

## Towards process-informed bias correction of climate change simulations

Maraun, Douglas; Shepherd, Theodore G. ; Widmann, Martin; Zappa, Giuseppe ; Walton, Daniel ; Gutiérrez, José M. ; Hagemann, Stefan; Richter, Ingo ; Soares, Pedro M. M.; Hall, Alex ; Mearns, Linda O.

DOI:  
[10.1038/nclimate3418](https://doi.org/10.1038/nclimate3418)

License:  
None: All rights reserved

*Document Version*  
Peer reviewed version

*Citation for published version (Harvard):*  
Maraun, D, Shepherd, TG, Widmann, M, Zappa, G, Walton, D, Gutiérrez, JM, Hagemann, S, Richter, I, Soares, PMM, Hall, A & Mearns, LO 2017, 'Towards process-informed bias correction of climate change simulations', *Nature Climate Change*, vol. 7, pp. 764–773. <https://doi.org/10.1038/nclimate3418>

[Link to publication on Research at Birmingham portal](#)

**Publisher Rights Statement:**  
Published online via [10.1038/nclimate3418](https://doi.org/10.1038/nclimate3418)

### General rights

Unless a licence is specified above, all rights (including copyright and moral rights) in this document are retained by the authors and/or the copyright holders. The express permission of the copyright holder must be obtained for any use of this material other than for purposes permitted by law.

- Users may freely distribute the URL that is used to identify this publication.
- Users may download and/or print one copy of the publication from the University of Birmingham research portal for the purpose of private study or non-commercial research.
- User may use extracts from the document in line with the concept of 'fair dealing' under the Copyright, Designs and Patents Act 1988 (?)
- Users may not further distribute the material nor use it for the purposes of commercial gain.

Where a licence is displayed above, please note the terms and conditions of the licence govern your use of this document.

When citing, please reference the published version.

### Take down policy

While the University of Birmingham exercises care and attention in making items available there are rare occasions when an item has been uploaded in error or has been deemed to be commercially or otherwise sensitive.

If you believe that this is the case for this document, please contact [UBIRA@lists.bham.ac.uk](mailto:UBIRA@lists.bham.ac.uk) providing details and we will remove access to the work immediately and investigate.



18 Climate scientists are confronted with a growing pressure to support adaptation decisions  
19 and face the dilemma of operationalising what is still foundational research<sup>1,2</sup>. The models  
20 often used to inform adaptation decisions - global coupled atmosphere ocean general circula-  
21 tion models (GCMs), potentially downscaled with regional climate models (RCMs) - have  
22 horizontal resolutions often far coarser than those demanded, and suffer from substantial bi-  
23 ases<sup>3,4</sup>. To reduce biases and to overcome the scale gap between the numerical model grid and  
24 the desired scale, climate model output is almost routinely post-processed by bias correction  
25 (often called bias adjustment) methods. A vast number of bias corrected national and global  
26 climate change projections has been published<sup>5,6,7,8,9,10,11,12,13</sup>, has served as input for impact  
27 studies<sup>14,15,10,16</sup> as well as assessment reports<sup>17,18,19</sup>, and has been made available through data  
28 portals<sup>20,21,13</sup>. A wide variety of bias correction methods is in use, ranging from simple adjust-  
29 ments of the mean to flexible, potentially multivariate, quantile mapping approaches<sup>22,23,24</sup>.  
30 Yet many problems related to bias correction have been identified<sup>25,8,26,27,28,29</sup>. Thus, even  
31 though bias correction is often considered a necessary step in climate impact modelling<sup>24</sup>, the  
32 approach is prone to misuse and best practice still needs to be established<sup>30</sup>. Some authors  
33 even question the very basis of bias correction<sup>31</sup>.

34 Current developments on bias correction have largely focused on improving statistical  
35 methodology: to better match variability and extremes<sup>24,32,33,34</sup>, the dependence between  
36 different climatic variables<sup>35,36</sup>, the location of features<sup>37</sup>, or to retain simulated trends<sup>6,32,11</sup>.  
37 This focus has ignored a major issue: a key requirement of climate model projections is  
38 credibility<sup>38,1,2</sup>. Here, we argue that current bias correction methods might improve the  
39 applicability of climate simulations, but in general cannot improve low model credibility.  
40 Indeed, bias correction may hide a lack of credibility or may even reduce credibility. The way  
41 bias correction is often applied and evaluated might ultimately lead to ill-informed adaptation  
42 decisions.

43 We start from the basic reason underlying the demand to bias correct: all models are  
44 substantial simplifications of a real system. Climate models are based on physical laws such  
45 as conservation of energy, mass and momentum, thermodynamic and radiation laws. But  
46 models have a limited spatial resolution, their topography is coarse and they will never re-  
47 solve nor represent all relevant processes from planetary waves down to turbulence. Sub-grid  
48 processes are simplified by parameterisations. As a consequence, many relevant atmospheric,  
49 oceanic and coupled processes are not realistically represented, with knock-on effects on other  
50 processes even far away from where the primary biases occur<sup>39</sup>. Biases in basic quantities  
51 such as mean and variance are therefore commonplace, even for something as fundamental  
52 as global-mean surface temperature<sup>3</sup>. Often, a realistic behaviour is only achieved by tuning  
53 the model<sup>3</sup>. In short, climate model biases are severe enough to in principle justify the use  
54 of bias correction techniques to render model output more useful for impact studies.

55 We therefore argue that bias correction should not be dismissed, but that a solid conceptual  
56 and process understanding of climate model biases is required to successfully apply bias  
57 correction. The extent to which biases can be mitigated by post-processing depends on  
58 their origin. We present several examples, discuss their correctability by state-of-the-art bias  
59 correction methods, and propose alternative approaches and future directions of research.

# 1 Bias correction in a nutshell

We define a bias as the systematic difference between a modelled property of the climate system and the corresponding real property<sup>40,41,25,31,42,43</sup>. Such properties could be mean temperature, variance or a 100-year return value. The term “systematic” refers to all differences that are not due to sampling uncertainty. Biases are typically assumed to be time-independent<sup>44,45,23,11</sup>, but in principle may vary in time<sup>41,40,25,42</sup>. Some authors define a bias as the time independent error component of a model<sup>24,46,47</sup>. The problems we discuss below occur irrespective of the specific bias definition.

As bias correction we consider all methods that calibrate an empirical transfer function between simulated and observed distributional parameters, and apply this transfer function to output simulated by the considered model. Bias correction according to this definition is a mere post-processing.

We focus on two different types of methods which are broadly representative of those commonly used: a simple adjustment of the mean, and quantile mapping. A simple mean bias correction would estimate a bias as the difference (or ratio for, e.g., precipitation) between simulated and observed mean over a reference period, and adjust the simulated time series over a scenario period by the estimated bias (by subtracting it, or rescaling). Quantile mapping individually adjusts each quantile. The transfer functions are then applied to climate change simulations under the assumption that they are time-invariant.

Bias correction relies on observational reference data, which should in many cases be considered a model product themselves. This holds true in particular for gridded data sets. Related issues are an important topic for bias correction, but are outside the scope of this article.

## 2 The evaluation problem

[Figure 1 about here.]

To begin with, we demonstrate the difficulties to evaluate the performance of bias correction. The evaluation of statistical models, e.g., in weather forecasting, is generally done by cross-validation: the model is calibrated to a subset of the available data only, the evaluation is carried out by assessing the prediction of the remaining (independent) data. Cross-validation is widely used for establishing skill of bias correction, often only for calibrated properties of the marginal distribution<sup>6,47,48,23,49</sup> (some exceptions evaluate temporal or spatial dependence<sup>24,27</sup>). Here we demonstrate that such an evaluation is not suitable to establish bias correction skill.

Consider the rather absurd setting of bias correcting simulated daily temperature from the Southern Ocean against observed daily precipitation over central Europe during boreal winter. The corresponding model grid boxes are simply taken from the exact opposite side of the globe. Whereas the temperature field over the Southern Ocean (mapped onto Europe) is very smooth (Fig. 1a,d), precipitation in Europe has a distinct pattern controlled by distance to sea and orography (b,e). But even though modelled temperature and observed

99 precipitation fields are essentially unrelated and both fields show different long-term changes,  
100 the quantile mapping looks reasonable for the validation period, for mean and high values  
101 (c,f). The residual bias (g) between corrected model and observation purely stems from the  
102 different trends in both regions. The problem is especially severe for non-parametric quantile  
103 mapping, as demonstrated for the grid box enclosing Venice (h): even though the tempera-  
104 ture and precipitation distributions have completely different shapes, and both distributions  
105 change substantially over time (mean precipitation +28%, mean temperature -0.29K in the  
106 corresponding Southern Ocean grid box), the QQ plot looks reasonable also for the validation  
107 period. In other words: cross-validation of calibrated climatological properties is not able  
108 to identify the absurdity of the chosen example, and is thus not sufficient to evaluate the  
109 performance of bias correction. The reason for the failure is that, in climate modelling, model  
110 and observations are not in synchrony and predictive skill cannot, as in weather forecasting,  
111 be established by cross-validation<sup>26</sup>. The evaluation is restricted to long term distributional  
112 aspects only, and provided the sampling is adequate, cross-validation will merely reproduce  
113 the long-term distribution. But in a non-synchronous setting it is still possible to evaluate  
114 non-calibrated aspects, in particular for the temporal and, if required, spatial dependence  
115 structure. Such an evaluation would yield essential and indispensable information about the  
116 appropriateness of a bias correction.

### 117 3 Bias correction under present conditions

118 [Figure 2 about here.]

119 Bias correction may introduce artefacts already for present climate conditions which are  
120 invisible to an evaluation of marginal distributional properties. As example, consider correc-  
121 tions of the drizzle effect, i.e., the fact that climate models often simulate too high a number  
122 of wet days with very low intensities. Quantile mapping adjusts the number of wet days by  
123 changing the least wet days into dry days. The adjustment in turn improves the representa-  
124 tion of dry spells of typically up to about 20 days<sup>50</sup>. But climate models have considerable  
125 deficiencies in representing temporal variability beyond the drizzle effect. Dry spells are often  
126 too short, e.g., because the persistence of blocking highs is typically underrepresented<sup>51</sup>, or  
127 because a dry valley may be represented as an exposed location by a typical climate model  
128 with coarse topography. Whereas the drizzle effect may indeed be correctable, an attempt to  
129 correct other, more fundamental errors in the spell length distribution may result in unwanted  
130 artefacts (Fig. 2). In many cases one may simply miss the long spells (a), in some cases one  
131 may by chance even combine short spells into long ones and therefore improve the overall  
132 spell length distribution (b). But in a substantial amount of cases, the wet-day adjustment  
133 might either produce too many short and medium-length spells (c) or even too long spells  
134 (d). This example highlights that bias correction is not a one-size-fits-all approach, but needs  
135 to be user-tailored: is the overall wet-day probability relevant or the representation of spell  
136 lengths? A careful decision needs to be drawn, and a sensible adjustment carried out. Other  
137 examples, where attempts to bias correct temporal structure might cause severely misleading  
138 results, are the diurnal cycle of precipitation or the onset of the rainy season<sup>8</sup>.

[Figure 3 about here.]

140 Bias correction may further be infeasible if the climate model variable does not capture  
 141 the relevant regional processes. Consider a GCM that simulates reasonable ENSO variability,  
 142 but does not reproduce the clustering of extreme precipitation in Peru during El Niño events  
 143 (Fig. 3, top and middle left panels). Quantile mapping trivially adjusts the distributions (right  
 144 panel), but still the result is meaningless as the wrong clustering is not improved (bottom  
 145 left panel). In this example, already a visual inspection of the resulting time series uncovers  
 146 the bias correction problem. When evaluating many grid boxes, an evaluation conditional on  
 147 El Niño events might be required. A similar representativeness problem may be caused by a  
 148 coarse model topography, which may act as an unrealistically strong meteorological divide<sup>28</sup>.

[Figure 4 about here.]

150 In many cases bias correction is used to downscale to a finer spatial resolution<sup>5,48,49,35,15,12</sup>.  
 151 Current approaches, however, are unable to generate unexplained subgrid day-to-day variabil-  
 152 ity and may even introduce artefacts, e.g., in the representation of extreme precipitation<sup>27</sup>.  
 153 But similar effects might also occur for temperature fields in complex terrain. Consider tem-  
 154 perature inversions, a common feature in the Central Valley, California (Fig. 4). A bias cor-  
 155 rected GCM will trivially reproduce the climatological temperature difference of 2 K between  
 156 a location in the valley and a nearby location higher up in the Sierra Nevada. But whereas  
 157 the actual day-to-day temperature difference has a broad distribution - with negative values  
 158 indicating inversions - the bias corrected difference is essentially constant (it varies slightly  
 159 because quantile mapping corrects different quantiles individually). Stochastic approaches  
 160 explicitly modelling unexplained sub-grid variability may thus be required in complex terrain  
 161 or for highly variable fields.

## 162 4 Bias correction under climate change conditions

163 Some artefacts of bias correction may only appear under changing climatic conditions and  
 164 may thus be invisible to evaluation against present observations.

165 One cause of such artefacts are GCMs biases in the large-scale atmospheric circulation<sup>52,53</sup>,  
 166 which themselves result from an insufficient resolution of the atmospheric model<sup>54</sup>, a coarse  
 167 topography<sup>55,56</sup> or from biases in the underlying sea surface temperature<sup>57,58,59</sup>. For instance,  
 168 over Europe the North Atlantic winter storm track is too zonal in most models and crosses  
 169 Europe too far south<sup>53</sup>. Such biases exert a strong control on regional climate<sup>26,60</sup>. They are  
 170 inherited by downscaling and are reflected in regional biases<sup>61</sup>.

[Figure 5 about here.]

172 It has been argued that biases in surface weather resulting from circulation biases cannot be  
 173 bias corrected<sup>26,30</sup>. For instance, when the frequency of circulation types is misrepresented,  
 174 bias correction may increase biases for specific circulation types<sup>29</sup>. Here we further show that

175 bias correction in the presence of substantial circulation biases may induce implausible future  
176 signals.

177 Consider precipitation projections based on a GCM with a substantial southward bias  
178 of the Atlantic storm track, such that the maximum of present day winter precipitation in  
179 Western Europe is shifted southwards by about  $20^\circ$  (Fig. 5 top). The GCM simulates a north-  
180 ward shift of the storm track. A mean bias correction of winter precipitation will perfectly  
181 align simulated present-day mean precipitation with observations, by damping precipitation  
182 over Southern Europe, and amplifying it over Central and Northern Europe. Applying this  
183 correction to the future simulation, however, the northward shift of the uncorrected precipita-  
184 tion peak - indicating a northward shift of the storm track - is transformed into a southward  
185 precipitation shift.

186 In other words: in the presence of major circulation biases, bias correction - even though  
187 the local climate change signal is preserved - might create implausible patterns of surface  
188 climate change. Such problems can be avoided by a careful climate model selection: for a  
189 GCM with a lower circulation bias, the precipitation bias correction preserves the northward  
190 precipitation shift consistent with the storm track shift (Fig. 5 right bottom panel).

191 Two approaches have been suggested to correct atmospheric circulation biases. First,  
192 to bias correct GCM fields prior to dynamical downscaling<sup>62</sup>; and second to spatially shift  
193 simulated fields<sup>37</sup>. Both approaches, trivially, correct biases in the climatological atmospheric  
194 fields. The first approach, however, introduces inconsistencies in the atmospheric dynamics:  
195 for instance, individual storms are - in the GCM - still generated at the wrong position of the  
196 polar front and then - in the RCM - interact with the corrected climatological polar front. The  
197 second approach ignores that the simulated position of circulation features is intricately linked  
198 to the model orography, simulated land-sea contrasts and sea surface temperature biases, and  
199 thus introduces inconsistencies with these model properties.

200 Another cause of artefacts is the modification of the climate change signal by variance-  
201 adjusting bias correction methods<sup>8,27,63</sup>. A debate has arisen whether these trend modifica-  
202 tions might actually improve or deteriorate the raw climate change signal<sup>40,64</sup>, and several  
203 trend preserving bias correction approaches have been developed<sup>32,11,65,66</sup>. We argue that this  
204 issue cannot be resolved based on purely statistical arguments. Again, one needs to refer to  
205 process understanding.

206 Obviously, a credibly simulated trend should not be altered by any postprocessing. In  
207 such a case, the assumption of a time invariant correction is fulfilled and a trend preserving  
208 bias correction is the method of choice. Often, however, climate model biases depend on the  
209 actual state of the climate system<sup>41,25,67</sup>, so in a changing climate they are not time-invariant.  
210 Two questions arise: first, in what situations are climate model trends implausible? And  
211 second, in which situations could bias correction methods like quantile mapping potentially  
212 improve such trends?

213 Many cases have been identified where climate models may simulate implausible changes  
214 of large-scale climatic phenomena, because the underlying processes are not realistically rep-  
215 resented. Prominent examples are the representation of ENSO feedbacks<sup>68,69</sup>, the Indian  
216 summer monsoon<sup>70,71,72</sup>, the influence of increased diabatic heating on the intensification of  
217 extratropical cyclones<sup>73</sup>, or European blocking<sup>51</sup>. Current bias correction methods will not

218 succeed in improving these changes, as they result from fundamental climate model errors<sup>30</sup>.  
219 At the regional scale, misrepresented land-surface interactions may result in implausi-  
220 ble climate change trends. For instance, models simulating unrealistically low summer soil  
221 moisture tend to over-represent summer temperature increases<sup>74,75</sup>; similarly the simulated  
222 increase of spring temperature is tightly linked to snow-albedo feedback strength<sup>74</sup>. Further-  
223 more trends may be implausible as a result of inadequately parameterised sub-grid processes.  
224 For instance, there is evidence that the response of summer convective precipitation extremes  
225 to global warming is mis-represented by regional climate models with parameterised convec-  
226 tion<sup>76,77</sup>.

227 In such situations, it has been argued that quantile mapping may improve implausible  
228 trends<sup>40,64</sup>, because its correction is value-dependent: a simulated value of, say, 25°C will be  
229 adjusted with a specific correction irrespective of the actual state of the climate system, i.e.,  
230 in present and future climate. The distributions typically adjusted by quantile mapping are  
231 mostly spanned by day-to-day variability, which is mainly caused by the passage of different  
232 types of airmasses. Under climate change, the properties of airmasses themselves will change.  
233 If a temperature of 25°C corresponds to a rare, sunshiny day in present climate, such a  
234 temperature might correspond to an overcast rainy day in a warmer climate. It is thus  
235 conceivable that the value dependence of biases found for present day climate<sup>40</sup> might be  
236 different in the future. The same reasoning can be made from a time-scale point of view: as  
237 bias correction is calibrated on daily time scales, also the modification of the climate change  
238 signal stems from the rescaling of modelled day-to-day variability<sup>27,63</sup>. Therefore, a trend  
239 modification by quantile mapping can only be sensible if - in a given context - the transfer  
240 function calibrated on short time scales can sensibly be applied to correct biases on long time  
241 scales.

242 [Figure 6 about here.]

243 We illustrate this issue with spring temperature trends in mountaineous terrain. Consider  
244 again the example from California (Fig. 6). A GCM misses the complex topography of the  
245 region and thus simulates a rather smooth temperature field for present climate (a). Quantile  
246 mapping trivially produces the correct present temperature fields (b). Similarly, a high reso-  
247 lution RCM simulates a realistic temperature field (c). The RCM also simulates a plausible  
248 climate change signal which varies systematically across topography (f): at high elevations,  
249 the warming is amplified by the snow-albedo feedback. The climate change signal of the GCM,  
250 however, is again unrealistically smooth (d); no elevation dependent warming is produced.  
251 A trend preserving bias correction would fully inherit this implausible climate change signal.  
252 Standard quantile mapping modifies the large-scale changes, but in an unsystematic way (e).  
253 We do not know whether the RCM simulation is correct, but the preserved and bias corrected  
254 GCM signals are highly implausible.

255 Thus, bias correction is trapped in a fundamental dilemma: in situations where the driving  
256 model simulates a credible change a trend preserving bias correction<sup>32,11</sup> is a sensible choice.  
257 In many cases, however, we may have strong evidence that the simulated regional climate  
258 change is implausible - we would like to improve the change. Standard quantile mapping  
259 modifies simulated trends. But as discussed above and demonstrated for the snow albedo



260 feedback, we know that these modifications may not be physically justified. Here, one would  
261 have to assess the raw and modified changes on a case-by-case basis, referring to the relevant  
262 climatic processes and their model representation.

## 263 5 Ways Ahead

264 We presented examples of artefacts that may occur when bias correction is applied without  
265 considering the underlying processes. These examples illustrate that bias correction is only  
266 recommended if, in a given context, the following assumptions hold: first, relevant processes  
267 are reasonably well captured by the chosen climate models, including the temporal structure  
268 (Figure 2) and location (Figure 5) of the large-scale circulation, as well as the regional response  
269 to large-scale processes (Figure 3) and local feedbacks (Figure 6). Second, the climate models  
270 resolve the local spatial-temporal variability (Figure 4) and climate change (Figure 6). Over  
271 areas where some of these assumptions are not valid, the bias corrected output should be  
272 handled with great care. To avoid the related artefacts, we advocate research along four major  
273 strands. Process understanding should inform bias correction already during the climate  
274 model selection, as part of the actual bias correction procedure, when evaluating the correction  
275 and when shifting to alternative approaches.

### 276 5.1 Understanding Model Biases

277 Any regional climate projection that is intended to serve for decision making relies on a  
278 realistic simulation of all relevant processes controlling climate change. It has thus to be  
279 recognised that the appropriateness of a bias correction is only partly a statistical issue, but  
280 importantly an issue of the credibility of the driving model. Thus it is important to understand  
281 the origins of model biases, from the large-scale circulation to regional-scale forcings and  
282 feedbacks.

283 Emergent constraints<sup>78</sup> are a promising approach to understand the influence of model  
284 biases in present climate on the climate change signal. The essence of this approach is to  
285 identify strong statistical relationships between (1) an observable feature of the simulated  
286 present climate and (2) a future climate change signal in a large ensemble of climate models.  
287 If the statistical relationship is associated with robust physics, then the most realistic models  
288 in the present climate can be declared to have the most credible future climate change signal.  
289 Basically, emergent constraints allow one to determine which present climate biases are most  
290 consequential for future climate change signals. Emergent constraints have already been  
291 applied extensively to global-scale processes and feedbacks. However, there is no reason  
292 they cannot be applied to regional-scale processes, either in ensembles of global models or  
293 associated downscaled data products. Examples are the influence of location biases in the  
294 large-scale atmospheric circulation on regional precipitation changes<sup>79</sup>, or the influence of  
295 biases in snow-albedo feedbacks on the regional warming signal<sup>80</sup>. We advocate searching  
296 for emergent constraints along these lines at the regional scale. This technique would exploit  
297 regional biases to improve the credibility of future climate change signals, instead of trying  
298 to get rid of them in some unphysical way.

299 As discussed above, a key issue is also to understand the relationship of biases across  
300 time-scales: how do biases in day-to-day or interannual variability translate into biases in the  
301 climate change signal? Identifying such linkages may help to judge the feasibility of trend  
302 modifications.

303 Given that fundamental model errors cannot be corrected by bias correction<sup>30</sup>, we advocate  
304 for a region-targeted selection of the driving GCMs prior to any downscaling exercise. The  
305 aim of such a procedure would neither be to identify the overall best performing GCM, nor  
306 to discard models simulating biased surface variables. Rather, it would be to discard those  
307 GCMs that unrealistically simulate the processes controlling the regional climate of interest,  
308 and those that have strong location biases in the large-scale atmospheric circulation (see  
309 Figure 5). The selection of course has to account in some manner for the range of uncertainty  
310 in global climate sensitivity.

311 There is realistic hope that further model improvements and increased model resolution  
312 may improve the representation of both local and large-scale processes<sup>81,54,82,58,83</sup>. The re-  
313 sulting reduction in location biases and the increase in credibility of future projections will  
314 render subsequent bias correction a more defensible approach.

## 315 5.2 New Bias Correction Approaches

316 We identified two major limitations of current bias correction methods: their difficulties  
317 in downscaling to finer spatial scales, and their inability to improve the local climate change  
318 signal. To address both these issues, we advocate the development of new methods, combining  
319 advanced statistical modelling with physical understanding.

320 The downscaling problem requires stochastic approaches which generate sub-grid spatial  
321 variability: to simulate fine-scale precipitation fields, or to simulate sub-grid temperature  
322 variations such as inversions. Recently it has been proposed to carry out the bias correction  
323 at the grid-box scale, and then to stochastically downscale to finer scales<sup>84</sup>. More realistic  
324 fields can be obtained by including process information, e.g., by conditioning the downscaling  
325 on the atmospheric circulation<sup>29</sup>.

326 As laid out above, a misrepresentation of regional feedbacks may result in an implausible  
327 regional climate change signal, and quantile mapping will likely not be able to improve it.  
328 Avenues should be explored to explicitly account for regional-scale processes and feedbacks  
329 for improving the climate change signal in the statistical postprocessing. One such avenue is,  
330 again, process-based bias correction. For instance, summer temperature biases may depend  
331 on temperature because of soil moisture feedbacks. Here it has been suggested to condition  
332 the correction on simulated soil moisture<sup>67</sup>. Another avenue are emulators of high-resolution  
333 RCMs, which simulate a credible climate change signal. For instance, local variations in the  
334 warming signal could be statistically expressed by covariates such as elevation, continentality  
335 or large-scale warming patterns. These expressions can be calibrated across a range of dy-  
336 namically downscaled GCMs, and then applied to statistically downscale the climate change  
337 signal of other GCMs<sup>85</sup>. Such emulators could also be developed for other regional processes  
338 such as convection: measures of stability and moisture convergence could serve as input to  
339 emulate high-resolution convection permitting models. Thereby the representation of extreme

340 events could be improved, a weak point of essentially all statistical post-processing methods  
341 so far.

### 342 **5.3 Evaluating Bias Correction**

343 None of the artefacts we presented would have been identified by a standard cross-validation  
344 of marginal aspects. Rigorous standards for evaluating bias correction methods need thus  
345 be developed. These should encompass temporal as well as process-oriented aspects<sup>86</sup>. For  
346 instance, an investigation of the spell length distribution (Figure 2), or an evaluation condi-  
347 tional on the state of the relevant climatic phenomenon (Figure 3) may help to reveal bias  
348 correction problems. In any case, the resulting bias corrected time series should be - at least  
349 for some selected grid boxes - visually inspected and compared with observational data. A  
350 useful indicator for an unphysical bias correction is the dis-similarity between modelled and  
351 observed distribution (Figure 1): major differences point to a misrepresentation of key pro-  
352 cesses, and a bias correction is unlikely to be sensible. In any case one should investigate the  
353 projected signals for implausible change (Figures 5 and 6). The use of pseudo realities for  
354 evaluating simulated trends<sup>86</sup> should further be explored.

### 355 **5.4 Alternative approaches**

356 Finally, we advocate to explore alternative approaches in any given context. In some cases,  
357 perfect prognosis statistical downscaling and change factor weather generators<sup>22</sup> may be more  
358 appropriate than bias correction. In other cases, response surfaces<sup>87</sup> with qualitative input  
359 of possible climate changes might suffice to obtain decision relevant information, or expert  
360 knowledge combined with raw climate model simulations might provide useful information.  
361 Location biases of the atmospheric circulation may be reduced by surrogate climate warming  
362 studies<sup>88</sup>. Finally, storyline simulations of how single but relevant past events might look in  
363 a warmer future may substantially improve the representation of local feedbacks: they reduce  
364 computational costs and thereby enable much higher model resolutions<sup>89</sup>.

## 365 **6 Final Remarks**

366 Bias correction is not a Swiss Army knife, many issues remain unresolved, and research is  
367 needed to understand its limitations and to develop new concepts for mitigating the effects of  
368 climate model biases. Bias correction is not a purely statistical problem and cannot overcome  
369 fundamental deficiencies in climate models.

370 We recommend carrying out any bias correction or downscaling based on solid knowledge  
371 about the relevant climatic phenomena and the ability of the employed climate models to  
372 simulate them. To identify implausible results, a successful bias correction thus requires a  
373 close collaboration with global and regional climate modellers as well as experts both in the  
374 relevant large scale climatic phenomena and the local weather and climate of the target region.  
375 We recommend a concerted action among all involved disciplines to build up the necessary  
376 knowledge and to develop best practice guidelines to make bias correction a rigorous science.

377 In any case, it is essential to disclose relevant expert decisions affecting the results and to  
378 transparently discuss the usefulness and limitations of the output with users, in particular as  
379 the use of climate model data by non-experts is more and more operationalised by climate  
380 service providers<sup>2</sup>.

## 381 References

- 382 [1] B.C. Hewitson, J. Daron, R.G. Crane, M.F. Zermoglio, and C. Jack. Interrogating  
383 empirical-statistical downscaling. *Clim. Change*, 122:539–554, 2014.
- 384 [2] P. Adams, E. Eitland, B. Hewitson, C. Vaughan, R. Wilby, and S. Zebiak. Toward an  
385 ethical framework for climate services. A white paper of the climate services partnership  
386 working group on climate services ethics, 2015.
- 387 [3] G. Flato, J. Marotzke, B. Abiodun, P. Braconnot, S.C. Chou, W. Collins, P. Cox, F. Dri-  
388 ouech, S. Emori, V. Eyring, C. Forest, P. Gleckler, E. Guilyardi, C. Jakob, V. Kattsov,  
389 C. Reason, and M. Rummukainen. *Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovern-  
390 mental Panel on Climate Change*, chapter Evaluation of Climate Models. Cambridge  
391 University Press, Cambridge, United Kingdom and New York, NY, USA, 2013.
- 392 [4] S. Kotlarski, K. Keuler, O.B. Christensen, A. Colette, M. Déqué, A. Gobiet, K. Goergen,  
393 D. Jacob, D. Lüthi, E. van Meijgaard, G. Nikulin, C. Schär, C. Teichmann, R. Vautard,  
394 K. Warrach-Sagi, and V. Wulfmeyer. Regional climate modelling on European scales: A  
395 joint standard evaluation of the EURO-CORDEX RCM ensemble. *Geosci. Model. Dev. Discuss.*, 7:217–293, 2014.
- 396 [5] E.P. Maurer. Fine-resolution climate projections enhance regional climate change impact  
397 studies. *EOS*, 88(47):504, 2007.
- 398 [6] H. Li, J. Sheffield, and E.F. Wood. Bias correction of monthly precipitation and tem-  
399 perature fields from Intergovernmental Panel on Climate Change AR4 models using  
400 equidistant quantile matching. *J. Geophys. Res.*, 115:D10101, 2010.
- 401 [7] A. Dosio, P. Paruolo, and R. Rojas. Bias correction of the ENSEMBLES high resolution  
402 climate change projections for use by impact models: Analysis of the climate change  
403 signal. *J. Geophys. Res. Atmos.*, 117(D17), 2012.
- 404 [8] S. Hagemann, C. Chen, J.O. Haerter, J. Heinke, D. Gerten, and C. Piani. Impact of  
405 a statistical bias correction on the projected hydrological changes obtained from three  
406 GCMs and two hydrology models. *J. Hydrometeorol.*, 12(4):556–578, 2011.
- 407 [9] A.M.K Stoner, K. Hayhoe, X. Yang, and D.J. Wuebbles. An asynchronous regional  
408 regression model for statistical downscaling of daily climate variables. *Int. J. Climatol.*,  
409 33:2473–2494, 2013.
- 410  
411

- 412 [10] E.H. Girvetz, E.P. Maurer, P.B. Duffy, A. Ruesch, B. Thrasher, and C. Zganjar. Making  
413 climate data relevant to decision making: the important details of spatial and temporal  
414 downscaling. The World Bank, 2013.
- 415 [11] S. Hempel, K. Frieler, L. Warszawski, J. Schewe, and F. Piontek. A trend-preserving  
416 bias correction - the ISI-MIP approach. *Earth Syst. Dynam.*, 4:219–236, 2013.
- 417 [12] E.P. Maurer, L. Brekke, T. Pruitt, B. Thrasher, J. Long, P. Duffy, M. Dettinger,  
418 D. Cayan, and J. Arnold. An enhanced archive facilitating climate impact and adaptation  
419 analysis. *Bull. Amer. Meteorol. Soc.*, 95(7):1011–1019, 2014.
- 420 [13] CORDEX. Bias-adjusted RCM data. [http://www.cordex.org/index.php?option=com\\_content&view=](http://www.cordex.org/index.php?option=com_content&view=article&id=123)  
421 2016.
- 422 [14] S. Gangopadhyay, T. Pruitt, L. Brekke, and D. Raff. Hydrologic projections for the  
423 Western United States. *EOS*, 92(48):441–442, 2011.
- 424 [15] S. Hagemann, C. Chen, D. Clark, S. Folwell, S.N. Gosling, I. Haddeland, N. Hannasaki,  
425 J. Heinke, F. Ludwig, F. Voss, and A. Wiltshire. Climate change impact on available  
426 water resources obtained using multiple global climate and hydrology models. *Earth*  
427 *Syst. Dynam.*, 4:129–144, 2013.
- 428 [16] L. Warszawski, K. Frieler, V. Huber, F. Piontek, O. Serdeczny, and J. Schewe. The Inter-  
429 Sectoral Impact Model Intercomparison Project (ISIMIP): Project framework. *Proc. Nat.*  
430 *Acad. Sci.*, 111(9):3228–3232, 2014.
- 431 [17] D. Cayan, K. Kunkel, C. Castro, A. Gershunov, J. Barsugli, A. Ray, J. Overpeck, M. An-  
432 derson, J. Russell, B. Rajagopalan, I. Rangwala, and P. Duffy. *Assessment of Climate*  
433 *Change in the Southwest United States: A Report Prepared for the National Climate*  
434 *Assessment*, chapter Future climate: Projected average., pages 101–125. Island Press,  
435 2013.
- 436 [18] World Bank. Turn Down the Heat: Climate Extremes, Regional Impacts, and the Case  
437 for Resilience. A report for the World Bank by the Potsdam Institute for Climate Impact  
438 Research and Climate Analytics, 2013.
- 439 [19] A. Georgakakos, P. Fleming, M. Dettinger, C. Peters-Lidard, T.C. Richmond, K. Reck-  
440 how, K. White, and D. Yates. *Climate Change Impacts in the United States: The Third*  
441 *National Climate Assessment*, chapter Water Resources, pages 69–112. U.S. Global  
442 Change Research Program, 2014.
- 443 [20] Worldbank. Climate change knowledge portal. <http://sdwebx.worldbank.org/climateportal/>.
- 444 [21] prepdata. Partnership for Resilience and Preparedness (PREP). [www.prepdata.org](http://www.prepdata.org).
- 445 [22] D. Maraun, F. Wetterhall, A.M. Ireson, R.E. Chandler, E.J. Kendon, M. Widmann,  
446 S. Brienen, H.W. Rust, T.Sauter, M.Thiemeßl, V.K.C. Venema, K.P. Chun, C.M. Good-  
447 ess, R.G. Jones, C. Onof, M. Vrac, and I. Thiele-Eich. Precipitation downscaling under

- 448 climate change. Recent developments to bridge the gap between dynamical models and  
449 the end user. *Rev. Geophys.*, 48:RG3003, 2010.
- 450 [23] C. Teutschbein and J. Seibert. Bias correction of regional climate model simulations for  
451 hydrological climate-change impact studies: Review and evaluation of different methods.  
452 *J. Hydrol.*, 456:12–29, 2012.
- 453 [24] C. Piani, J.O. Haerter, and E. Coppola. Statistical bias correction for daily precipitation  
454 in regional climate models over Europe. *Theor. Appl. Climatol.*, 99(1-2):187–192, 2010.
- 455 [25] S. Vannitsem. Bias correction and post-processing under climate change. *Nonlin. Proc.*  
456 *Geophys.*, 18:911–924, 2011.
- 457 [26] J. Eden, M. Widmann, D. Grawe, and S. Rast. Skill, correction, and downscaling of  
458 GCM-simulated precipitation. *J. Climate*, 25:3970–3984, 2012.
- 459 [27] D. Maraun. Bias correction, quantile mapping and downscaling: Revisiting the inflation  
460 issue. *J. Climate*, 26:2137–2143, 2013.
- 461 [28] D. Maraun and M. Widmann. The representation of location by a regional climate model  
462 in complex terrain. *Hydrol. Earth Syst. Sci.*, 19:3449–3456, 2015.
- 463 [29] N. Addor, M. Rohrer, R. Furrer, and J. Seibert. Propagation of biases in climate models  
464 from the synoptic to the regional scale: Implications for bias adjustment. *J. Geophys.*  
465 *Res.*, 121(5):2075–2089, 2016.
- 466 [30] T.F. Stocker, Q. Dahe, G.-K. Plattner, and M. Tignor. IPCC Workshop on Re-  
467 gional Climate Projections and their Use in Impacts and Risk Analysis Studies.  
468 [https://www.ipcc.ch/pdf/supporting-material/RPW\\_WorkshopReport.pdf](https://www.ipcc.ch/pdf/supporting-material/RPW_WorkshopReport.pdf), 2015.
- 469 [31] U. Ehret, E. Zehe, V. Wulfmeyer, K. Warrach-Sagi, and J. Liebert. Should we apply  
470 bias correction to global and regional climate model data? *Hydrol. Earth Syst. Sci.*, 16,  
471 2012.
- 472 [32] J.O. Haerter, S. Hagemann, C. Moseley, and C. Piani. Climate model bias correction  
473 and the role of timescales. *Hydrol. Earth Syst. Sci.*, 15(3):1065–1079, 2011.
- 474 [33] F. Johnson and A. Sharma. A nesting model for bias correction of variability at multiple  
475 time scales in general circulation model precipitation simulations. *Wat. Resour. Res.*,  
476 48:W01504, 2012.
- 477 [34] P.-A. Michelangeli, M. Vrac, and H. Loukos. Probabilistic downscaling approaches: Ap-  
478 plication to wind cumulative distribution functions. *Geophys. Res. Lett.*, 36(11), 2009.
- 479 [35] C. Piani and J.O. Haerter. Two dimensional bias correction of temperature and precip-  
480 itation copulas in climate models. *Geophys. Res. Lett.*, 39(20):L20401, 2012.
- 481 [36] M. Vrac and P. Friederichs. Multivariate-intervariable, spatial, and temporal-bias cor-  
482 rection. *J. Climate*, 28(1):218–237, 2015.

- 483 [37] A.A.L. Levy, W. Ingram, M. Jenkinson, C. Huntingford, F.H. Lambert, and M. Allen.  
484 Can correcting feature location in simulated mean climate improve agreement on pro-  
485 jected changes? *Geophys. Res. Lett.*, 40:354–358, 2013.
- 486 [38] J.J. Barsugli, G. Guentchev, R. M. Horton, A. Wood, L.O. Mearns, X.-Z. Liang, J.A.  
487 Winkler, K. Dixon, K. Hayhoe, R.B. Rood, L. Goddard, A. Ray, L. Buja, and C. Am-  
488 mann. The practitioner’s dilemma: How to assess the credibility of downscaled climate  
489 projections. *EOS*, 94(46):424–425, 2013.
- 490 [39] C. Wang, L. Zhang, S.-K. Lee, L. Wu, and C.R. Mechoso. A global perspective on CMIP5  
491 climate model biases. *Nat. Clim. Change*, 4:201–205, 2014.
- 492 [40] J.H. Christensen, F. Boberg, O.B. Christensen, and P. Lucas-Picher. On the need for  
493 bias correction of regional climate change projections of temperature and precipitation.  
494 *Geophys. Res. Lett.*, 35:L20709, 2008.
- 495 [41] C.M. Buser, H.R. Künsch, D. Lüthi, M. Wild, and C. Schär. Bayesian multi-model projec-  
496 tion of climate: bias assumptions and interannual variability. *Clim. Dynam.*, 33:849–868,  
497 2009.
- 498 [42] D. Maraun. Nonstationarities of regional climate model biases in European seasonal  
499 mean temperature and precipitation sums. *Geophys. Res. Lett.*, 39:L06706, 2012.
- 500 [43] C. Teutschbein and J. Seibert. Is bias correction of regional climate model (rcm) simula-  
501 tions possible for non-stationary conditions? *Hydrol. Earth Syst. Sci.*, 17(12):5061–5077,  
502 2013.
- 503 [44] F. Johnson and A. Sharma. Accounting for interannual variability: A comparison of  
504 options for water resources climate change impact assessments. *Wat. Resour. Res.*,  
505 47:W04508, 2011.
- 506 [45] P. Berg, H. Feldmann, and H.-J. Panitz. Bias correction of high resolution regional  
507 climate model data. *J. Hydrol.*, 448-449:80–92, 2012.
- 508 [46] C. Chen, J.O. Haerter, S. Hagemann, and C. Piani. On the contribution of statistical  
509 bias correction to the uncertainty in the projected hydrological cycle. *Geophys. Res.*  
510 *Lett.*, 38:L20403, 2011.
- 511 [47] A. Dosio and P. Paruolo. Bias correction of the ENSEMBLES high resolution climate  
512 change projections for use by impact models: Evaluation on the present climate. *J.*  
513 *Geophys. Res. Atmos.*, 116(D16), 2011.
- 514 [48] M. J. Themeßl, A. Gobiet, and A. Leuprecht. Empirical-statistical downscaling and  
515 error correction of daily precipitation from regional climate models. *Int. J. Climatol.*,  
516 31:1530–1544, 2011.

- 517 [49] L. Gudmundson, J.B. Bremnes, J.E. Haugen, and T. Engen-Skaugen. Technical Note:  
518 Downscaling RCM precipitation to the station scale using statistical transformations a  
519 comparison of methods. *Hydrol. Earth Syst. Sci.*, 16:3383–3390, 2012.
- 520 [50] J. Rajczak, S. Kotlarski, and C. Schär. Does quantile mapping of simulated precipitation  
521 correct for biases in transition probabilities and spell lengths? *J. Climate*, 29:1605–1615,  
522 2016.
- 523 [51] G. Masato, B.J. Hoskins, and T. Woollings. Winter and summer northern hemisphere  
524 blocking in CMIP5 models. *J. Climate*, 26:7044–7059, 2013.
- 525 [52] T. Woollings. Dynamical influences on European climate: An uncertain future. *Phil.*  
526 *Trans. R. Soc. A*, 368:3733–3756, 2010.
- 527 [53] G. Zappa, L.C. Shaffrey, and K.I. Hodges. The ability of CMIP5 models to simulate  
528 North Atlantic extratropical cyclones. *J. Climate*, 26:5379–5396, 2013.
- 529 [54] P. Davini, J. von Hardenberg, S. Corti, H.M. Christensen, S. Juricke, A. Subramanian,  
530 P.A.G. Watson, A. Weisheimer, and T.N. Palmer. Climate SPHINX: evaluating the  
531 impact of resolution and stochastic physics parameterisations in the EC-Earth global  
532 climate model. *Geosci. Model Dev.*, 10(3):1383–1402, 2017.
- 533 [55] F. Pithan, T.G. Shepherd, G. Zappa, and I. Sandu. Climate model biases in jet streams,  
534 blocking and storm tracks resulting from missing orographic drag. *Geophys. Res. Lett.*,  
535 43:7231–7240, 2016. doi:10.1002/2016GL069551.
- 536 [56] A. van Niekerk, J.F. Scinocca, and T.G. Shepherd. The modulation of stationary waves,  
537 and their response to climate change, by parameterized orographic drag. *J. Atmos. Sci.*,  
538 74:2557–2574, 2017. doi:10.1175/JAS-D-17-0085.1.
- 539 [57] M. Ashfaq, C.B. Skinner, and N.S. Diffenbaugh. Influence of SST biases on future climate  
540 change projections. *Clim. Dynam.*, 36:1303–1319, 2011.
- 541 [58] A.A. Scaife, D. Copsey, C. Gordon, C. Harris, T. Hinton, S. Keeley, A. O’Neill,  
542 M. Roberts, and K. Williams. Improved Atlantic winter blocking in a climate model.  
543 *Geophys. Res. Lett.*, 38:L23703, 2011.
- 544 [59] S.P.E. Keeley, R.T. Sutton, and L.C. Shaffrey. The impact of North Atlantic sea surface  
545 temperature errors on the simulation of North Atlantic European region climate. *Q. J.*  
546 *R. Meteorol. Soc.*, 138:1774–1783, 2012.
- 547 [60] T.G. Shepherd. Atmospheric circulation as a source of uncertainty in climate change  
548 projections. *Nat. Geosci.*, 7:703–708, 2014.
- 549 [61] A. Hall. Projecting regional change. *Science*, 346(6216):1461–1462, 2014.
- 550 [62] A. Colette, R. Vautard, and M. Vrac. Regional climate downscaling with prior statistical  
551 correction of the global climate forcing. *Geophys. Res. Lett.*, 39(13):L13707, 2012.



- 552 [63] E.P. Maurer and D.W. Pierce. Bias correction can modify climate model simulated  
553 precipitation changes without adverse effect on the ensemble mean. *Hydrol. Earth Syst.*  
554 *Sci.*, 18(3):915–925, 2014.
- 555 [64] A. Gobiet, M. Suklitsch, and G. Heinrich. The effect of empirical-statistical correction  
556 of intensity-dependent model errors on the temperature climate change signal. *Hydrol.*  
557 *Earth Syst. Sci.*, 19:4055–4066, 2015.
- 558 [65] D.W. Pierce, D.R. Cayan, E.P. Maurer, J.T. Abatzoglou, and K.C. Hegewisch. Im-  
559 proved Bias Correction Techniques for Hydrological Simulations of Climate Change. *J.*  
560 *Hydrometeorol.*, 16(6):2421–2442, 2015.
- 561 [66] M.B. Switanek, P.A. Troch, C.L. Castro, A. Leuprecht, H.-I. Chang, R. Mukherjee, and  
562 E.M.C. Demaria. Scaled distribution mapping: a bias correction method that preserves  
563 raw climate model projected changes. *Hydrol. Earth Syst. Sci.*, 21:2649–2666, 2017.
- 564 [67] O. Bellprat, S. Kotlarski, D. Lüthi, and C. Schär. Physical constraints for temperature  
565 biases in climate models. *Geophys. Res. Lett.*, 40:4042–4047, 2013.
- 566 [68] H. Bellenger, E. Guilyardi, J. Leloup, M. Lengaigne, and J. Vialard. ENSO representation  
567 in climate models: from CMIP3 to CMIP5. *Clim. Dynam.*, 42:1999–2018, 2014.
- 568 [69] L. Chen, T. Li, and Y. Yu. Causes of strengthening and weakening of ENSO amplitude  
569 under global warming in four CMIP5 models. *J. Climate*, 28:3250–3274, 2015.
- 570 [70] M. Collins, R. Knutti, J. Arblaster, J.-L. Dufresne, T. Fichefet, P. Friedlingstein, X. Gao,  
571 W.J. Gutowski, T. Johns, G. Krinner, M. Shongwe, C. Tebaldi, A.J. Weaver, and  
572 M. Wehner. *Climate Change 2013: The Physical Science Basis. Contribution of Work-*  
573 *ing Group I to the Fifth Assessment Report of the Intergovernmental Panel on Cli-*  
574 *mate Change*, chapter Long-term Climate Change: Projections, Commitments and Ir-  
575 *reversibility*. Cambridge University Press, Cambridge, United Kingdom and New York,  
576 NY, USA, 2013.
- 577 [71] H. Annamalai, J. Hafner, K.P. Sooraj, and P. Pillai. Global warming shifts the monsoon  
578 circulation, drying South Asia. *J. Climate*, 26:2701–2718, 2013.
- 579 [72] M.K. Roxy, K. Ritika, P. Terray, R. Murtugudde, K. Ashok, and B.N. Goswami. Drying  
580 of indian subcontinent by rapid indian ocean warming and a weakening land-sea thermal  
581 gradient. *Nat. Commun.*, 6:7423, 2015.
- 582 [73] J. Willison, W.A. Robinson, and G.M. Lackmann. North atlantic storm-track sensitivity  
583 to warming increases with model resolution. *J. Climate*, 28:4513–4524, 2015.
- 584 [74] A. Hall, X. Qu, and J.D. Neelin. Improving predictions of summer climate change in the  
585 united states. *Geophys. Res. Lett.*, 35:L01702, 2008.
- 586 [75] J.H. Christensen and F. Boberg. Temperature dependent climate projection deficiencies  
587 in CMIP5 models. *Geophys. Res. Lett.*, 39:L24705, 2012.

- 588 [76] E.J. Kendon, N.M. Roberts, H.J. Fowler, M.J. Roberts, S.C. Chan, and C.A. Senior.  
589 Heavier summer downpours with climate change revealed by weather forecast resolution  
590 model. *Nat. Clim. Change*, 4:570–576, 2014.
- 591 [77] E.P. Meredith, D. Maraun, V.A. Semenov, and W. Park. Evidence for added value of  
592 convection permitting models for studying changes in extreme precipitation. *J. Geophys.*  
593 *Res. Atmos.*, 120:12,500–12,513, 2015.
- 594 [78] M. Collins, R.E. Chandler, P.M. Cox, J.M. Huthnance, J. Rougier, and D.B. Stephenson.  
595 Quantifying future climate change. *Nat. Clim. Change*, 2(6):403–409, 2012.
- 596 [79] I.R. Simpson, R. Seager, M. Ting, and T.A. Shaw. Causes of change in northern hemi-  
597 sphere winter meridional winds and regional hydroclimate. *Nat. Clim. Change*, 6(1):65–  
598 70, 2016.
- 599 [80] X. Qu and A. Hall. On the persistent spread in snow-albedo feedback. *Clim. Dynam.*,  
600 42(1-2):69, 2014.
- 601 [81] A. Dawson, T.N. Palmer, and S. Corti. Simulating regime structures in weather and  
602 climate prediction models. *Geophys. Res. Lett.*, 39(21), 2012.
- 603 [82] P. Liu and Coauthors. An MJO simulated by the NICAM at 14- and 7-km resolutions.  
604 *Mon. Wea. Rev.*, 137:3254–3268, 2009.
- 605 [83] A.F. Prein, W. Langhans, G. Fossier, A. Ferrone, N. Ban, K. Goergen, M. Keller, M. Tlle,  
606 O. Gutjahr, F. Feser, E. Brisson, S. Kollet, J. Schmidli, N.P.M. van Lipzig, and R. Le-  
607 ung. A review on regional convection-permitting climate modeling: Demonstrations,  
608 prospects, and challenges. *Rev. Geophys.*, 53(2):323–361, 2015.
- 609 [84] C. Volosciuk, D. Maraun, M. Vrac, and M. Widmann. A combined statistical bias  
610 correction and stochastic downscaling method for precipitation. *Hydrol. Earth Syst. Sci.*,  
611 21(3):1693–1719, 2017.
- 612 [85] D.B. Walton, F. Sun, A. Hall, and S. Capps. A hybrid dynamical-statistical downscaling  
613 technique. part i: Development and validation of the technique. *J. Climate*, 28(12):4597–  
614 4617, 2015.
- 615 [86] D. Maraun, M. Widmann, J.M. Gutierrez, S. Kotlarski, R.E. Chandler, E. Hertig,  
616 J. Wibig, R. Huth, and R.A.I. Wilcke. VALUE: A framework to validate downscaling  
617 approaches for climate change studies. *Earth’s Future*, 3:1–14, 2015.
- 618 [87] C. Prudhomme, R.L. Wilby, S. Crooks, A.L. Kay, and N.S. Reynard. Scenario-neutral  
619 approach to climate change impact studies: Application to flood risk. *J. Hydrol.*, 390:198–  
620 209, 2010.
- 621 [88] C. Schär, C. Frei, D. Lüthi, and H.C. Davies. Surrogate climate-change scenarios for  
622 regional climate models. *Geophys. Res. Lett.*, 23(6):669–672, 1996.

- 623 [89] W. Hazeleger, B.J.J.M. van den Hurk, E. Min, G.J. van Oldenborgh, A.C. Petersen, D.A.  
624 Stainforth, E. Vasileiadou, and L.A. Smith. Tales of future weather. *Nat. Clim. Change*,  
625 5:107–113, 2015.
- 626 [90] A.M.G. Klein Tank, J.B. Wijngaard, G.P. Können, R. Böhm, G. Demarée, A. Gocheva,  
627 M. Mileta, S. Pashiardis, L. Hejkrlik, C. Kern-Hansen, R. Heino, P. Bessemoulin,  
628 G. Müller-Westermeier, M. Tzanakou, S. Szalai, T. Pálsdóttir, D. Fitzgerald, S. Ru-  
629 bin, M. Capaldo, M. Maugeri, A. Leitass, A. Bukantis, R. Aberfeld, A.F.V. van Engelen,  
630 E. Forland, M. Miletus, F. Coelho, C. Mares, V. Razuvaev, E. Nieplova, T. Cegnar,  
631 J.A. López, B. Dahlström, A. Moberg, W. Kirchhofer, A. Ceylan, O. Pachaliuk, L.V.  
632 Alexander, and P. Petrovic. Daily dataset of 20th-century surface air temperature and  
633 precipitation series for the European Climate Assessment. *Int. J. Climatol.*, 22(12):1441–  
634 1453, 2002.
- 635 [91] J.M. Gutiérrez, R. Cano, A.S. Cofiño, and C. Sordo. Analysis and downscaling multi-  
636 model seasonal forecasts in Peru using self-organizing maps. *Tellus A*, 57(3):435–447,  
637 2005.
- 638 [92] E.P. Maurer, A.W. Wood, J.C. Adam, D.P. Lettenmaier, and B. Nijssen. A long-term  
639 hydrologically-based data set of land surface fluxes and states for the conterminous United  
640 States. *J. Climate*, 15(22):3237–3251, 2002.
- 641 [93] D.P. van Vuuren, J. Edmonds, M. Kainuma, K. Riahi, A. Thomson, K. Hibbard,  
642 G. Hurtt, T. Kram, V. Krey, J.-F. Lamarque, T. Masui, M. Meinshausen, N. Naki-  
643 cenovic, S.J. Smith, and S.K. Rose. The representative concentration pathways: an  
644 overview. *Clim. Change*, 109(1):5–31, 2011.

645  
646 **Corresponding author**

647 Douglas Maraun (douglas.maraun@uni-graz.at)

648  
649 **Acknowledgements**

650 Many of the ideas laid out in this paper have been developed during discussions within the EU  
651 COST Action ES1102 “VALUE”, in particular during the workshop “Global climate model  
652 biases: causes, consequences and correctability” held at the Max Planck Institute for Meteo-  
653 rology, Hamburg, October 2014.

654  
655 **Author contributions**

656 The paper is a result of a workshop organised by D.M. and S.H. D.M. wrote the first draft  
657 of the manuscript with inputs from all authors. D.M., G.Z., J.M.G and D.W. contributed  
658 analyses underlying the figures. All authors discussed the content of the manuscript.

# 659 List of Figures

660 1 **Cross-validation problem** | Quantile mapping from ERA40 daily boreal winter (DJF) temperature [ $^{\circ}\text{C}$ , Southern Ocean, 45S-55S, 175W-163W] to E-OBS daily precipitation [mm/day, Central Europe, 45N-55N, 5E-17E], calibrated over 1961-1980. **a-c**, mean and **d-f**, 95th percentile over validation period (1981-2000). **a,d**, uncorrected ERA40, **b,e** observations, **c,f** corrected ERA40. **g**, histogram of biases across all grid boxes. **h** QQ-plot for grid box close to Venice (see cross in panel **a**). A QQ-plot plots the quantiles of two distributions against each other, i.e., for two time series, the values are sorted separately and then plotted against each other. The correction function is based on linear interpolation between empirical quantiles with a constant correction for new extreme values. . . . . 21

661  
662  
663  
664  
665  
666  
667  
668  
669  
670

671 2 **Unrealistic dry spell lengths** | Distribution of dry spell lengths (wet-day threshold 0.1 mm) at **a**, Tafjord (Norway;  $7.41^{\circ}\text{W}$ ,  $62.23^{\circ}\text{N}$ , winter), **b**, Constanta (Romania;  $28.63^{\circ}\text{E}$ ,  $44.22^{\circ}\text{N}$ , winter), **c**, Sion (Switzerland,  $7.33^{\circ}\text{E}$ ,  $46.22^{\circ}\text{N}$ , winter) and **d**, Rome (Italy,  $12.58^{\circ}\text{E}$ ,  $41.78^{\circ}\text{N}$ , summer) of MPI-ESM-LR downscaled with CLM to a horizontal resolution of  $0.44^{\circ}$ , 1971-2000. Black: observations (ECA-D<sup>90</sup>), blue: raw climate model, red: corrected climate model. Long dry spells are typically underrepresented even after a seasonal wet day correction (**a**), although in some cases the correction may improve the overall distribution (**b**). Often, artefacts are introduced for short (**c**) and long (**d**) spells. . . . . 22

672  
673  
674  
675  
676  
677  
678  
679  
680

681 3 **Non-representative model output** | Daily precipitation bias correction for the GISS-E2-R model against station data at Piura, Peru<sup>91</sup> from 1976-2000. **a**, observations; **b**, raw GCM data; **c**, quantile mapped GCM data; **d**, QQ plot. Grey shading: El Niño events. As the GCM is run in climate mode, simulated events are not synchronised with observations. Even though the quantile mapping perfectly adjusts the simulated distribution, the result is meaningless, as the GCM does not correctly capture the clustering of extreme precipitation during El Niño events. . . . . 23

682  
683  
684  
685  
686  
687  
688

689 4 **Missing temperature inversions** | Distribution of spring (MAM) daily mean temperature differences between Fresno ( $\sim 90\text{m}$ ) and Three Rivers (70 km towards the southeast, at  $\sim 400\text{m}$ ) in California, US, 1981-2000. Blue: observations ( $1/8^{\circ}$  gridded data<sup>92</sup>), orange: GFDL-CM3 GCM after quantile mapping against observations (scaled by 1/4). In reality, temperature inversions ( $\Delta T < 0$ ) in the Central Valley occur on about 7% of the days. The coarse-resolution GCM does not simulate such inversions. Quantile mapping provides the correct climatological temperature difference, but is by construction unable to produce sub-grid inversions. The correction function was based on parametric Gaussian distributions. . . . . 24

690  
691  
692  
693  
694  
695  
696  
697  
698

699	5	<b>Large-scale circulation problems   a-c</b> , FGOALS-g2; <b>d-f</b> , MPI-ESM-MR.	
700		<b>a,d</b> , simulated (colour shading, mm/day) and observed (contour lines at 4	
701		and 6 mm/day) mean winter precipitation 1976-2005. <b>b,e</b> , uncorrected mean	
702		precipitation averaged over 10W to 20E (vertical red lines in <b>a</b> and <b>d</b> ) from	
703		present and future (2070-2099, RCP8.5 <sup>93</sup> ) simulations. <b>c,f</b> , corresponding cor-	
704		rected simulations (the black line by construction equals observed winter pre-	
705		cipitation). Precipitation is bias corrected relative to the GPCP climatology	
706		(1980-2013). In FGOALS-g2, the storm track is unrealistically far south. As a	
707		result, even though the storm track shifts northwards in the future simulation,	
708		the corrected precipitation shifts southwards. For MPI-ESM-MR the circula-	
709		tion bias is low, avoiding an unphysical inconsistency between circulation and	
710		precipitation shift. The correction function multiplicatively adjusts long-term	
711		mean biases. . . . .	25
712	6	<b>Implausible sub-grid climate change signal   Spring (MAM) daily mean</b>	
713		temperature [°C] in the Sierra Nevada and Central Valley, California, US. <b>a-c</b> ,	
714		present climate (1981-2000 average); <b>d-f</b> , simulated change (2081-2100 aver-	
715		age minus 1981-2000 average, RCP8.5 scenario <sup>93</sup> ). <b>a,d</b> , GFDL-CM3 GCM,	
716		bilinearly interpolated to 8km grid; <b>b,e</b> , corrected GCM (for present by con-	
717		struction identical with observations at 8km horizontal resolution <sup>92</sup> ); <b>c,f</b> , WRF	
718		RCM at 3km horizontal resolution, driven with GFDL-CM3 climate change sig-	
719		nal <sup>85</sup> . Whereas the RCM simulates plausible strong elevation-dependent warm-	
720		ing (the strongest temperature increase in the Sierra Nevada mountains), the	
721		bias correction modulates the GCM change unsystematically and not related	
722		to elevation. . . . .	26

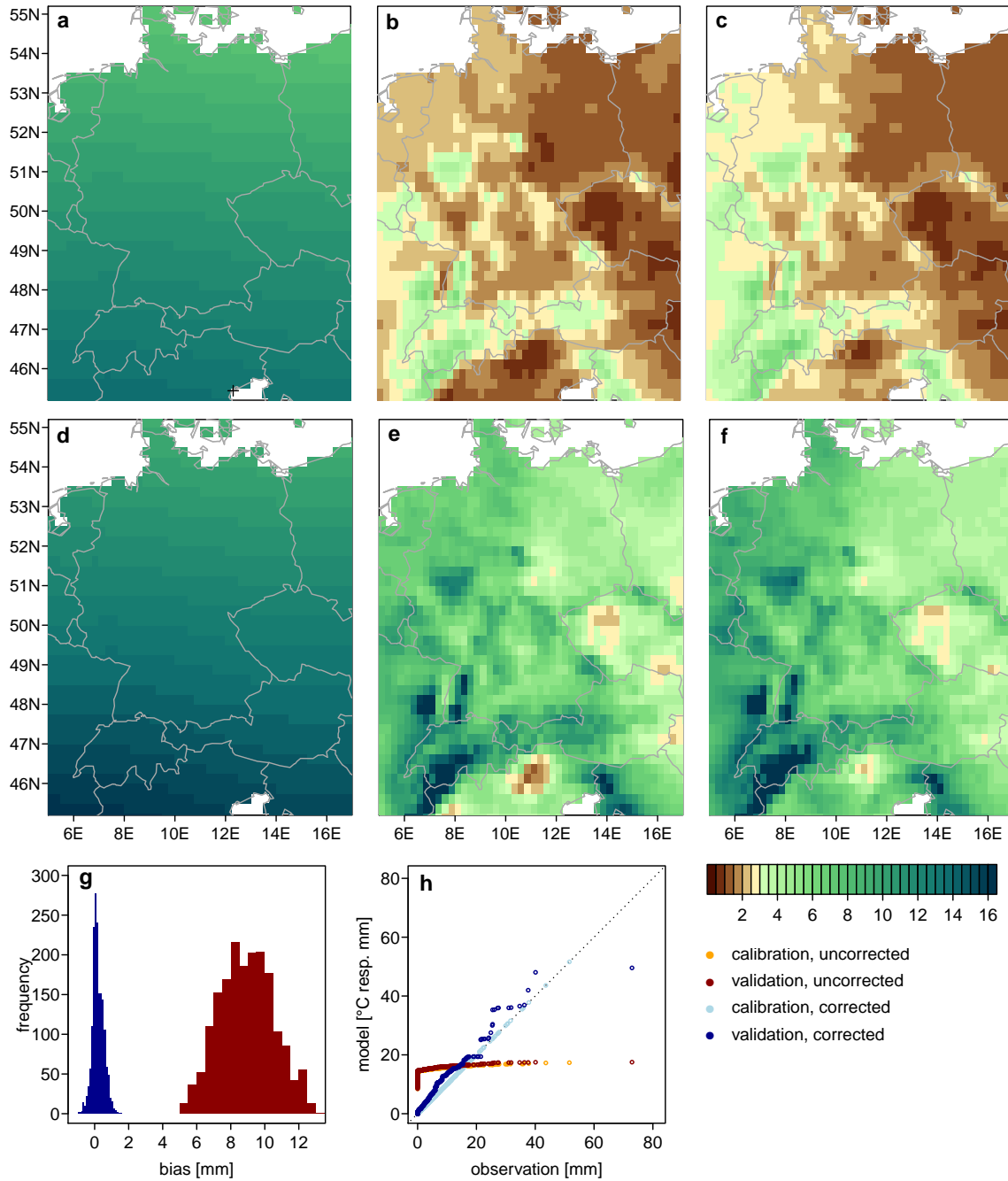


Figure 1: **Cross-validation problem** | Quantile mapping from ERA40 daily boreal winter (DJF) temperature [ $^{\circ}\text{C}$ , Southern Ocean, 45S-55S, 175W-163W] to E-OBS daily precipitation [mm/day, Central Europe, 45N-55N, 5E-17E], calibrated over 1961-1980. **a-c**, mean and **d-f**, 95th percentile over validation period (1981-2000). **a,d**, uncorrected ERA40, **b,e** observations, **c,f** corrected ERA40. **g**, histogram of biases across all grid boxes. **h** QQ-plot for grid box close to Venice (see cross in panel **a**). A QQ-plot plots the quantiles of two distributions against each other, i.e., for two time series, the values are sorted separately and then plotted against each other. The correction function is based on linear interpolation between empirical quantiles with a constant correction for new extreme values.

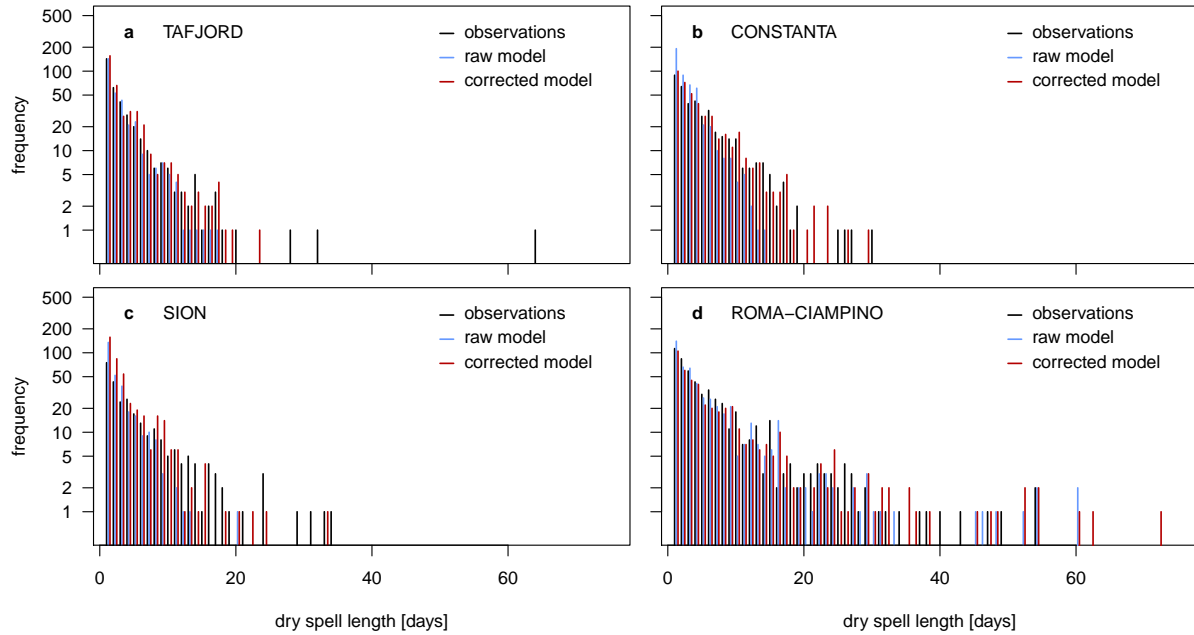


Figure 2: **Unrealistic dry spell lengths** | Distribution of dry spell lengths (wet-day threshold 0.1 mm) at **a**, Tafjord (Norway; 7.41° W, 62.23°N, winter), **b**, Constanta (Romania; 28.63° E, 44.22°N, winter), **c**, Sion (Switzerland, 7.33° E, 46.22°N, winter) and **d**, Rome (Italy, 12.58° E, 41.78°N, summer) of MPI-ESM-LR downscaled with CLM to a horizontal resolution of 0.44°, 1971-2000. Black: observations (ECA-D<sup>90</sup>), blue: raw climate model, red: corrected climate model. Long dry spells are typically underrepresented even after a seasonal wet day correction (**a**), although in some cases the correction may improve the overall distribution (**b**). Often, artefacts are introduced for short (**c**) and long (**d**) spells.

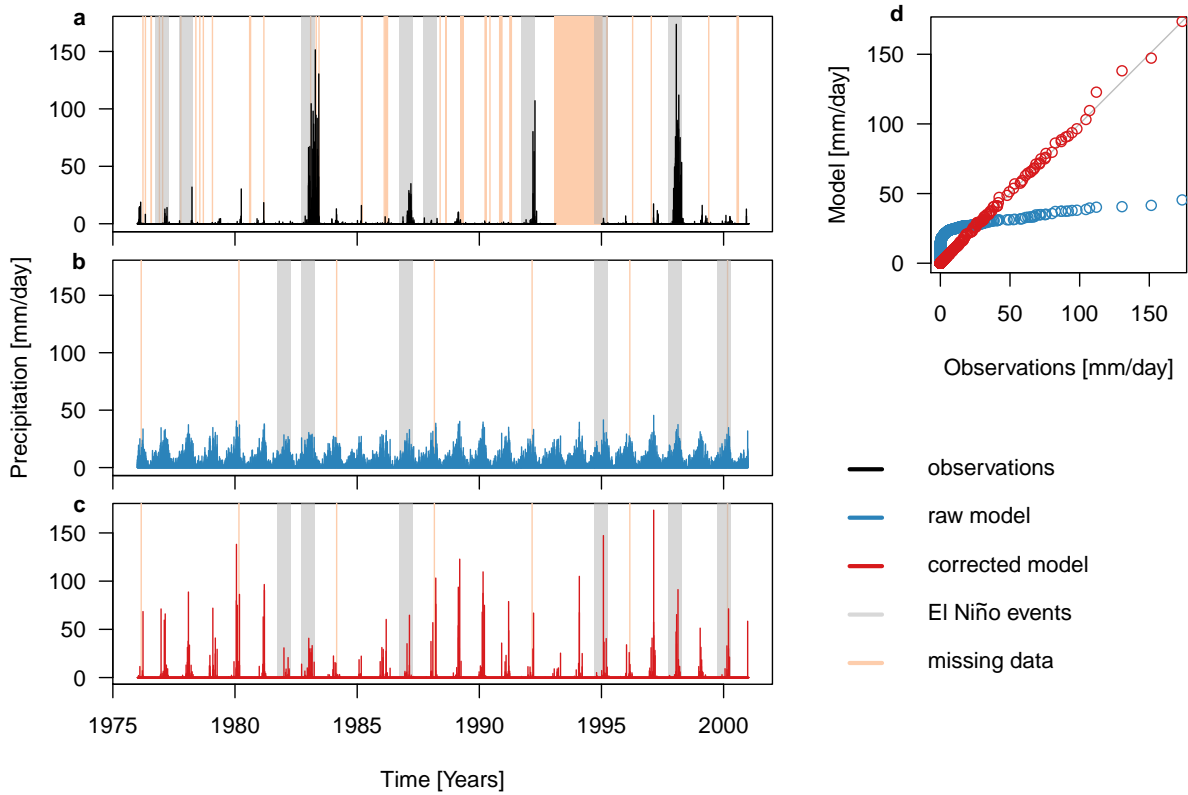


Figure 3: **Non-representative model output** | Daily precipitation bias correction for the GISS-E2-R model against station data at Piura, Peru<sup>91</sup> from 1976-2000. **a**, observations; **b**, raw GCM data; **c**, quantile mapped GCM data; **d**, QQ plot. Grey shading: El Niño events. As the GCM is run in climate mode, simulated events are not synchronised with observations. Even though the quantile mapping perfectly adjusts the simulated distribution, the result is meaningless, as the GCM does not correctly capture the clustering of extreme precipitation during El Niño events.



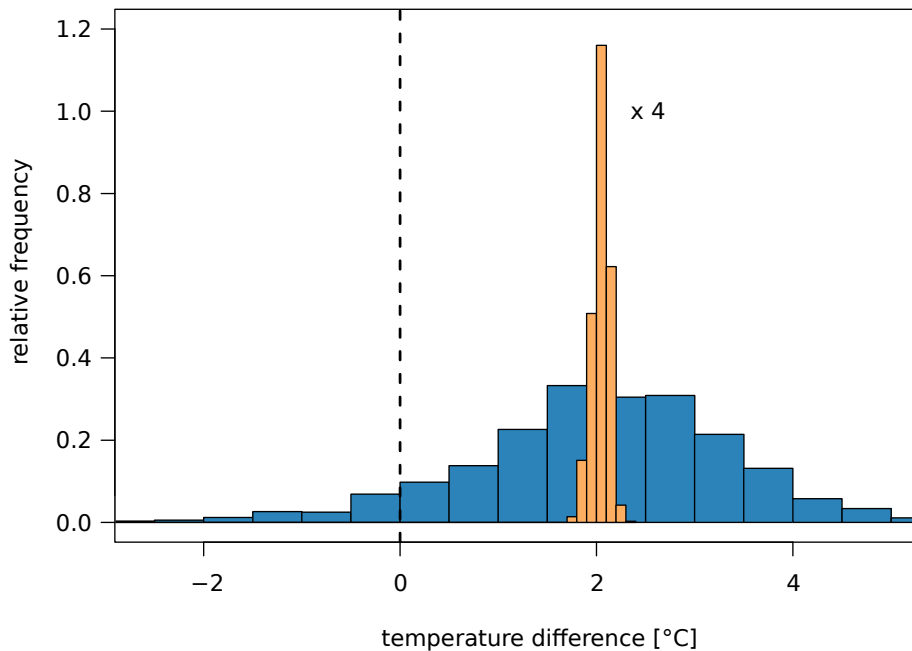


Figure 4: **Missing temperature inversions** | Distribution of spring (MAM) daily mean temperature differences between Fresno ( $\sim 90\text{m}$ ) and Three Rivers (70 km towards the south-east, at  $\sim 400\text{m}$ ) in California, US, 1981-2000. Blue: observations ( $1/8^\circ$  gridded data<sup>92</sup>), orange: GFDL-CM3 GCM after quantile mapping against observations (scaled by  $1/4$ ). In reality, temperature inversions ( $\Delta T < 0$ ) in the Central Valley occur on about 7% of the days. The coarse-resolution GCM does not simulate such inversions. Quantile mapping provides the correct climatological temperature difference, but is by construction unable to produce sub-grid inversions. The correction function was based on parametric Gaussian distributions.

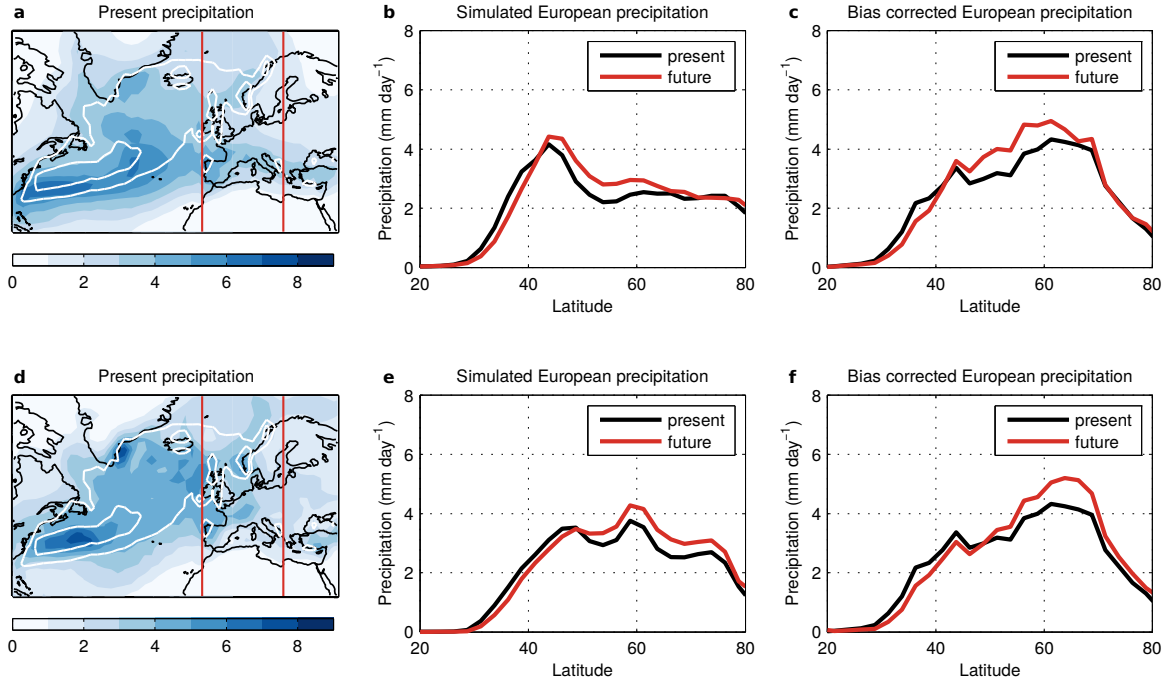


Figure 5: **Large-scale circulation problems** | a-c, FGOALS-g2; d-f, MPI-ESM-MR. a,d, simulated (colour shading, mm/day) and observed (contour lines at 4 and 6 mm/day) mean winter precipitation 1976-2005. b,e, uncorrected mean precipitation averaged over 10W to 20E (vertical red lines in a and d) from present and future (2070-2099, RCP8.5<sup>93</sup>) simulations. c,f, corresponding corrected simulations (the black line by construction equals observed winter precipitation). Precipitation is bias corrected relative to the GPCP climatology (1980-2013). In FGOALS-g2, the storm track is unrealistically far south. As a result, even though the storm track shifts northwards in the future simulation, the corrected precipitation shifts southwards. For MPI-ESM-MR the circulation bias is low, avoiding an unphysical inconsistency between circulation and precipitation shift. The correction function multiplicatively adjusts long-term mean biases.

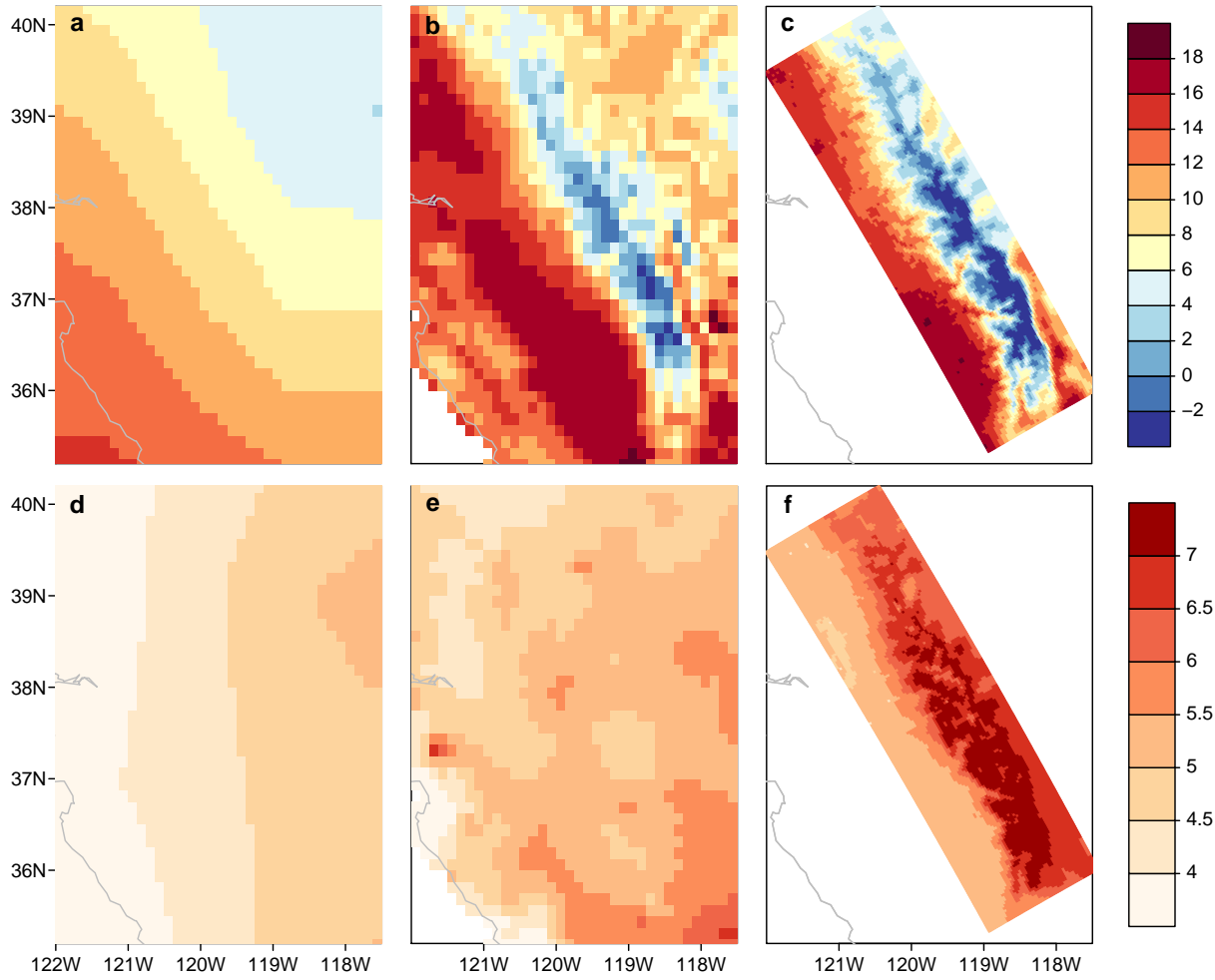


Figure 6: **Implausible sub-grid climate change signal** | Spring (MAM) daily mean temperature [ $^{\circ}\text{C}$ ] in the Sierra Nevada and Central Valley, California, US. **a-c**, present climate (1981-2000 average); **d-f**, simulated change (2081-2100 average minus 1981-2000 average, RCP8.5 scenario<sup>93</sup>). **a,d**, GFDL-CM3 GCM, bilinearly interpolated to 8km grid; **b,e**, corrected GCM (for present by construction identical with observations at 8km horizontal resolution<sup>92</sup>); **c,f**, WRF RCM at 3km horizontal resolution, driven with GFDL-CM3 climate change signal<sup>85</sup>. Whereas the RCM simulates plausible strong elevation-dependent warming (the strongest temperature increase in the Sierra Nevada mountains), the bias correction modulates the GCM change unsystematically and not related to elevation.