

Manuscript prepared for Hydrol. Earth Syst. Sci. Discuss.  
with version 4.1 of the L<sup>A</sup>T<sub>E</sub>X class copernicus\_discussions.cls.  
Date: 10 September 2013

# Precipitation bias correction of very high resolution regional climate models

**D. Argüeso, J. P. Evans, and L. Fita**

Climate Change Research Centre, University of New South Wales, Sydney, NSW 2052, Australia

ARC Centre of Excellence for Climate System Science, University of New South Wales, Sydney NSW 2052, Australia

Correspondence to: D. Argüeso  
([d.argueso@unsw.edu.au](mailto:d.argueso@unsw.edu.au))

## Abstract

Regional climate models are prone to biases in precipitation that are problematic for use in impact models such as hydrology models. A large number of methods have already been proposed aimed at correcting various moments of the rainfall distribution. They all require that the model produce the same or a higher number of rain days than the observational datasets, which are usually gridded datasets. Models have traditionally met this condition because their spatial resolution was coarser than the observational grids. But recent climate simulations use higher resolution and the models are likely to systematically produce fewer rain days than the gridded observations.

In this study, model outputs from a simulation at 2-km resolution are compared with gridded and in-situ observational datasets to determine whether the new scenario calls for revised methodologies. The gridded observations are found to be inadequate to correct the high-resolution model at daily timescales, because they are subjected to too frequent low intensity precipitation due to spatial averaging. A histogram equalisation bias correction method was adapted to the use of station, alleviating the problems associated with relative low-resolution observational grids. The wet-day frequency condition might not be satisfied for extremely dry biases, but the proposed approach substantially increases the applicability of bias correction to high-resolution models. The method is efficient at bias correcting both seasonal and daily characteristic of precipitation, providing more accurate information that is crucial for impact assessment studies.

## 1 Introduction

Regional Climate Models (RCMs) are an outstanding tool to study the mechanisms of climate at scales that are not yet resolved by General Circulation Models (GCMs). RCMs' higher spatial resolution and specifically designed parameterisations improve the representation of many aspects of climate (Feser et al., 2011; Giorgi, 2006), especially variables that are essentially local such as precipitation (Di Luca et al., 2011; Evans and McCabe, 2010; Tselioudis et al.,

2012). Hydrological models, which are especially sensitive to precipitation, thus benefit from high-resolution RCM outputs (Maraun et al., 2010).

However, RCMs are still prone to biases and the simulated climate is not always fully consistent with the observations, which is critical in climate change impact research (Portoghese et al., 2011). Previous authors (Christensen et al., 2008; Déqué et al., 2007) have advocated the use of bias correction to reduce model systematic deviations and provide more reliable outputs. It is arguable that bias correction should not be regarded as a definitive solution and its application is also matter of criticism (Ehret et al., 2012). Despite the fact that efforts should indeed be devoted towards the development of better models, there are no feasible alternatives to bias correction in terms of improving current simulations. Furthermore, some authors have examined the uncertainty introduced by bias correction and its impact on climate change estimates and overall recommended the use of non-linear methods to provide better projections (Chen et al., 2011; Themeßl et al., 2012).

Several methodologies have been recently proposed and evaluated, mostly focused on precipitation and temperature (Berg et al., 2012; Bordoy and Burlando, 2013; Haerter et al., 2011; Piani et al., 2010a; Terink et al., 2010). The physical characteristics of precipitation make it more difficult to correct and most studies tend to concentrate on precipitation correction. Methods of different complexity have been put forward, aimed at correcting various moments of the rainfall distribution. They range from relatively simple linear methods (Hay et al., 2000; Lenderink et al., 2007) to distribution-based algorithms, either using empirical distributions (Themeßl et al., 2011) or theoretical functions (Piani et al., 2010a). Hydrological impact studies require at least accurate daily data and hence methods that correct higher moments are preferable (Portoghese et al., 2011). Several evaluations (Lafon et al., 2013; Teutschbein and Seibert, 2012; Themeßl et al., 2011) have shown that distribution-mapping methods generally outperform others.

Bias correction is normally performed towards gridded datasets such as E-OBS (Haylock et al., 2008) or AWAP (Jones et al., 2009), because they cover the entire spatial domain and have complete timeseries. The spatial compatibility of gridded observations and model outputs of similar horizontal resolution has also been a reason to use gridded products in model bias

correction. Both the model outputs and the gridded dataset are reduced to the same grid and the transfer functions are calculated grid-point by grid-point.

No matter which method is selected among the myriad proposed, they all impose a common limitation to provide accurate corrections: they assume that the model produces the same or a higher number of rain days, independently from how these are defined. Distribution-based methods do not strictly require an equal or larger number of wet days, but if the model is generating too few rain days, the method might fail to adequately correct the model outputs. If any method is to be applied to model output with fewer rain days, it is necessary to introduce additional precipitation events (e.g. through Frequency Adaptation as in Themeßl et al. (2012)) otherwise daily intensity might be unrealistically corrected to match, for example, the monthly means.

So far this situation has rarely arisen and RCMs have traditionally met the aforementioned condition, partly because their spatial resolution is coarser than the observational gridded dataset to which they are compared. Models often display an increase in low intensity precipitation as a function of spatial resolution due to various reasons (i.e., model physics, spatial representativeness), producing more frequent but less intense precipitation than the station measurement as the resolution decreases. Gridded observations are also affected by similar behaviour, mostly due to the gridding methodology that involves some sort of spatial averaging. RCM simulations that exceed the spatial resolution of most gridded products have become possible due to improvements in computational resources. Such RCMs are likely to produce systematically less rain days than the gridded observations and thus the existing bias correction methodologies have to be revised.

In this paper, we analysed a RCM simulation at 2-km spatial resolution and compared it with both gridded and station based in-situ observational datasets to determine whether increasing resolution has implications in terms of the bias correction. We propose an alternative approach to the use of gridded observations for this purpose.

## 2 Model and observational data

### 2.1 Model description and setup

The Weather Research and Forecasting (WRF) model version 3.3.1 (Skamarock et al., 2009) was selected to simulate the recent climate (1990–2009) over the Sydney region (Figure 1a). The model was configured following (Evans and McCabe, 2010, 2013; Evans and Westra, 2012), where thorough evaluations of WRF over the region are provided covering time scales from sub-daily to inter-annual. The original simulation comprised two domains at 50- and 10-km spatial resolution covering southeastern Australia. The boundary conditions were obtained from the National Centers for Environmental Prediction (NCEP)/National Center for Atmospheric Research (NCAR) Reanalysis Project (NNRP1) and are provided every 6 hours to the model.

An additional 2-km spatial resolution domain (Figure 1) that covers the Sydney region was added to the original configuration (Argüeso et al., 2013). The convective parameterisation was switched off in this domain, while the microphysics parameterisation was changed to the more complex Thompson scheme (Thompson et al., 2006).

### 2.2 Observational data

The first observational reference dataset used in this study was a gridded dataset derived from observations and generated by the Australian Water Availability Project (AWAP) as described in (Jones et al., 2009). The resulting grid covers Australia at a  $0.05^\circ$  by  $0.05^\circ$  ( $\sim 5\text{km}$  by  $5\text{km}$ ) spatial resolution and provides values for different surface variables including precipitation, which was used here. The precipitation grid was constructed by interpolating surface station measurements that amounted to between 6000 and 7000 stations for the period 1990–2009. The dataset has been evaluated using in-situ observations over the region in previous studies (Jones et al., 2009; King et al., 2012). The second dataset to represent the Sydney region climate was obtained from the Global Historical Climatological Network (GHCN) database (Menne et al., 2012) and comprised 362 rain gauges distributed within the 2-km domain that had at least 10

years of valid data over the period 1990-2009 (Figure 1c). All the stations used in this study had lower measurement limit of 0.1 mm.

### 3 A methodology for the new paradigm

Existing methods usually perform the bias correction grid-point by grid-point and assume that the model produces too many rain events. However, the respective number of wet days, defined as days with any precipitation registered, tends to decrease with resolution as evidenced by Figure 2 and thus that assumption is unlikely to be valid for the increasingly high-resolution simulations being performed now and in the future. Indeed, the 2-km WRF simulation produces many less rain days than AWAP and therefore using the gridded dataset to correct the 2-km model outputs is problematic. As mentioned before, introducing new rain days to match the observed frequency poses a number of problems (i.e., when to introduce them, what is their intensity, how to keep spatial coherence) that encourages the proposal of alternatives.

Because the number of rain days decreases with increasing resolution the question that arises is: why is station data not used directly to correct very high-resolution model outputs? There are two major obstacles that explain why bias correction has not traditionally been carried out based on in-situ measurements: 1) Spatial and temporal coverage and, 2) Discrepancies in the spatial scale represented by models and stations. Spatial discrepancies are reduced with higher resolution, but it remains a burden when comparing stations and model outputs. The coverage is still an issue regardless of the model spatial resolution and hence a good quality network is necessary. However, even the use of in-situ observation does not guarantee that the model will have higher rain day frequency and the biases in the number of wet days must still be calculated to ensure that the assumption is valid.

A completely new method is not necessary given the large number of bias correction methodologies that have already been proposed and proven to provide satisfactory results. Instead, we suggest here an alternative approach aimed at overcoming the two obstacles above, which consists in adapting an existing method based on histogram equalisation (Piani et al., 2010a,b) to the use of stations as observational reference. This method was chosen among a wide range of

options available because it is widely adopted (Lafon et al., 2013; Piani and Haerter, 2012; Rojas et al., 2011; Schoetter et al., 2012), corrects high moments of the distribution and performs generally better than others (Berg et al., 2012; Teutschbein and Seibert, 2012). Here we call attention to a problem that is likely to emerge in future simulations as resolution increases, and offer a solution.

The original method proposed in Piani et al. (2010a) is a distribution-based algorithm, which assumes that the probability distribution of both the observed and the simulated daily rainfall could be approximated by a theoretical function, a gamma distribution. In particular, the algorithm calculates the cumulative probability from each of the theoretical distributions (i.e., from the model and the observations) at every grid-point. It then corrects each of the modelled rainfall intensities towards the observed value to match their respective cumulative probabilities (Figure 3). Therefore, provided that  $F_m$  and  $F_o$  are the gamma functions that approximate the model and the observations at a particular location, for a given event in the model ( $M^i$ ) the theoretical cumulative probability ( $CP_m^i$ ) is calculated as:

$$CP_m^i = F_m(M^i) \quad (1)$$

then the inverse gamma function of the observations ( $F_o^{-1}$ ) to determine which observed intensity ( $O^i$ ) has the same probability as  $M^i$  in  $F_o$ :

$$O^i = F_o^{-1}(CP_m^i) \quad (2)$$

and  $M^i$  is replaced by  $O^i$  in the bias corrected output.

In this study, the method has been modified such that the 5 nearest stations to each model grid-point are selected to correct its precipitation instead of a gridded dataset. Therefore, for each model grid-point and each day there will be 5 possible corrections ( $O_s^i$  with  $s=1\dots 5$ ) and not only one as occurs in the original method. These 5 corrected values are averaged using an inverse distance squared weighting ( $W_s$ ). The obstacle of not having a unique associated station with a complete timeseries for each of the model locations is hence circumvented. Also, the

stations are aggregated and the spatial scales of the observations and the model are now more comparable. A similar approach was also proposed by Gutjahr and Heinemann (2013).

In addition, the area is divided into different regions (Figure 1c) of climatological affinity that were identified using a multi-step regionalisation (Argüeso et al., 2011). It consists of three successive steps (Principal Component Analysis, an agglomerative clustering and a non-hierarchical clustering) that are applied to daily precipitation. In this case it was applied to AWAP daily precipitation due to its spatial and temporal coverage, which let us identify 5 different regions with similar precipitation characteristics according to the observations. The monthly climatologies of AWAP precipitation averaged over the grid points from each of the regions are illustrated in Figure 4 to show how different their rainfall regimes are, particularly during the first half of the year. A comparison between AWAP and GHCN monthly was also conducted to verify their consistency (supplementary material). Using the regionalisation, we are able to give larger weight to stations that belong to the same region as the model grid-point than those that are likely to have different precipitation regimes. A penalisation factor ( $P_s$ ) of 0.5 is applied to the stations weights when they are located in a region different to that of the grid point; otherwise the factor is 1. The sensitivity of the method to different values of the penalisation (0.1, 0.5 and 0.9) and different number of stations (1, 3, 5, and 7) was conducted finding no major impact on the performance in terms of mean absolute error, but the spatial structure measured through the pattern correlation was better reproduced when using 3 to 7 stations than using a single one (supplementary material).

At every grid point, the corrected value using the adapted method ( $BC^i$ ) for each event in the model ( $M^i$ ) is obtained as:

$$BC^i = \frac{\sum_{s=1}^5 O_s^i * W_s * P_s}{\sum_{s=1}^5 W_s * P_s} \quad (3)$$

The parameters of the gamma distributions are calculated using only rain days, which are usually defined as days with precipitation above a certain thresholds. Here we define any day with precipitation larger than 0.0 mm as wet, although in the case of observations this is equivalent to 0.1 mm day<sup>-1</sup> due to gauge precision. This is certainly not the only possible choice



and other authors selected different thresholds to define wet days from 0 mm (Teutschbein and Seibert, 2012) to 1 mm (Maraun, 2013). However, we decided to use the lowest possible limit in order to include as many rain days as possible. In addition, Berg et al. (2012) found that a similar histogram equalisation method was not sensitive to the choice of the threshold within the 0-1mm range. Modelled rain days are defined in a more flexible way, using a calibrated precipitation threshold as proposed in Schmidli et al. (2006) to adjust the potential excess of wet-day frequencies. Otherwise it is kept to 0 mm day<sup>-1</sup>.

The bias correction was originally designed to use all available values at once and generate a single gamma function for each grid point. The substantial differences in the mechanisms that drive precipitation throughout the year could result in different biases for each of the seasons, which motivated us to apply the bias correction seasonally and thus calculate the gamma parameters for each of the seasons separately. In this study, we applied the bias correction methodology over a single 20-yr period. The aim here was to call the attention to an issue that will become increasingly frequent and put forward a method to reduce its impacts, thus a single period is enough to exemplify the procedure. Future applications of the method will require two calibration periods for calibration and validation purposes.

## 4 Results

The use of stations does not necessarily mean that the model produces more rain days than the observations, since it could be affected by very strong biases that are not compensated by light precipitation events due to spatial averaging or the 'drizzle' effect (Gutowski Jr et al., 2003). Figure 5 shows the seasonal biases in the number of rain days with respect to stations (calculated using the same weighting approach described in the previous section) and AWAP. This figure evidences the step forward in terms of the wet-day frequency disparity problem which is substantially reduced using stations, although there are still regions (region 5) where there are limitations in the bias correction even using in-situ observations. Results in region 5 might be regarded as an example of the method limitations and inadequate corrected values might be expected in this area. The suitability of the observational network or bias correction

purposes and the assessment of the method performance is investigated through comparison of both the original and the bias corrected model outputs with gridded and station observations.

The seasonal deviations of the corrected and non-corrected model outputs with respect to both datasets are illustrated in Figure 6. The biases of the original model outputs show that it overestimates the precipitation induced by orography, generating too much precipitation in the mountains and amplifying the orographic blocking of fronts coming from the ocean, thus leading to underestimation of rainfall towards the interior. This spatial distribution of the biases suggest that the model is overestimating the topographic effect on precipitation at this resolution.

Figure 6 also shows that the bias correction methodology is efficient at seasonal timescales since most of the systematic errors are reduced or even removed with respect to both observational datasets. Indeed, seasonal deviations are reduced to below  $10 \text{ mm month}^{-1}$  over most of the domain. Although there are areas where biases still exist after histogram equalisation, the improvement by the bias correction is noteworthy since the original model estimates were strongly affected by deviations in these areas (e.g, inner west and mountains. The westernmost region is an interesting example because, as it was shown in 5, it is affected by very dry biases and thus the applicability of the method is restricted. However, as already mentioned before, seasonal and longer timescales are still corrected satisfactorily even if the condition of higher number of wet days is not met. For instance, in this region the bias with respect to stations is reduced on average from  $-46.3 \text{ mm month}^{-1}$  to  $-13.5 \text{ mm month}^{-1}$  in winter and from  $-45.2 \text{ mm month}^{-1}$  to  $-21.1 \text{ mm month}^{-1}$  in spring. For the rest of the regions and seasons, the condition fulfilled and the method performs much better providing larger improvements in the seasonal means (Figure 6).

The spatial patterns of the biases with respect to both stations and AWAP are very similar, but the magnitude of the seasonal deviations differs in some areas (e.g., positive biases along the coast in JJA and SON). The agreement between seasonal precipitation from corrected and non-corrected WRF outputs indicates that both observational datasets are appropriate for bias correction of high-resolution models as far as monthly or seasonal timescales are concerned.

However, the differences in the magnitudes indicate that they might not be equally adequate at shorter timescales.

Indeed, impact assessment studies strongly rely on accurate daily precipitation. The correct distribution of rainfall events according to their intensity as well as their occurrence is crucial to evaluate the risks and characterise their possible impact. The probability distribution of precipitation events is examined for AWAP, the stations and the two model outputs to assess the performance of the bias correction at daily timescale and evaluate the potential benefits of using stations to correct high-resolution climate simulations.

The contribution to total precipitation by events of different intensity is used instead of the traditional Probability Distribution Function (PDF). Unlike the PDF, this alternative view of the probability distribution makes it easier to evaluate the relative importance of the errors through the entire rainfall spectrum and includes information relative to the bias in the number of rain days for each of this intensities, which makes it preferable to this particular study.

Figure 7 summarises the contribution from rainfall events of different intensity in the 5 regions. This figure complements the information provided by the monthly climatologies (Figure 4) and emphasises the differences amongst regions. Also, the comparison between distributions from observational datasets yields important differences in all regions, especially for precipitation events below  $10 \text{ mm day}^{-1}$ , which are systematically overestimated by AWAP. As for more intense events, AWAP tends to underestimate their contribution to total precipitation in most regions; though, in the northeast (region 3) there is a clear overestimation. These differences are related to the difference in the spatial scales the observational products represent and suggest that AWAP, and more generally the observation-based grids, are not suitable to correct model outputs with finer resolution. Maraun (2013) found that the use of stations to correct coarser RCMs (25km-by-25km) tend to inflate the variability of the model. We have looked into the quantile-quantile distribution, which is basically an alternative representation of the probability distribution, to determine whether this occurs in our higher resolution simulation and found no evidence of such an artificial inflation of the variability (supplemental material).

In most regions, WRF produces too much light precipitation ( $0\text{-}2 \text{ mm day}^{-1}$ ), underestimates moderate events ( $2\text{-}20 \text{ mm day}^{-1}$ ) and generates too many extreme events ( $>20 \text{ mm day}^{-1}$ ),

which altogether results in the positive biases shown in Figure 6. The behaviour of the model is different in region 5, which was already identified as a problematic region and represents an example of the method limitations. Unlike seasonal biases, the daily precipitation distribution is not corrected adequately with this method due to the negative biases in the number of rain days and both the original and the corrected model outputs show a persistent negative deviation across the entire spectrum of events, except for the very light events that are also overestimated. In this region, the validity of the assumption of a wetter model is compromised and thus the method might produce inaccurate results.

However, while the bias correction is not able to completely remove the errors, particularly in regions 4 and 5, it succeeds in providing a much better representation of the events distribution compared to the stations, which is a good indicator of the method's skills. To quantify this improvement, the similarity between different PDFs was measured using the skill score (SS) proposed in Perkins et al. (2007), which calculates the common area shared by two PDFs. The SS confirms that the bias correction significantly improves the rain events distribution in the model over the entire domain. Ordered by regions, the non-corrected model outputs and the in-situ observations share 80.3%, 70.2%, 74.0%, 76.1% and 54.5% of their precipitation PDFs, whereas the bias correction increases these percentages to 97.1%, 95.1%, 96.7%, 96.6% and 94.0%, respectively. The improvement is generally observed over the entire spectrum of intensities, but it might also be partly explained by rainfall events in the range 0-0.1 mm/day that were not captured in the observations and did exist in the model outputs.

In the station dataset, the precipitation with the largest contribution occur in the range between 2-6 mm day<sup>-1</sup>, whereas rain events below 2 mm day<sup>-1</sup> make a smaller contribution. This characteristic of rainfall distribution usually goes unnoticed in RCMs and is not captured by gridded datasets either, but it is better reproduced in the bias-corrected model outputs. This is a feature of daily precipitation that could play an important role in the hydrological cycle and thus represents a noteworthy improvement.

## 5 Conclusions

Bias correction has traditionally relied on the assumption that models produce more rain days than the reference observations, which are usually gridded datasets due to their spatial and temporal characteristics. However, climate simulations are currently being completed at spatial resolutions that make this assumption no longer valid. A histogram equalisation method (Piani et al., 2010a) was adapted to be used with stations, which are not subjected to more frequent low intensity precipitation due to spatial averaging. Although the use of stations do not completely overcome the bias correction limitations in areas of very strong dry biases, it makes the wet-day condition more likely to be fulfilled and thus significantly lessens the problem of rain-day frequency disparities. The stations were aggregated to bypass the two major obstacles for their use in bias correction, the differences in the model and stations spatial scales, and the completeness and sparseness of the timeseries.

The method has been proven to substantially reduce the seasonal biases of precipitation when compared to both gridded and station datasets, even in areas where the wet-day condition was not met. Gridded datasets are also appropriate to correct high-resolution model at seasonal or even monthly timescales, but it has been shown here that they are not adequate to correct daily features of precipitation anywhere in the domain. Indeed, the major contribution of this study is the efficient bias correction of the daily precipitation probability distributions of very high-resolution models. A much better representation of the frequency of the rainfall events is achieved after bias correction for all regions, especially in those where rainfall is overestimated. In areas where biases are markedly dry, the method provides an improvement but the results are not as good due to differences in the wet-day occurrence. In general, the relative importance of moderate event with respect to very light ones is also better reproduced, which could also have important implications for impact assessment studies. We acknowledge that the availability of a high-quality observation network is required to apply this bias correction method, but the generation of reliable gridded datasets also need such a network. In this study, we have addressed issues related to precipitation, which is one of the most problematic variables in terms of bias

correction, but the method is applicable to other variables using different fitting functions and thus reduce biases due to misrepresentation of local features such as the orography.

*Acknowledgements.* This work was made possible by funding from the NSW Environment Trust (RM08603), as well as the NSW Office of Environment and Heritage, and the Australian Research Council as part of the Future Fellowship FT110100576. This work was supported by an award under the Merit Allocation Scheme on the NCI National Facility at the ANU.

## References

Argüeso, D., Hidalgo-Muñoz, J. M., Gámiz-Fortis, S. R., Esteban-Parra, M. J., Castro-Díez, Y., and Dudhia, J.: Evaluation of WRF parameterizations for climate studies over Southern Spain using a multi-step regionalization, *Journal of Climate*, 24, 5633–5651, 2011.

Argüeso, D., Evans, J. P., Fita, L., and Bormann, K. J.: Temperature response to future urbanization and climate change, *Climate Dynamics*, 2013.

Berg, P., Feldmann, H., and Panitz, H. J.: Bias correction of high resolution RCM data, *Journal of Hydrology*, 448-449, 80–92, 2012.

Bordoy, R. and Burlando, P.: Bias Correction of Regional Climate Model Simulations in a Region of Complex Orography, *Journal of Applied Meteorology and Climatology*, 52, 82–101, 2013.

Chen, C., Haerter, J. O., Hagemann, S., and Piani, C.: On the contribution of statistical bias correction to the uncertainty in the projected hydrological cycle, *Geophysical Research Letters*, 38, n/a–n/a, 2011.

Christensen, J. H., Boberg, F., Christensen, O. B., and Lucas-Picher, P.: On the need for bias correction of regional climate change projections of temperature and precipitation, *Geophysical Research Letters*, 35, 2008.

Déqué, M., Rowell, D. P., Lüthi, D., Giorgi, F., Christensen, J. H., Rockel, B., Jacob, D., Kjellstrom, E., de Castro, M., and Hurk, B. V. D.: An intercomparison of regional climate simulations for Europe: assessing uncertainties in model projections, *Climatic Change*, 81, 53–70, 2007.

Di Luca, A., de Elia, R., and Laprise, R.: Potential for added value in precipitation simulated by high-resolution nested Regional Climate Models and observations, *Climate Dynamics*, 38, 1229–1247, 2011.

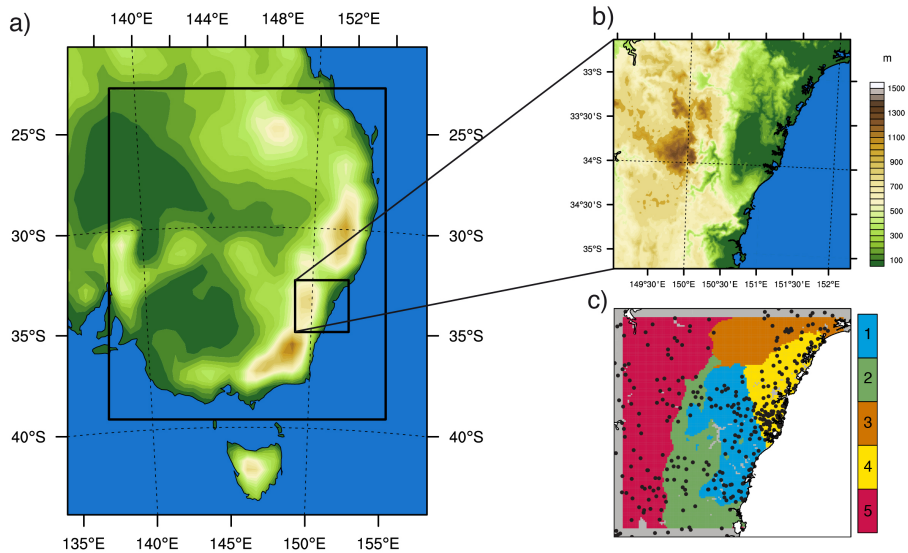
Ehret, U., Zehe, E., Wulfmeyer, V., Warrach-Sagi, K., and Liebert, J.: Should we apply bias correction to global and regional climate model data?, *Hydrology and Earth System Sciences*, 16, 3391–3404, 2012.

- Evans, J. P. and McCabe, M.: Regional climate simulation over Australia's Murray-Darling basin: A multitemporal assessment, *Journal of Geophysical Research*, 115(D14114), 2010.
- Evans, J. P. and McCabe, M. F.: Effect of model resolution on a regional climate model simulation over southeast Australia, *Climate Research*, 56, 131–145, 2013.
- 5 Evans, J. P. and Westra, S.: Investigating the Mechanisms of Diurnal Rainfall Variability Using a Regional Climate Model, *Journal of Climate*, 25, 7232–7247, 2012.
- Feser, F., Rockel, B., von Storch, H., Winterfeldt, J., and Zahn, M.: Regional Climate Models Add Value to Global Model Data: A Review and Selected Examples, *Bulletin Of The American Meteorological Society*, 92, 1181–1192, 2011.
- 10 Giorgi, F.: Regional climate modeling: Status and perspectives, *J. Phys. IV*, 139, 101–118, 2006.
- Gutjahr, O. and Heinemann, G.: Comparing precipitation bias correction methods for high-resolution regional climate simulations using COSMO-CLM, *Theoretical and Applied Climatology*, doi:10.1007/s00704-013-0834-z, 2013.
- Gutowski Jr, W., Decker, S., Donavon, R., Pan, Z., Arritt, R., and Takle, E.: Temporal-spatial scales of observed and simulated precipitation in central U.S. climate, *Journal of Climate*, 16, 3841–3847, 2003.
- 15 Haerter, J. O., Hagemann, S., Moseley, C., and Piani, C.: Climate model bias correction and the role of timescales, *Hydrology and Earth System Sciences*, 15, 1065–1079, 2011.
- Hay, L., Wilby, R., and Leavesley, G.: A comparison of delta change and downscaled GCM scenarios for three mountainous basins in the United States, *Journal of the American Water Resources Association*, 36, 387–397, 2000.
- 20 Haylock, M. R., Hofstra, N., Tank, A. M. G. K., Klok, E. J., Jones, P. D., and New, M.: A European daily high-resolution gridded data set of surface temperature and precipitation for 1950–2006, *Journal of Geophysical Research-Atmospheres*, 113, 1–12, 2008.
- 25 Jones, D. A., Wang, W., and Fawcett, R.: High-quality spatial climate data-sets for Australia, *Australian Meteorological and Oceanographic Journal*, 58, 233–248, 2009.
- King, A. D., Alexander, L. V., and Donat, M. G.: The efficacy of using gridded data to examine extreme rainfall characteristics: a case study for Australia, *International Journal of Climatology*, 33, 2376–2387, 2012.
- 30 Lafon, T., Dadson, S., Buys, G., and Prudhomme, C.: Bias correction of daily precipitation simulated by a regional climate model: a comparison of methods, *International Journal of Climatology*, 33, 1367–1381, 2013.

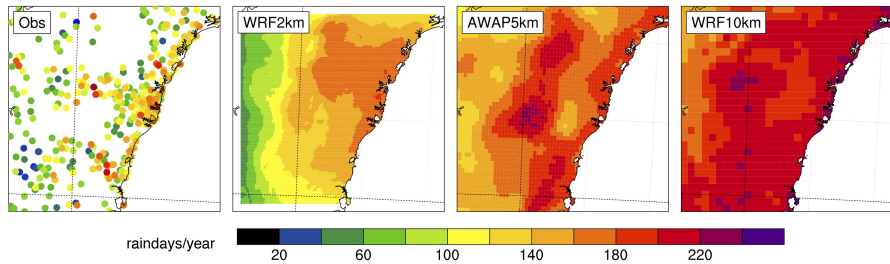
- Lenderink, G., Buishand, A., and Van Deursen, W.: Estimates of future discharges of the river Rhine using two scenario methodologies: direct versus delta approach, *Hydrology and Earth System Sciences*, 11, 1145–1159, 2007.
- Maraun, D.: Bias Correction, Quantile Mapping, and Downscaling: Revisiting the Inflation Issue, *Journal of Climate*, 26, 2137–2143, doi:10.1175/JCLI-D-12-00821.1, 2013.
- Maraun, D., Wetterhall, F., Ireson, A. M., Chandler, R. E., Kendon, E. J., Widmann, M., Brienen, S., Rust, H. W., Sauter, T., Themeßl, M., Venema, V. K. C., Chun, K. P., Goodess, C. M., Jones, R. G., Onof, C., Vrac, M., and Thiele-Eich, I.: Precipitation downscaling under climate change: Recent developments to bridge the gap between dynamical models and the end user, *Reviews of Geophysics*, 48, RG3003, 2010.
- Menne, M. J., Durre, I., Vose, R. S., Gleason, B. E., and Houston, T. G.: An Overview of the Global Historical Climatology Network-Daily Database, *Journal of Atmospheric and Oceanic Technology*, 29, 897–910, 2012.
- Perkins, S. E., Pitman, A. J., Holbrook, N. J., and McAneney, J.: Evaluation of the AR4 climate models' simulated daily maximum temperature, minimum temperature, and precipitation over Australia using probability density functions, *Journal of Climate*, 20, 4356–4376, 2007.
- Piani, C. and Haerter, J. O.: Two dimensional bias correction of temperature and precipitation copulas in climate models, *Geophysical Research Letters*, 39, L20401, 2012.
- Piani, C., Haerter, J., and Coppola, E.: Statistical bias correction for daily precipitation in regional climate models over Europe, *Theoretical and Applied Climatology*, 99, 187–192, 2010a.
- Piani, C., Weedon, G. P., Best, M., Gomes, S. M., Viterbo, P., Hagemann, S., and Haerter, J. O.: Statistical bias correction of global simulated daily precipitation and temperature for the application of hydrological models, *Journal of Hydrology*, 395, 199–215, 2010b.
- Portoghese, I., Bruno, E., Guyennon, N., and Iacobellis, V.: Stochastic bias-correction of daily rainfall scenarios for hydrological applications, *Natural Hazards and Earth System Science*, 11, 2497–2509, 2011.
- Rojas, R., Feyen, L., Dosio, A., and Bavera, D.: Improving pan-European hydrological simulation of extreme events through statistical bias correction of RCM-driven climate simulations, *Hydrology and Earth System Sciences*, 15, 2599–2620, 2011.
- Schmidli, J., Frei, C., and Vidale, P. L.: Downscaling from GCM precipitation: a benchmark for dynamical and statistical downscaling methods, *International Journal of Climatology*, 26, 679–689, 2006.



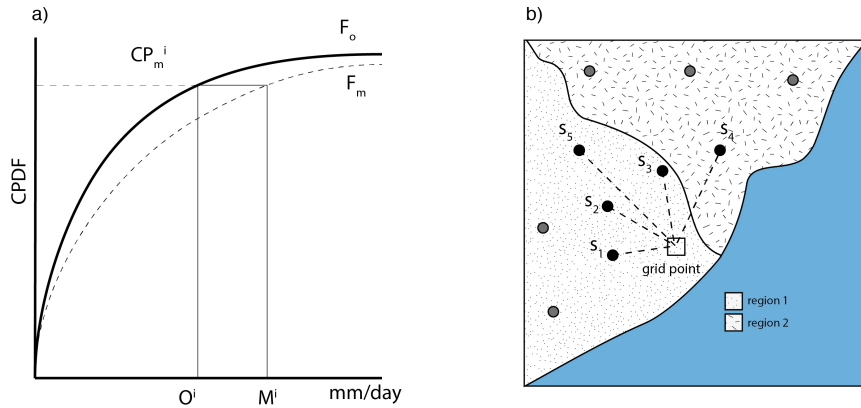
- Schoetter, R., Hoffmann, P., Rechid, D., and Schlünzen, K. H.: Evaluation and Bias Correction of Regional Climate Model Results Using Model Evaluation Measures, *Journal of Applied Meteorology and Climatology*, 51, 1670–1684, 2012.
- 5 Skamarock, W. C., Klemp, J. B., Dudhia, J., Gill, D. O., Barker, D. M., Duda, M. G., Huang, X.-Y., Wang, W., and Powers, J. G.: A Description of the Advanced Research WRF Version 3, NCAR/TN-475+STR NCAR TECHNICAL NOTE, p. 125, 2009.
- Terink, W., Hurkmans, R. T. W. L., Torfs, P. J. J. F., and Uijlenhoet, R.: Evaluation of a bias correction method applied to downscaled precipitation and temperature reanalysis data for the Rhine basin, *Hydrology and Earth System Sciences*, 14, 687–703, 2010.
- 10 Teutschbein, C. and Seibert, J.: Bias correction of regional climate model simulations for hydrological climate-change impact studies: Review and evaluation of different methods, *Journal of Hydrology*, 456–457, 12–29, 2012.
- Themeßl, M. J., Gobiet, A., and Leuprecht, A.: Empirical-statistical downscaling and error correction of daily precipitation from regional climate models - Jakob Themeßl - 2010 - *International Journal of Climatology* - Wiley Online Library, *International Journal of Climatology*, 31, 1530–1544, 2011.
- 15 Themeßl, M. J., Gobiet, A., and Heinrich, G.: Empirical-statistical downscaling and error correction of regional climate models and its impact on the climate change signal, *Climatic Change*, 112, 449–468, 2012.
- Thompson, G., Field, P. R., Hall, W. D., and Rasmussen, R. M.: A new bulk microphysical parameterization for WRF (& MM5), *Proc. Seventh Weather Research and Forecasting Model Workshop*, pp. 1–11, 2006.
- 20 Tselioudis, G., Douvis, C., and Zerefos, C.: Does dynamical downscaling introduce novel information in climate model simulations of precipitation change over a complex topography region?, *International Journal of Climatology*, 32, 1572–1578, 2012.



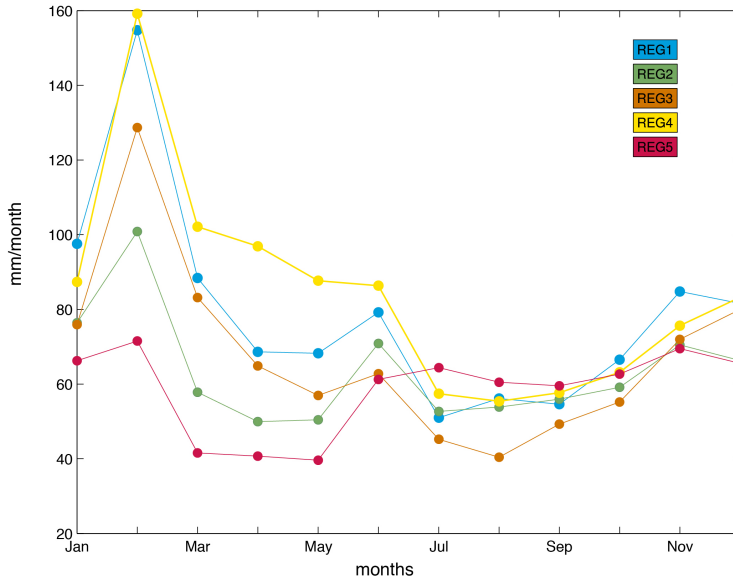
**Fig. 1.** (a) Topography and location of all domains of the simulation, (b) topography and extension of the inner domain, and (c) location of the stations (black dots) and the 5 different precipitation regions (colored areas) within the model domain.



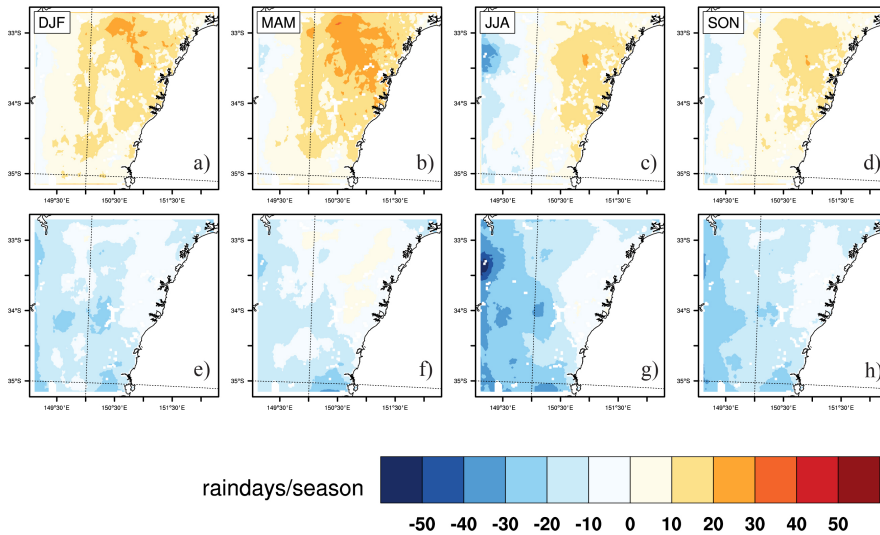
**Fig. 2.** Annual mean number of rain days over the period 1990-2009 for stations, the WRF simulation at 2-km resolution, the AWAP dataset and the intermediate WRF domain at 10-km resolution used to provide the boundary conditions.



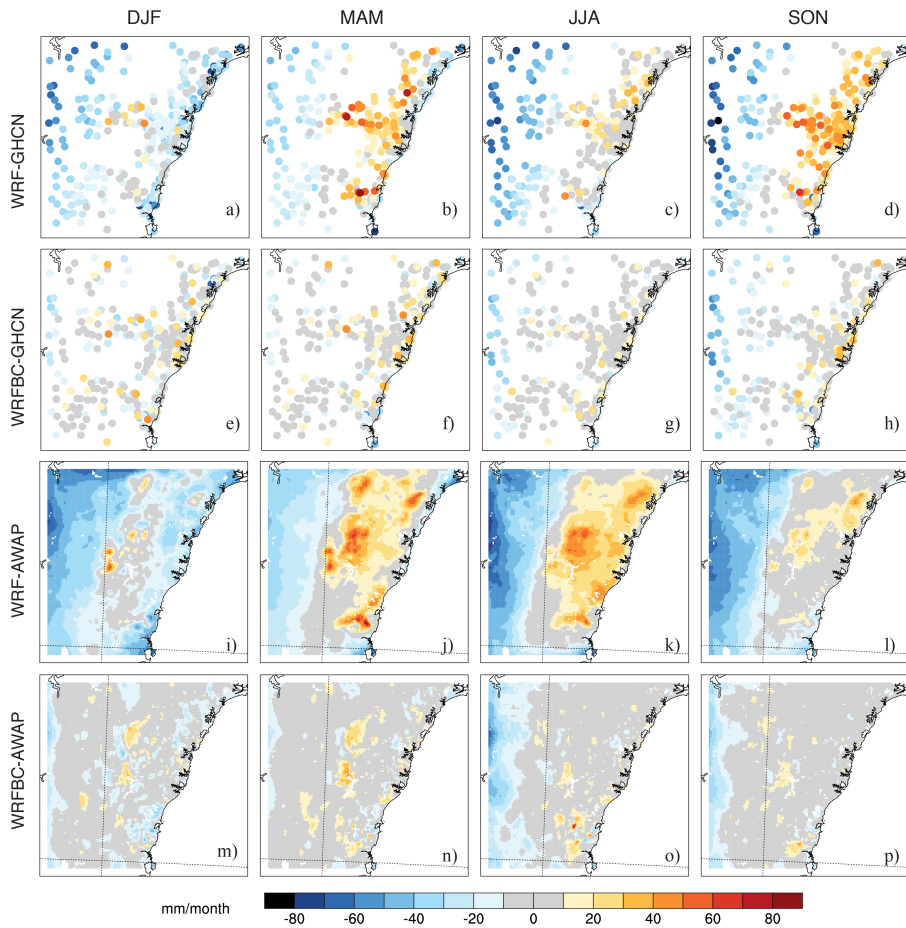
**Fig. 3.** (a) Schematic of the bias correction proposed by Piani et al. (2010a).  $M_i$  is the intensity of an event in the model and  $O_i$  is intensity of an observed event with the same cumulative probability ( $CP_m^i$ ) as defined by  $F_m$  and  $F_o$ , which are the cumulative probability functions for the model and the observations. (b) Schematic of the adaptation of the bias-correction method using stations and regions.



**Fig. 4.** Monthly climatologies of precipitation for each of the regions obtained by averaging all AWAP grid points that belong to each of the divisions.

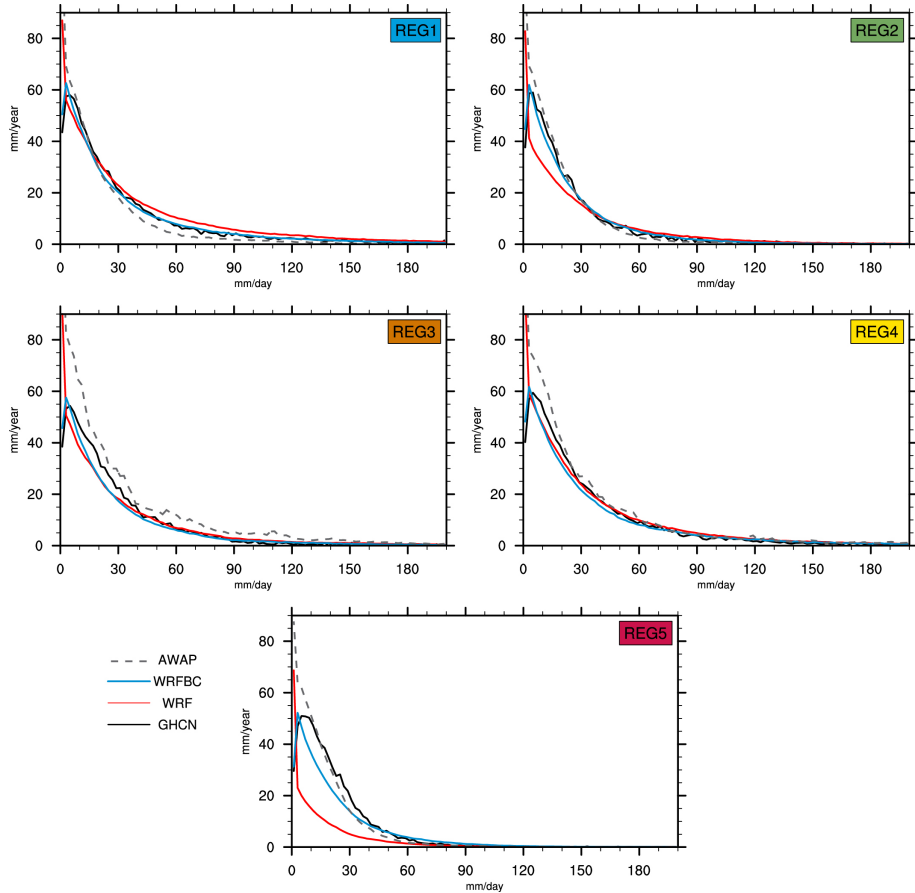


**Fig. 5.** Seasonal biases in the number of wet days in the WRF simulation at 2km with respect to GHCN (a-d) and AWAP (e-h).



**Fig. 6.** Precipitation seasonal biases of non-corrected WRF with respect to the stations (a-d), bias-corrected WRF with respect to the stations (e-h), non-corrected WRF with respect to AWAP (i-l) and bias-corrected WRF with respect to AWAP (m-p) over the period 1990-2009.





**Fig. 7.** Contribution to total annual precipitation by rainfall events of different intensity in the 5 precipitation regions for AWAP, bias-corrected and non-corrected WRF outputs, and GHCN stations.