Martínková, M., 2026. Bias correction of daily air temperature from ALADIN/CNRM-ESM2-1 over the Czech Republic. In: Tolasz, R., Šustková, V. (Ed.), Proceedings of the 2nd PERUN Project Conference (TA ČR, SS02030040). CHMI, Praha, pp. 46–53, ISBN 978-80-7653-095-9.
https://doi.org/10.59984/978-80-7653-095-9.06

# Bias correction of daily air temperature from ALADIN/CNRM-ESM2-1 over the Czech Republic

## (Korekce biasu teploty z modelu ALADIN/CNRM-ESM2-1 pro území ČR)

Martínková, M.

Marta Martínková, Institute of Atmospheric Physics CAS, marta@ufa.cas.cz

https://doi.org/10.59984/978-80-7653-095-9.06

***Abstract:*** Daily air-temperature outputs from the ALADIN regional climate model driven by CNRM-ESM2-1 were post-processed within the PERUN project to assess three univariate bias-correction approaches relevant for future SSP5-8.5 applications: mean bias correction, empirical quantile mapping and quantile delta mapping. The methods were evaluated on the 1990–2014 historical overlap against synthetic station-observation time series aggregated to the model grid. The raw simulation reproduces the domain mean warming trend well, with annual trends of 0.359 °C·decade$^{-1}$ in observations and 0.374 °C·decade$^{-1}$ in the raw model. However, grid-cell trend differences are asymmetric and seasonal discrepancies are substantial, especially in late winter, spring, early summer and December. Distributional diagnostics show pronounced month-dependent biases in the median and tails of daily temperature. Mean bias correction efficiently adjusts the central tendency but leaves residual tail errors. Quantile mapping and quantile delta mapping reduce the empirical biases to near zero in the reference period diagnostics. For climate change applications, quantile delta mapping is the preferred compromise because it combines distributional correction with explicit preservation of modelled changes in quantiles.

***Keywords:*** bias correction – quantile mapping – quantile delta mapping – ALADIN – Czech Republic – daily temperature

***Abstrakt:*** Teplota vzduchu jako výstup z regionálního klimatického modelu ALADIN řízeného modelem CNRM-ESM2-1 byla v rámci projektu PERUN následně zpracována s cílem posoudit tři jednorozměrné přístupy ke korekci systematické chyby relevantní pro budoucí aplikace scénáře SSP5-8.5: korekci průměrné odchylky, empirické kvantilové mapování a kvantilové delta mapování. Metody byly hodnoceny pro historické překryvné období 1990–2014 vůči syntetickým časovým řadám staničních pozorování agregovaným do modelové mřížky. Nekorigovaná simulace dobře reprodukuje trend oteplování průměrovaný přes celé území, přičemž roční trendy činí 0,359 °C za dekádu v pozorováních a 0,374 °C za dekádu v nekorigovaném modelu. Rozdíly trendů na úrovni jednotlivých gridových buněk jsou však asymetrické a sezónní odchylky jsou značné, zejména na konci zimy, na jaře, na začátku léta a v prosinci. Distribuční diagnostika ukazuje výrazné, měsíčně proměnlivé odchylky v mediánu i v chvostech rozdělení denní teploty. Korekce průměrné odchylky účinně upravuje centrální tendenci, avšak ponechává reziduální chyby v chvostech rozdělení. Kvantilové mapování a kvantilové delta mapování snižují empirické odchylky v diagnostikách referenčního období téměř na nulu. Pro aplikace v oblasti změny klimatu představuje kvantilové delta mapování preferovaný kompromis, protože kombinuje korekci rozdělení s explicitním zachováním modelovaných změn v kvantilech.

***Klíčová slova:*** korekce biasu – kvantilové mapování – kvantilové delta mapování – ALADIN – Česká republika – denní teplota

## Introduction

Regional climate-model output is the main source of physically consistent climate information at spatial and temporal scales required by impact studies, but it is rarely used directly. Systematic errors arise from the driving global model, limited spatial resolution, parameterized processes and representation errors between a grid-cell average and local observations. These errors affect not only the average temperatures but also seasonal variability and distribution tails, which are important for heat stress, frost, ecosystem and hydrological applications.

Bias correction is therefore a standard post-processing step. Its purpose is not to improve the dynamics of the climate model but to transform simulated time series so that selected statistical properties agree with a reference observational data set. Reviews emphasize that the usefulness of such correction depends on the credibility of the modelled climate-change signal and on the assumption that the diagnosed model error is transferable beyond the reference period (Maraun 2016; Maraun et al. 2017). Recent work also shows that downscaling and bias correction themselves can add substantial uncertainty to local climate projections (Lafferty and Sriver 2023).

This contribution focuses on the ALADIN/CNRM-ESM2-1 temperature simulation prepared in the PERUN project. ALADIN is a limited-area modelling system widely used in weather and climate applications (Termonia et al. 2018), while CNRM-ESM2-1 is a CMIP6-generation Earth-system model that represents coupled atmosphere-ocean-land-carbon-cycle processes (Séférian et al. 2019). The objective is to compare three methods that are frequently used in climate-impact workflows: mean bias correction, standard quantile mapping (QM) and quantile delta mapping (QDM).

The key question is practical. If future daily temperature series under SSP5-8.5 are to be corrected for downstream impact models in the Czech Republic, which method gives the best balance between correcting the historical distribution and preserving the raw model's projected warming signal? To investigate this question, we evaluate annual trends, grid-cell trends, monthly trends and distributional diagnostics for the lower tail, median and upper tail of daily temperature.

## Data

The analysis uses daily air-temperature time series from the ALADIN regional simulation driven by the CNRM-ESM2-1 global model. The project processing chain is designed for future SSP5-8.5 time series, while the present diagnostic evaluation uses the historical overlap 1990–2014, where model output can be compared with synthetic station-observation time series. All temperatures are treated in degrees Celsius, and all corrections are performed as additive corrections.

The processing pipeline works on complete grid-cell time series rather than on time slices. This is important because quantile-based corrections require the empirical distribution within each grid cell and calendar month. Observations and model data are standardised to date, temperature and grid identifiers, matched on common dates, grouped by grid cell and month, corrected by the three methods and then evaluated over the full overlapping reference period. The evaluated metrics include mean bias, MAE, RMSE, correlation, standard-deviation ratio and biases of the 0.05, 0.50 and 0.95 quantiles.

## Methods

The processing pipeline works on complete grid-cell time series rather than on time slices. This is important because quantile-based corrections require the empirical distribution within each grid cell and calendar month. Observations and model data are standardised to date, temperature and grid identifiers, matched on common dates, grouped by grid cell and month, corrected by the three methods and then evaluated over the full overlapping reference period. The evaluated metrics include mean bias, MAE, RMSE, correlation, standard-deviation ratio and biases of the 0.05, 0.50 and 0.95 quantiles.

The overview of applied methods is presented in Tab. 1. All three approaches are applied independently for each grid cell and calendar month.

*Table 1. Bias-correction methods compared in this study.*

| **Method** | Strength | Main caveat |
|---|---|---|
| Mean bias correction | Simple and straightforward; removes mean offset. | Does not alter distribution shape or variance. |
| Quantile mapping | Corrects the full empirical distribution in the reference period. | May distort climate-change signal outside the reference period. |
| Quantile delta mapping | Corrects distribution while preserving modelled quantile changes. | Still assumes that reference-period bias information is relevant for future periods. |

Mean bias correction estimates one additive offset from the reference overlap and applies it to all daily values in the corresponding grid cell and month. It is robust and easy to communicate, but it cannot correct a model whose cold tail, warm tail or variance is wrong. It can therefore be sufficient for applications that depend mainly on monthly or annual means, but it is less suitable for threshold indices and extremes.

Empirical quantile mapping transforms each modelled value according to its percentile in the modelled reference distribution and then replaces it with the observed value at the same percentile. This approach is effective when the aim is to match the reference distribution, and it is widely used in climate-impact studies. However, because the future value is mapped through a fixed historical transfer function, the resulting future change may differ from the raw model change, particularly in distribution tails (Cannon et al. 2015; Tong et al. 2021).

Quantile delta mapping modifies QM by retaining the modelled change at each percentile. For additive variables such as temperature, the correction adds the modelled quantile anomaly relative to the modelled reference distribution to the observed reference quantile. This makes QDM and related change-preserving bias-adjustment methods especially relevant where the corrected series must be consistent with the modelled climate-change signal (Cannon et al. 2015, Lange 2019, Switanek et al. 2017).

Linear trends are estimated by ordinary least squares (OLS) regression and expressed in °C per decade. Distributional performance is evaluated using density curves, quantile-quantile plots and monthly signed biases of the median, Q05 and Q95. The use of the same full historical overlap for estimating and evaluating the correction means that the distributional results should be interpreted only as an in-sample diagnostic. This is useful for diagnosing method behaviour but should be complemented by split-sample or cross-validation tests before operational use.

## Discussion

At the domain-mean annual scale, the raw model closely matches the observed warming trend over 1990–2014. The observed annual mean temperature trend for 1990–2014 is 0.359 °C·decade$^{-1}$, while the raw ALADIN/CNRM-ESM2-1 trend is 0.374 °C·decade$^{-1}$. The difference of 0.015 °C·decade$^{-1}$ is small compared with interannual variability, and it indicates that the modelled large-scale warming rate is plausible for the historical reference period (Fig. 1).

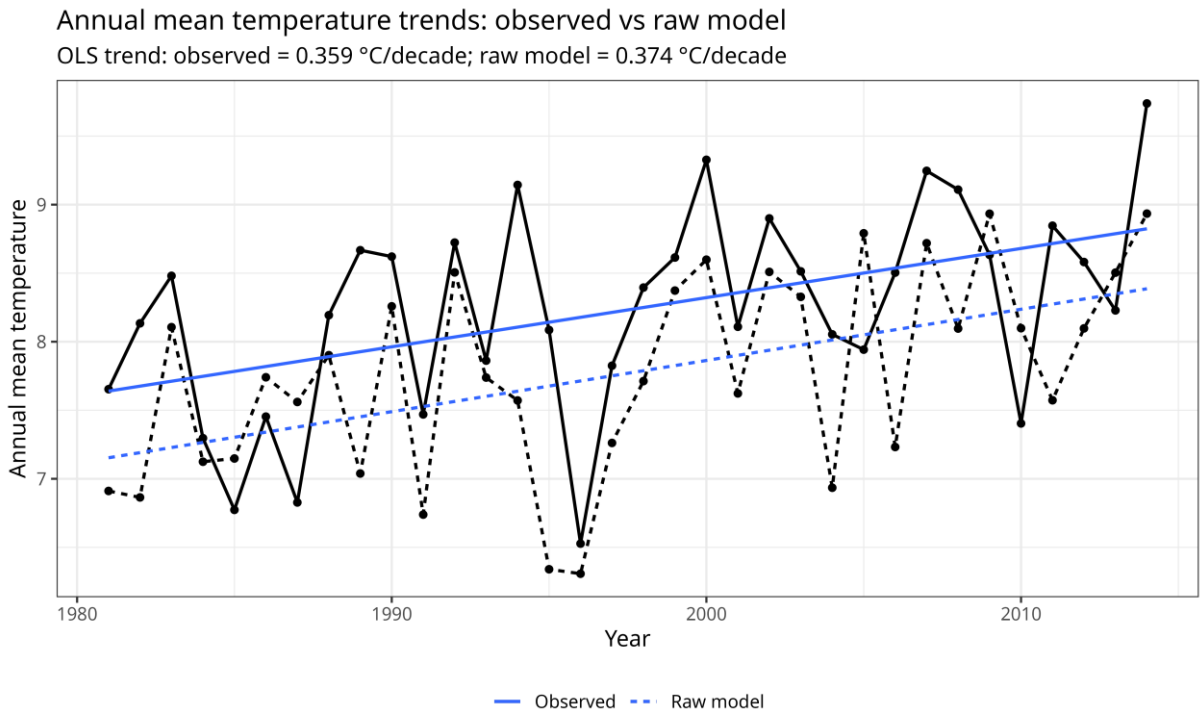



*Fig. 1. Observed and raw-model annual mean temperature trends over the evaluation period. The raw simulation reproduces the domain-mean trend closely, but interannual differences remain substantial.*

The apparent agreement in the domain mean does not imply that the model reproduces local trends everywhere. Grid-cell diagnostics reveal a broader distribution of raw-minus-observed trend differences, with most values near zero but an asymmetric positive tail. In practical terms, some grid cells show an overestimate of historical warming even though the spatial mean trend over the whole domain is close to observations (Fig. 2).

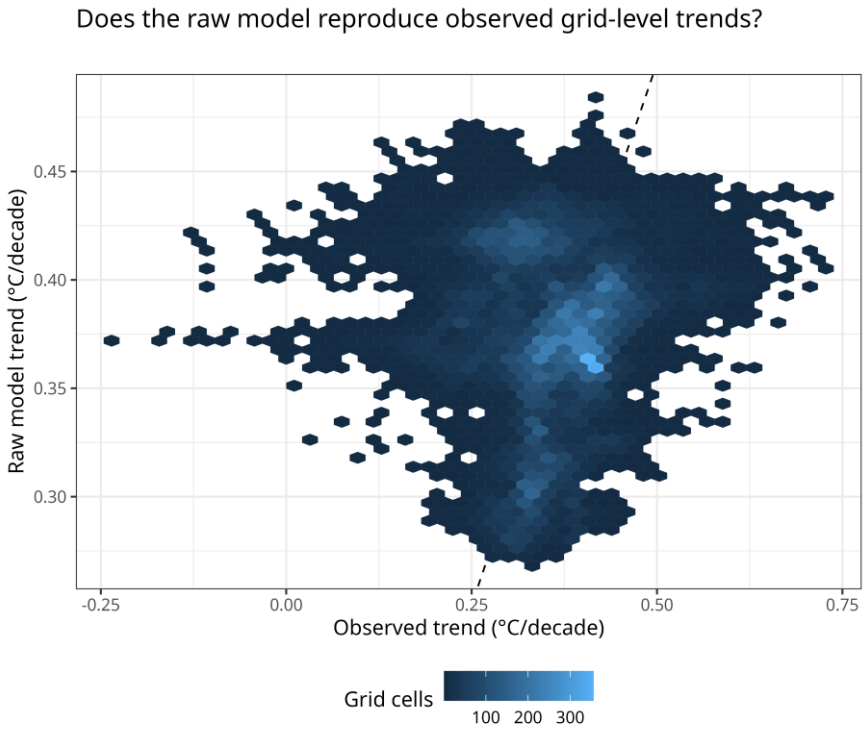



*Fig. 2. Grid-cell comparison of annual mean temperature trends in the observational reference and the raw ALADIN/CNRM-ESM2-1 simulation over 1990–2014. Trends are estimated by ordinary least-squares regression and expressed in °C decade⁻¹. The colour scale shows the number of grid cells in each bin, and the dashed line marks the one-to-one relationship. Values above the line indicate grid cells where the raw model overestimates the observed warming trend, whereas values below the line indicate underestimated warming.*

Monthly trends show even stronger structure. Both observed and raw model series are generally warming, but the correspondence varies sharply among months. The largest mismatches occur in February, March, April, June, October and December. In March, the raw model shows a negative trend while observations show a positive trend, whereas in February and December the model trend exceeds the observed trend (Fig. 3).

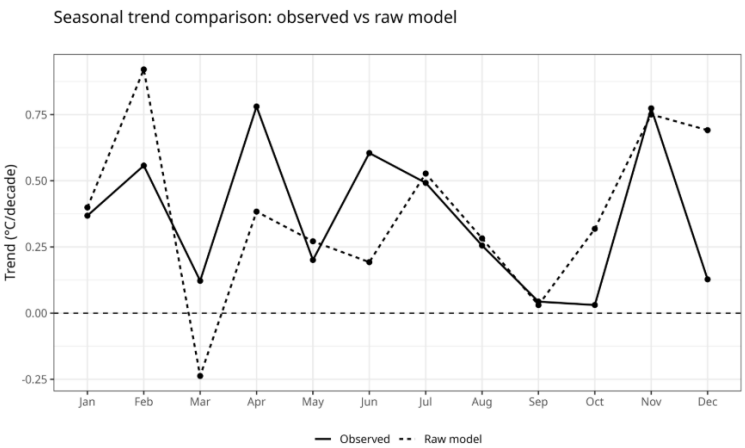



*Fig. 3. Monthly trend comparison. Seasonal discrepancies are substantially larger than the annual-domain mean difference, demonstrating the need for month-specific evaluation.*

The raw model also has clear distributional biases. Median temperatures are close to observations in January but become too cold in spring and early summer, with median biases of approximately −0.81 °C

in March, −1.54 °C in April, −2.18 °C in May and −1.53 °C in June. The upper tail is also too cold in several months; for example, the raw Q95 bias reaches approximately −3.40 °C in June. These errors matter for derived temperature indices because a mean correction alone can leave threshold exceedances and tail behaviour biased.

Mean bias correction removes much of the central offset but leaves residual tail errors. This behaviour is expected because the method shifts the entire distribution by one monthly offset. For example, after mean bias correction the cold tail remains too warm in some winter months and the warm tail remains too cold in June and July. The method is therefore attractive as a transparent baseline but not as a complete correction for daily temperature distributions.

Quantile mapping and QDM almost eliminate the empirical biases in the Q05, median and Q95 during the reference overlap. The monthly signed biases for QM and QDM are approximately zero across the displayed quantiles. The Q-Q plots confirm this result: the QM and QDM curves lie nearly on the one-to-one line, while the raw model and mean-bias-corrected series retain visible deviations in the distribution tails (Fig. 4).

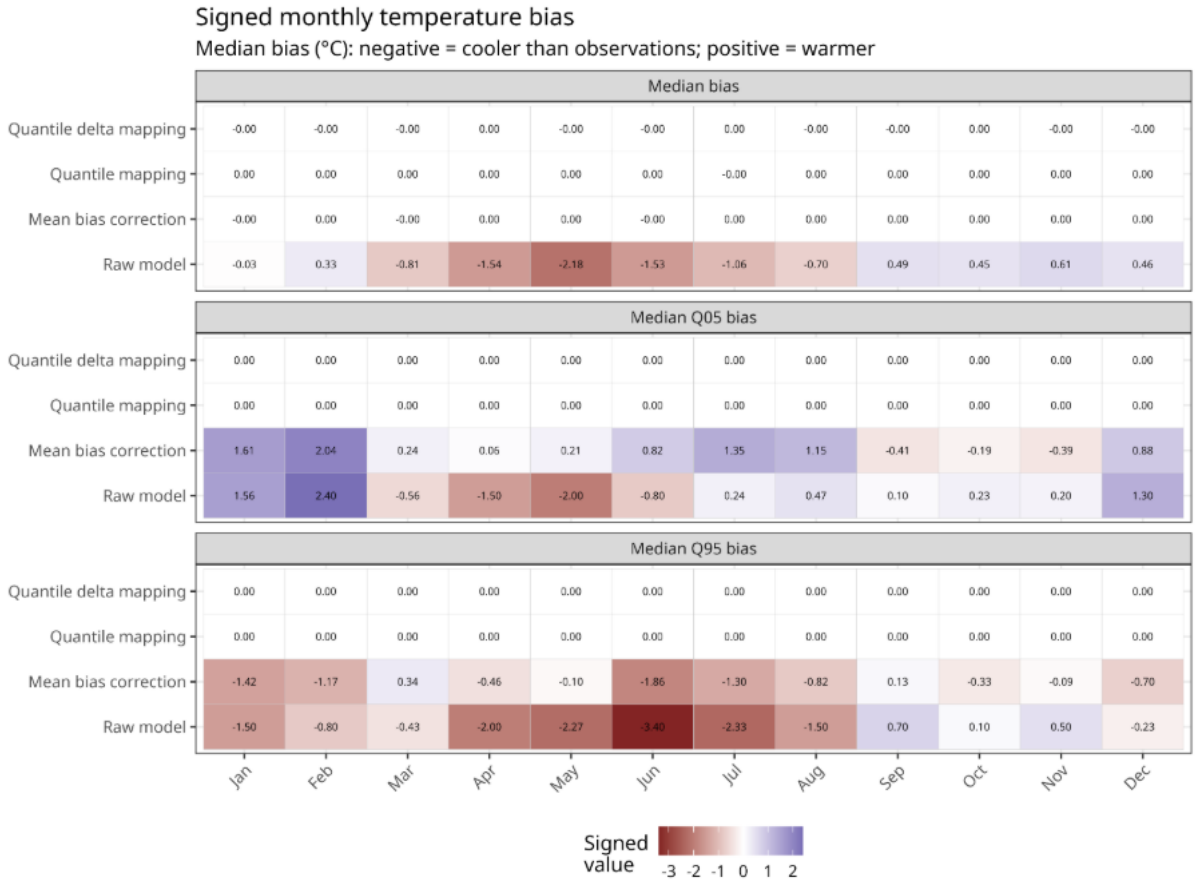



*Fig. 4. Monthly signed temperature biases for the raw model and three correction methods. Mean bias correction improves the centre of the distribution, whereas QM and QDM align the empirical quantiles in the reference period.*

For future SSP5-8.5 projections, the method choice is not determined only by historical fit. Standard QM gives excellent historical distributional agreement, but it can modify the modelled climate-change signal because future values are forced through a transfer function derived from the historical overlap. This limitation is not only theoretical: previous studies have shown that QM can alter changes in quantiles and extremes, while QDM was introduced to preserve such changes more explicitly (Cannon et al. 2015; Tong et al. 2021).

Among the three tested univariate methods, QDM is the most suitable default for future SSP5-8.5 daily-temperature series in this application, because it combines reference-period distributional adjustment with explicit preservation of modelled quantile changes. This recommendation remains conditional on the plausibility of the raw model's projected temperature-change signal and should be complemented by out-of-sample validation. It retains the strengths of empirical distribution correction and is better aligned with the central requirement of climate-change impact studies: corrected future time series should not erase or artificially reshape the modelled warming signal. This does not mean that QDM is universally optimal. It remains a univariate marginal correction and does not by itself correct temporal persistence, spatial dependence or physical inconsistencies among variables. Multivariate approaches can address some of these issues but may also introduce trade-offs between marginal fit and dependence structure (François et al. 2020).

The presented diagnostics use the full 1990–2014 overlap to estimate and evaluate the transfer functions. This design is appropriate for a whole-period comparison and maximizes the sample size within each month, but it cannot quantify out-of-sample robustness.

## Conclusion

The ALADIN/CNRM-ESM2-1 historical simulation closely reproduces the domain-mean annual warming trend over the Czech Republic during 1990–2014, with 0.374 °C·decade$^{-1}$ in the raw model compared with 0.359 °C·decade$^{-1}$ in the observational reference. However, this domain-mean agreement masks spatial and seasonal discrepancies, including an asymmetric positive tail in grid-cell trend errors and pronounced monthly differences in February, March, April, June, October and December.

The raw daily-temperature distribution contains month-dependent biases in the median and in both tails. Mean bias correction is useful as a simple baseline because it reduces the central offset, but it does not correct the distribution shape. Standard quantile mapping and quantile delta mapping provide the strongest historical distributional correction, with near-zero Q05, median and Q95 biases in the reference-period diagnostics.

For climate-change applications under SSP5-8.5, QDM is recommended as the primary method among the three tested approaches. It corrects the observed historical distribution while preserving modelled quantile changes, which is essential when corrected data are used to assess future warming impacts. The recommendation should be treated as conditional on additional validation, because no univariate bias-correction method can compensate for an implausible driving climate signal or guarantee stable biases under future conditions.

## Acknowledgments

The research was partially funded by the Technology Agency of the Czech Republic (grant no. SS02030040, PERUN).

## References


Cannon, A. J., Sobie, S. R., Murdock, T. Q., 2015. Bias correction of GCM precipitation by quantile mapping: How well do methods preserve changes in quantiles and extremes? *Journal of Climate*, **28**(17), pp. 6938–6959. https://doi.org/10.1175/JCLI-D-14-00754.1

François, B., Vrac, M., Cannon, A. J., Robin, Y., Allard, D., 2020. Multivariate bias corrections of climate simulations: Which benefits for which losses? *Earth System Dynamics*, **11**, pp. 537–562. https://doi.org/10.5194/esd-11-537-2020

Lafferty, D. C., Sriver, R. L., 2023. Downscaling and bias-correction contribute considerable uncertainty to local climate projections in CMIP6. *npj Climate and Atmospheric Science*, **6**(158). https://doi.org/10.1038/s41612-023-00486-0

Lange, S., 2019. Trend-preserving bias adjustment and statistical downscaling with ISIMIP3BASD (v1.0). *Geoscientific Model Development*, **12**, pp. 3055–3070. https://doi.org/10.5194/gmd-12-3055-2019

Maraun, D., 2016. Bias correcting climate change simulations – a critical review. *Current Climate Change Reports*, **2**, pp. 211–220. https://doi.org/10.1007/s40641-016-0050-x

Maraun, D., Shepherd, T. G., Widmann, M., Zappa, G., Walton, D., et al., 2017. Towards process-informed bias correction of climate change simulations. *Nature Climate Change*, **7**, pp. 764–773. https://doi.org/10.1038/nclimate3418

Séférian, R., Nabat, P., Michou, M., Saint-Martin, D., Voldoire, A., et al., 2019. Evaluation of CNRM Earth System Model, CNRM-ESM2-1: Role of Earth system processes in present-day and future climate. *Journal of Advances in Modeling Earth Systems*, **11**, pp. 4182–4227. https://doi.org/10.1029/2019MS001791

Switanek, M. B., Troch, P. A., Castro, C. L., Leuprecht, A., Chang, H. I., et al., 2017. Scaled distribution mapping: A bias correction method that preserves raw climate model projected changes. *Hydrology and Earth System Sciences*, **21**, pp. 2649–2666. https://doi.org/10.5194/hess-21-2649-2017

Termonia, P., Fischer, C., Bazile, E., Bouyssel, F., Brožková, R., et al., 2018. The ALADIN System and its canonical model configurations AROME CY41T1 and ALARO CY40T1. *Geoscientific Model Development*, **11**, pp. 257–281. https://doi.org/10.5194/gmd-11-257-2018

Tong, Y., Gao, X., Han, Z., Xu, Y., Xu, Y., Giorgi, F., 2021. Bias correction of temperature and precipitation over China for RCM simulations using the QM and QDM methods. *Climate Dynamics*, **57**, pp. 1425–1443. https://doi.org/10.1007/s00382-020-05447-4