

[Click here to view linked References](#)

Climate Dynamics manuscript No. (will be inserted by the editor)
--

Process-conditioned bias correction for seasonal forecasting: A case-study with ENSO in Peru

R. Manzananas · J. M. Gutiérrez

Received: date / Accepted: date

Abstract This work assesses the suitability of a first simple attempt for process-conditioned bias correction in the context of seasonal forecasting. To do this, we focus on the northwestern part of Peru and bias correct one- and four-month lead seasonal predictions of boreal winter (DJF) precipitation from ECMWF System4 forecasting system for the period 1981-2010. In order to include information about the underlying large-scale circulation which may help to discriminate between precipitation affected by different processes, we introduce here an empirical quantile-quantile mapping method which runs conditioned on the state of the Southern Oscillation Index (SOI), which is accurately predicted by System4 and is known to affect the local climate.

Beyond the reduction of model biases, our results show that the SOI-conditioned method yields better ROC Skill Scores and reliability than the raw model output over the entire region of study, whereas the standard unconditioned implementation provides no added value for any of these metrics. This suggests that conditioning the bias correction on simple but well-simulated large-scale processes relevant to the local climate may be a suitable approach for seasonal forecasting. Yet, further research on the suitability of the application of similar approaches to the one considered here for other regions, seasons and/or variables is needed.

Keywords Bias correction, process-conditioning, seasonal forecasting, precipitation, ENSO, SOI, Peru

1 Introduction

Flood and drought episodes triggered by ENSO pose serious economic, social and environmental concerns in many tropical countries like Peru, especially during strong El Niño events such as the one occurred in 2015-2016 (see, e.g., Zhai et al, 2016; Emerton et al, 2017; Sulca et al, 2017). The recent advances achieved in the predictability of ENSO (see, e.g., Barnston et al, 2015; Zheng et al, 2016) may help

R. Manzananas (✉) · J. M. Gutiérrez
Meteorology Group. Institute of Physics of Cantabria (IFCA), CSIC-University of Cantabria.
Santander, 39005, Spain. E-mail: rmanzananas@ifca.unican.es

to mitigate these adverse effects by allowing for adequate planning several months in advance. In this context, various initiatives under the umbrella of the Global Framework for Climate Services —such as CLIMANDES-2 for Peru— are focused on the development of climate services which bridge the gap between science and policy by providing key climate information at time-scales that can be relevant for decision-making (for instance, seasonal forecasts especially tailored for different socio-economic sectors such as agriculture, energy or transport).

However, the still limited horizontal resolution of the state-of-the-art seasonal forecasts obtained with General Circulation Models (GCMs) prevents from their direct use in many practical applications, which typically require local climate information (see Doblas-Reyes et al, 2013, and references therein). As indicated in Manzananas et al (2017b), different statistical downscaling methods have been developed since the early 1990s (see, e.g., von Storch et al, 1993) to bridge the gap between these coarse-resolution outputs and the local scale. Under the Model Output Statistics (MOS) approach, these methods rely on empirical/statistical models which link the local observed predictands of interest (precipitation in this work) with explicative large-scale GCM predictors over the area of interest. In the context of seasonal forecasting, different types of MOS have been applied, from linear regression and Canonical Correlation Analysis (see, e.g., Landman and Tennant, 2000; Sinha et al, 2013) to more sophisticated ensemble MOS (EMOS) corrections which take into account the spread-skill relationship (see, e.g., Gneiting et al, 2005; Torralba et al, 2017; Zhao et al, 2017). However, all these methods rely on the temporal correspondence between GCM predictors and observed predictands (they operate at a time-series level) and, therefore, they can only be applied at a monthly/seasonal time-scale —seasonal forecasts do not provide day-to-day correspondence with observations beyond a few days after being issued (see, e.g., Johansson, 2007).— For this reason, simple alternatives based on distributional Bias Correction (BC) methods such as the quantile-quantile mapping (see, e.g., Maraun, 2013) applied in this work have become increasingly popular during the last decade (see, e.g., Themeßl et al, 2012), since they provide a straightforward way to adjust/correct the daily model predictions so that their statistical properties (in terms of daily distributions) are similar to those from the observations. Nevertheless, several problems have been identified which prevent from the uninformed use of BC methods (see, e.g., Ehret et al, 2012; White and Toumi, 2013; Maraun, 2016; Maraun et al, 2017b). For instance, they inherit the model circulation biases —e.g., errors in the position of the inter-tropical convergence zone,— which can lead to meaningless results (Maraun et al, 2017b).

In seasonal forecasting, an important limitation of BC methods derives from the fact that they do not rely on the temporal match between predictors and predictands (they operate at a distributional level, not at a time-series level). As a result, these methods can introduce arbitrary temporal changes (Maraun, 2013) which can deteriorate the interannual variability of the raw predictions (Maraun et al, 2017a). This may be particularly the case for some regions where the local climatology is mostly dominated by local processes poorly resolved by the GCMs such as the tropics (Manzananas et al, 2017b). In this regard, some recent alternatives propose a two step BC adjustment-reshuffling approach to mimic the observed temporal variability (see, e.g., Vrac and Friederichs, 2014). However, this is done at the cost of destroying the temporal consistency with the driving model. Moreover,

the reshuffling strategy followed in these methods prevents from their use in a cross-validation or operational setup.

To partially alleviate these problems, IPCC (2015) and Maraun et al (2017b) advocated the development of process-informed BC methods, combining the statistical modeling with the knowledge about the relevant processes for the problem under study. In this regard, weather types offer a simple and practical solution to define representative circulation regimes and to analyze conditional model biases (Addor et al, 2016). Although BC methods conditioned to a number of weather types obtained for the region of interest have been already applied to correct climate change simulations (see, e.g., Bellprat et al, 2013), they should be carefully taken due to the possible shifts that may appear in the future atmospheric configurations (see, e.g., Wetterhall et al, 2012). Nevertheless, and despite this is not a relevant issue for the case of seasonal forecasts due to their shorter predictive horizon, the application of process-based (e.g. weather-type conditioned) BC methods remains unexplored yet for this particular type of predictions.

In this work we focus on this matter and assess the suitability of a first simple attempt for process-conditioned BC over northern Peru. In particular, we focus on boreal winter (DJF) precipitation and introduce a quantile-quantile mapping technique conditioned on the state of ENSO. We compare the performance of this conditioned method against the standard (unconditioned) implementation, which has been recently shown to be inappropriate for reproducing the clustered temporal precipitation structures (driven by ENSO) characteristic of this region (Maraun et al, 2017b). Beyond the expected reduction in model bias, the conditioned method is found to appropriately modify the non-representative temporal structure of the raw model output, providing more realistic local time-series, which results in improved forecast association, accuracy and reliability.

The paper is organized as follows. In Section 2 we introduce the region of study, the data used and the BC methods that are applied. The results obtained are presented through Section 3. Finally, the most important conclusions are given in Section 4.

2 Data and Methods

2.1 Region of Study and Observations

In this work we focus on precipitation over the northwestern part of Peru, which, besides the strong influence of ENSO (see, e.g., Bazo et al, 2013; Sanabria et al, 2017) and the ITCZ (Garreaud, 2009), is affected by local forcing related to the presence of the Andes and the complex land-sea contrast. As a result, precipitation over this region exhibits a large local variability and provides therefore an appropriate test-bed for downscaling studies (see, e.g., Horel and Cornejo-Garrido, 1986). We used daily precipitation from 71 gauges owned and maintained by SENAMHI (the national meteorological service) which cover 1981-2010, the period for which seasonal forecast data was available (see next section). Figure 1(a) shows two homogeneous clusters of stations which were obtained by applying the Ward's minimum variance method (Ward, 1963) to the interannual time-series of accumulated precipitation for boreal winter (DJF, the target season of this work), which are shown in panels (b) and (c). The highlighted series correspond to San Miguel and

Magunchal, two representative stations which will be used in Section 3 to illustrate some of the results obtained.

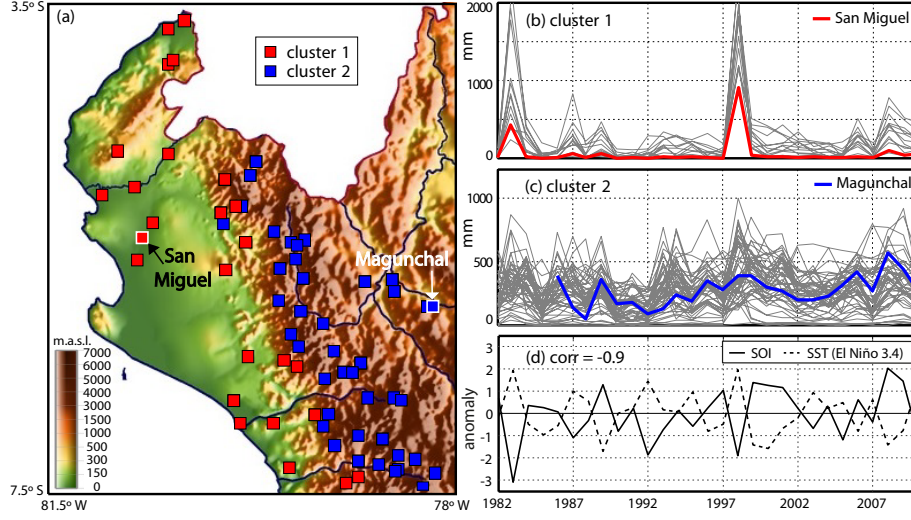


Fig. 1 (a) Orography of the region under study (in meters above sea level) and location of the available stations, classified in two clusters obtained by applying the Ward's minimum variance method on the interannual time-series of accumulated precipitation for the boreal winter (DJF). The interannual series for all the stations in clusters 1 (red) and 2 (blue) are shown in panels (b) and (c), respectively. Two representative stations (San Miguel and Magunchal) are highlighted. (d) Interannual anomalies of observed SOI (solid line) and SST averaged over the El Niño 3.4 region (dashed line). Correlation between both is given in the upper part of the panel.

The solid line in Figure 1(d) shows the interannual anomalies of the observed Southern Oscillation Index (SOI) for DJF. The SOI (see, e.g., Trenberth, 1984) is a standardized index computed as the difference between mean sea level pressure anomalies at Tahiti and Darwin, in Australia (these data were obtained from <http://www.cgd.ucar.edu/cas/catalog/limind/soi.html>). For comparison purposes, the dashed line shows the observed SST anomalies averaged over the El Niño 3.4 region (data obtained from <http://www.cpc.ncep.noaa.gov/data/indices/sstoi.indices>). In both cases, anomalies were computed with respect to the whole period of study, 1981-2010. The high negative correlation found between SOI and El Niño 3.4 SSTs (-0.91) indicates that both indices are mostly equal representations of the same underlying phenomenon, ENSO. Therefore, as seasonal forecasts of SST were not available for this work, we considered the SOI as a proxy for ENSO. Note that, given the strong connection between SOI and El Niño 3.4 SSTs, results are expected to be very similar in both cases.

To quantitatively assess the existing relationship between ENSO and the local climate, Figure 2 shows the interannual correlation between the observed SOI and precipitation at the 71 stations, in terms of the Pearson and the Spearman coefficients —significant ($\alpha = 0.05$) values are marked with a black dot.— In

agreement with previous studies (see, e.g., Sulca et al, 2016), significant positive correlations are found for both coefficients in cluster 2, indicating that high/low precipitation is received during high/low SOI episodes (i.e., during La Niña/El Niño conditions). However, for cluster 1, whereas close to zero correlations are found in terms of Spearman, significant negative ones are found for Pearson. This is due to the particular effect of ENSO over this area, where only very strong El Niño conditions trigger high precipitation episodes (see, e.g., Rau et al, 2017) —see, for instance, the observed peaks in San Miguel in years 1982/83 and 1997/98, which are shown in Figure 1(b).— Note that, these extreme conditions have a stronger impact in the calculation of the (Gaussian-based) Pearson coefficient than in the (rank-based) Spearman one. Therefore, both correlations are needed in order to properly assess the effect of ENSO on the local precipitation (teleconnections) over the entire study region. However, for validation purposes, and for simplicity, we only use the Spearman coefficient in the following. Moreover, note that tercile-based scores (such as the ROC Skill Score; see Section 2.4) are more suitable for validation since they provide more meaningful information.

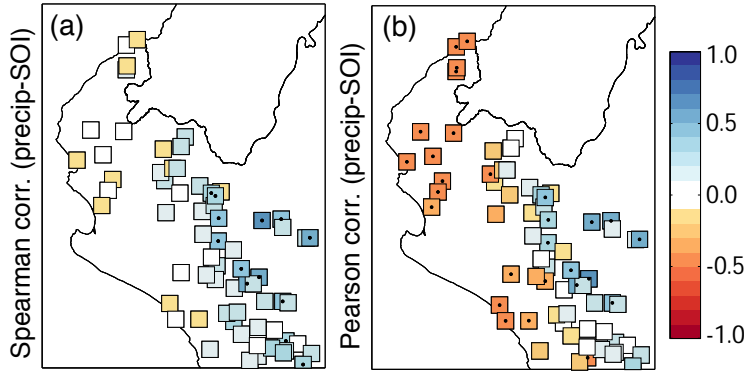


Fig. 2 Interannual correlation between observed local precipitation and SOI, in terms of (a) Spearman and (b) Pearson coefficients.

2.2 Seasonal Forecasts

One- and four-month lead (i.e. initialization of November and August, respectively) retrospective seasonal forecasts of sea level pressure —which is used to compute the predicted SOI— and surface precipitation for DJF were used for this work. They come from the System4 (Molteni et al, 2011) of the European Centre for Medium-Range Weather Forecasts (ECMWF), which is based on the atmospheric model IFS (cycle 36r4) and the oceanic model NEMO. In particular, we first considered the seasonal experiment, in which 15 equiprobable members were run at the beginning of each month for the period 1981-2010, providing seven month-long predictions. Additionally, in order to assess the importance of the ensemble size, we also considered the annual experiment, in which 51 members were run four

times a year (the first of February, May, August and November), providing twelve month-long forecasts.

For the illustrative case of one-month lead forecasts and 15 members, Figure 3 shows the interannual anomalies for the observed (black) and predicted (red) SOI over the period of study. A correlation of 0.91 is found for the ensemble mean (dark red line). Such a good model performance for predicting this index, along with the significant correlations found between the index itself and local precipitation (Figure 2) may provide an opportunity for operational seasonal forecasting in the studied region based on the state of SOI. This possibility is addressed here.

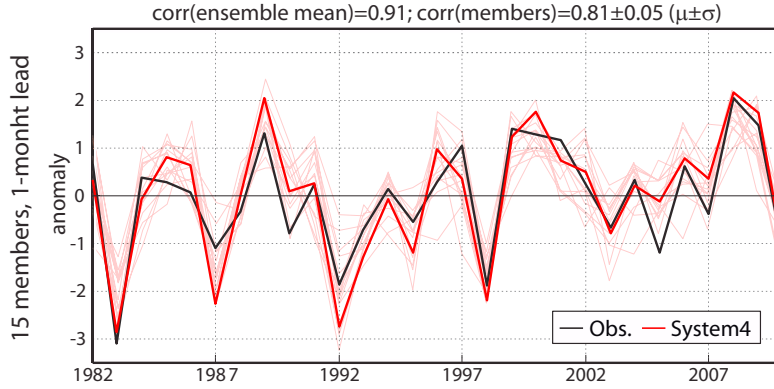


Fig. 3 Interannual anomalies of observed (black) and predicted SOI for DJF over the period of study. For the latter, one-month lead forecasts (i.e. initialization of November) from the seasonal experiment of the System4 are used. The 15 available members (ensemble mean) are shown in light (dark) red. Correlations between the observed and the predicted index (as given by the different members and the ensemble mean) are given in the upper part of the plot.

2.3 Bias Correction Methods

In this work we consider an empirical quantile-quantile mapping method participating in the VALUE downscaling intercomparison initiative (Maraun et al, 2017a) which has been recently applied to correct seasonal precipitation forecasts (Manzananas et al, 2017b). As described in the latter reference, this method consists of calibrating the predicted empirical probability density function (PDF) by adjusting a number of quantiles based on the empirical observed PDF (see, e.g., Déqué, 2007). In particular, we adjusted percentiles 1 to 99 and linearly interpolated every two consecutive percentiles inside this range. Outside this range, a constant extrapolation (using the correction obtained for the 1st or 99th percentile) was applied. Negative values (if any) were set to zero. Moreover, in cases when the predicted frequency of dry days is larger than the observed one, the frequency adaptation proposed by Themeßl et al (2012) was applied. To avoid the artificial skill derived from model over-fitting, we applied a leave-one-out cross validation

scheme (Lachenbruch and Mickey, 1968) in which each year was separately considered for test, whilst the remaining ones were kept for training. Note that this cross-validation framework is the most adequate to test the potential usefulness of the methods for operational seasonal forecasting.

Besides the standard unconditioned implementation of the above described method (denoted simply as BC henceforward), we also considered a circulation-conditioned one, which takes into account the state of the underlying ENSO (as described by the SOI). This implementation, which will be referred to as SOI-BC hereafter, is explained in Figure 4. For each test year (for instance 1998), we first found out the tercile in which the predicted SOI (as given by the ensemble mean) fell—the threshold values identifying the different terciles are computed based on all the remaining years.— As shown in Figure 4(1), let’s assume it was the low tercile (T1). Instead of considering the whole training dataset, the above described quantile-quantile mapping is then fitted considering raw model and observed precipitation in those years for which the observed SOI fell in the same tercile, T1 (2 in the figure)—again, the threshold values identifying the different terciles are computed based on all years except 1998.— Once the parameters of the mapping are found (3), they are used to correct the raw model precipitation for the test year (4). This process is repeated year by year in order to get the final corrected time-series for the entire period of study.

In both BC and SOI-BC, the quantile-quantile mapping was independently fitted/applied for each of the available members in the System4 (member-wise approach). We also tested an ensemble-wise approach in which all members were pooled together and a unique set of adjusting parameters (based on the joint distribution) was then applied to each individual member. In agreement with Manzananas et al (2017b), the results obtained were very similar in both cases, so only the member-wise alternative is considered henceforward.

Note that, although more sophisticated process-conditioned corrections could also be applied, we test here whether or not this simple approach may help to discriminate between precipitation affected by different processes (as characterized by SOI/ENSO). Note also that, in order to refine the correction, more than three categories could be considered for the conditioning. For instance, we also tested the suitability of using five categories (i.e. quintiles), obtaining similar results. Hence, terciles were finally considered as a compromise between getting good validation results and retaining a sufficiently large sample size that allows for a robust statistical fitting.

2.4 Validation Measures

The validation of seasonal forecasts is a multi-faceted problem which requires the use of different scores that allow to properly analyze different quality aspects. Some of these scores (e.g. continuous ranked probability score) are sensitive to changes in the mean and, therefore, can be easily improved by using BC methods as a result of the model bias reduction. In this work, we are interested in the added value of these methods beyond this model bias reduction. Thus, we focus on two validation metrics which are not sensitive to changes in the mean and/or the variance: the ROC Skill Score (ROCSS) and reliability.

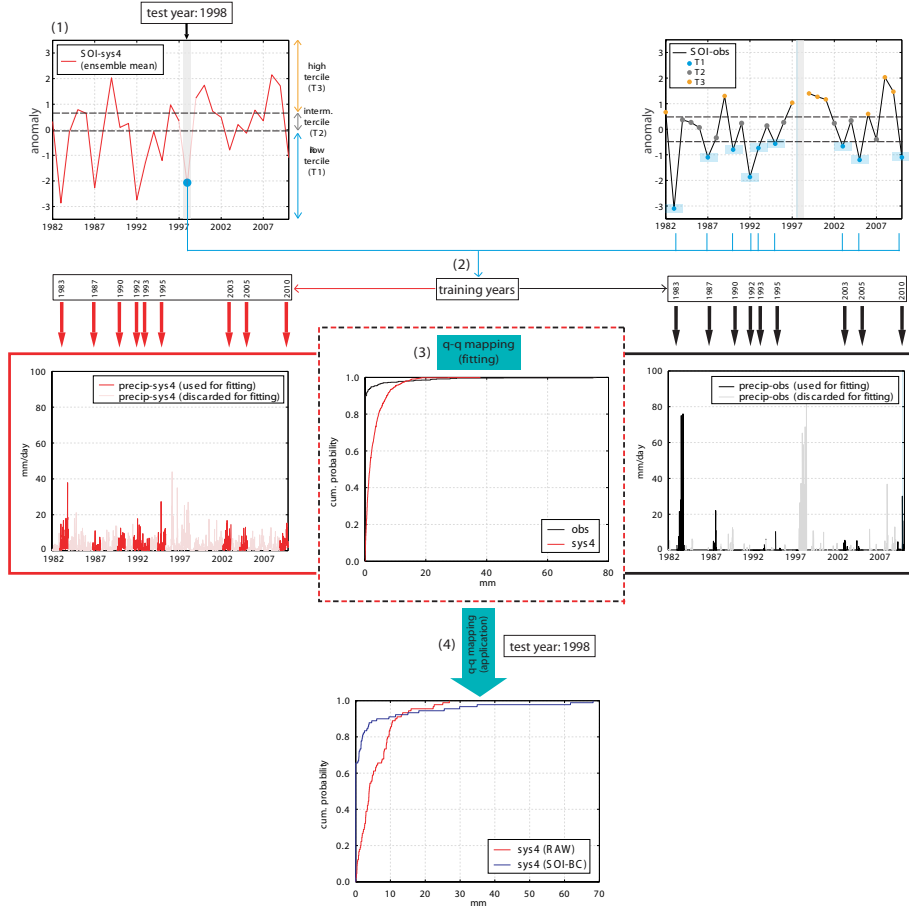


Fig. 4 Diagram illustrating the implementation of the SOI-BC method. For each test year (for instance, 1998): (1) We found out the tercile in which the predicted SOI fell; T1 in this case. (2) The quantile-quantile mapping is fitted considering raw model and observed precipitation in those years for which the observed SOI fell in the same tercile. (3) The parameters of the mapping are found and (4) they are used to correct the raw model precipitation for the test year. This process is repeated year by year in order to get the corrected time-series for the entire period of study. See the text for more details.

The ROCSS measures the accuracy of probabilistic forecasts of different categories —terciles here: dry (T1), normal (T2) and wet (T3).— As described in Manzananas et al (2014), it is computed as $2A - 1$, where A is the area under the ROC curve, so it ranges between 1 ($A = 1$: perfect forecast system) and -1 ($A = 0$: perfectly bad forecast system), with a zero value ($A = 0.5$) indicating no skill compared with a climatological prediction. This metric is recommended by the Lead Centre for the Standardized Verification System of Long Range Forecasts (<http://www.bom.gov.au/wmo/lrfvs/index.html>) for the verification of seasonal forecasts and is a reasonable first choice to communicate the value of a forecast to the end-users. Reliability measures how closely the forecast probabilities of a cer-

tain event correspond to the actual chance of observing that event (terciles here). Reliability diagrams (Weisheimer and Palmer, 2014) have been traditionally used to assess reliability. These diagrams (see next section) plot the observed frequency of the event considered (e.g. T1, T2 or T3) as a function of its forecast probability, as represented by a determined number of bins (see, e.g., Doblas-Reyes et al, 2008, for details).

3 Results

Figure 5 provides an illustrative example of the application of the two BC methods presented for the particular case of one-month lead predictions from the seasonal experiment of the System4 (15 members). Interannual time-series of observed (black) and predicted precipitation are shown for San Miguel and Magunchal (left and right column, respectively), the two representative stations marked in Figure 1. Among the predictions, the raw System4 output, the unconditioned and the SOI-conditioned method are shown in different rows. Light (dark) blue corresponds to each of the available members (the ensemble mean). The ROCSS for the dry and wet terciles, along with the interannual Spearman correlation are given in the upper part of the plots.

Beyond the expected reduction of the model bias achieved by both BC methods—which is particularly visible in Magunchal; compare panels (d) and (f) with panel (b)—the conditioned SOI-BC method is shown to modify the temporal structure given by the raw model output, improving both the ROCSS and the interannual correlation in the two stations (regardless of their different interannual precipitation regimes). Note that, in the case of San Miguel, the standard BC method is unable to reproduce the observed clustered precipitation around strong El Niño years (panel (c)), which has recently reported by Maraun et al (2017b) as one of the cases for which standard BC methods fail due to non-representative model output. However, the SOI-BC method introduced in this work provides more realistic time-series (panel (e)), with precipitation more adequately distributed along El Niño/non El Niño years. For the case of Magunchal, it is also clear that the SOI-BC method improves the temporal structure given by the standard BC method (compare panels (d) and (f)), which basically follows the raw model output (panel (b)).

In order to further analyze how SOI-BC modifies the temporal structure given by the System4, Figure 6 shows the tercile plots (see, e.g., Manzananas et al, 2017a) for the raw model output, the standard and the conditioned BC method (from top to bottom) for the two stations (again, one-month lead forecasts are considered). Each plot displays, year by year, the predicted probabilities (as obtained from the frequencies of the 15-member ensemble) for each of the three terciles (T1, T2 and T3) in a white-to-orange colored scale, along with the observed tercile (black circles). Overall, as a consequence of the conditioning followed, the SOI-BC method exhibits higher resolution (probabilities far from the climatological value: $1/3$) than both the raw model output and the unconditioned method, which lead to very similar results. In particular, whereas the standard method maintains the temporal sequence of probabilities given by the System4, the SOI-conditioned version can introduce notable changes for particular years. For instance, the SOI-BC significantly increases the probability of the wet (dry) tercile for 2008 towards

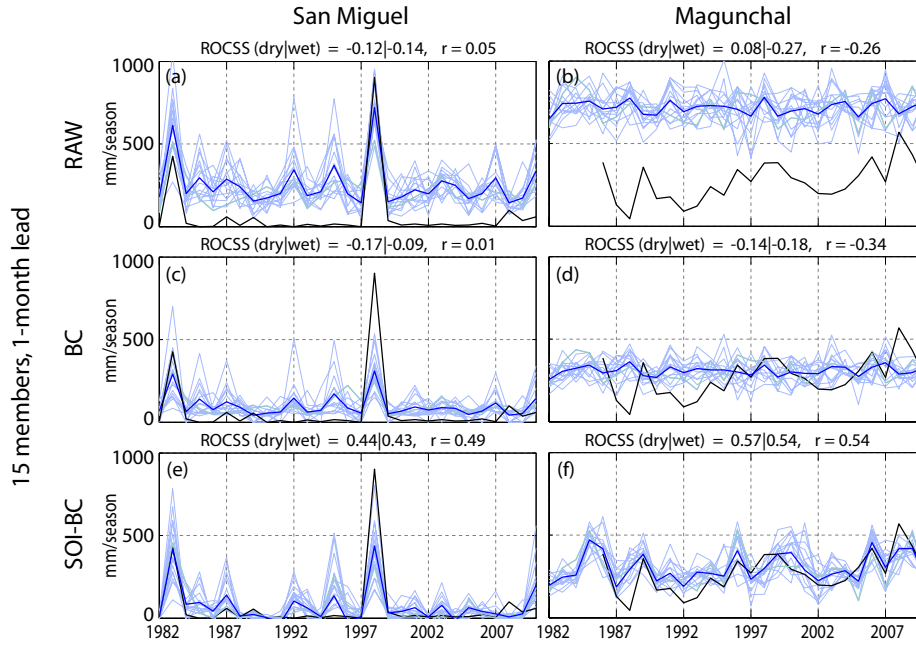


Fig. 5 Interannual time-series for San Miguel and Magunchal (left and right column, respectively), as given by the raw System4 output and the two BC methods applied (in rows) at one month lead-time. The light blue lines correspond to each of the 15 members, whereas the dark blue one represents the ensemble mean. The ROCSS for the dry and wet terciles, along with the interannual Spearman correlation are shown in the plots. In all cases, observations are displayed in black.

the observed value in Magunchal (San Miguel). Alternatively, there are also cases for which it can wrongly modify the prediction of the global model (e.g. the wet tercile for 1998 in Magunchal).

The above results found for San Miguel and Magunchal are representative of the overall performance of the two BC methods applied for the entire study region. Figure 7 shows the ROCSS for the dry and wet terciles (in rows) for raw System4 precipitation (the 15-member experiment), the unconditioned and the SOI-BC method (in columns) for the 71 stations considered. Black dots indicate significant — $\alpha = 0.05$, marked with a black dot— values. As in (Manzananas et al, 2014), significance was computed by means of bootstrapping (Mason and Graham, 2002) with 1000 samples; i.e., by generating 1000 time-series of probabilistic forecasts by randomly resampling the original series. For the case of one-month lead predictions (left panel), whereas the standard BC method provides no improvement with respect to the raw model output, the SOI-BC version yields better ROCSS over most of the stations, particularly in cluster 2, where System4 (and the standard BC method) exhibits negative ROCSS. Moreover, although there is a general decrease in skill for the case of four-months lead predictions (right panel), the SOI-BC method still provides better results than both the raw model and the unconditioned BC method at this longer lead-time, which points out the potential usefulness of process-conditioned BC methods for operational decision-making.

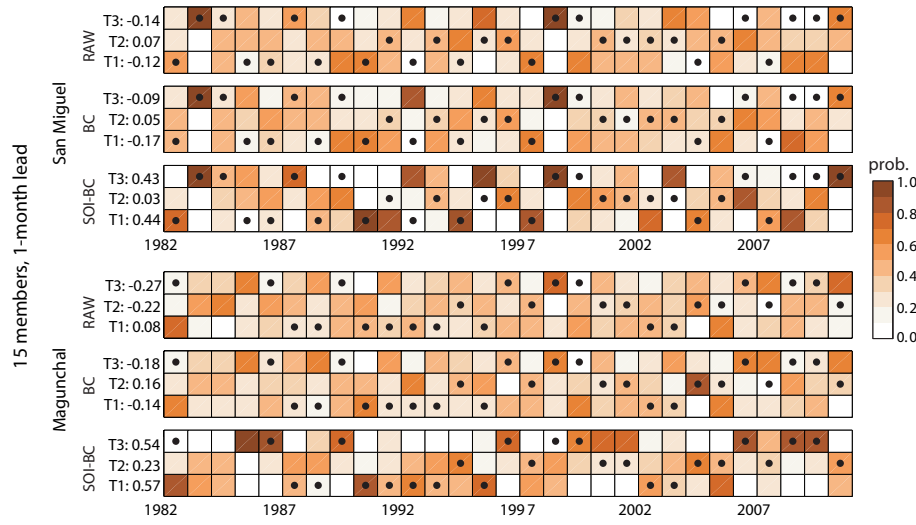


Fig. 6 Tercile plots for the raw System4 output —for the illustrative case of one-month lead forecasts from the 15-member experiment— and the two BC methods applied (from top to bottom), in San Miguel and Magunchal. Each plot displays, year by year, the predicted probabilities (white-to-orange colored scale) for each of the three terciles (T1, T2 and T3), along with the observed tercile (black circles). Numbers on the left indicate the ROCSS for each tercile.

Figure 8 shows the reliability diagrams obtained for the dry and wet terciles (left and right column, respectively) for cluster 1 and 2 —note that the joined series of the different stations falling within each cluster are considered.— Black, blue and red lines correspond to the raw model precipitation (the 15-member experiment), the unconditioned and the SOI-conditioned BC methods, respectively. Note that, for a perfectly reliable forecasting system, the curve obtained would match the diagonal. Points falling within the so-called skill region (in gray), i.e., the region contained between the no-resolution line (which indicates the expected frequency of the event: 1/3 for terciles) and the no-skill line (halfway between the no-resolution line and the diagonal) positively contribute to the forecast skill (Brier Skill Score > 0). With the exception of the wet tercile in cluster 1, the results found for this metric are in agreement with those obtained for the ROCSS, with the SOI-BC method improving the poor reliability (closeness to the diagonal) exhibited by both the raw model predictions and the BC method (which provide very similar results), either at one or four months lead-time (top and bottom panel, respectively). Aside from the lack of reliability, raw model forecasts and the standard BC method also show a lack of resolution (closeness to the horizontal climatological line). To some extent, this limitation can be also overcome by the SOI-BC method. Note the importance of these results, since reliable and resolute seasonal forecasts are essential for any forecast-based decision-making (Weisheimer and Palmer, 2014).

In order to assess the effect of the ensemble size on these results, all the above validation (ROCSS and reliability) was also performed for the case of the annual experiment of the System4, for which 51 members were available. The results

341 obtained for this larger ensemble were very similar for the two metrics considered
 342 (not shown for brevity).

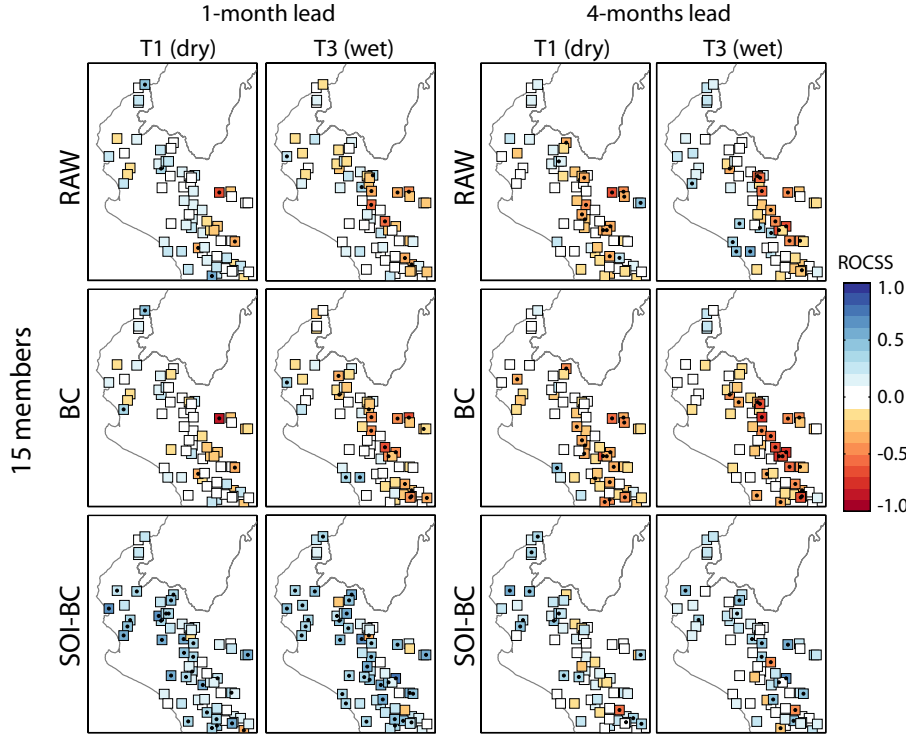


Fig. 7 ROCSS obtained for the dry and wet terciles (in columns) for raw System4 precipitation and the two BC methods applied (in rows). Left (right) panel corresponds to the case of one-month (four-months) lead predictions from the 15-member experiment of the System4.

343 An explanation of the above results comes from the fact that although the
 344 seasonal forecasting model is able to accurately forecast the interannual variabil-
 345 ity of ENSO (Figure 3), it fails to predict and reproduce the local impact of this
 346 phenomenon over the region under study. For instance, for the case of one-month
 347 lead predictions from the 15-member experiment, Figure 9 shows the interannual
 348 correlation between the observed SOI and the predicted local precipitation (left).
 349 The model fails to reproduce the observed SOI teleconnections (shown in Figure
 350 2(a)). In particular, System4 does not differentiate between those stations
 351 positively correlated with SOI (cluster 2) and those not significantly correlated
 352 (cluster 1), exhibiting a uniform pattern of negative correlations over the entire
 353 area of study. Moreover, this same pattern is also returned by the unconditioned
 354 BC method (center) since, as discussed, it does not significantly alter the temporal
 355 structure of the raw model output and, therefore, the correlation with SOI is pre-
 356 served in the corrected series. Differently, as a result of the changes in the temporal
 357 structure introduced by the SOI-BC method, part of the observed teleconnections
 358 is properly recovered in this conditioned implementation, although reinforced for

the stations in cluster 1 (right). Note that this is due to the simple conditioning setup used in this work (we only considered three categories of the SOI), which may not be the optimum solution for the stations in cluster 1, where only the strongest El Niño events have an impact in local precipitation (see, for instance, the case of San Miguel in Figure 2(b)).

4 Conclusions

In this work we have assessed the suitability of a first simple attempt for process-conditioned bias correction in the context of seasonal forecasting. To do this, we have focused on the northwestern part of Peru and bias corrected one- and four-month lead seasonal forecasts of boreal winter (DJF) precipitation from the ECMWF System4 (15- and 51-member experiments have been analyzed) for the period 1981-2010. With the aim of including some information about the underlying large-scale circulation, we have introduced an empirical quantile-quantile mapping which runs conditioned on the state of the Southern Oscillation Index (SOI). In this method, for each test year, the quantile-quantile mapping is trained using only those years for which the observed SOI lied in the same category (terciles are used here) the predicted SOI fell into —SOI is used here as a proxy for ENSO, which is known to strongly affect the climate of the study region.— This SOI-conditioned method was compared against a standard unconditioned implementation in which it was directly applied over the entire period of study.

Our results show that the unconditioned method broadly preserves the temporal structure of the raw model precipitation and, as a consequence, does not improve its unskillful predictions (beyond correcting the mean biases). Contrarily, the SOI-conditioned version can modify the temporal sequence of the raw model output, providing more realistic local time-series, and yielding better ROC Skill Scores (ROCCS) and reliability over the entire study area. Nevertheless, despite this general skill improvement, it is important to note that this conditioned method should not be expected to properly capture precipitation due to small-scale processes (e.g. convection) or other local features which are not directly related to the SOI and, therefore, not taken into account in this implementation.

The results obtained in this work suggest the potential usefulness of this new, SOI-conditioned method, especially for those regions where the local climate variability is largely driven by SOI/ENSO. Noticeably, similar approaches as the one applied here could be also valuable for other regions of the world affected by other large-scale phenomena (e.g. teleconnection patterns), as long as the models used are good at reproducing such phenomena, but do fail in predicting the associated local variability. Further investigation on the application of process-conditioned BC methods for those regions is still needed.

Acknowledgements This work has received funding from the MULTI-SDM project (MINECO/FEDER, CGL2015-66583-R). The authors are grateful to SENAMHI for the observational data, which are publicly available from <http://www.senamhi.gob.pe/?p=data-historica>, and to the European Center for Medium-Range Weather Forecast (ECMWF), for the access to the System4 seasonal forecasting hindcast. Also, the authors want to acknowledge the two anonymous reviewers for their useful comments, which have 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
- Barnston AG, Tippett MK, van den Dool HM, Unger DA (2015) Toward an improved multimodel ENSO prediction. *Journal of Applied Meteorology and Climatology* 54(7):1579–1595, DOI 10.1175/JAMC-D-14-0188.1
- Bazo J, Lorenzo MN, Porfirio da Rosa R (2013) Relationship between monthly rainfall in NW Peru and tropical Sea Surface Temperature. *Advances in Meteorology* (152875), DOI 10.1155/2013/152875, URL <https://www.hindawi.com/journals/amete/2013/152875/>
- Bellprat O, Kotlarski S, Lüthi D, Schär C (2013) Physical constraints for temperature biases in climate models. *Geophysical Research Letters* 40(15):4042–4047, DOI 10.1002/grl.50737
- 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, URL <http://www.sciencedirect.com/science/article/pii/S0921818106002748>
- Doblas-Reyes FJ, Coelho CAS, Stephenson DB (2008) How much does simplification of probability forecasts reduce forecast quality? *Meteorological Applications* 15(1):155–162, DOI 10.1002/met.50, URL <http://dx.doi.org/10.1002/met.50>
- Doblas-Reyes FJ, García-Serrano J, Lienert F, Biescas AP, Rodrigues LRL (2013) Seasonal climate predictability and forecasting: Status and prospects. *Wiley Interdisciplinary Reviews: Climate Change* 4(4):245–268, DOI 10.1002/wcc.217
- Ehret U, Zehe E, Wulfmeyer V, Warrach-Sagi K, Liebert J (2012) HESS Opinions: "Should we apply bias correction to global and regional climate model data?". *Hydrology and Earth System Sciences Discussions* 9(4):5355–5387, DOI 10.5194/hessd-9-5355-2012
- Emerton R, Cloke HL, Stephens EM, Zsoter E, Woolnough SJ, Pappenberger F (2017) Complex picture for likelihood of ENSO-driven flood hazard. *Nature Communications* 8:14,796, DOI 10.1038/ncomms14796, URL <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC5355947>
- Garreaud RD (2009) The Andes climate and weather. *Advances in Geosciences* 22:3–11, DOI 10.5194/adgeo-22-3-2009, URL <http://www.adv-geosci.net/22/3/2009>
- Gneiting T, Raftery AE, Westveld AH, Goldman T (2005) Calibrated probabilistic forecasting using Ensemble Model Output Statistics and minimum CRPS estimation. *Monthly Weather Review* 133(5):1098–1118, DOI 10.1175/MWR2904.1
- Horel JD, Cornejo-Garrido AG (1986) Convection along the coast of northern Peru during 1983: Spatial and temporal variation of clouds and rainfall. *Monthly Weather Review* 114(11):2091–2105
- IPCC (2015) Report of the Intergovernmental Panel on Climate Change workshop on regional climate projections and their use in impacts and risk analysis studies. University of Bern

- Johansson A (2007) Prediction skill of the NAO and PNA from daily to seasonal time scales. *Journal of Climate* 20(10):1957–1975, DOI 10.1175/JCLI4072.1, URL <https://doi.org/10.1175/JCLI4072.1>, <https://doi.org/10.1175/JCLI4072.1>
- Lachenbruch PA, Mickey MR (1968) Estimation of error rates in discriminant analysis. *Technometrics* 10(1):1–11, DOI 10.2307/1266219, URL <http://www.jstor.org/stable/1266219>
- Landman WA, Tennant WJ (2000) Statistical downscaling of monthly forecasts. *International Journal of Climatology* 20(13):1521–1532, DOI 10.1002/1097-0088(20001115)20:13<1521::AID-JOC558>3.0.CO;2-N, URL [http://onlinelibrary.wiley.com/DOI/10.1002/1097-0088\(20001115\)20:13<1521::AID-JOC558>3.0.CO;2-N/abstract](http://onlinelibrary.wiley.com/DOI/10.1002/1097-0088(20001115)20:13<1521::AID-JOC558>3.0.CO;2-N/abstract)
- Manzanas R, Frías MD, Cofiño AS, Gutiérrez JM (2014) Validation of 40 year multimodel seasonal precipitation forecasts: The role of ENSO on the global skill. *Journal of Geophysical Research: Atmospheres* 119(4):1708–1719, DOI 10.1002/2013JD020680, URL <http://onlinelibrary.wiley.com/doi/10.1002/2013JD020680/abstract>
- Manzanas R, Gutiérrez JM, Fernández J, van Meijgaard E, Calmanti S, Magariño ME, Cofiño AS, Herrera S (2017a) Dynamical and statistical downscaling of seasonal temperature forecasts in Europe: Added value for user applications. *Climate Services* DOI 10.1016/j.cliser.2017.06.004, URL <http://www.sciencedirect.com/science/article/pii/S2405880717300067>
- Manzanas R, Lucero A, Weisheimer A, Gutiérrez JM (2017b) Can bias correction and statistical downscaling methods improve the skill of seasonal precipitation forecasts? *Climate Dynamics* 50(3):1161–1176, DOI 10.1007/s00382-017-3668-z, URL <https://link.springer.com/article/10.1007/s00382-017-3668-z>
- Maraun D (2013) Bias correction, quantile mapping, and downscaling: Revisiting the inflation issue. *Journal of Climate* 26(6):2137–2143, DOI 10.1175/JCLI-D-12-00821.1, URL <http://dx.doi.org/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
- Maraun D, Huth R, Gutiérrez JM, Martn DS, Dubrovsky M, Fischer A, Hertig E, Soares PMM, Bartholy J, Pongrcz R, Widmann M, Casado MJ, Ramos P, Bedia J (2017a) The VALUE perfect predictor experiment: Evaluation of temporal variability. *International Journal of Climatology* DOI 10.1002/joc.5222, URL <http://onlinelibrary.wiley.com/doi/10.1002/joc.5222/abstract>
- Maraun D, Shepherd TG, Widmann M, Zappa G, Walton D, M GJ, Hagemann S, Richter I, Soares PMM, Hall A, Mearns LO (2017b) Towards process-informed bias correction of climate change simulations. *Nature Climate Change* 7:764–773, DOI 10.1038/nclimate3418
- Mason SJ, Graham NE (2002) Areas beneath the relative operating characteristics (ROC) and relative operating levels (ROL) curves: Statistical significance and interpretation. *Quarterly Journal of the Royal Meteorological Society* 128:2145–2166, DOI 10.1256/003590002320603584
- Molteni F, Stockdale T, Balmaseda M, Balsamo G, Buizza R, Ferranti L, Magnusson L, Mogensen K, Palmer T, Vitart F (2011) The new ECMWF seasonal forecast system (System 4). European Centre for Medium-Range Weather Forecasts, URL http://climate.ncas.ac.uk/people/allan/Fire_Risk_Insurance_Papers/Molteni%20etal%202011.pdf

- Rau P, Bourrel L, Labat D, Melo P, Dewitte B, Frappart F, Lavado W, Felipe O (2017) Regionalization of rainfall over the Peruvian Pacific slope and coast. *International Journal of Climatology* 37(1):143–158, DOI 10.1002/joc.4693, URL <http://dx.doi.org/10.1002/joc.4693>
- Sanabria J, Bourrel L, Dewitte B, Frappart F, Rau P, Solís O, Labat D (2017) Rainfall along the coast of Peru during strong El Niño events. *International Journal of Climatology* pp n/a–n/a, DOI 10.1002/joc.5292, URL <http://dx.doi.org/10.1002/joc.5292>
- Sinha P, Mohanty UC, Kar SC, Dash SK, Robertson AW, Tippett MK (2013) Seasonal prediction of the indian summer monsoon rainfall using Canonical Correlation Analysis of the NCMRWF global model products. *International Journal of Climatology* 33(7):1601–1614, DOI 10.1002/joc.3536, URL <http://onlinelibrary.wiley.com/DOI/10.1002/joc.3536/abstract>
- von Storch H, Zorita E, Cubasch U (1993) Downscaling of global climate change estimates to regional scales: An application to Iberian rainfall in wintertime. *Journal of Climate* 6(6):1161–1171, DOI 10.1175/1520-0442(1993)006<1161:DOGCCE>2.0.CO;2
- Sulca J, Vuille M, Silva Y, Takahashi K (2016) Teleconnections between the Peruvian Central Andes and Northeast Brazil during extreme rainfall events in austral summer. *Journal of Hydrometeorology* 17(2):499–515, DOI 10.1175/JHM-D-15-0034.1, URL <https://doi.org/10.1175/JHM-D-15-0034.1>
- Sulca J, Takahashi K, Espinoza JC, Vuille M, Lavado Casimiro W (2017) Impacts of different ENSO flavors and tropical Pacific convection variability (ITCZ, SPCZ) on austral summer rainfall in South America, with a focus on Peru. *International Journal of Climatology* pp n/a–n/a, DOI 10.1002/joc.5185
- 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
- Torralba V, Doblas-Reyes FJ, MacLeod D, Christel I, Davis M (2017) Seasonal climate prediction: A new source of information for the management of wind energy resources. *Journal of Applied Meteorology and Climatology* 56(5):1231–1247, DOI 10.1175/JAMC-D-16-0204.1
- Trenberth KE (1984) Signal versus noise in the Southern Oscillation. *Monthly Weather Review* 112(2):326–332, DOI 10.1175/1520-0493(1984)112<0326:SVNITS>2.0.CO;2
- Vrac M, Friederichs P (2014) Multivariate -intervariable, spatial, and temporal- bias correction. *Journal of Climate* 28(1):218–237, DOI 10.1175/JCLI-D-14-00059.1
- Ward JH (1963) Hierarchical grouping to optimize an objective function. *Journal of the American Statistical Association* 58(301):236–244, URL <http://www.jstor.org/stable/2282967>
- Weisheimer A, Palmer TN (2014) On the reliability of seasonal climate forecasts. *Journal of the Royal Society Interface* 11(96), DOI 10.1098/rsif.2013.1162
- Wetterhall F, Pappenberger F, He Y, Freer J, Cloke HL (2012) Conditioning model output statistics of regional climate model precipitation on circulation patterns. *Nonlinear Processes in Geophysics* 19(6):623–633, DOI 10.5194/npg-19-623-2012, URL <http://www.nonlin-processes-geophys.net/19/623/2012/>

- White RH, Toumi R (2013) The limitations of bias correcting regional climate model inputs. *Geophysical Research Letters* 40(12):2907–2912, DOI 10.1002/grl.50612
- Zhai P, Yu R, Guo Y, Li Q, Ren X, Wang Y, Xu W, Liu Y, Ding Y (2016) The strong El Niño of 2015/16 and its dominant impacts on global and China’s climate. *Journal of Meteorological Research* 30(3):283–297, DOI 10.1007/s13351-016-6101-3
- Zhao T, Bennett JC, Wang QJ, Schepen A, Wood AW, Robertson DE, Ramos MH (2017) How suitable is quantile mapping for postprocessing GCM precipitation forecasts? *Journal of Climate* 30(9):3185–3196, DOI 10.1175/JCLI-D-16-0652.1
- Zheng Z, Hu ZZ, L’Heureux M (2016) Predictable components of ENSO evolution in real-time multi-model predictions. *Scientific Reports* 6:35,909, DOI 10.1038/srep35909, URL <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC5075933/>

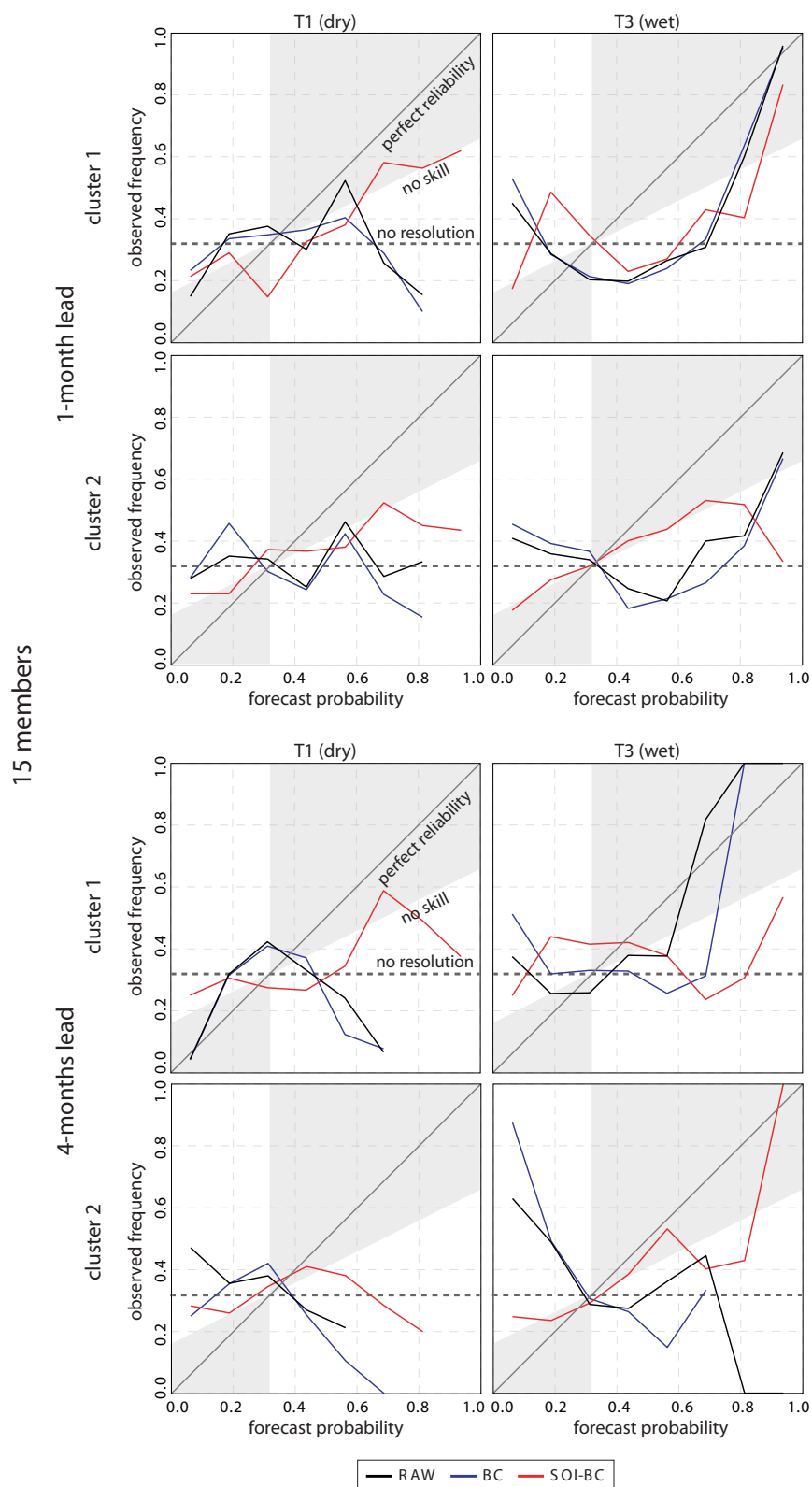


Fig. 8 Reliability diagrams for the dry and wet terciles (left and right column, respectively) for cluster 1 and 2 (top and bottom rows). Black/blue/red line corresponds to the raw model precipitation/BC/SOI-BC. Top (bottom) panel corresponds to the case of one-month (four-months) lead predictions from the 15-member experiment of the System4.

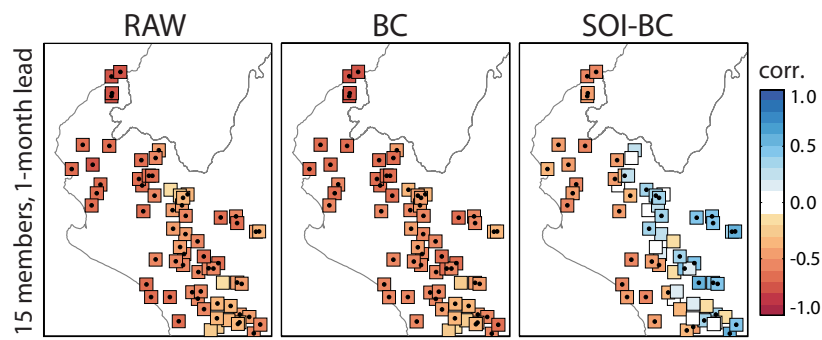


Fig. 9 Interannual Spearman correlation between the observed SOI and the raw and corrected model precipitation (from left to right), at one month lead-time. In all cases, the ensemble mean resulting from the 15-member experiment of the System4 is considered. Significant ($\alpha = 0.05$) values are marked with a black dot.