Evaluation of regional climate models ALARO-0 and REMO2015 at 0.22° resolution over the CORDEX Central Asia domain

Sara Top[1,2], Lola Kotova[3], Lesley De Cruz[4], Svetlana Aniskevich[5], Leonid Bobylev[6], Rozemien De Troch[4], Natalia Gnatiuk[6], Anne Gobin[7,8], Rafiq Hamdi[4], Arne Kriegsmann[3], Armelle Reca Remedio[3],
5 Abdulla Sakalli[9], Hans Van De Vyver[4], Bert Van Schaeybroeck[4], Viesturs Zandersons[5], Philippe De Maeyer[1], Piet Termonia[2,4], Steven Caluwaerts[2]

[1]Department of Geography, Ghent University (UGent), Ghent, 9000, Belgium
[2]Department of Physics and Astronomy, Ghent University (UGent), Ghent, 9000, Belgium
[3]Climate Service Center Germany (GERICS), Helmholtz Zentrum Geesthacht, Hamburg, 20095, Germany
10 [4]Royal Meteorological Institute of Belgium (RMIB), Brussels, 1180, Belgium
[5]Latvian Environment, Geology and Meteorology Centre (LEGMC), Riga, LV - 1019, Latvia
[6]Nansen International Environmental and Remote Sensing Centre (NIERSC), St. Petersburg, 199034, Russia
[7]Flemish Institute for Technological Research (VITO), Mol, 2400, Belgium
[8]Department of Earth and Environmental Sciences, Faculty of BioScience Engineering, Heverlee, 3001, Belgium
15 [9]Iskenderun Technical University, Iskenderun, 31200, Turkey

Correspondence to: Sara Top (sara.top@ugent.be)

Abstract. To allow for climate impact studies on human and natural systems high-resolution climate information is needed. Over some parts of the world plenty of regional climate simulations have been carried out, while in other regions hardly any high-resolution climate information is available. This publication aims at addressing one of these regional gaps by presenting 20 an evaluation study for two regional climate models (RCMs) (REMO and ALARO-0) at a horizontal resolution of 0.22° (25 km) over Central Asia. The output of the ERA-Interim driven RCMs is compared with different observational datasets over the 1980-2017 period. The choice of the observational dataset has an impact on the scores but in general one can conclude that both models reproduce reasonably well the spatial patterns for temperature and precipitation. The evaluation of minimum and maximum temperature demonstrates that both models underestimate the daily temperature range. More detailed studies of 25 the annual cycle over subregions should be carried out to reveal whether this is due to an incorrect simulation in cloud cover, atmospheric circulation or heat and moisture fluxes. In general, the REMO model scores better for temperature whereas the ALARO-0 model prevails for precipitation. This publication demonstrates that the REMO and ALARO-0 RCMs can be used to perform climate projections over Central Asia and that the produced climate data can be applied in impact modelling.

1 Introduction

30 There is a strong need for climate information at the regional-to-local scale to allow for impact studies on human and natural systems (Giorgi et al., 2009). In order to accommodate for this, the World Climate Research Program (WCRP) Coordinated Regional Climate Downscaling Experiment (CORDEX) was initiated with the aim to design and gather several high-resolution

experiments over prescribed spatial domains across the globe. CORDEX creates a framework to perform both dynamical and statistical downscaling, to evaluate these regional climate downscaling techniques and to characterize uncertainties of regional

35   climate change projections by producing ensemble projections (Giorgi and Gutowski, 2015). Within CORDEX there are large ensembles of model simulations available at different resolutions for Africa (Nikulin et al., 2012; Nikulin et al., 2018), Europe (Jacob et al., 2014; Kotlarski et al., 2014) and the Mediterranean (Ruti et al., 2015) CORDEX regions (Gutowski et al., 2016). In addition, a new ensemble of climate and climate change simulations covering all major inhabited regions with a spatial resolution of about 25 km, within the WCRP CORDEX COmmon Regional Experiment (CORE) Framework, has been

40   established in support of the growing demands for climate services (Remedio et al., 2019).

While high-resolution ensembles (up to 12.5 km spatial resolution) are available for certain regions, e.g. EURO-CORDEX (Jacob et al., 2014), for other regions such as Australasia (Di Virgilio et al., 2019) and the Antarctic (Souverijns et al., 2019) the first experiments were performed only recently. For the CORDEX Central Asia (CAS-CORDEX) domain only a single climate run with the regional climate model (RCM) HadRM3P (Gordon et al., 2000) of the Met Office Hadley Centre (MOHC)

45   at a resolution of 0.44° was publicly available through the Earth System Grid Federation (ESGF) archive until 2019. In addition, climate projections with the RegCM model at 0.44° resolution for the 2071-2100 period and different emission scenarios were reported in Ozturk et al. (2012, 2016), however they are not available through the ESGF archive. Moreover, this resolution is insufficient for impact modelling and environmental assessment applications and thus higher-resolution climate data over the CAS-CORDEX region is needed (Kotova et al., 2018). Recently, Russo et al. (2019) presented model

50   evaluation results of the COSMO-CLM 5.0 model ran at 0.22° or 25 km resolution over the CAS-CORDEX region. The current study significantly extends our knowledge over the CAS-CORDEX domain by validating two different RCMs based on multiple scores for temperature (mean, minimum and maximum) and precipitation over a much longer period

In order to fill the knowledge gap over Central Asia two RCMs, ALARO-0 and REMO, were run over this region at 0.22° resolution in line with the CORDEX-CORE protocol (CORDEX Scientific Advisory Team, consulted on 01/03/2019). Here

55   we present the model evaluation through the use of so-called "perfect boundary conditions" taken from reanalysis data and by comparing the downscaled results to observed data for the period 1980-2017. Such a validation study is necessary in order to gain confidence in the RCM downscaling procedure before its application in the context of climate projections where the RCM is driven by a GCM (Giorgi and Mearns, 1999). The methodology for validation is partially based on Kotlarski et al. (2014) and Giot et al. (2016), that compared a large ensemble of RCMs over the EURO-CORDEX region with the high-resolution E-

60   OBS observational dataset (Hofstra et al., 2009). However, in this study a slightly different approach is necessary due to 1) the absence of an ensemble of RCM runs over Central Asia, and 2) the absence of a reliable observational dataset over this region. While the Central Asian region is a vast area, the network of measurement stations is unevenly and sparsely distributed, especially the latter is problematic for several large subregions within the domain, that are sparsely populated. Therefore, the quality of gridded observational datasets, constructed through interpolation or area-averaging of station observations suffers,

65   in some regions, from the small number of stations that leads to over-smoothing especially of more extreme values (Hofstra et al., 2010). Additionally, the measurements at existing stations may not be representative for their large-scale environments, in

particular in orographically complex regions such as the Himalayas. In order to account for the lack of a model ensemble and reliable observations, this study compares the model simulations with different gridded observational datasets and reanalysis data. The model biases are compared with the differences among the observational datasets where the latter could be seen as

70   estimates of the observational uncertainty (New et al., 1999). For instance, spatially-similar bias patterns among the two models could be caused by observational errors that might be revealed by large differences between the observational datasets.

This study contains two assets: for the first time an in-depth evaluation of the RCMs ALARO-0 and REMO, ran at 0.22° resolution, is performed over the CAS-CORDEX domain and in addition we reflect on the impact of choice of the observational dataset on the model validation. Such an analysis is a prerequisite in order to be able to use the climate data in a sound way for

75   later impact studies, e.g. for investigating climate change impacts on crop yields and biomass production in forest ecosystems, which will be done in the framework of the AFTER project (Kotova et al., 2018).

In the following section we describe the applied methodology for this study. This section contains details about the study area, the model description, datasets used for the evaluation and the methodology of the analysis. In Sect. 3, we describe seasonal and annual means, biases and variability of mean, minimum and maximum surface air temperature and precipitation. Further,

80   we evaluate and provide a discussion of some remarkable anomalies in Sect. 4, complemented a brief outlook of the future plans of the ALARO-0 and REMO simulations. In the final Sect. 5 we summarize the conclusions.

2 Methods

Coordinated sets of experiments were recently performed or are still ongoing for CORDEX regions such as South America (Solman et al., 2013), Central America (Fuentes-Franco et al., 2015), North America (Wang and Kotamarthi, 2015; Diaconescu

85   et al., 2016; Whan and Zwiers, 2017), South Asia (Ghimire et al., 2018), East Asia (Zou et al., 2016), Southeast Asia (Tangang et al., 2018; Tangang et al., 2019; Tuyet et al., 2019), Australasia (Di Virgilio et al., 2019), Arctic (Koenigk et al., 2015), Antarctic (Souverijns et al., 2019) and Middle East North Africa (Almazroui et al., 2016; Bucchignani et al., 2018). In this paper we discuss the results of the simulations over the CAS-CORDEX domain.

2.1 CORDEX Central Asia domain

90   The CAS-CORDEX domain as shown in Fig. 1 contains Eastern Europe, a large part of the Middle East (including: Saudi-Arabia, Jordania, Syria, Iraq, Iran) and Central Asia (including: Kazakhstan, Uzbekistan, Turkmenistan, Afghanistan, Pakistan, Tajikistan, Kyrgyzstan and Mongolia). The majority of Russia and China (excluding the most eastern provinces) and the northern part of India are included as well. This domain is an exceptional CORDEX domain in the sense that it barely covers any ocean or sea. It contains several important mountain ranges e.g. Ural, Caucasus, Altay and Himalaya, and deserts e.g.

95   Arabian, Karakum, Thar, Taklamakan and Gobi desert. Mountainous environments are of special interest for regional climate modelling since global climate models do not resolve the mountain ranges and hence RCMs may have an added value here (Torma et al., 2015). In addition, the CAS-CORDEX domain contains a wide range of climatic and bioclimatic zones, with in

the north permafrost and snow-driven processes and in the south extremely hot regions (e.g. Arabian Peninsula) and monsoon-driven climates with excessive convection linked to the Inter-Tropical Convergence Zone (ITCZ) passing.

100   In order to obtain simulations that are comparable, the CORDEX initiative prescribes the minimum inner domain of each CORDEX region that the RCM has to cover. While REMO uses the exact rotated lat-lon CAS-CORDEX grid (Jacob et al., 2007) described by the CORDEX community, ALARO-0 has adopted a conformal Lambert projection (Giot et al., 2016), which implies that the non-rotated boundary box should be applied in order to define the domain. The grids were set up in such a way that the CAS-CORDEX domain is completely covered by the non-coupling zone. The CAS-CORDEX 0.22°

105   ALARO-0 inner domain encompasses 333 and 223 grid boxes, while REMO circumscribes 309 and 201 grid boxes in the east-west direction and north-south direction, respectively. The outer domain consists of the inner domain plus a coupling zone of eight grid points in each direction.

The CAS-CORDEX domain overlaps with eight other CORDEX domains, including the ones covering Europe, the Arctic, East Asia, South East Asia, South Asia, Africa/MENA and the Mediterranean. Both RCMs used in this study, ALARO-0 and

110   REMO, were already run and validated over the EURO-CORDEX region (Kotlarski et al., 2014; Giot et al., 2016) and additionally, REMO has been validated over five other overlapping CORDEX regions (Remedio et al., 2019).
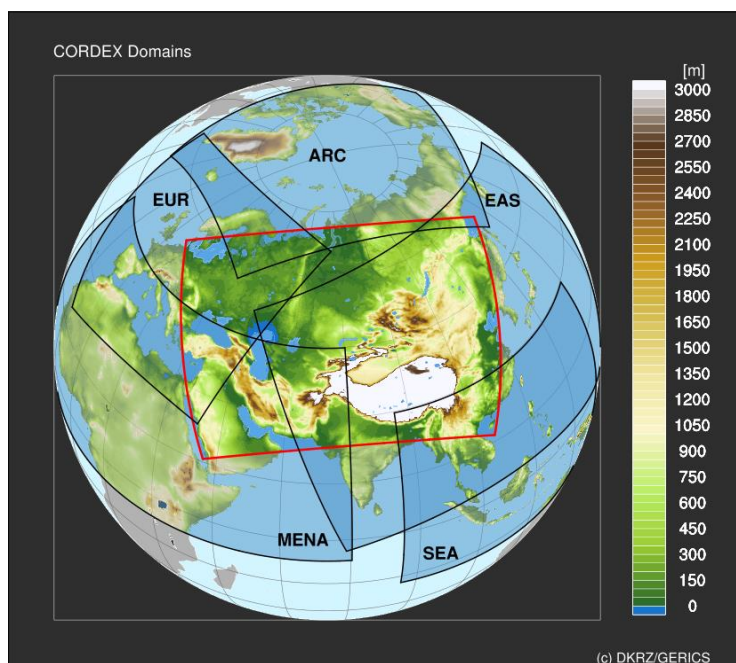
Figure 1: The CAS-CORDEX domain demarcated by a red contour and main overlapping CORDEX domains (black contour lines):
115   **Europe (EUR), Arctic (ARC), South East Asia (SEA), East Asia (EAS) and MENA projected upon the topography of Eurasia.**

2.2 Model description and experimental design

REMO and ALARO-0 are hydrostatic atmospheric circulation models aimed to run over limited areas. The ALARO-0 model is a configuration of the ALADIN model (ALADIN international team, 1997; Termonia et al. 2018a) which is developed, maintained and used operationally by the 16 countries of the ALADIN consortium. The dynamical core of the ALADIN model
120 is based on a spectral spatial discretization and a semi-implicit semi-Lagrangian time stepping algorithm. The ALARO-0 configuration is based on the physics parameterization scheme 3MT (Modular Multiscale Microphysics and Transport (Gerard et al. 2009)), which handles convection, turbulence and microphysics. ALARO-0 has been used and validated for regional climate studies (Hamdi et al., 2012; De Troch et al., 2013; Giot et al., 2016; Termonia et al. 2018b).

The REMO model is based on the Europa Model, the former NWP model of the German Weather Service (Jacob, 2001). The
125 model development was initiated by the Max-Planck-Institute for Meteorology and is further maintained and extended by the German Institute for Climate Services (HZG-GERICS). The physical parameterization originates from the global circulation model ECHAM4 (Roeckner et al., 1996), but there have been many further developments (Hagemann, 2002; Semmler et al., 2004; Pfeifer, 2006; Pietikäinen et al., 2012; Wilhelm et al., 2014). REMO is used in its most recent hydrostatic version, REMO 2015, and the dynamical core has a leap-frog time stepping with semi-implicit correction and Asselin-filter. For both
130 RCMs, the vertical levels are based on hybrid normalized pressure coordinates which follow the orography at the lowest levels. For the ALARO-0 experiment 46 levels were used whereas the REMO run employs 27 levels. More details on the general setup of ALARO-0 can be found in Giot et al. (2016) and for REMO we refer to Jacob et al. (2001) and Jacob et al. (2012). An overview of the model specifications is given in Table S1 of the supplementary material.

In order to validate both RCMs, a validation run driven by a large-scale forcing taken from the ERA-Interim global reanalysis
135 (Dee et al., 2011) is undertaken for the period 1980-2017. A nesting strategy is applied to dynamically downscale the ERA-Interim data, having a horizontal resolution of about 0.70° (approximately 80 km), to a high-resolution over the CAS-CORDEX domain (Denis et al., 2002). The ERA-Interim forcing data is prescribed at the lateral boundaries using the Davies (1976) relaxation scheme and the downscaling is performed to a horizontal resolution of 0.22° (approximately 25 km). Both model experiments are continuous runs initialised on the 1st of January 1979 and then forced every 6 hours at the boundaries up to
140 December 31st 2017. Following the methodology of Giot et al. (2016), constant climatological fields for some parameters are used and updated monthly. These include sea surface temperatures (SSTs), surface roughness length, surface albedo, surface emissivity and vegetation parameters. A spin-up period is needed to allow the models and their surface fields to adjust to the forcing and internal model physics (Giot et al., 2016). The model was spun-up for 30 years from 1979 to 2008 for REMO to produce an equilibrium for the soil temperature and soil moisture and these soil fields were then used as initial soil conditions
145 when restarting the model from 1979, while for ALARO-0 the year 1979 was taken as spin-up year. Therefore, 1979 will not be used for the analysis in the subsequent sections. The data produced by both models have been uploaded to the ESGF data nodes (website: http://esgf.llnl.gov/).

2.3 Reference datasets

In order to validate the model results, monthly, seasonally and annually averaged values for temperature and precipitation are
150   compared with different reference datasets. A multitude of datasets were considered to estimate the reliability of the gridded observational temperature and precipitation (New et al., 1999). When these datasets show large differences amongst each other, then the obtained model biases could be in part attributed to the observational uncertainty. The reference datasets are briefly presented in Table 1 and in the next sections we give a more detailed overview of the different datasets used in this study.

155   #### 2.3.1 Climatic Research Unit TS dataset

The gridded Climatic Research Unit (CRU) TS dataset (version 4.02) contains ten climate related variables for the period 1901-2017 (Harris et al., 2014) at a grid resolution of 0.50° covering the complete global land mass (excluding Antarctica) (New et al., 1999; New et al., 2000; Harris et al., 2014). Monthly values of minimum, maximum and mean near surface air temperature and precipitation are used in the current study. This dataset is widely used all over the world and in a wide range
160   of disciplines (Harris et al., 2014), however, there are also some issues reported. Main concerns include sparse coverage of measurement stations over certain regions, e.g. the North of Russia (New et al., 2002) and the dissimilarities in measurement methods that are used between and within different countries (New et al., 1999). New et al. (1999) indicate as well that the interpolation method is likely to produce warmer temperatures in sparsely covered mountainous areas and Hu et al. (2018) reported that the precipitation is underestimated in the centre of the CAS-CORDEX domain, especially in the mountainous
165   areas.

2.3.2 Matsuura and Willmott gridded dataset

The Matsuura and Willmott (MW) (version 5.01) gridded dataset of the University of Delaware contains monthly values at a 0.5° resolution based on temperature and precipitation station observations. The main differences with the CRU dataset are the use of different measurement station networks and spatial interpolation methods (Willmott et al., 1985; Willmott and
170   Matsuura, 1995; Willmott and Robeson, 1995). It is known that the MW dataset underestimates the precipitation especially during spring (Hu et al., 2018).

2.3.3 Global Precipitation Climatology Centre dataset

The Global Precipitation Climatology Centre (GPCC) (version 2018) of the Deutscher Wetterdienst is a monthly land surface precipitation dataset at 0.25° resolution based on rain gauge measurements. The GPCC full data monthly product version 2018
175   contains globally regular gridded monthly precipitation totals. This updated version is using "climatological infilling" to avoid interpolation artefacts for regions where an entire 5° grid is not covered by any station data (Schneider et al., 2018). Hu et al. (2018) concluded that GPCC is more in line with the observed station data in Central Asia compared to CRU and MW,

however, precipitation is underestimated in mountainous areas and seasonal precipitation is underestimated during spring. In addition, the GPCC has no similar dataset for other variables and thus, only precipitation can be validated with this dataset.

180 **2.3.4 ERA-Interim**

Reanalysis products like ERA-Interim are more continuous in space and time than station data, but they do contain biases as well. The ERA-Interim reanalysis of the European Centre for Medium-Range Weather Forecasts (ECMWF) is available from 1979 onwards. Total monthly precipitation at a spatial resolution of 0.25° was obtained from the Monthly Means of Daily Forecast Accumulations dataset by taking the mean over the precipitation amounts that are available for two time steps: 00:00

185 and 12:00. The Monthly Means of Daily Means data of 2 m temperature at 0.25° is used to study the difference between gridded datasets and reanalysis data. In addition, the temperature of ERA-Interim can reveal if the deviations between the RCMs and observational datasets are due to initial errors in the boundary conditions or not, since the RCMs were driven by ERA-Interim. This is not the case for precipitation since the RCMs are not using the ERA-Interim precipitation as forcing. They simulate precipitation based on other variables which are forced by ERA-Interim such as temperature and specific

190 humidity. Several studies have shown that ERA-Interim tends to have a warm bias in the northern part over the CAS-CORDEX region, especially during winter (Ozturk et al., 2012 and 2016). Ozturk et al. (2012) relates this to the insufficient ability of ERA-Interim to produce a snow cover in winter. Additionally, Ozturk et al. (2016) showed that ERA-Interim tends to have a dry bias over the CAS-CORDEX region.

2.4 Analysis methods

195 Gridded datasets are based on interpolated station data and used instead of station observations to overcome the scale difference between the model and observation field (Tustison et al., 2001). Nevertheless, the grids of the observational and reanalysis datasets generally differ from the model grid. Therefore, an interpolation to one common grid is needed in order to compare them (Kotlarski et al., 2014). As the CRU dataset has the lowest spatial resolution, the other datasets (both modelled and gridded) are upscaled to this grid. For the interpolation to the CRU grid, bilinear interpolation is used.

200 For ALARO-0 and REMO, hourly values of 2 m temperature and convective and stratiform rain and snow are available. The precipitation variables were added up in order to obtain the hourly total precipitation which in turn was used to calculate monthly totals and seasonal and annual means. The hourly temperature data are used to compute the daily minimum, mean and maximum temperatures. These daily values were then used to create monthly, seasonal and annual means of the mean, minimum and maximum temperature. Additionally, a height correction was performed for mean, minimum and maximum

205 temperature using the topography of the CRU database and assuming a uniform temperature lapse rate of 0.0064 K m^{-1}.

The model evaluation is done by calculating different evaluation metrics over the CAS-CORDEX domain for the 1980-2017 period. We computed the bias for the seasonal and annual climatological means of the evaluated variables based on the monthly means of the datasets to get maps that visualise the spatial patterns of the differences between the RCM or reference dataset and the CRU dataset. The relative bias for precipitation is computed by subtracting the CRU value from the RCM value and

210   dividing it by the CRU value. These climatological means and biases were spatially averaged to obtain one mean value over

the complete domain. Additionally, Taylor diagrams were produced in order to study the model performance for the different

seasons and for annual means. Taylor diagrams supplement the bias analysis by visualizing in a concise way information about

the correlation, centered root mean square error (RMSE) and ratio of spatial variability (RSV) between the model and the

observational dataset (Taylor, 2001). The RSV is defined as the ratio of the model standard deviation and the standard deviation

215   of the reference dataset, here CRU, over the spatial grid domain. In this study the Taylor diagrams represent the spatial pattern

correlation between model and reference data, which is obtained by calculating correlations across the grid points of the CAS-

CORDEX domain. For the used formulas we refer to appendix A of Kotlarski et al. (2014).

3 Results

In this section, the model validation results are presented with a focus on evaluation metrics of seasonal means in mean,

220   minimum and maximum near surface air temperature (henceforth denoted as temperature) and seasonal mean precipitation

(henceforth precipitation). Limitations of the observational datasets should be kept in mind when interpreting the evaluation

results (Kotlarski et al., 2014). These are investigated by comparing the different observational datasets and their implications

for the validation as will be described in Sect. 4.

3.1 Mean temperature

225   In Fig. 2, the mean seasonal and annual temperature observations of CRU and the model biases with respect to CRU are shown

for the 1980-2017 period. Both RCMs are producing similar mean annual temperature patterns since they have similar biases

with respect to CRU, except for the northeastern part of the domain, where REMO has a limited positive bias and ALARO-0

a limited negative one. Apart from the orographically complex regions, annual biases vary between -3°C and 3°C for both

RCMs. On the seasonal timescale this range is exceeded by the ALARO-0 data with a significant warm bias in winter and cold

230   bias in spring in the northern part of the domain. Mean temperature biases are for both RCMs largest in the eastern half of the

domain and are most outspoken for the ALARO-0 model. For the winter (DJF) period the REMO model shows a significant

warm bias over Mongolia and the eastern part of the domain, whereas for ALARO-0 the warm bias is concentrated over Russia

and Kazakhstan. A similar warm bias during winter was found over Scandinavia in the EURO-CORDEX runs with ALARO-0

(Giot et al., 2016). Giot et al. (2016) suggested this could be due to the strong synoptic scale forcing in winter and stable

235   boundary layer issues. A warm bias during winter in the northeastern part of the domain was found as well by Russo et al.

(2019) and Ozturk et al. (2012 and 2016) for the COSMO-CLM 5.0 and RegCM models, respectively. Furthermore, Fig. 2

shows that the REMO model has a cold bias in northeastern Europe during winter, a feature previously found for REMO over

different domains that include this region (Pietikäinen et al., 2018). During spring (MAM) only a modestly cold bias is found

for REMO in the northern part of the domain, while ALARO-0 has a very strong cold bias. For the summer (JJA) season,

240   biases are limited over most of the domain for the REMO model but for the ALARO-0 model there are warm biases, except

for the cold biases in the northwest and over the mountain ranges. Similar biases to those of ALARO-0 in summer were found by Russo et al. (2019) with the RCM COSMO-CLM 5.0. In spring and summer both RCMs show a pronounced warm bias
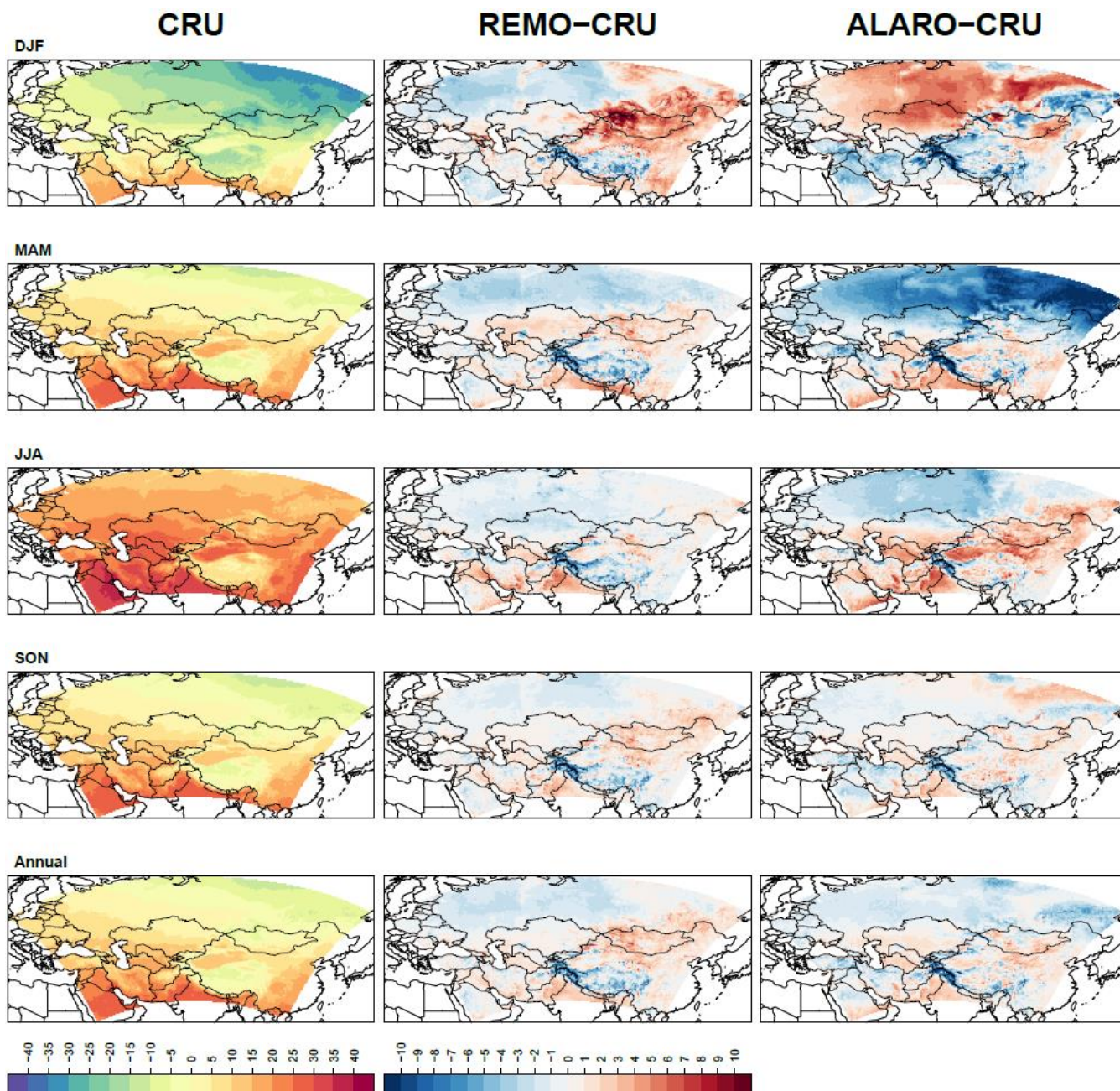
Figure 2: On the left: mean air temperature (°C) at 2 m height over the CAS-CORDEX domain based on the observational CRU dataset for the 1980-2017 period for winter (DJF), spring (MAM), summer (JJA) and autumn (SON). In the middle: temperature difference (°C) between the simulated REMO mean temperature and the CRU mean temperature. On the right: temperature difference (°C) between the simulated ALARO-0 mean temperature and the CRU mean temperature.

over Pakistan and the northern part of India and there is also a north-south gradient from cold to warm biases over the Arabian Peninsula. The outcomes of both RCMs for mean temperature agree well with the CRU data in autumn (SON). Biases in the main high-altitude regions are largely persistent throughout the seasons. More specifically, both ALARO-0 and REMO have large negative biases over the Pamir Mountains (Tadjikistan) and the Himalayas, while they also feature negative biases over the Tibetan Plateau, although this is to a lesser extent the case for ALARO-0 where this is only clearly visible for the winter season. Additionally, REMO contains large positive biases over the Altai, especially in winter, while this is not the case for ALARO-0. As mentioned before these biases should be placed in perspective since there are uncertainties in the observational dataset as well, especially in the mountainous regions where observations are sparse.

The spatially averaged mean temperatures of CRU for the different seasons during the 1980-2017 period are given in Table 2, accompanied by the mean bias over the domain for the RCMs. In agreement with Fig. 2 the biases are very small for both RCMs during autumn. Furthermore, it is clear from Table 2 that the strong cold bias during spring in the north for the ALARO-0 model has a larger negative impact on the spatially averaged bias than the warm bias during winter.

Figure 3 shows a spatial Taylor diagram for the mean temperature of both RCMs for the different seasons and for the annual mean value. Both models have in general a good model performance for temperature over the different seasons and on the annual level since the spatial correlation between the model output and the CRU data is high ($> 90\%$), while the centred RMSE is small (< 0.5) and the normalized RSV is mostly close to 1. Based on Fig. 3, both RCMs perform best during autumn and the spatial correlation is lower during summer for ALARO-0. However, the biases during summer are for both RCMs smaller than during winter and spring (Table 2 and Fig. 2). This is related to the smaller spatial range in temperatures during summer compared to the other seasons, as can be seen in Fig. 2 for CRU. An equal deviation in temperature for each season will lead to a less good correlation in summer due to the smaller spatial variability in temperature during summer. During autumn and winter, both RCMs do simulate the normalized standard deviation of the temperature very well. However, there was a clear warm bias observed during winter (Table 2 and Fig. 2) but indicates that the RCMs capture the spatial variability well. During spring cold bias in the north is limited for the REMO model but not for ALARO-0, which leads to a clear overestimation of the normalized RSV during spring. Both RCMs overestimate the normalized RSV during summer and spring, while in winter they underestimate it slightly. The underestimation of the spatial variation by the RCMs in winter is due to the warmer temperatures in the northern part of the domain, where the coldest temperatures are observed for CRU (Fig. 2 and 3). In spring and summer, the spatial variation is overestimated since colder temperatures are simulated by the RCMs in the coldest part of the domain.

The small mean bias for ALARO-0 during summer (JJA) (Table 2) is obtained by averaging the warm biases in the south and the cold biases in the north (Fig. 2) and does not result in a very good overall performance of the modelled temperature (Fig. 3). Based on Fig. 2, Fig. 3 and Table 2, ALARO-0 has a slightly better performance during autumn than REMO. Comparing the metrics of the RCMs shows that REMO is better in simulating the variability in temperature, except for autumn, and has smaller biases compared to ALARO-0 except for the autumn. On the other hand ALARO-0 better captures spatial temperature patterns since the spatial pattern correlation is slightly higher than for REMO, except during summer.
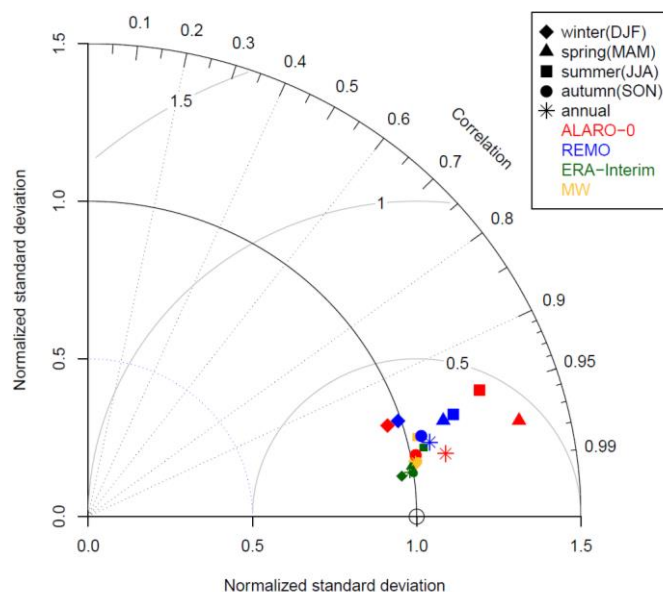
Figure 3: Normalized Taylor diagram representing the performance of mean temperature for seasonal and annual means for both RCMs (ALARO-0 and REMO), the ERA-Interim reanalysis and MW observational data with respect to CRU.

3.2 Minimum temperature

Similar as for the mean temperature, the modelled daily minimum temperature averaged over the different seasons and years during 1980-2017 is compared with the observational CRU data. Spatially averaged biases are larger for the minimum temperature than those of the mean temperature, except for the spring season, indicating that the model outputs are deviating more from the CRU data (Tables 2 and 3). This is due to the fact that both RCMs produce seasonal and annual means over the domain which are generally warmer for the minimum temperature than it was the case for the mean temperature. This causes a stronger warm bias in winter for the minimum temperature, which is especially visible in the northern part of the domain for the ALARO-0 model (Fig. 4). The REMO model also shows warmer biases over Mongolia during winter and spring when compared to the mean temperature (Fig. 2 and 4). Moreover, the cold bias in the north during spring for the ALARO-0 model is weaker for the minimum temperature than it was the case for the mean temperature. During the summer season the biases for the REMO model are small, while the ALARO-0 model output has a cold bias in the northwestern part of Russia and warm bias in the other regions (Fig. 4). Following the main trend, these warm biases have a larger magnitude for minimum temperature when compared to the mean temperature. In autumn, both models have a warm bias over almost the entire domain, which was not the case for mean temperature. The warm minimum temperatures of the RCMs indicate that they underestimate the coldest diurnal temperatures or that the observational CRU dataset overestimates them. Although the magnitude of the biases is different for mean and minimum temperature, the spatial patterns are maintained for each of the RCMs. This means that these two variables are spatially highly correlated with each other in both, models and observations.
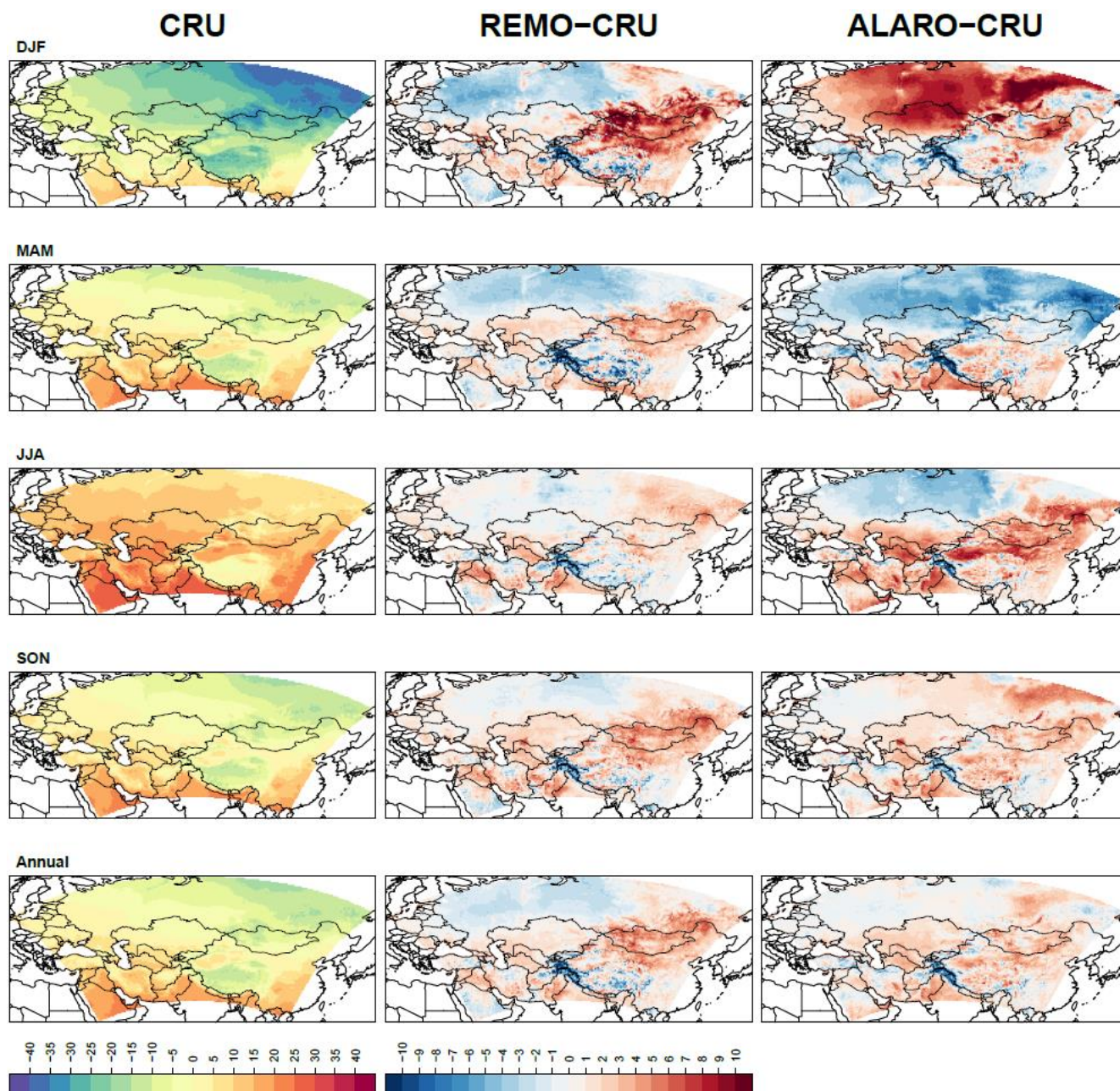
Figure 4: On the left: minimum air temperature (°C) at 2 m height over the CAS-CORDEX domain based on the observational CRU dataset for the 1980-2017 period for winter (DJF), spring (MAM), summer (JJA) and autumn (SON). In the middle: temperature difference (°C) between the simulated REMO minimum temperature and the CRU minimum temperature. On the right: temperature difference (°C) between the simulated ALARO-0 minimum temperature and the CRU minimum temperature.

The metrics in Fig. 5 show that the RCMs simulate the minimum temperature spatially well for annual and seasonal means. When comparing them to those of the mean temperature (Fig. 3), then it is seen that the metrics of both variables are similar

for both RCMs during the different seasons. Similar as was found for the mean temperature, ALARO-0 has on annual level a slightly better spatial pattern correlation with the minimum temperatures of the CRU dataset when compared to REMO, except for the summer for which the correlation deviates even more for minimum temperature (Fig. 5). On the other hand, REMO simulates better the variability and mean minimum temperature (Fig. 5 and Table 3). Similar as for mean temperature, 315 ALARO-0 simulates for the minimum temperature the variability less good during summer and spring (Fig. 4 and 5).
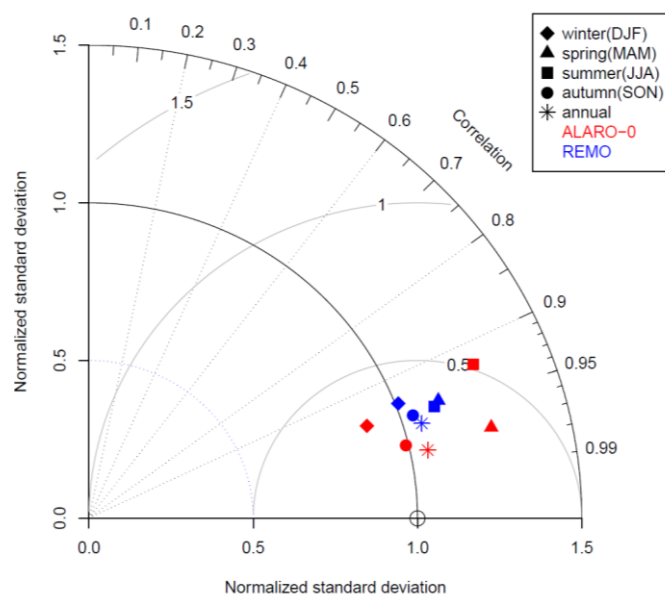
Figure 5: Normalized Taylor diagram representing the model performance of the minimum temperature for seasonal and annual means for both RCMs (ALARO-0 and REMO) with respect to CRU.

320 **3.3 Maximum temperature**

For the maximum temperature (Fig. 6), similar spatial patterns are found in the biases as for the mean temperature and the minimum temperature (Fig. 2 and 4) over the different seasons and for the annual mean. However, the biases are generally colder than it was the case for the variables mean temperature and minimum temperature (Tables 2, 3 and 4). This underestimation of the maximum temperatures is more pronounced for ALARO-0 than for REMO. During winter it counters 325 the warm bias that was obtained for mean and minimum temperature, resulting in a negative spatially averaged bias for the mean maximum temperature of ALARO-0 and a small positive one for REMO (Table 4). In Fig. 4 it is seen that he cold bias present in the northern part of the domain during spring is more pronounced for both RCMs due to the underestimation in maximum temperatures, which results especially for ALARO-0 in a strong deviation from the observational data. In autumn the smallest range in biases is obtained for both RCMs, which was the case as well for minimum and mean temperature.
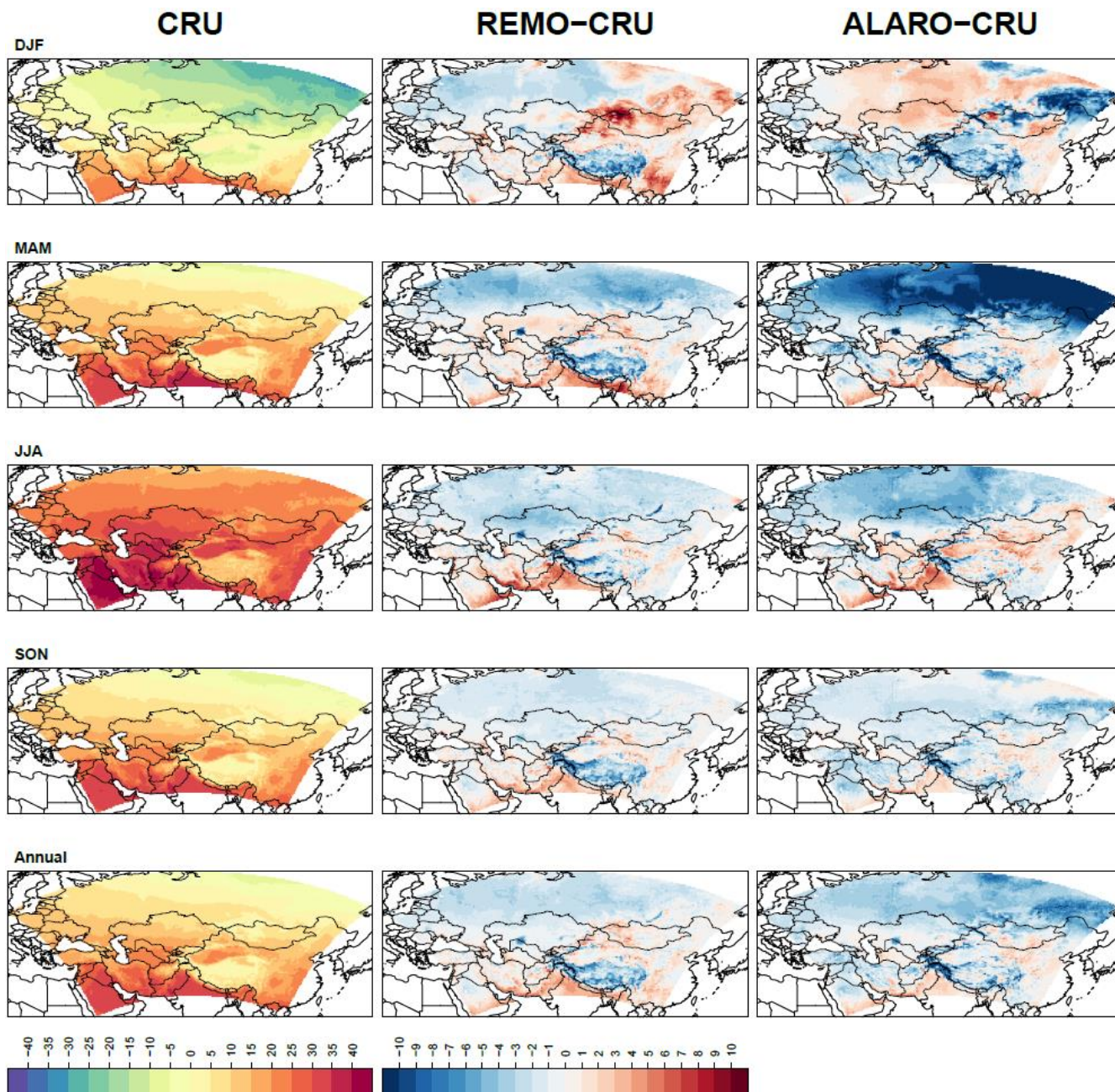
Figure 6: On the left: maximum air temperature (°C) at 2 m height over the CAS-CORDEX domain based on the observational CRU dataset for the 1980-2017 period for winter (DJF), spring (MAM), summer (JJA) and autumn (SON). In the middle: temperature difference (°C) between the simulated REMO maximum temperature and the CRU maximum temperature. On the right: temperature difference (°C) between the simulated ALARO-0 maximum temperature and the CRU maximum temperature.

Figure 7 shows that both models have an acceptable model performance over the CAS domain, since the spatial pattern correlation is high and the normalized RSV is mostly close to 1. Additionally, it is seen that both RCMs overestimate the

normalized RSV of the maximum temperature (Fig. 7). This differs from the mean temperature where both models underestimated the normalized RSV during winter (Fig. 3). Based on Fig. 6 and 7, both RCMs simulate best the maximum temperature during autumn.

340   The strong warm bias in the mean temperature over Russia for ALARO-0 during winter (Fig. 2) is mostly caused by the warm bias in minimum temperatures (Fig. 4), since the warm bias is larger for minimum temperatures than for maximum and mean temperatures (Fig. 6 and 2). This means that ALARO-0 fails to reproduce the low nocturnal temperatures. On the other hand, the large negative bias in spring over Russia is mostly caused by the cold bias in maximum temperatures (Fig. 6), meaning that ALARO-0 fails to reproduce the daytime temperatures in spring. In general the minimum temperature (Table 3 and Fig.

345   4) shows warmer biases than the mean temperature (Table 2 and Fig. 2) and the maximum temperature (Table 4 and Fig. 6) shows colder biases compared with the mean temperature over the different seasons. From this can be concluded that both cold and warm extreme temperatures are simulated less extremely by the models over most of the domain compared to the extreme temperatures in the observational CRU dataset. In other words, the daily temperature range is generally underestimated by both RCMs.

350
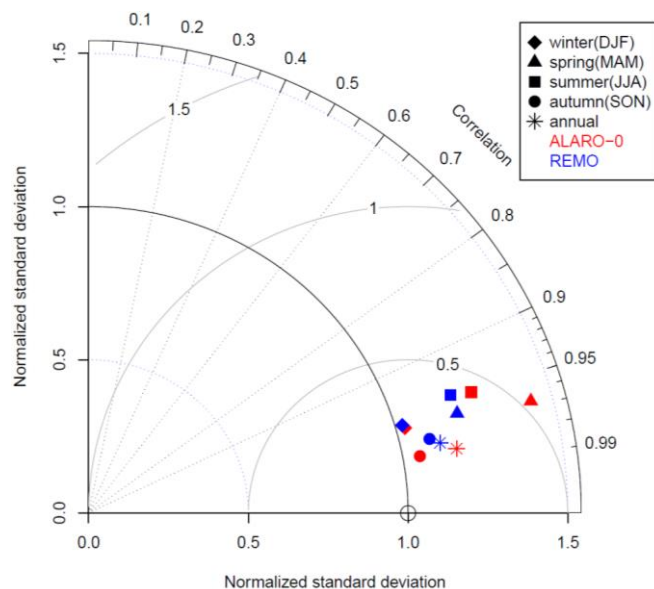
Figure 7: Normalized Taylor diagram representing the model performance of the maximum temperature for seasonal and annual means for both RCMs (ALARO-0 and REMO) with respect to CRU.

3.4 Precipitation

355   In Table 5, the spatially averaged precipitation over the 1980-2017 period is given for CRU and the relative biases of the RCMs with respect to CRU during the different seasons and on annual level are presented as well. For both RCMs the overall bias for precipitation is dry, except for REMO in spring. In Fig. 8, it is shown that this wet bias for REMO during spring is caused
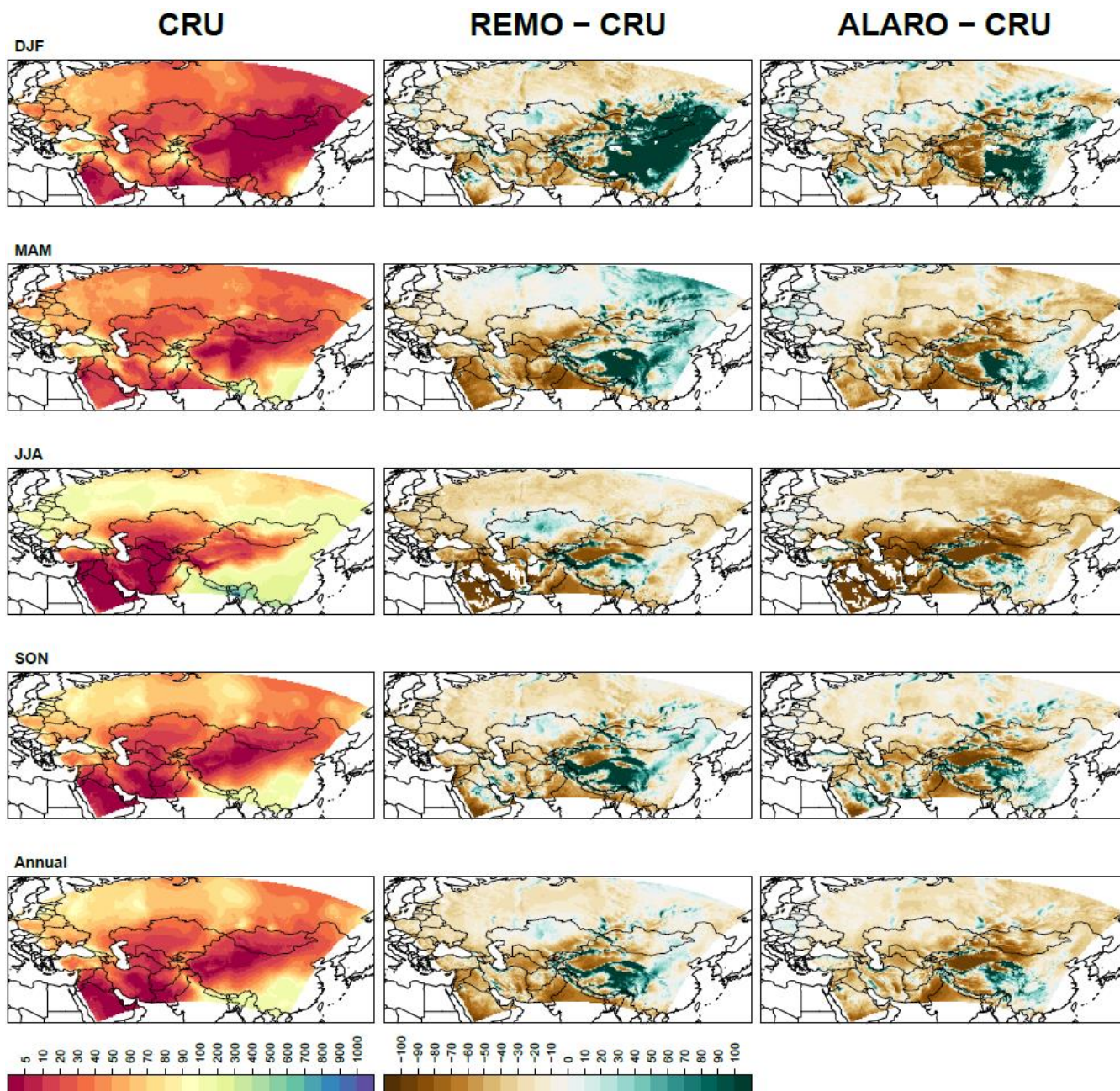
Figure 8: Relative difference between the average seasonal and annual CRU precipitation (mm month⁻¹) and the precipitation simulated by REMO and ALARO-0 over the 1980-2017 period.

by an overestimation of precipitation in the eastern part of the CAS-CORDEX domain. During spring, ALARO-0 shows only an extended wet bias in the southeastern part of the domain. A very strong relative wet bias is present during winter for both RCMs over this southeastern region and for REMO this wet bias even extents further up north to the Russian-Mongolian border. Russo et al. (2019) found a similar spatial pattern of a wet bias with their COSMO-CLM model as presented here for

365 REMO. However, this large relative wet bias during the winter is partly due to the low rainfall quantities in the observational CRU dataset. The largest relative biases can be found in relatively dry regions, therefore the absolute biases are presented in the supplementary material Fig. S1 and Table S2. When the absolute bias during winter is examined (supplementary material Fig. S1), then it is seen that REMO does not simulate a large absolute overestimation in precipitation in Mongolia and the northern part of China, but both RCMs do overestimate the precipitation in the Southeast Asian monsoon region during winter

370 and spring. This wet bias over the southeastern monsoon region during winter and spring is almost completely turned into a weak dry bias in summer, when most rain falls, except for Northern India and Pakistan where there is even a strong dry bias (Fig. 8 and S1). The wet bias during winter and spring and dry bias in summer between CRU and both RCMs over the southeastern part of the CAS-CORDEX domain can be linked to an early onset of the monsoon. However, this should be further investigated with an annual cycle over the region. This feature is more pronounced for REMO and was already

375 highlighted by Remedio et al. (2019), who saw the same shift for REMO with different CORDEX experiments over the subtropical region where the Asian monsoon takes place.
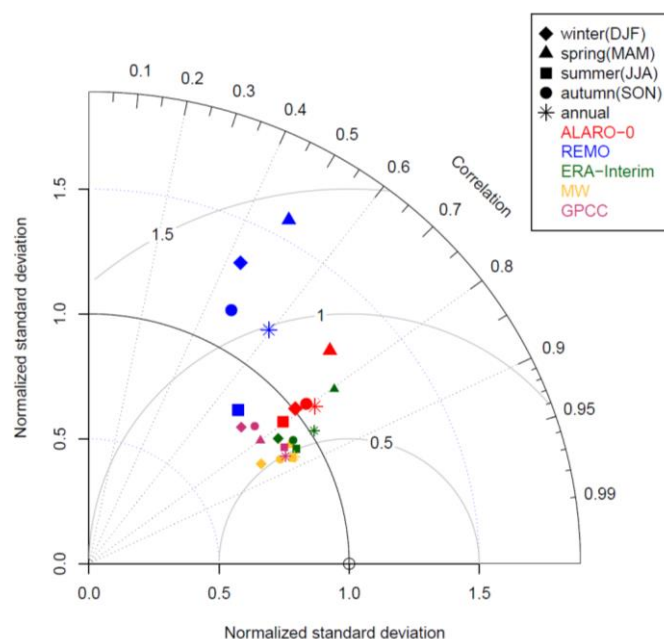
Figure 9: Normalized Taylor diagram representing the model performance of precipitation for seasonal and annual means for both
380 **RCMs (ALARO-0 and REMO), gridded observational datasets (MW, GPCC) and the ERA-Interim reanalysis data with respect to CRU.**

Next to these biases in the monsoon region, both models show during spring and summer dry biases over the Tarim basin and the southwestern part of the domain, where respectively the Taklamakan and Arabian desert are located, which are already dry regions in the CRU dataset and therefore show a strong dry bias in Fig. 8. The absolute biases over this region are less

385 pronounced in Fig. S1. In addition, both RCMs have a dry bias in the northern part of the domain during summer, which is the

strongest dry bias in this region over the different seasons in absolute precipitation deficiency, causing the strongest dry spatially averaged bias for this season (Table 5).

From Fig. 9 can be deduced that ALARO-0 is better than REMO in capturing the seasonal variation in precipitation since the RSVs are closer to 1. Additionally, ALARO-0 captures for all seasons better the spatial patterns since the correlations are

390 larger than those for REMO. The dry biases for ALARO-0 in Table 5 are thus caused by the simulation of systematically less precipitation than the precipitation amounts in the CRU data. Both models are worse in simulating the spatial correlation of precipitation (Fig. 9) compared to the mean, minimum and maximum temperature (Fig. 3, 5 and 7). The lower accuracy of simulated precipitation is due to the fact that precipitation is less systematic affected by land cover and topography compared to temperature (Kotlarski et al., 2014). Both RCMs show the largest error in normalized RSV during spring. This

395 overestimation of the spatial variation is due to the overestimation and underestimation of the precipitation amount in respectively the wettest and driest areas of the domain (Fig. 8). During summer both RCMs overestimate the variability in temperature (Fig. 3), while they underestimate the variability in precipitation (Fig. 9).

4 Discussion

4.1 Temperature

400 The underestimation of the diurnal range over the CAS-CORDEX domain was also observed by Russo et al. (2019) for the winter and summer season. Their RCM produced smaller diurnal ranges compared to different observational datasets and the comparison between the observational datasets pointed out that CRU overestimates the diurnal range in the northeastern part of the domain. This explains why both RCMs show the largest shift between biases in minimum and maximum temperature over this area. Hence, the RCMs underestimate the diurnal range, which is similar to the findings over other regions (Laprise

405 et al., 2003; Kyselý and Plavcová 2012), but in the northeast of the Central Asia domain the more pronounced underestimation is due to an overestimation by CRU. For the Czech Republic (Europe), Kyselý and Plavcová (2012) stated that this underestimation is probably caused by an incorrect simulation of atmospheric circulation, cloud cover or heat and moisture fluxes between land surface and atmosphere.

When we compare the above results for temperature with the other reference datasets, then the normalized standard deviation

410 of ERA-Interim and MW deviate less from CRU than the RCMs do during spring and summer (Fig. 3). This implies that the deviation in spatial variation of temperature between the RCMs and CRU cannot be completely explained by the observational uncertainty, meaning that the data of the RCMs deviates from the observations and can be improved. The spatial correlations between CRU and ERA-Interim or MW are close to those between CRU and the RCMs, which indicates that the RCMs are able to reproduce the spatial temperature patterns very well, even though they were slightly deviating from the spatial

415 temperature patterns in the CRU data. When we compare the mean spatial biases for the 1980-2017 period (Fig. 2, Fig. 10 and Table 2), then it is seen that the differences between MW and CRU are smaller than the differences between the RCMs and CRU, except for the autumn and for REMO on the annual level. From this we conclude that both RCMs are able to simulate

temperatures in the autumn that are within the range of observational uncertainty. During winter, spring and summer none of the validated RCMs are able to reproduce temperature means that can be completely explained by the observational

420 uncertainty. Hence, in winter both models are producing on average temperatures that are too warm and in spring they are too cold. Figure 10 shows that the driving force ERA-Interim has a warm bias in winter over the northeastern part of the domain and thus, the warm bias that is produced by both RCMs in winter can be assigned to this forcing. It must be noted that the spatial pattern of the warm bias in the ERA-Interim data is more similar to the warm bias pattern created by REMO. This warm bias in winter for the driving force ERA-Interim and the reflection of it in the RCM data was also found by Ozturk et al. (2012

425 and 2016). They related this warm bias already to shortcomings in the simulation of snow. Contrary to Ozturk et al. (2016) but similar to Ozturk et al. (2012), the warm bias in the ERA-Interim forcing during winter is amplified by the two RCMs evaluated in this study. Additionally to the influence of the warm forcing New et al. (1999) mentioned that CRU contains colder temperatures in winter over Russia, although this is not the main reason of the warm bias over Russia since there are only patches of warm biases observed between the two observational datasets MW and CRU, namely over the mountain ranges

430 Yablonovyy and Stanovoy in the southeastern part of Russia (Fig. 10). Hence, we conclude that the warm forcing is the main reason for the warm bias over Eastern Russia during winter. In contrast to winter, a cold bias is obtained in the northeast during spring for both RCMs (Fig. 2), although a weak warm bias is still present in the ERA-Interim forcing (Fig. 10). This feature was also presented for RCMs at 0.50° horizontal resolution in Ozturk et al. (2012 and 2016), but they did not explain it. The warm bias during winter and cold bias during spring in the northeastern part of the domain could be due to an incorrect

435 simulation of snow related processes or a delay in the simulation snow cover. This should be further investigated by looking if there is a temperature delay in the annual cycle over Eastern Russia and if there is a similar delay in one of the processes. Russo et al. (2019) found, however, for their RCM that changes in the snow scheme did not affect the simulation results significantly and it did not reduce the warm bias in the northeast during winter. Cloud cover is another process that might explain the pronounced temperature biases in the north. Ozturk et al. (2012) obtained significant better temperature results in

440 the northern part of the CAS-CORDEX domain when using a cloud cover correction. Hamdi et al. (2012) found a strong correlation between a warm bias and cloud cover representation over Belgium (Europe) for ALARO-0, so this could be the reason why there are some large temperature biases in the north especially for ALARO-0. From Fig. 4 and 6 was deduced that ALARO-0 overestimates the nocturnal winter temperatures, while the diurnal temperatures in spring are underestimated. Both could be due to too much cloud cover and this could explain as well why the RCMs underestimate the diurnal range. Therefore,

445 the relation between the temperature biases in the north and the cloud cover should be further investigated by studying this specific region more comprehensively. Another possibility is that the RCMs calculate the temperature incorrectly during stable circumstances.

New et al. (1999) found that CRU overestimates the temperatures in summer over Russia. When ERA-Interim and MW are both compared to CRU, then it is seen that these two datasets contain lower temperatures over Western Russia during all

450 seasons except for winter (Fig. 10) and thus, the weak cold bias over Western Russia during these seasons for both RCMs (Fig. 2) can be attributed to temperatures in the CRU dataset that are too warm. However, the cold bias in the northwest during

winter for REMO cannot be explained by this feature since a small warm bias is found between ERA-Interim and CRU during

winter. As mentioned before this feature was already described by Pietikäinen et al. (2018). The north-south gradient in the
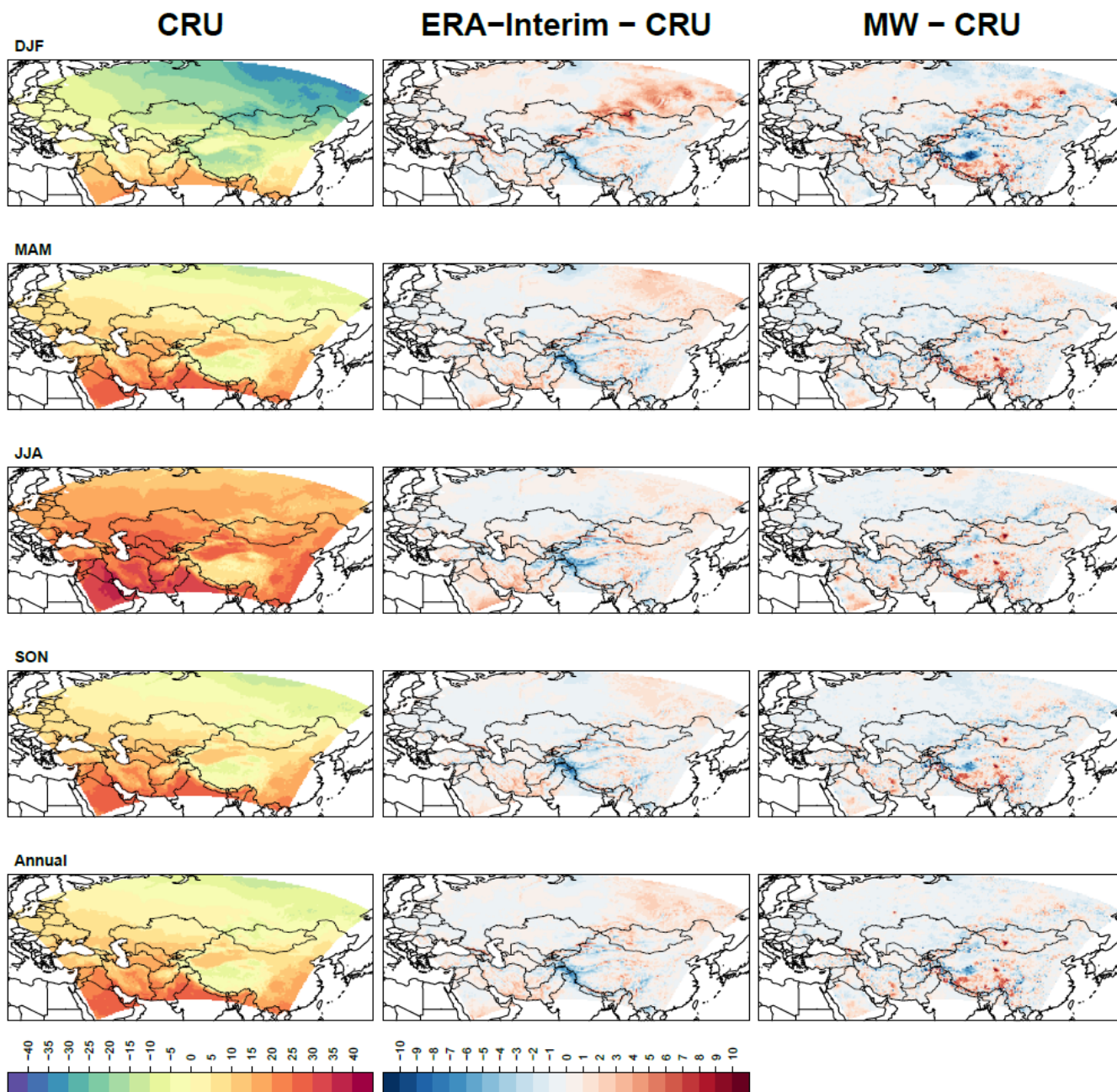
Figure 10: CRU temperature (left), difference between ERA-Interim temperature and CRU temperature (middle) and difference between MW temperature and CRU temperature (right) for the winter (DJF), spring (MAM), summer (JJA), autumn (SON) and annual (ANN) mean of the 1980 - 2017 period over the CAS-CORDEX domain.

temperature bias over the Arabian Peninsula during spring and summer for both RCMs (Fig. 2) can be explained by a sparse coverage of observational stations in the CRU dataset over this region (New et al., 1999), since for both ERA-Interim and MW

460   a similar bias is found with respect to CRU (Fig. 10). The warm bias over Pakistan and Northern India, present for both RCMs during spring, summer and autumn, cannot be explained by the ERA-Interim forcing or differences between the observational datasets and thus a process in the RCMs is likely to overestimate the temperatures in this region. There is as well a significant cold bias between the ERA-Interim and CRU data over the Himalayas and Tibetan Plateau during the different seasons. The latter was also observed by Ozturk et al. (2012 and 2016) and is due to the fact that gridded data is based on measurements of

465   meteorological stations in the valleys (New et al., 1999). This is the case as well for the gridded observational data of WM. Similar as Russo et al. (2019) concluded for COSMO-CLM 5.0, the cold bias of the RCMs over the Himalayas and Tibetan Plateau is mainly due to the gridded observations that are less reliable. The amplification of the biases over the mountainous regions for the RCMs can be attributed to the used assumption of the spatially and temporally uniform lapse rate of 0.0064 K m^{-1} for the elevation correction (Kotlarski et al., 2014) or by an amplification induced by the RCMs.

470   There is the same tendency as mentioned by Kotlarski et al. (2014) for the European domain that the RCMs underestimate the spatial variation slightly during winter and overestimate it during. In Fig. 3, it is seen that the larger RSVs of the RCMs during summer are due to an underestimation of the variability in the CRU dataset since the ERA-Interim and MW data show both a slight overestimation compared to CRU. In addition it is seen that the spatial patterns during summer are not completely captured by the CRU data since, the two other reference datasets both show a lower spatial correlation with CRU during

475   summer, compared to the other seasons. The lower performance of the RCMs during summer can thus be explained by the uncertainty in spatial variation of temperatures within the observational CRU dataset. As mentioned before this is more pronounced for the summer season since the spatial variation in temperature is lower during this season. Ozturk et al. (2016) found as well a lower spatial correlation during summer with their RCM RegCM4.3.5 at 0.50° horizontal resolution. Additionally, similar high spatial correlations are obtained during the different seasons for ALARO-0 and REMO at 0.22°

480   horizontal resolution when compared to the results of Ozturk et al. (2016).

4.2 Precipitation

Table 5 and Fig. 11 show that CRU overestimates the precipitation amounts since the two other observational datasets, MW and GPCC, have a strong dry bias over almost the complete domain when compared to CRU. This can explain the systematic dry bias that was found for ALARO-0 during all seasons (Table 5). The small patches with wet biases in the southeastern part

485   of the domain for these two gridded datasets, however, do not explain the extensive wet bias in the southeast during winter and spring which was observed for both RCMs (Fig. 8). This wet bias is present in the ERA-Interim data (Fig. 11) and thus the RCMs might produce this wet bias due to an overestimation in specific humidity of the ERA-Interim forcing, although ALARO-0 is able to reduce the excessive amount of precipitation visible in the ERA-Interim data to a certain extent (Fig. 8, 11, S1 and S2). The biases of REMO tend more towards those of ERA-Interim, although REMO and ERA-Interim parameterize

490   precipitation differently, and the biases of ALARO-0 tend more towards those of MW and GPCC. For example, the weak wet

bias which was observed in the northeastern part of the domain during spring for REMO and not for ALARO-0 is also visible in the ERA-Interim data, but not in the WM and GPCC data. This difference between ALARO-0 and REMO is related to the 3MT cloud microphysics scheme of ALARO-0, which is known for its good performance (Giot et al., 2016). Another similarity between the ERA-Interim data and the output of the RCMs is seen in Fig. 9, where both RCMs are worst in simulating the

495   spatial variation during spring and the ERA-Interim data has a similar overestimation in spatial variation during spring when compared to CRU. Except for ERA-Interim in spring, the other reference datasets have a lower spatial variation in precipitation during all seasons, which means that CRU generally overestimates the spatial variation in precipitation (Fig. 9).
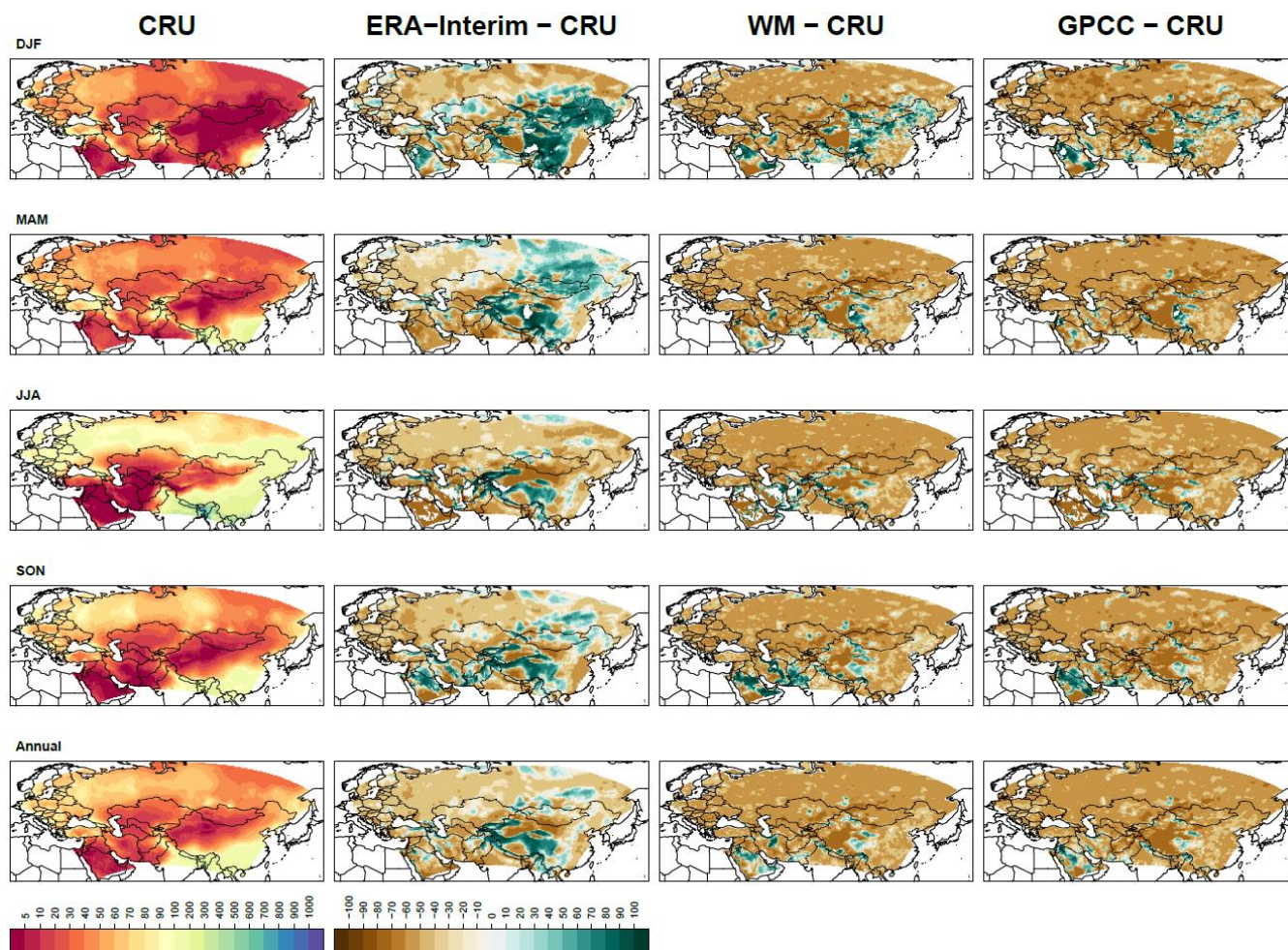
Figure 11: Relative difference between the average seasonal and annual CRU precipitation (mm month⁻¹) and the precipitation in the ERA-Interim, MW and GPCC datasets over the 1980-2017 period.

In Table 5 and Fig. 11, it is seen that the gridded observational datasets, CRU, MW and GPCC, show a drier environment than the ERA-Interim reanalysis dataset during spring, which is a known feature (Hu et al., 2018). The observed relative wet bias in the east for the ERA-Interim data during winter is in absolute values not that outspoken as the wet bias in spring, which is

505   due to the low precipitation quantities over this region during winter, as was mentioned before (Fig. 11 and S2). The wet bias in winter for ERA-Interim is not reflected by a positive value for the spatial mean bias in Table 5, since it is completely compensated by a dry bias in the northwestern part of the domain (Table 5 and Fig. 11 and S2). The dry bias in the southwest of the domain during spring and summer, which was observed for both RCMs, is also seen for the ERA-Interim, MW and GPCC data. From this is concluded that this dry bias is due to a small overestimation in precipitation by CRU which leads to

510   large relative biases since the precipitation quantities are low (Fig. 13). Harris et al. (2013) mentioned that the Middle East is sparsely covered with precipitation measurements, which leads to uncertainties and errors in the CRU data. Both RCMs have the driest spatially mean bias compared to CRU in summer due to a dry bias over Russia and Northern India (Table 5 and Fig. 8 and S1). Similar patterns are found for the observational datasets MW and GPCC when looking to the absolute differences with CRU (Fig. S2). It is known that CRU data shows higher precipitation rates at most of the grid points in eastern Russia

515   due to poor station coverage (New et al., 1999). The dry biases over this regions during summer for the RCMs and observational datasets MW and GPCC are thus due to an overestimation of precipitation in the CRU data (Fig. 8, 11, S1 and S2). This overestimation of precipitation in the CRU data causes a larger spread in variability, which explains why the RCMs underestimated the spatial variation only during summer (Fig. 9). The overestimation in precipitation by both RCMs over the eastern part of Tibetan Plateau and the Altay, Tianshan and Kunlun Mountains on the annual level is according to Zhu et al.

520   (2015) due to the fact that that gridded datasets underestimate the precipitation over these mountainous regions. It is a known feature that the accuracy of gridded precipitation datasets decreases with elevation, especially when the altitude of 1500 m is reached (Zhu et al., 2015). Table 5, Fig. 9 and Fig. 11 show that the observational gridded datasets and ERA-Interim deviate more from CRU than it was the case for temperature, resulting in a larger observational uncertainty for precipitation. Russo et al. (2019) showed additionally that the influence of observational data sets on the RSV is larger for precipitation than for

525   temperature.

     Ozturk et al. (2012 and 2016) and Russo et al. (2019) obtained similar seasonal patterns in precipitation, with their model simulations at a horizontal resolution of 0.50° and 0.22°, respectively. An excess of precipitation was simulated over the mountainous areas of the Asian monsoon region during winter, spring and autumn, while in summer a dry bias was observed. Additionally, they obtained as well a dry bias in summer over the northwestern and southwestern part of the domain. The

530   ALARO-0 and REMO models produce at a horizontal resolution of 0.22° smaller spatially averaged precipitation biases over the CAS-CORDEX region than was obtained with the RegCM4.3.5 model at a resolution of 0.50° (Ozturk et al., 2016). ALARO-0 and REMO have similar values for spatial correlations of precipitation (Fig. 9) as for regions in the EURO-CORDEX domain which range between 40 % and 90 % (Kotlarski et al., 2014). The spatial correlations between CRU and REMO are similar to the values obtained with RegCM4.3.5, except for winter where REMO has a higher spatial correlation.

535   ALARO-0 obtains higher values for the spatial correlations and they are close to those of the other observational datasets (Fig. 9). Although the observational uncertainty is quite large, we can conclude that REMO simulates the precipitation fairly well and ALARO-0 performs very well. However, the uncertainty range and error in the observational products should be restricted to improve the evaluation of precipitation.

The warm temperatures obtained with REMO in winter and the cold temperatures in spring over the northeastern part of the
540   domain can be linked with the dry and wet bias in winter and spring respectively. This strengthens our hypothesis that there is
a delay by REMO in simulating snow or snow cover. As stated before for temperature this should be further analysed by
plotting the annual cycles of precipitation and temperature for this region. For ALARO-0 this link between an overestimation
(underestimation) of temperature and an underestimation (overestimation) in precipitation during winter (spring) is not seen.
Therefore, it is likely that some processes affecting the temperature are not simulated well by ALARO-0 over the northeastern
545   part of the domain. The persistent warm bias over Pakistan and Northern India of both RCMs can be explained by the persistent
underestimation in simulated precipitation over this region by both RCMs.

When we compare the results of temperature (Fig. 3) and precipitation (Fig. 9) with other domains e.g. Fig. 9 and 10 in
Kotlarski et al. (2014), then we can conclude that these RCMs have a similar model performance as the RCMs have over other
domains. However, one should be aware that the CAS-CORDEX domain as a whole is a larger domain and thus the result
550   might be more smoothed because of the larger amount of grid points which was taken into account to create the Taylor diagram.

4.3 Outlook

In the near future a similar evaluation over several subregions will be undertaken, since this evaluation over the large domain
highlighted some specific regions where there might be deficiencies in the RCMs e.g. the warm bias in winter over Eastern
Russia and wet bias over the East Asian monsoon region. By looking into more detail to subregions we hope to understand
555   which processes in the RCMs cause the deficiencies e.g. shift in snow related processes and monsoon. In addition, we ran both
RCMs up to 2100 driven by different GCMs under the scenarios of representative concentration pathways (RCPs) 2.6, 4.5
(only for ALARO) and 8.5, which will be used to investigate the climate sensitivity over Central Asia and to study the evolution
of extreme events. Further, we plan to perform a bias adjustment on the model data by using observations. To select the optimal
bias adjustment method, a comparison of different approaches will be made. This will enable impact modellers to optimally
560   use our climate data in their models for crop production, biomass production, etc.

5 Conclusion

The first validation results over the CAS-CORDEX domain of ALARO-0 and REMO, ran at 0.22° resolution, showed that
both RCMs reproduced realistic spatial patterns for temperature and precipitation with biases within an acceptable range,
except for the temperature of ALARO-0 during spring. However, there are large biases in several regions during several
565   seasons e.g. a warm bias in the north during winter and a wet bias in spring over the Asian monsoon region. The comparison
between CRU, ERA-Interim and the other gridded observational datasets showed that the warm bias in winter is induced by
the warm ERA-Interim forcing and a delay in the simulation of snow and snow cover for REMO. For ALARO-0 the
temperature delay could not be explained by a delay in precipitation and thus it is likely that some processes which affect the
temperature in this region are not captured well by ALARO-0. A similar validation over subregions should be done, to examine

570 the shift in the annual cycle and the processes that are lacking or simulated incorrectly over those particular regions where a less good performance was found. Negative precipitation biases for both RCMs during all seasons are due to an overestimation of precipitation in the CRU data since the other reference datasets show dry biases. For all variables large biases are observed over the mountainous areas but these are mainly attributed to the observational error.

Both RCMs perform very well during the autumn, showing biases within the range of observational uncertainty for temperature
575 and precipitation. Additionally, the values for spatial variation and pattern correlation of both RCMs are very close to the values obtained with other reference datasets for the mean temperature. For precipitation these metrics indicated a less good performance of the RCMs since they deviated more from the reference datasets than it was the case for temperature. However, the different reference datasets deviated more for precipitation from CRU, than for temperature which indicates that there is a larger uncertainty in the spatial patterns of precipitation. The precipitation biases of both RCMs are within the range of
580 observational uncertainty and the precipitation is simulated similar for REMO or better for ALARO-0 when compared to other CORDEX simulations. REMO is better than ALARO-0 in reproducing the temperatures based on the biases and spatial variability, except during autumn, while ALARO-0 is very good in estimating the precipitation.

The evaluation of minimum and maximum temperatures showed that the RCMs simulate these variables less pronounced over most of the domain compared to the observational CRU dataset, which is generally caused by an underestimation of the daily
585 temperature range. This shows the advantage of taking more evaluation variables into account than only the ordinary mean temperature and precipitation. These findings are important for regional impact modelling. Since the RCMs perform as well over the CAS-CORDEX domain as other RCMs do, we finally conclude that these RCMs can be used to perform climate projections and the produced climate data can be applied in impact modelling.

Code availability

590 The R code used for the analysis is available through: http://doi.org/10.5281/zenodo.3659717 (Top et al., 2020).

For the code of the ALARO-0 model we refer to the Code availability section in Termonia et al. (2018). More information about the REMO model is available on request by contacting the Climate Service Center Germany (contact@remo-rcm.de).

Data availability

The climate data produced by ALARO-0 and REMO2015 have been uploaded to the ESGF data nodes (website:
595 http://esgf.llnl.gov/). In order to obtain the data, one of the nodes must be chosen. Thereafter, click on 'CORDEX' or search for 'CORDEX' and then select the domain 'CAS-22' and the RCM model in the left column. The exact identifiers can be found in Table S2 of the supplementary material.

The CRU data is available through (http://www.cru.uea.ac.uk). The MW data is freely available at: http://climate.geog.udel.edu/~climate/html_pages/download.html and NetCDF files can be found here:

600 https://www.esrl.noaa.gov/psd/data/gridded/data.UDel_AirT_Precip.html: air.mon.mean.v501.nc and
precip.mon.total.v501.nc. The GPCC data can be accessed through: doi: 10.5676/DWD_GPCC/FD_M_V2018_025.

Author contribution

Modelling and performing simulations: C.S., D.C.L., D.T.R., K.L., K.A, R.R.A. ;Post-processing: D.C.L., D.T.R., K.L., K.A,
R.R.A. ;Visualization: K.L., T.S. ;Writing - original draft: T.S. ;Writing - review & editing: A.S., B.L., C.S., D.C.L., D.M.P.,
605 D.T.R., G.N., G.A., H.R., K.A, K.L., R.R.A., S.A., T.P., T.S., V.D.V.H., V.S.B., Z.V. ;Supervision: C.S., D.M.P., T.P.
;Funding acquisition: A.S., B.L., D.M.P., G.A., K.L., T.P.

Acknowledgement

The AFTER project is granted by the ERA.Net RUS Plus Initiative, ID 166. HZG-GERICS received funding from the Federal
Ministry of Education and Research (BMBF). Ghent University and VITO received funding from the Research Foundation -
610 Flanders (FWO), grant G0H6117N. NIERSC received funding from the Russian Foundation for Basic Research (RFBR) -
grant № 18-55-76004. LEGMC received funding from the State Education Development Agency (SEDA) and ISTE received
funding from the scientific and technological research council of Turkey (TUBITAK agreement nr: 2017O394).
The computational resources and services for the ALARO-0 regional climate simulations were provided by the Flemish
Supercomputer Center (VSC), funded by the Research Foundation - Flanders (FWO) and the Flemish Government department
615 EWI. The CORDEX-CORE REMO simulations were performed under the GERICS/HZG share at the German Climate
Computing Centre (DKRZ).

Competing interests

The authors declare that they have no conflict of interest.

References

620 ALADIN International Team: The ALADIN project: Mesoscale modelling seen as a basic tool for weather forecasting and
atmospheric research, WMO bull., 46, 317–324, 1997.
Almazroui, M., Islam, M. N., Alkhalaf, A. K., Saeed, F., Dambul, R. and Rahman, M. A.: Simulation of temperature and
precipitation climatology for the CORDEX-MENA/Arab domain using RegCM4, Arab. J. of Geosci., 9, 13,
doi:10.1007/s12517-015-2045-7, 2016.

625 Bucchignani, E., Mercogliano, P., Panitz, H. J. and Montesarchio, M.: Climate change projections for the Middle East–North Africa domain with COSMO-CLM at different spatial resolutions, Advances in Climate Change Research, 9, 66–80, doi:10.1016/j.accre.2018.01.004, 2018.

CORDEX Scientific Advisory Team: The WCRP CORDEX Coordinated Output for Regional Evaluations (CORE) Experiment Guidelines, Available online: http://www.cordex.org/experiment-guidelines/cordex-core (accessed on 1 March 630 2019).

Davies, H.C.: A lateral boundary formulation for multi-level prediction models, Quart. J. R. Meteor. Soc., 102, 405–418, doi:10.1002/qj.49710243210, 1976.

Dee, D.P., Uppala, S.M., Simmons, A.J., Berrisford, P., Poli, P., Kobayashi, S., Andrae, U., Balmaseda, M.A., Balsamo, G., Bauer, P., Bechtold, P., Beljaars, A.C., van de Berg, L., Bidlot, J., Bormann, N., Delsol, C., Dragani, R., Fuentes, M., Geer, 635 A.J., Haimberger, L., Healy, S.B., Hersbach, H., Hólm, E.V., Isaksen, L., Kållberg, P., Köhler, M., Matricardi, M., Mcnally, A.P., Monge-Sanz, B.M., Morcrette, J.J., Park, B.K., Peubey, C., de Rosnay, P., Tavolato, C., Thépaut, J.N. and Vitart, F.: The ERA-Interim reanalysis: Configuration and performance of the data assimilation system, Quarterly Journal of the Royal Meteorological Society, 137, 553–597, doi:10.1002/qj.828, 2011.

De Troch, R., Hamdi, R., Van de Vyver, H., Geleyn, J. F., and Termonia, P.: Multiscale performance of the ALARO-0 model 640 for simulating extreme summer precipitation climatology in Belgium, Journal of Climate, 26, 8895–8915, doi:10.1175/JCLI-D-12-00844.1, 2013.

Denis, B., Laprise, R., Caya, D. and Côté, J.: Downscaling ability of one-way nested regional climate models: the Big-Brother Experiment. Climate Dynamics, 18, 627–646, doi:10.1007/s00382-001-0201-0, 2002.

Diaconescu, E. P., Gachon, P., Laprise, R. and Scinocca, J. F.: Evaluation of precipitation indices over North America from 645 various configurations of regional climate models, Atmosphere-Ocean, 54, 418–439, doi:10.1080/07055900.2016.1185005, 2016.

Di Virgilio, G., Evans, J. P., Di Luca, A., Olson, R., Argüeso, D., Kala, J., Andrys, J., Hoffmann, P., Katzfey, J. J. and Rockel, B.: Evaluating reanalysis-driven CORDEX regional climate models over Australia: model performance and errors, Climate Dynamics, 53, 2985–3005, doi:10.1007/s00382-019-04672-w, 2019.

650 Fuentes-Franco, R., Coppola, E., Giorgi, F., Pavia, E. G., Diro, G. T. and Graef, F.: Inter-annual variability of precipitation over Southern Mexico and Central America and its relationship to sea surface temperature from a set of future projections from CMIP5 GCMs and RegCM4 CORDEX simulations. Climate Dynamics, 45, 425-440, doi:10.1007/s00382-014-2258-6, 2015.

Gerard, L., Piriou, J. M., Brožková, R., Geleyn, J. F. and Banciu, D.: Cloud and precipitation parameterization in a meso-655 gamma-scale operational weather prediction model, Monthly Weather Review, 137, 3960–3977, doi:10.1175/2009MWR2750.1, 2009.

Ghimire, S., Choudhary, A. and Dimri, A. P.: Assessment of the performance of CORDEX-South Asia experiments for monsoonal precipitation over the Himalayan region during present climate: part I, Climate dynamics, 50, 2311–2334, doi:10.1007/s00382-015-2747-2, 2018.

660   Giorgetta, M. and Wild, M.: The water vapor continuum and its representation in ECHAM4, MPI for Meterolo., report no. 162, Hamburg, 1995.

Giorgi, F., Jones, C. and Asrar, G. R.: Addressing climate information needs at the regional level: the CORDEX framework, World Meteorological Organization (WMO) Bulletin, 58, 175, 2009.

Giorgi, F. and Gutowski Jr, W. J.: Regional dynamical downscaling and the CORDEX initiative, Annual Review of
665   Environment and Resources, 40, 467–490, 2015.

Giorgio, F. and Mearns, L. O.: Introduction to special section: Regional climate modeling revisited, J. Geophys. Res., 104, 6335–6352, doi:10.1029/98JD02072, 1999.

Giot, O., Termonia, P., Degrauwe, D., De Troch, R., Caluwaerts, S., Smet, G., Berckmans, J., Deckmyn, A., De Cruz, L., De Meutter, P., Duerinckx, A., Gerard, L., Hamdi, R., Van den Bergh, J., Van Ginderachter, M. and Van Schaeybroeck, B.:
670   Validation of the ALARO-0 model within the EURO-CORDEX framework, Geosci. Model Dev., 9, 1143–1152, doi:10.5194/gmd-9-1143-2016, 2016.

Gordon, C., Cooper, C., Senior, C. A., Banks, H., Gregory, J. M., Johns, T. C., Mitchell, J. F. B. and Wood, R. A.: The simulation of SST, sea ice extents and ocean heat transports in a version of the Hadley Centre coupled model without flux adjustments, Climate dynamics, 16, 147–168, doi:10.1007/s003820050010, 2000.

675   Gutowski Jr., W. J., Giorgi, F., Timbal, B., Frigon, A., Jacob, D., Kang, H.-S., Raghavan, K., Lee, B., Lennard, C., Nikulin, G., O'Rourke, E., Rixen, M., Solman, S., Stephenson, T., and Tangang, F.: WCRP COordinated Regional Downscaling EXperiment (CORDEX): a diagnostic MIP for CMIP6, Geosci. Model Dev., 9, 4087–4095, doi:10.5194/gmd-9-4087-2016, 2016.

Hagemann, S.: An improved land surface parameter data set for global and regional climate models, Max Planck Institute for
680   Meteorology report series, Hamburg, Germany, Report No. 336, 2002.

Hamdi, R., Van de Vyver, H. and Termonia, P.: New cloud and microphysics parameterisation for use in high-resolution dynamical downscaling: application for summer extreme temperature over Belgium, Int. J. Climatol., 32, 2051–2065, doi:10.1002/joc.2409, 2012.

Harris, I., Jones, P.D., Osborn, T.J. and Lister, D.H.: Updated high-resolution grids of monthly climatic observations – the
685   CRU TS3.10 Dataset, Int. J. Climatol., 34, 623–642, doi:10.1002/joc.3711, 2014.

Hofstra, N., Haylock, M., New, M. and Jones, P. D.: Testing E-OBS European high-resolution gridded data set of daily precipitation and surface temperature, Journal of Geophysical Research: Atmospheres, 114, doi:10.1029/2009JD011799, 2009.

Hofstra, N., New, M. and McSweeney, C.: The influence of interpolation and station network density on the distributions and trends of climate variables in gridded daily data, Climate dynamics, 35, 841–858, doi:10.1007/s00382-009-0698-1, 2010.

690 Hu, Z., Zhou, Q., Chen, X., Li, J., Li, Q., Chen, D., Liu, W. and Yin, G.: Evaluation of three global gridded precipitation data sets in central Asia based on rain gauge observations, International Journal of Climatology, 38, 3475–3493, doi:10.1002/joc.5510, 2018.

Jacob, D.: A note to the simulation of the annual and inter-annual variability of the water budget over the Baltic Sea drainage basin. Meteorol. Atmos. Phys., 77, 61–73, doi:10.1007/s007030170017, 2001.

695 Jacob, D., Bärring, L., Christensen, O. B., Christensen, J. H., De Castro, M., Déqué, M., Giorgi, F., Hagemann, S., Hirschi, M., Jones, R., Kjellström, E. Lenderink, G., Rockel, F., Sánchez, E., Schär, C., Seneviratne, S. I., Somot, S., van Ulden, A. and van den Hurk, B.: An inter-comparison of regional climate models for Europe: model performance in present-day climate, Climatic change, 81, 31–52, doi:10.1007/s10584-006-9213-4, 2007.

Jacob, D., Elizalde, A., Haensler, A., Hagemann, S., Kumar, P., Podzun, R., Rechid, D., Remedio, A. R., Saeed, F., Sieck, K.,
700 Teichmann, C. and Wilhelm, C.: Assessing the transferability of the regional climate model REMO to different coordinated regional climate downscaling experiment (CORDEX) regions, Atmosphere, 3, 181-199, doi:10.3390/atmos3010181, 2012.

Jacob, D., Petersen, J., Eggert, B., Alias, A., Christensen, O. B., Bouwer, L. M., Braun, A., Colette, A., Déqué, M., Georgievski, G., Georgopoulou, E., Gobiet, A., Menut, L., Nikulin, G., Haensler, A., Hempelmann, N., Jones, C., Keuler, K., Kovats, S., Kröner, N., Kotlarski, S., Kriegsmann, A., Martin, E., van Meijgaard, E., Moseley, C., Pfeifer, S., Preuschmann, S.,
705 Radermacher, C., Radtke, K., Rechid, D., Rounsevell, M., Samuelsson, P., Somot, S., Soussana, J.-F., Teichmann, C., Valentini, R., Vautard, R., Weber, B. and Yiou, P.: EURO-CORDEX: new high-resolution climate change projections for European impact research, Regional environmental change, 14, 563–578, doi:10.1007/s10113-013-0499-2, 2014.

Karger, D. N., Conrad, O., Böhner, J., Kawohl, T., Kreft, H., Soria-Auza, R. W., Zimmermann, N. E., Linder, H. P. and Kessler, M.: Climatologies at high resolution for the earth's land surface areas. Scientific data, 4, 170122,
710 doi:10.1038/sdata.2017.122, 2017.

Kotlarski, S.: A subgrid glacier parameterisation for use in regional climate modelling, PhD thesis, Reports on Earth System Science, Max Planck Institute for Meteorology, Hamburg, 2007.

Kotlarski, S., Keuler, K., Christensen, O. B., Colette, A., Déqué, M., Gobiet, A., Goergen, K., Jacob, D., Lüthi, D., van Meijgaard, E., Nikulin, G., Schär, C., Teichmann, C., Vautard, R., Warrach-Sagi, K., and Wulfmeyer, V.: Regional climate
715 modeling on European scales: a joint standard evaluation of the EURO-CORDEX RCM ensemble, Geosci. Model Dev., 7, 1297–1333, doi:10.5194/gmd-7-1297-2014, 2014.

Kotova L., Aniskevich S., Bobylev L., Caluwaerts S., De Cruz L., De Troch R., Gnatiuk N., Gobin A., Hamdi R., Sakalli A., Sirin A., Termonia P., Top S., Van Schaeybroeck B. and Viksna A.: A new project AFTER investigates the impacts of climate change in the Europe-Russia-Turkey region, Climate Services, 12, 64–66. doi:10.1016/j.cliser.2018.11.003, 2018.

720 Koenigk, T., Berg, P. and Döscher, R.: Arctic climate change in an ensemble of regional CORDEX simulations, Polar Research, 34, 24603, doi:10.3402/polar.v34.24603, 2015.

Kyselý, J., and Plavcová, E.: Biases in the diurnal temperature range in Central Europe in an ensemble of regional climate models and their possible causes, Climate dynamics, 39, 1275–1286, doi:10.1007/s00382-011-1200-4, 2012.

Laprise, R., Caya, D., Frigon, A. and Paquin, D.: Current and perturbed climate as simulated by the second-generation
725   Canadian Regional Climate Model (CRCM-II) over northwestern North America, Climate Dynamics, 21, 405–421, doi:
10.1007/s00382-003-0342-4, 2003.

Lohmann, U. and Roeckner, E.: Design and performance of a new cloud microphysics scheme developed for the ECHAM
general circulation model. Climate Dynamics, 12, 557–572, doi:10.1007/BF00207939, 1996.

Morcrette, J.-J., Smith, L. and Fouquart, Y.: Pressure and temperature dependence of the absorption in longwave radiation
730   parameterizations, Atmos. Phys., 59, 455–469, 1986.

New, M., Hulme, M. and Jones, P.: Representing twentieth-century space–time climate variability. Part I: Development of a
1961–90 mean monthly terrestrial climatology, Journal of climate, 12, 829–856, doi:10.1175/1520-
0442(1999)012<0829:RTCSTC>2.0.CO;2, 1999.

New, M., Hulme, M., and Jones, P.: Representing twentieth-century space–time climate variability. Part II: Development of
735   1901–96 monthly grids of terrestrial surface climate, Journal of climate, 13, 2217–2238, doi:10.1175/1520-
0442(2000)013<2217:RTCSTC>2.0.CO;2, 2000.

Nikulin, G., Jones, C., Giorgi, F., Asrar, G., Büchner, M., Cerezo-Mota, R., Christensenf, O. B., Déquég, M., Fernandezh, J.,
Hänsleri, A., van Meijgaardj, E., Samuelssona, P., Syllab, M. B. and Sushamak, L.: Precipitation climatology in an ensemble
of CORDEX-Africa regional climate simulations, Journal of Climate, 25, 6057–6078, doi:10.1175/JCLI-D-11-00375.1, 2012.

740   Nikulin, G., Lennard, C., Dosio, A., Kjellström, E., Chen, Y., Hänsler, A., Kupiainen, M., Laprise, R., Mariotti, L., Fox Maule,
C., van Meijgaard, E., Panitz, H.-J., Scinocca, J. F. and Somot, S.: The effects of 1.5 and 2 degrees of global warming on
Africa in the CORDEX ensemble. Environ. Res. Lett., 13, 065003, doi:10.1088/1748-9326/aab1b1, 2018.

Ozturk, T., Altinsoy, H., Türkeş, M. and Kurnaz, M. L.: Simulation of temperature and precipitation climatology for the Central
Asia CORDEX domain using RegCM 4.0, Climate Research, 52, 63–76, doi:10.3354/cr01082, 2012.

745   Ozturk, T., Turp, M. T., Türkeş, M., and Kurnaz, M. L.: Projected changes in temperature and precipitation climatology of
Central Asia CORDEX Region 8 by using RegCM4. 3.5, Atmospheric Research, 183, 296–307,
doi:10.1016/j.atmosres.2016.09.008, 2016.

Pfeifer, S.: Modeling cold cloud processes with the regional climate model REMO, PhD thesis, Reports on Earth System
Science, Max Planck Institute for Meteorology, Hamburg, 2006.

750   Pietikäinen, J.-P., O'Donnell, D., Teichmann, C., Karstens, U., Pfeifer, S., Kazil, J., Podzun, R., Fiedler, S., Kokkola, H.,
Birmili, W., O'Dowd, C., Baltensperger, U., Weingartner, E., Gehrig, R., Spindler, G., Kulmala, M., Feichter, J., Jacob, D.,
and Laaksonen, A.: The regional aerosol-climate model REMO-HAM, Geosci. Model Dev., 5, 1323–1339, doi:/10.5194/gmd-
5-1323-2012, 2012..

Pietikäinen, J.-P., Markkanen, T., Sieck, K., Jacob, D., Korhonen, J., Räisänen, P., Gao, Y., Ahola, J., Korhonen, H.,
755   Laaksonen, A. and Kaurola, J.: The regional climate model REMO (v2015) coupled with the 1-D freshwater lake model FLake
(v1): Fenno-Scandinavian climate and lakes, Geosci. Model Dev., 11, 1321–1342, doi:10.5194/gmd-11-1321-2018, 2018.

Rechid, D.: On biogeophysical interactions between vegetation phenology and climate simulated over Europe, PhD thesis, Reports on Earth System Science, Max Planck Institute for Meteorology, Hamburg, 2009.

Remedio, A.R., Teichmann, C., Buntemeyer, L., Sieck, K., Weber, T., Rechid, D., Hoffmann, P., Nam, C., Kotova, L. and

760   Jacob, D.: Evaluation of New CORDEX Simulations Using an Updated Köppen–Trewartha Climate Classification, Atmosphere, 10, 726, doi:10.3390/atmos10110726, 2019.

Roeckner, E., Arpe, K., Bengtsson, L., Christoph, M., Claussen, M., Dümenil, L., Esch, M., Giorgetta, M., Schlese, U. and Schulzweida, U.: The Atmospheric General Circulation Model Echam-4: Model Description and Simulation of the Present Day Climate, Report No. 218, Max-Planck-Institute for Meteorology: Hamburg, Germany, 1996.

765   Russo, E., Kirchner, I., Pfahl, S., Schaap, M., and Cubasch, U.: Sensitivity studies with the regional climate model COSMO-CLM 5.0 over the CORDEX Central Asia Domain, Geosci. Model Dev., 12, 5229–5249, doi:/10.5194/gmd-12-5229-2019, 2019.

Ruti, P. M., Somot, S., Giorgi, F., Dubois, C., Flaounas, E., Obermann, A., Dell'Aquila, A., Pisacane, G., Harzallah, A., Lombardi, E., Ahrens, B., Akhtar, N., Alias, A., Arsouze, T., Aznar, R., Bastin, S., Bartholy, J., Béranger, K., Beuvier, J.,

770   Bouffies-Cloché, S., Brauch, J., Cabos, W., Calmanti, S., Calvet, J.-C., Carillo, A., Conte, D., Coppola, E., Djurdjevic, V., Drobinski, P., Elizalde-Arellano, A., Gaertner, M., Galàn, P., Gallardo, C., Gualdi, S., Goncalves, M., Jorba, O., Jordà, G., L'Heveder, B., Lebeaupin-Brossier, C., Li, L., Liguori, G., Lionello, P., Maciàs, D., Nabat, P., Önol, B., Raikovic, B., Ramage, K., Sevault, F., Sannino, G., Struglia, M. V., Sanna, A., Torma, C. and Vervatis, V.: MED-CORDEX initiative for Mediterranean climate studies, Bulletin of the American Meteorological Society, 97, 1187–1208, doi:10.1175/BAMS-D-14-

775   00176.1, 2016.

Schneider, U., Becker, A. Finger, P. Meyer-Christoffer, A. and Ziese, M.: GPCC Full Data Monthly Product Version 2018 at 0.25°: Monthly Land-Surface Precipitation from Rain-Gauges built on GTS-based and Historical Data, doi: 10.5676/DWD_GPCC/FD_M_V2018_025, 2018.

Semmler, T., Jacob, D., Schlünzen, K. H. and Podzun, R.: Influence of sea ice treatment in a regional climate model on

780   boundary layer values in the Fram Strait region, Mon. Weather Rev., 132, 985–999, doi:10.1175/1520-0493(2004)132<0985:IOSITI>2.0.CO;2, 2004.

Solman, S. A., Sanchez, E., Samuelsson, P., da Rocha, R. P., Li, L., Marengo, J., Pessacg, N. L., Remedio, A. R. C., Chou, S. C., Berbery, H., Le Treut, H., de Castro, M. and Jacob, D.: Evaluation of an ensemble of regional climate model simulations over South America driven by the ERA-Interim reanalysis: model performance and uncertainties, Climate Dynamics, 41,

785   1139–1157, doi:10.1007/s00382-013-1667-2, 2013.

Souverijns, N., Gossart, A., Demuzere, M., Lenaerts, J. T. M., Medley, B., Gorodetskaya, I. V., Vanden Broucke, S. and van Lipzig, N. P. M.: A New Regional Climate Model for POLAR-CORDEX: Evaluation of a 30-Year Hindcast with COSMO-CLM2 Over Antarctica, Journal of Geophysical Research: Atmospheres, 124, 1405–1427, doi:10.1029/2018JD028862, 2019.

Tangang, F., Santisirisomboon, J., Juneng, L., Salimun, E., Chung, J., Cruz, F., Ngai, S. T., Ngo-Duc, T., Singhruck, P.,

790   Narisma, G., Santisirisomboon, J., Wongsaree, W., Promjirapawat, K., Sukamongkol, Y., Srisawadwong, R., Setsirichok, D.,

Phan-Van, T., Gunawan, D., Aldrian, E., Nikulin, G. and Yang, H.: Projected future changes in mean precipitation over Thailand based on multi-model regional climate simulations of CORDEX Southeast Asia, Int. J. Climatol., 39, 5413–5436, doi:10.1002/joc.6163, 2019.

Tangang, F., Supari, S., Chung, J. X., Cruz, F., Salimun, E., Ngai, S. T., Juneng, L., Santisirisomboon, J., Santisirisomboon,
795 J., Ngo-Duc, T., Phan-Van, T., Narisma, G., Singhruck, P., Gunawan, D., Aldrian, E., Sopaheluwakan, A., Nikulin, G., Yang, H., Remedio, A.R.C., Sein, D. and Hein-Griggs, D.: Future changes in annual precipitation extremes over Southeast Asia under global warming of 2 C. APN Science Bulletin, 8, 3–8, doi:10.30852/sb.2018.436, 2018.

Taylor, K. E.: Summarizing multiple aspects of model performance in a single diagram, Journal of Geophysical Research: Atmospheres, 106, 7183–7192, doi:10.1029/2000JD900719, 2001.

800 Termonia, P., Fischer, C., Bazile, E., Bouyssel, F., Brožková, R., Bénard, P., Bochenek, B., Degrauwe, D., Derková, M., El Khatib, R., Hamdi, R., Mašek, J., Pottier, P., Pristov, N., Seity, Y., Smolíková, P., Španiel, O., Tudor, M., Wang, Y., Wittmann, C., and Joly, A.: The ALADIN System and its canonical model configurations AROME CY41T1 and ALARO CY40T1, Geosci. Model Dev., 11, 257–281, doi:10.5194/gmd-11-257-2018, 2018a.

Termonia, P., Van Schaeybroeck, B., De Cruz, L., De Troch, R., Caluwaerts, S., Giot, O., Hamdi, R., Vannitsem, S., Duchêne,
805 F., Willems, P., Tabari, H., Van Uytven, E., Hosseinzadehtalaei, P., Van Lipzig, N., Wouters, H., Vanden Broucke, S., van Ypersele, J.-P., Marbaix, P., Villanueva-Birriel, C., Fettweis, X., Wyard, C., Scholzen, C., Doutreloup, S., De Ridder, K., Gobin, G., Lauwaet, D., Stavrakou, T., Bauwens, M., Müller, J.-F., Luyten, P., Ponsar, S., Van den Eynde, D. and Pottiaux, E.: The CORDEX.be initiative as a foundation for climate services in Belgium, Climate Services, 11, 49–61, doi:10.1016/j.cliser.2018.05.001, 2018b.

810 Top, S., Kotova, L., De Cruz, L., Aniskevich, S., Bobylev, L., De Troch, R., Gnatiuk, N., Gobin, A., Hamdi, R., Kriegsmann, A., Remedio, A. R., Sakalli, A., Van De Vyver, H., Van Schaeybroeck, B., Zandersons, V., De Maeyer, P., Termonia, P. and Caluwaerts, S.: R code validation analysis ALARO-0 and REMO2015 climate data Central Asia Top et al. 2020, Zenodo, doi:10.5281/zenodo.3659717.

Torma, C., Giorgi, F. and Coppola, E.: Added value of regional climate modeling over areas characterized by complex
815 terrain—Precipitation over the Alps, Journal of Geophysical Research: Atmospheres, 120, 3957–3972, doi:10.1002/2014JD022781, 2015.

Tustison, B., Harris, D. and Foufoula-Georgiou, E.: Scale issues in verification of precipitation forecasts. Journal of Geophysical Research: Atmospheres, 106, 11775–11784, doi:10.1029/2001JD900066, 2001.

Tuyet, N. T., Thanh, N. D., and van Tan, P.: Performance of SEACLID/CORDEX-SEA multi-model experiments in simulating
820 temperature and rainfall in Vietnam, Vietnam Journal of Earth Sciences, 41, 374–387, doi:10.15625/0866-7187/41/4/14259, 2019.

Wang, J. and Kotamarthi, V. R.: High-resolution dynamically downscaled projections of precipitation in the mid and late 21st century over North America, Earth's Future, 3, 268–288, doi:10.1002/2015EF000304, 2015.

Whan, K. and Zwiers, F.: The impact of ENSO and the NAO on extreme winter precipitation in North America in observations

825   and regional climate models, Climate Dynamics, 48, 1401–1411, doi:10.1007/s00382-016-3148-x, 2017.

Wilhelm, C., Rechid, D., and Jacob, D.: Interactive coupling of regional atmosphere with biosphere in the new generation regional climate system model REMO-iMOVE, Geosci. Model Dev., 7, 1093–1114, doi:10.5194/gmd-7-1093-2014, 2014.

Willmott, C.J. and Matsuura, K.: Smart interpolation of annually averaged air temperature in the United States, Journal of Applied Meteorology, 34, 2577–2586, doi:10.1175/1520-0450(1995)034<2577:SIOAAA>2.0.CO;2, 1995.

830   Willmott, C.J. and Robeson S.M.: Climatologically aided interpolation (CAI) of terrestrial air temperature, International Journal of Climatology, 15, 221–229, doi:10.1002/joc.3370150207, 1995.

Willmott, C.J., Rowe, C.M. and Philpot, W.D.: Small-scale climate maps: a sensitivity analysis of some common assumptions associated with grid-point interpolation and contouring. American Cartographer, 12, 5–16, doi:10.1559/152304085783914686, 1985.

835   Zhu, X., Zhang, M., Wang, S., Qiang, F., Zeng, T., Ren, Z. and Dong, L.: Comparison of monthly precipitation derived from high-resolution gridded datasets in arid Xinjian, central Asia, Quaternary International, 358, 160-170, doi:10.1016/j.quaint.2014.12.027, 2015.

Zou, L., Zhou, T. and Peng, D.: Dynamical downscaling of historical climate over CORDEX East Asia domain: A comparison of regional ocean-atmosphere coupled model to stand-alone RCM simulations, Journal of Geophysical Research:

840   Atmospheres, 121, 1442–1458, doi:10.1002/2015JD023912, 2016.

Table 1: Overview of the used reference datasets.

Dataset	Short name	Type	Resolution	Used variables	Frequency	Temporal coverage	Domain
gridded Climatic Research Unit TS dataset (version 4.02)	CRU	gridded station data	0.50°	2 m mean air temperature, 2 m maximum air temperature, 2 m minimum air temperature, precipitation	monthly	1901 - 2017	global land mass (excluding Antarctica)
Matsuura and Willmot, University of Delaware (version 5.01)	MW	gridded station data	0.50°	2 m mean air temperature, precipitation	monthly	1900 - 2017	global land mass
Global Precipitation Climatology Centre gridded dataset (version 2018)	GPCC	gridded station data	0.25°	precipitation	monthly	1891 - 2016	global land mass (excluding Antarctica)
ERA-Interim	ERA-Interim	reanalysis data	0.25°	2 m mean air temperature, precipitation	monthly	1979 - 2017	global

Table 2: Climatological mean CRU temperature (°C) for the 1980-2017 period over the CAS-CORDEX domain and biases (°C) of
845 **the RCMs (REMO and ALARO-0) and the other reference datasets (ERA-Interim and MW) against those CRU means.**

	DJF	MAM	JJA	SON	Annual
CRU	-9.35	5.87	19.23	5.72	5.44
REMO - CRU	0.48	-0.56	-0.33	0.01	-0.11
ALARO - CRU	0.83	-3.19	0.02	-0.03	-0.60
ERA-Interim – CRU	0.42	0.21	0.16	-0.02	0.19
MW - CRU	-0.41	-0.19	-0.09	-0.43	-0.28

Table 3: Spatial average over the CAS-CORDEX domain of climatological mean CRU minimum temperature (°C) for the 1980-2017 period and biases (°C) against those CRU means for REMO and ALARO-0.

	DJF	MAM	JJA	SON	Annual
CRU	-14.43	-0.22	13.18	0.40	-0.20
REMO - CRU	0.77	-0.25	0.60	1.09	0.55
ALARO - CRU	2.85	-1.71	1.10	1.42	0.90

850

Table 4: Spatial average over the CAS-CORDEX domain of climatological mean CRU maximum temperature (°C) for the 1980-2017 period and biases (°C) against those CRU means for REMO and ALARO-0.

	DJF	MAM	JJA	SON	Annual
CRU	-4.29	11.97	25.34	11.06	11.09
REMO - CRU	0.08	-1.24	-1.07	-0.71	-0.74
ALARO - CRU	-0.77	-4.84	-1.46	-1.24	-2.08

855

Table 5: Climatological mean CRU precipitation (mm month⁻¹) for the 1980-2017 period over the CAS-CORDEX domain and relative biases (%) against those CRU means for the RCMs (REMO and ALARO-0), and the other reference datasets (ERA-Interim, MW and GPCC).

	DJF	MAM	JJA	SON	Annual
mean CRU	30.38	43.46	87.03	47.72	52.26
REMO - CRU	-4	3	-23	-11	-12
ALARO - CRU	-9	-11	-25	-9	-16
ERA-Interim - CRU	-10	3	-11	-10	-8
MW - CRU	-30	-28	-28	-27	-28
GPCC - CRU	-31	-32	-27	-30	-29

860