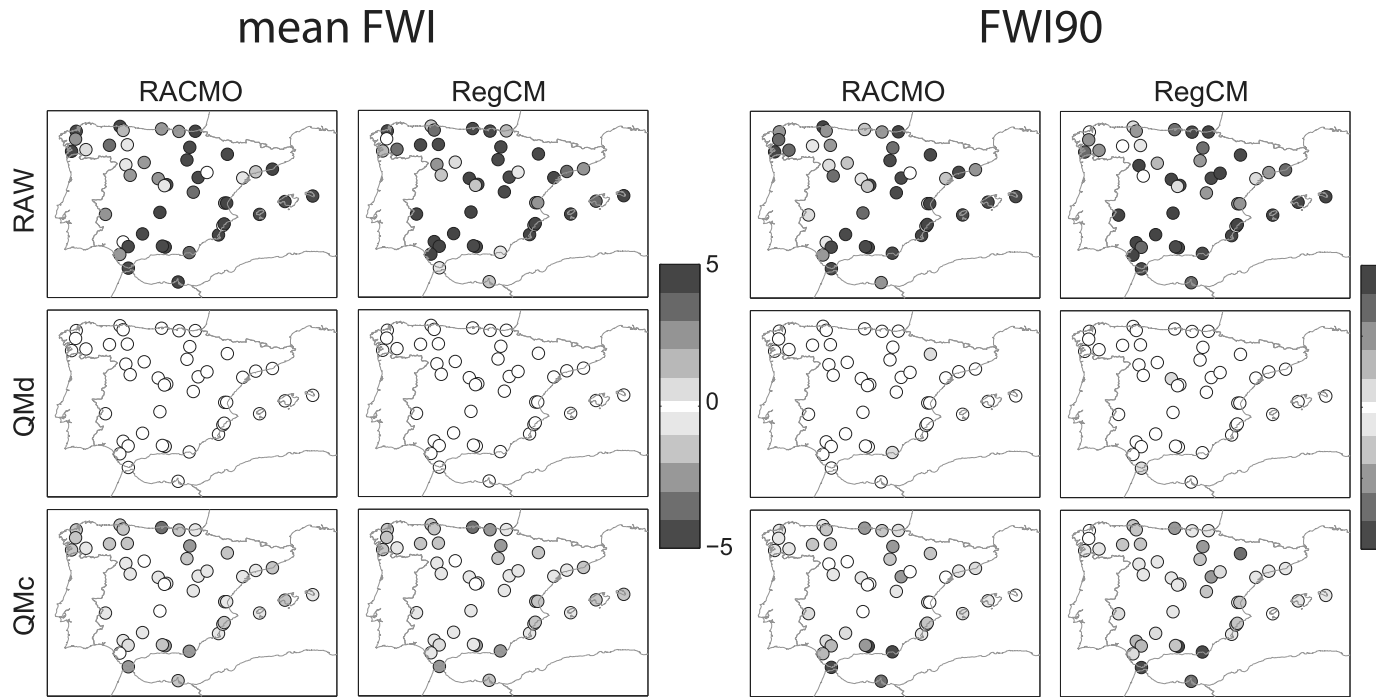


Fig. S1 Spatial distribution of the mean JJAS biases for mean FWI (left panel) and FWI90 (right panel), in the raw RCMs (first row), the QMd (second row) and the QMc (third row) for the two RCMs (columns).

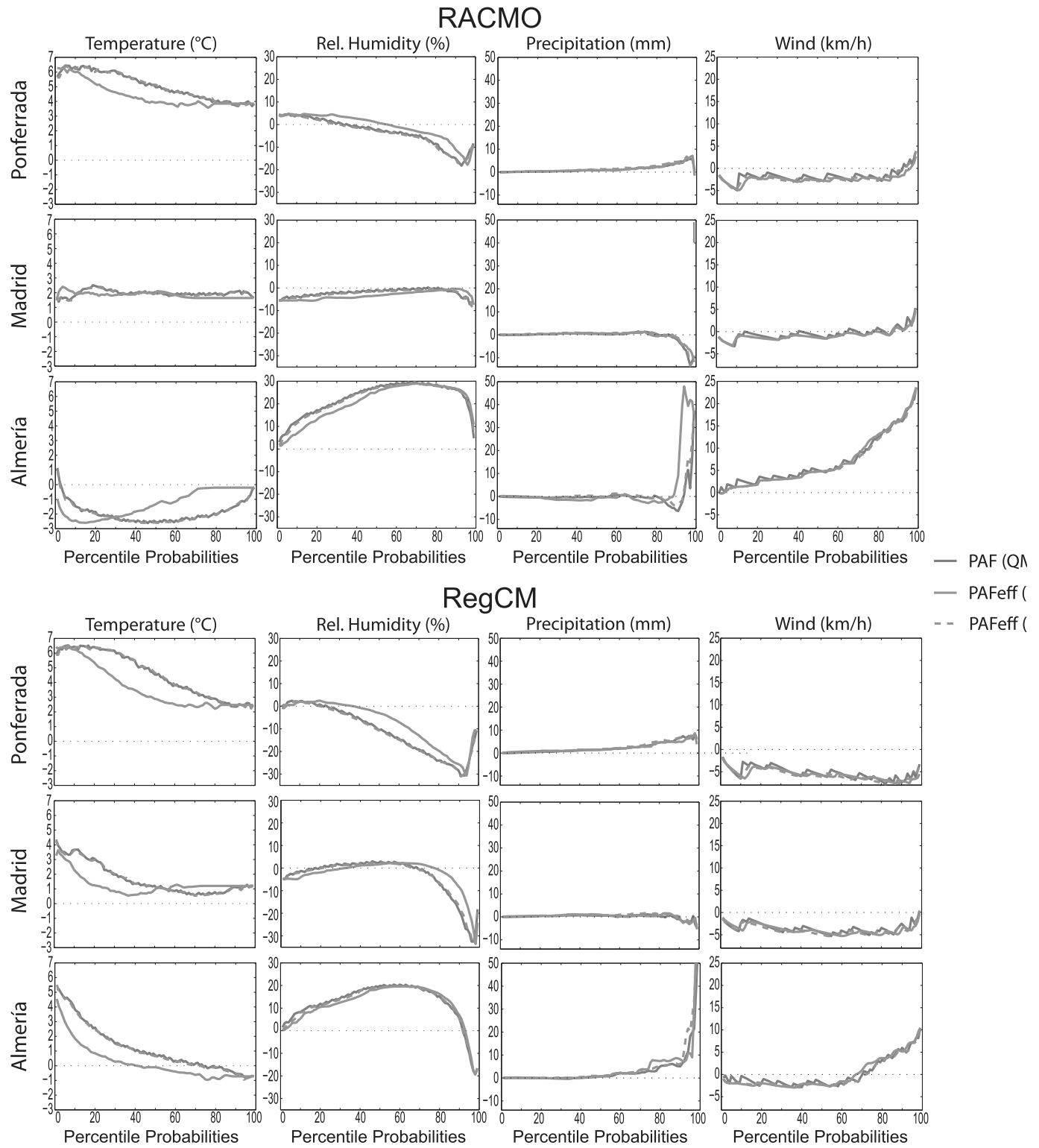


Fig. S2 Percentile adjustment functions (PAF, blue lines) calibrated by the QM for the individual variables in three exemplary locations and the two RCMs. Effective percentile adjustment functions (PAFeff) in the reference (dashed green) and future (solid green lines) periods are also shown.