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], 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
[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
[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 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 spread between the observational datasets 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 REMO model scores better for temperature whereas the ALARO-0 model prevails for precipitation. Studying annual cycles over specific subregions enables to get deeper insight into the strengths and weaknesses of both RCMs over the CAS-CORDEX domain. The evaluation of minimum and maximum temperature demonstrates that both models underestimate the daily temperature range. This publication demonstrates that the REMO and ALARO-0 RCMs can be used to perform climate projections over Central Asia and that the climate data can be used for in impact studies taking into account a bias correction for those regions where significant biases have been identified.

1 Introduction

There is a strong need for climate information at the regional-to-local scale that is useful and usable 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 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 the Africa (Nikulin et al., 2012; Nikulin et al., 2018), Europe (Jacob et al., 2014; Kotlarski et al., 2014), Mediterranean (Ruti et al., 2016) and North America (Diaconescu et al., 2016; Whan and Zwiers, 2017; Gibson, 2019) CORDEX regions (Gutowski et al., 2016). These large ensembles consist of more than ten different GCM-RCM combinations. In order to provide such ensembles over all CORDEX regions, 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; Cabos et al., 2019), South Asia (Ghimire et al., 2018), East Asia (Zou et al., 2016), South-East Asia (Tangang et al., 2018; Tangang et al., 2019; Tuyet et al., 2019), Australasia (Di Virgilio et al., 2019), Arctic (Koenigk et al., 2015; Akperov et al., 2018), Antarctic (Souverijns et al., 2019) and Middle East North Africa (Almazroui et al., 2016; Bucchignani et al., 2018). In addition, a new ensemble of climate change simulations covering all major inhabited regions with a spatial resolution of about 25 km has been established within the WCRP CORDEX COmmon Regional Experiment (CORE) Framework to support the growing demands for climate services (Remedio et al., 2019). Furthemore, a number of high-resolution global simulations at climatic timescales, with resolutions of at least 50 km in the atmosphere and 28 km in the ocean, have been performed within the Coupled Model Intercomparison Project 6 (CMIP6) (Haarsma et al., 2016).

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) 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 evaluation results of the COSMO-CLM 5.0 model run at 0.22° or 25 km resolution over the CAS-CORDEX region. The current study significantly extends our knowledge of the CAS-CORDEX domain by evaluating 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 we present the model evaluation through the use of so-called “perfect boundary conditions” taken from the reanalysis data and by comparing the downscaled results to observed data for the period 1980-2017. Such a study is necessary 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 evaluation 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-OBS observational dataset (Hofstra et al., 2009). However, in this study a slightly different approach is necessary due to the absence of an ensemble of RCM runs over Central Asia. Additionally, in some regions the quality of gridded observational datasets, constructed through interpolation or area-averaging of station observations, is poor due to over-smoothing of extreme values (Hofstra et al., 2010) and/or because of station observations that are nonrepresentative for their large-scale environments. This is particularly the case for orographically complex regions such as the Himalayas. The current study compares the model simulations with different gridded observational datasets and reanalysis data. When the different datasets show large deviations and a large spread, then their uncertainty is high and no robust conclusions can be drawn (Collins et al., 2013; Russo et al., 2019).

This study contains two assets: for the first time an in-depth evaluation of the RCMs ALARO-0 and REMO is performed at 0.22° spatial resolution over the CAS-CORDEX domain and we reflect on the impact of the observational datasets on the model evaluation. Such an analysis is a prerequisite in order to be able to use the climate data in a sound way for 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 (Sect. 2). 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 the annual cycle, seasonal and annual means, biases and variability of mean, minimum and maximum surface air temperature and precipitation. Further, we evaluate and provide a discussion of some remarkable anomalies in Sect. 4. In the final Sect. 5 we summarize the conclusions.

2 Methods

2.1 CORDEX Central Asia domain and subdomains

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 open ocean. It contains several important mountain ranges e.g. Ural, Caucasus, Altay and Himalaya, and deserts e.g. 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 with a spatial resolution less than 50 km 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.

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° 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 for both RCMs consists of the inner domain plus a relaxation zone of eight grid points at every boundary.

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 REMO, were already run and evaluated 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).

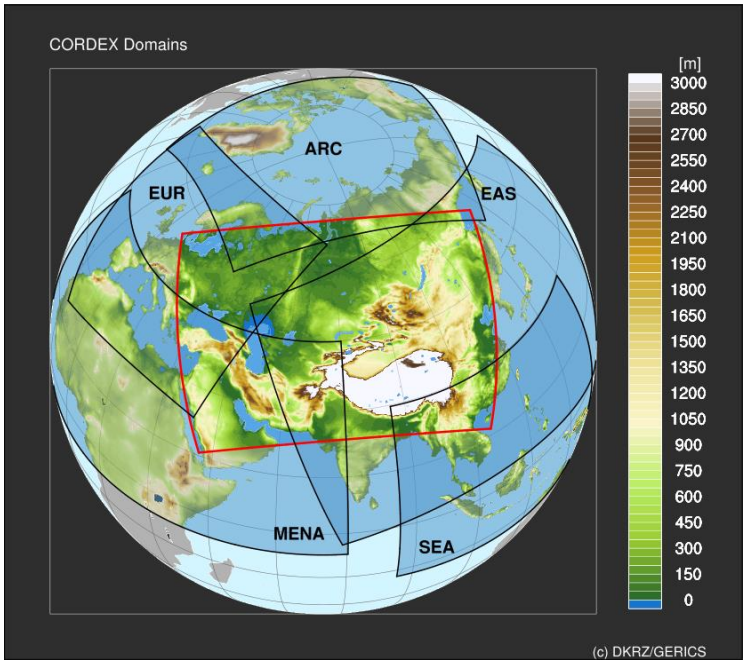

Figure 1: The CAS-CORDEX domain demarcated by a red contour and main overlapping CORDEX domains (black contour lines): Europe (EUR), Arctic (ARC), South East Asia (SEA), East Asia (EAS) and MENA projected upon the topography of Eurasia(geopotential height [m] of the GTOPO30 global digital elevation model (DEM) 3).

The CAS-CORDEX domain was further subdivided into five subdomains according to the IPCC reference regions (Iturbide et al., 2020) named as: East Europe, West Siberia, East Siberia, West Central Asia and Tibetan Plateau. These subdomains,

visualized in Fig. S1 of the supplementary material, were applied to evaluate the spatial differences in the study area and to investigate whether there were differences in the simulation of subcontinental processes.

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 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 model development was initiated by the Max-Planck-Institute for Meteorology and is further maintained and extended by the Climate Service Center Germany (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 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 evaluate both RCMs, a run driven by a large-scale forcing taken from the ERA-Interim global reanalysis (Dee et al., 2011) is undertaken for the period 1980-2017. A one-way nesting strategy is applied to dynamically downscale the ERA-Interim data, having a horizontal resolution of about 0.70° (approximately 79 km), to a higher 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 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). While for ALARO-0 the year 1979 was taken as spin-up year, REMO was spun-up for 10 years to produce an equilibrium for the soil temperature and soil moisture. These soil fields were then used as initial soil conditions when restarting the model from 1979. 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 compared with different reference datasets. Gridded datasets are based on interpolated station data and are used instead of station observations to overcome the scale difference between the model and observation field (Tustison et al., 2001). A multitude of datasets were considered to estimate the reliability of the gridded observational temperature and precipitation, since all gridded datasets are characterized by uncertainties (Gómez-Navarro et al., 2012). 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.

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 1901-2018 at a grid resolution of 0.50° covering the complete global land mass (excluding Antarctica) (Harris et al., 2020). 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 of disciplines, however, some issues have been reported (Harris et al., 2020). Main concerns include sparse coverage of measurement stations over certain regions, e.g. the Northern Russia and the dissimilarities in measurement methods that are used by different countries (Harris et al., 2020).

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 and Matsuura, 1995; Harris et al., 2020). It is known that the MW dataset generally underestimates the precipitation in the central part of the CAS-CORDEX domain but especially during spring (Hu et al., 2018). The MW dataset contains globally up to 0.4 °C warmer temperatures for the latest decades compared to CRU (Harris et al., 2020).

2.3.3 Global Precipitation Climatology Centre dataset

The Global Precipitation Climatology Centre (GPCC) (version 2018) of the German Weather Service is a monthly land surface precipitation dataset at 0.25° resolution based on rain gauge measurements. The GPCC full data monthly product version 2018 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 for the central part of our domain 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, especially during spring. In addition, the GPCC has no similar dataset for other variables and thus, only precipitation can be validated with this dataset.

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. The spatial resolution of the dataset is approximately 79 km (T255 spectral) with 60 levels in the vertical direction from the surface up to 0.1 hPa (Dee et al., 2011). The ERA-interim data have been further interpolated and used as forcing for both RCMs at a spatial resolution of 0.25°. Total monthly precipitation 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 and 12:00. The Monthly Means of Daily Means data of 2 m temperature are used to study the spread between observational gridded datasets and reanalysis data. 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

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). The output of the RCMs was upscaled and bilinearly interpolated to the 0.50° resolution grid of the observational gridded datasets.

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 diurnal temperature range was obtained by subtracting the minimum temperature from the maximum temperature and a height correction was performed for mean, minimum and maximum temperature 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 monthly, seasonal and annual climatological means of the evaluated variables to obtain graphs of the annual cycle maps that visualise the spatial patterns of the bias between the RCMs and reference datasets. The relative bias for precipitation is computed by subtracting the CRU value from the RCM or any other reference dataset and 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 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 results of the model evaluation are presented with a focus on evaluation metrics of seasonal means of mean, 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 limitations are investigated by comparing the different observational datasets and their implications for the evaluation will be described in Sect. 4.

3.1 Mean temperature

3.1.1 Annual and seasonal means over CAS-CORDEX domain

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. Moreover, the spread between the reference datasets (ERA-Interim, MW and CRU) is shown in the column at the right. Both RCMs are producing similar mean annual temperature patterns since they have similar biases with respect to CRU. At the same time a dipole pattern arises in the temperature bias of ALARO-0 between north and south and for REMO between east and west, with a peak in positively biased temperatures over north-western Mongolia. Annual biases vary between -3°C and 3°C for both RCMs, apart from the orographically complex regions and some areas in North and East Siberia for ALARO-0. These regions exhibit a spread of 3 °C and more between the observational datasets and thus it is difficult to evaluate the models accurately in those regions.

On the seasonal timescale, biases over larger areas are mainly pronounced in winter (DJF) and spring (MAM), particularly for ALARO-0 with strong biases up to 10 °C and -15 °C respectively in the north-eastern part of the domain. In winter the most pronounced bias is found for REMO over the north-western part of Mongolia in the Altai mountains. Additionally, the REMO model has a cold bias in the western part of Russia during winter, while ALARO-0 shows a warm bias. During spring, cold biases are found for both models in the northern part of the domain, but the biases of ALARO-0 are more pronounced than those of REMO. For the summer (JJA) season, warm biases occur over the southern part of the domain and cold biases are more dominant in the north. These biases in summer are more pronounced for ALARO-0. Both models show modest bias patterns in autumn (SON), with in particular warm biases over the eastern part of the domain.

Biases in the high-altitude regions are largely persistent throughout the seasons. More specifically, both RCMs 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. As mentioned before and visualised in Fig. 2, the biases in mountainous regions should be placed in perspective to the significant observational uncertainties that are typical over such complex orography.

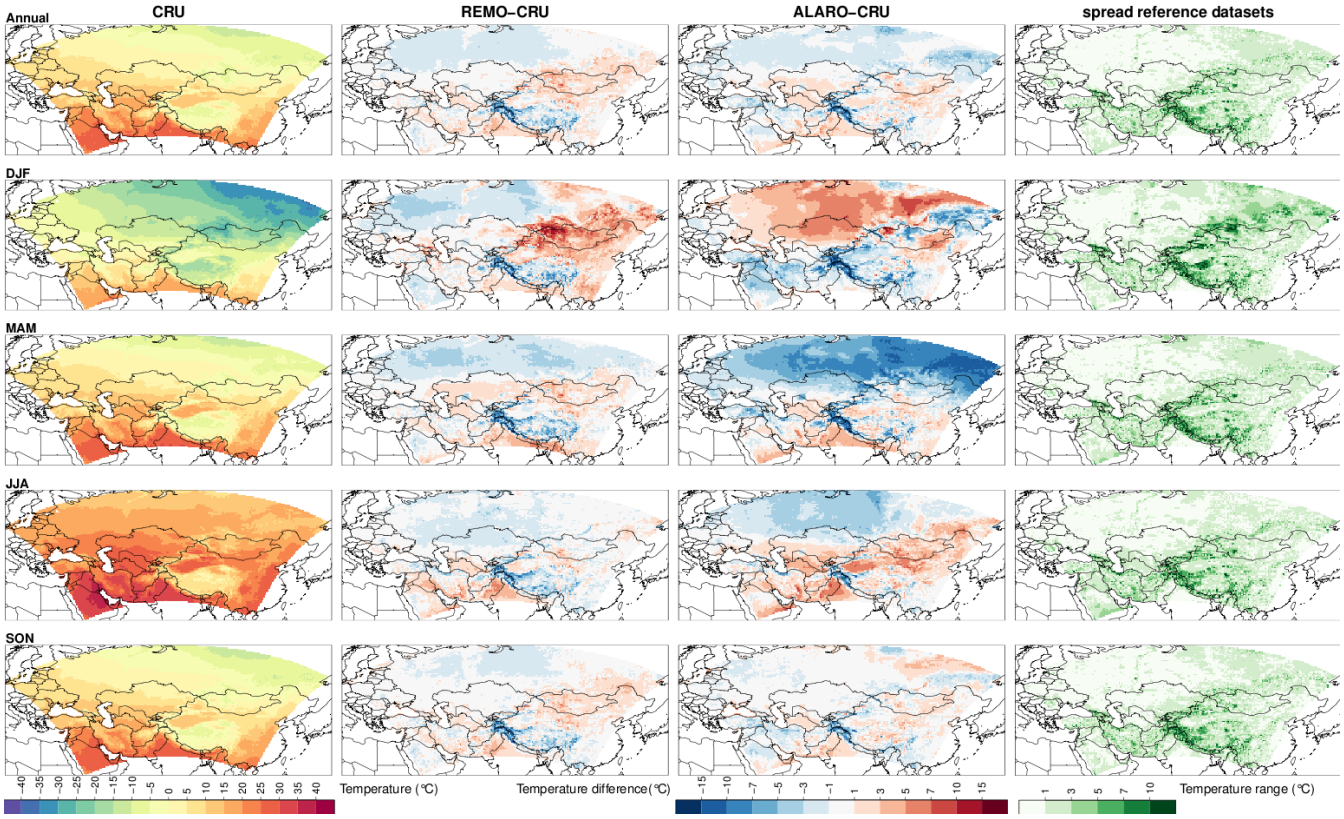

Figure 2: On the left: mean air temperature (°C) at 2 m height over the CAS-CORDEX domain based on the observational CRU dataset for the 1980-2017 period on annual level and for winter (DJF), spring (MAM), summer (JJA) and autumn (SON). In the middle columns: temperature difference (°C) between the simulated REMO mean temperature and the CRU mean temperature, and temperature difference (°C) between the simulated ALARO-0 mean temperature and the CRU mean temperature. On the right: the range in temperature (°C) between the different reference datasets (CRU, MW and ERA-Interim).

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. Figure 3 shows a 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 CAS-CORDEX domain for 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 lowest during summer for ALARO-0 while, the biases during summer are smaller than during winter and spring for both RCMs (Table 2 and Fig. 2). This is related to less spatial variability in temperatures during summer compared to the other seasons, as can be seen in Fig. 2 for CRU. An equal bias in temperature for each season would 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, although there was a clear warm bias observed during winter (Table 2 and Fig. 2). During spring the cold bias in the

north is limited to -5 °C 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). Comparing the metrics of the RCMs (Fig. 2, Fig. 3 and Table 2) shows that REMO is better in simulating the variability in temperature 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.

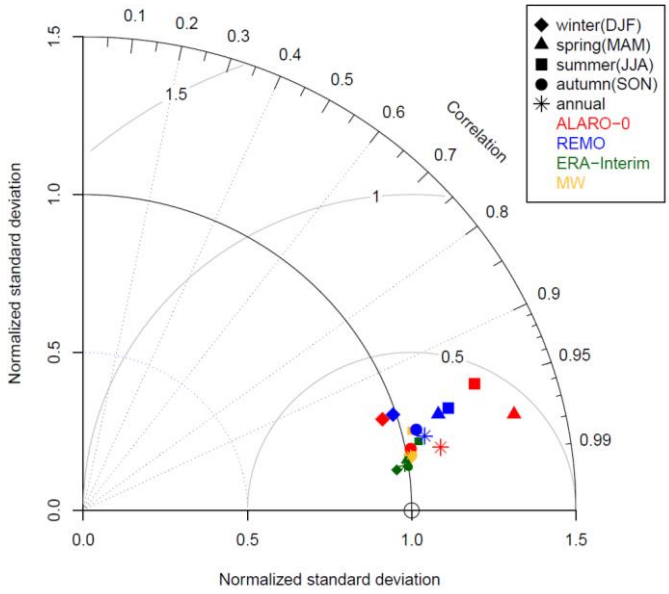

Figure 3: Normalized Taylor diagram representing the performance of mean temperature for seasonal and annual means for both RCMs (ALARO-0 and REMO), the ERA-Interim reanalysis and MW observational data with respect to CRU.

3.1.2 Annual cycles over subdomains

When analysing the seasonal cycle of the mean temperature for the different subdomains (Fig. 4), it is indeed observed that the RCMs simulate the mean temperature very well during the autumn months (months 9, 10 and 11). In the northern subdomains East Europe and West Siberia, there is on average a strong warm bias in December and January for ALARO-0, reaching a maximum of respectively 4.1 °C and 5.8 °C during December. During winter months (months 12, 1 and 2) REMO simulates temperatures within the uncertainty range for West Siberia and underestimates the temperatures on average by 1.4 °C in January over East Europe. REMO simulates warm biases around 2 °C in December and January over East Siberia. On

average there is no strong warm bias observed for ALARO-0 during the winter months in East Siberia due to the compensation effect of cold biases, both in time (Fig.4) and space (Fig. 2). Furthermore, there is a remarkable cold bias observed for ALARO-0 during spring (months 3, 4 and 5) and June in the northern subdomains East Europe, West Siberia and East Siberia, reaching up to -7.3 °C over East Siberia during April. REMO is performing well during spring months over the northern subdomains. Compared to the northern subdomains, ALARO-0 simulates the annual cycle better for the southern subdomains West Central Asia and Tibetan Plateau but slightly overestimates the amplitude of the annual temperature cycle. REMO simulates the mean temperature very well over the West Central Asian subdomain with only a slight overestimation of the temperatures in July and August. In the mountainous area of the Tibetan Plateau REMO underestimates the temperatures, except for January and December. The better results in spring, summer and autumn for ALARO-0 over the subdomain Tibetan Plateau are due to spatial averaging of cold biases in the northern Himalayas and warm biases over the Taklamakan Desert and the opposite is true for REMO during winter (Fig. 2).

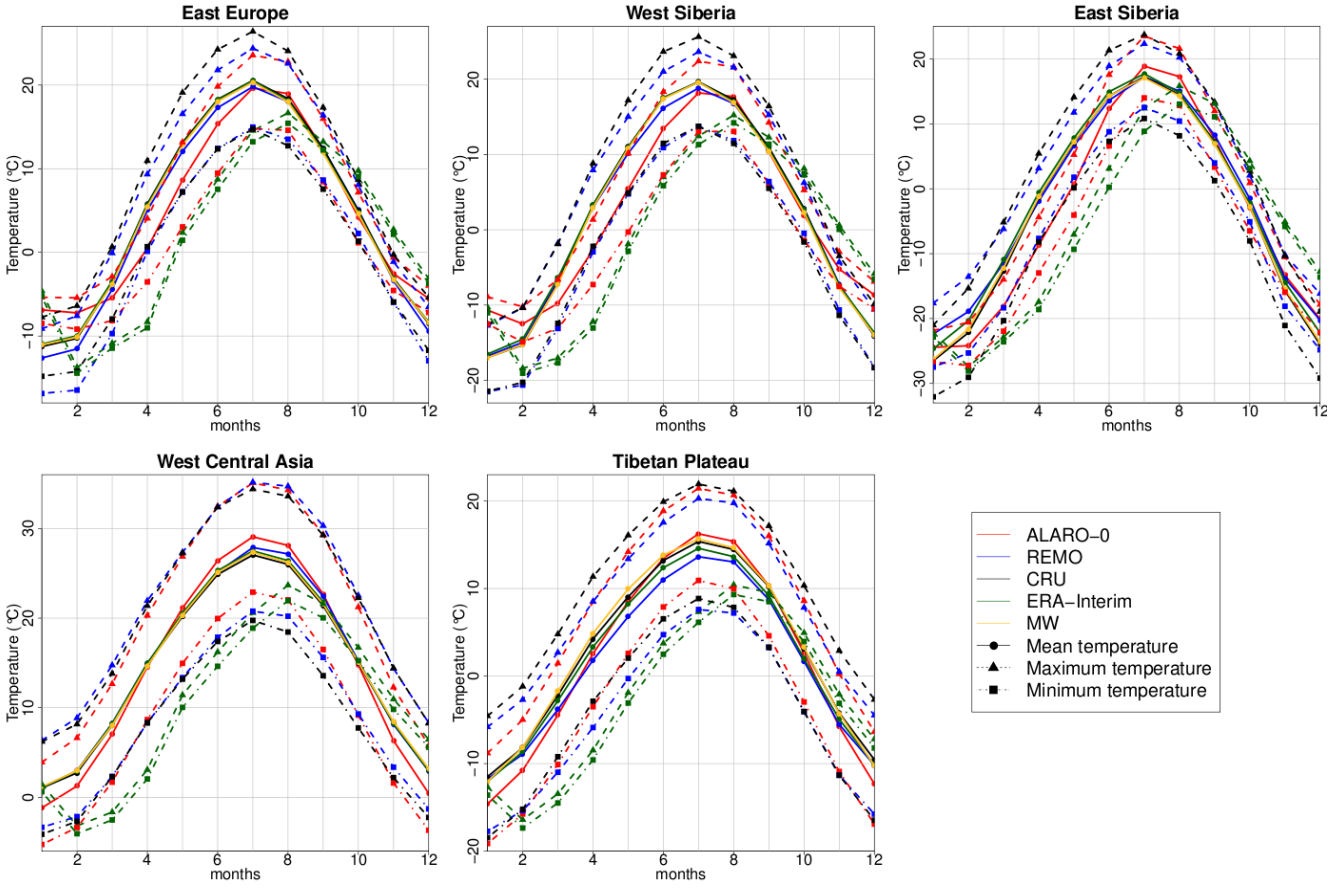

Figure 4: Annual cycles of the mean, minimum and maximum temperature for both RCMs (ALARO-0 and REMO) compared to the ERA-Interim reanalysis, MW and CRU observational data over five subdomains.

3.2 Diurnal temperature range

3.2.1 Annual and seasonal means over CAS-CORDEX domain

The diurnal temperature range is found by subtracting the minimum temperature from the maximum temperature. Therefore, minimum and maximum temperature are first discussed and then conclusions for the diurnal temperature range are deduced. Similar to 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. At the annual scale, the bias of the minimum temperature ranges mostly between -3 °C and 3 °C for REMO and between 0 °C and 5 °C for ALARO-0 (Fig. 5). Compared to ALARO-0, the REMO model shows larger warm biases over Mongolia during all seasons, except for summer. These warm biases are most pronounced during winter. ALARO-0 shows as well large biases up to 15 °C, but they cover the northern part of the domain while the warm biases for REMO cover the eastern part of the domain. Moreover, strong cold biases are present in the north during spring for both models, but they are more pronounced for the ALARO-0 model with biases up to -10 °C in the north-eastern part of the domain. During the summer season the biases for the REMO model are limited between -5 °C and 7 °C except for the Himalayan mountain range, while the ALARO-0 model output has, except for the Himalayas, a cold bias up to -7 °C in the north-western part of Russia and warm bias up to 10 °C in the southern and eastern part of the domain (Fig. 5). In autumn, both models have a warm bias over almost the entire domain, except for the cold biases in the mountainous areas, the Arabian Peninsula, northern Iran and for REMO also in the central northern part of the domain. The increased minimum temperatures obtained with the RCMs indicate that they do not capture the coldest diurnal temperatures.

The metrics in Fig. 6 show that the RCMs simulate the minimum temperature spatially well for annual and seasonal means. ALARO-0 has at annual and seasonal scale, except for summer, a slightly better spatial pattern correlation with the minimum temperatures of the CRU dataset than REMO. On the other hand, REMO better simulates the variability and mean minimum temperature, except for autumn where ALARO-0 simulates the variability better (Fig. 6 and Table 3).

The maximum temperatures are underestimated by both RCMs and this underestimation is more pronounced for ALARO-0 than for REMO at the annual scale and for all seasons (Fig. 7 and Table 4). Figure 7 shows that the cold bias is especially present in the northern part of the domain during spring and to a lesser extent during summer for both RCMs. In autumn the cold bias in the north is limited to -3 °C, but some stronger biases up to -7 °C appear in the north-east for the ALARO-0 model. The warm biases during autumn are limited to 5 °C and excluding the Himalayas, the smallest range in biases is obtained for both RCMs during this season. During winter, a negative spatially averaged bias of -0.77 °C is obtained for the mean maximum temperature of ALARO-0 and a small positive bias of 0.08 °C for REMO (Table 4). These limited spatial biases are obtained by biases with an opposite sign in different parts of the domain. REMO has cold biases in the north-west and warm biases in the east, except for the Tibetan Plateau, while ALARO-0 produces warm biases in the north and cold biases in the south-west and north-east.

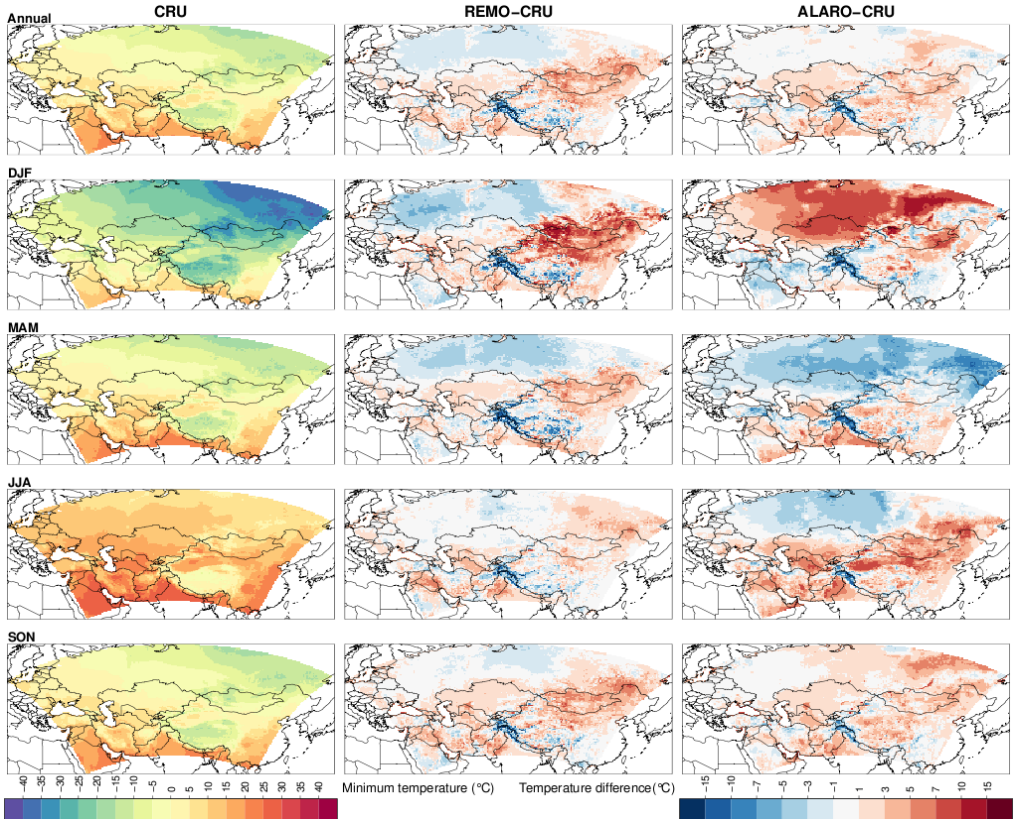

Figure 5: On the left: minimum air temperature (°C) at 2 m height over the CAS-CORDEX domain based on the observational CRU dataset for the 1980-2017 period on annual level and for winter (DJF), spring (MAM), summer (JJA) and autumn (SON). In the middle: temperature difference (°C) between the simulated 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.

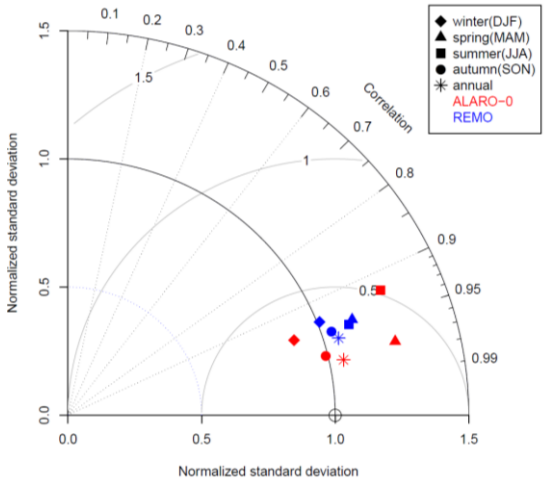

Figure 6: Normalized Taylor diagram representing the model performance of the minimum temperature for seasonal and annual means for both RCMs (ALARO-0 and REMO) with respect to CRU.

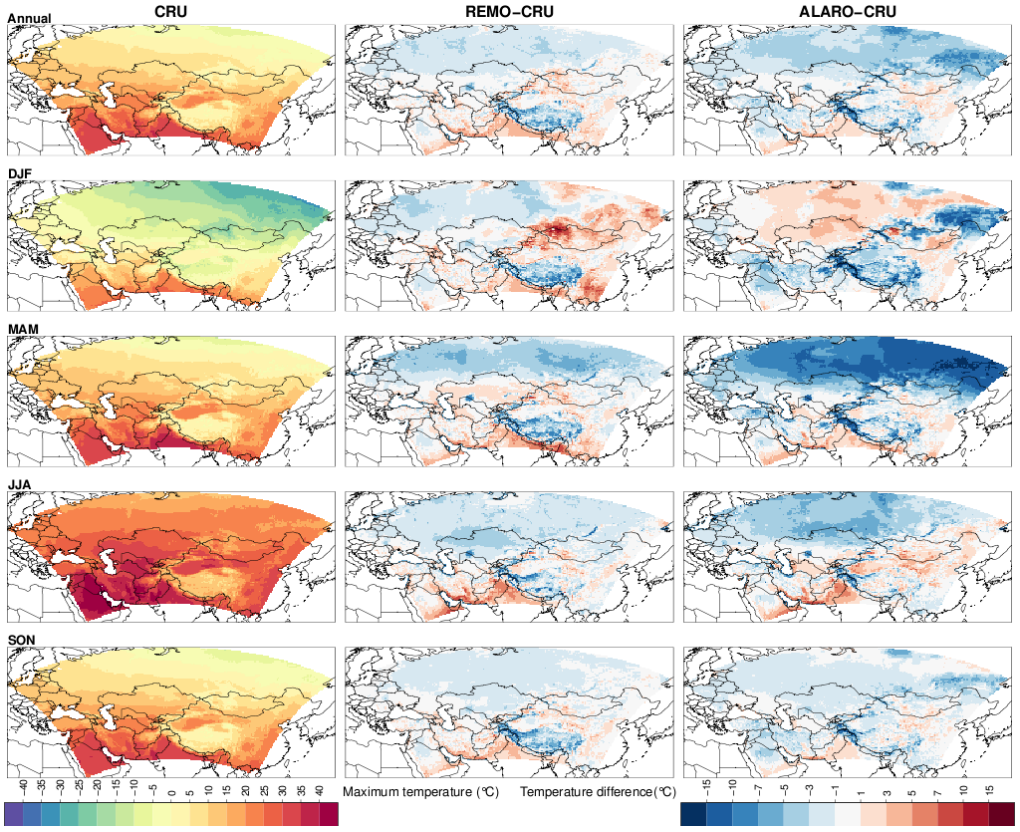

Figure 7: On the left: maximum air temperature (°C) at 2 m height over the CAS-CORDEX domain based on the observational CRU dataset for the 1980-2017 period on annual level and for winter (DJF), spring (MAM), summer (JJA) and autumn (SON). In the middle: temperature difference (°C) between the simulated 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.

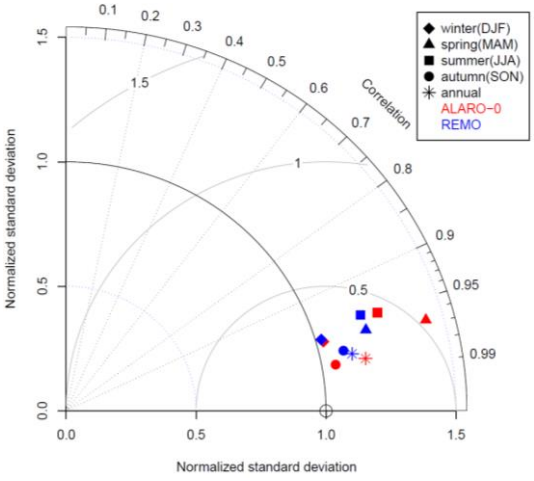

Figure 8: Normalized Taylor diagram representing the model performance of the maximum temperature for seasonal and annual means for both RCMs (ALARO-0 and REMO) with respect to CRU.

Figure 8 shows that both models have an acceptable model performance for maximum temperature 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. 8). Based on Fig. 7 and 8, both RCMs simulate the maximum temperature best during autumn.

In general the minimum temperature (Table 3 and Fig. 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 minimum and maximum temperatures are simulated less extremely by the models over most of the domain compared to the observational CRU dataset. In other words, the daily temperature range is generally underestimated by both RCMs.

3.2.2 Annual cycles over subdomains

Moreover, the annual cycles in Fig. 4 show that both minimum and maximum temperatures are overestimated by ALARO-0 during winter in the northern part of the domain, while they are underestimated during spring. In summer the model is able to restore its balance and to simulate temperatures as they are observed. For REMO the maximum temperature is underestimated during winter, spring and summer in East Europe, while the minimum temperature is only strongly underestimated during winter. REMO overestimates the minimum temperatures during the complete annual cycle for East Siberia, while the maximum temperatures in East Siberia are only overestimated during winter and underestimated during spring and summer. Both RCMs underestimate the maximum temperatures for the entire annual cycle over the Tibetan Plateau subregion. ALARO-0 underestimates the minimum temperatures during the winter months and overestimates them during the summer months, while REMO slightly overestimates winter and underestimates summer minimum temperatures.

3.3 Precipitation

3.3.1 Annual and seasonal means over CAS-CORDEX domain

In Table 5, the spatially averaged precipitation over the 1980-2017 period is given for CRU. 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. Figure 9 shows that the annual precipitation for both models lies mostly within the spread of the different reference datasets. Furthermore, a strong wet bias is present during winter for both RCMs over the south-eastern region and for REMO this wet bias extends even further up north to the Russian-Mongolian border. This large wet bias during the winter is partly due to the low precipitation quantities in several regions e.g. less than 5 mm per month in the Taklamakan and Gobi desert regions. The largest relative biases can be found in relatively dry regions and therefore the absolute biases are presented in the supplementary material Fig. S2 and Table S2. When the absolute bias during winter is examined, then it is seen that REMO only simulates a very small absolute overestimation in precipitation over Mongolia and the northern part of China, but both RCMs do overestimate the precipitation in the South-East Asian monsoon

region during winter and spring (Fig. 9). The wet bias of the ALARO-0 model in the south-eastern CAS-CORDEX region is situated within the spread of the different reference datasets (Fig. 9 and 11). In summer, when most rain falls due to the East Asian Monsoon, a dry bias is present (Fig. 9 and S2).

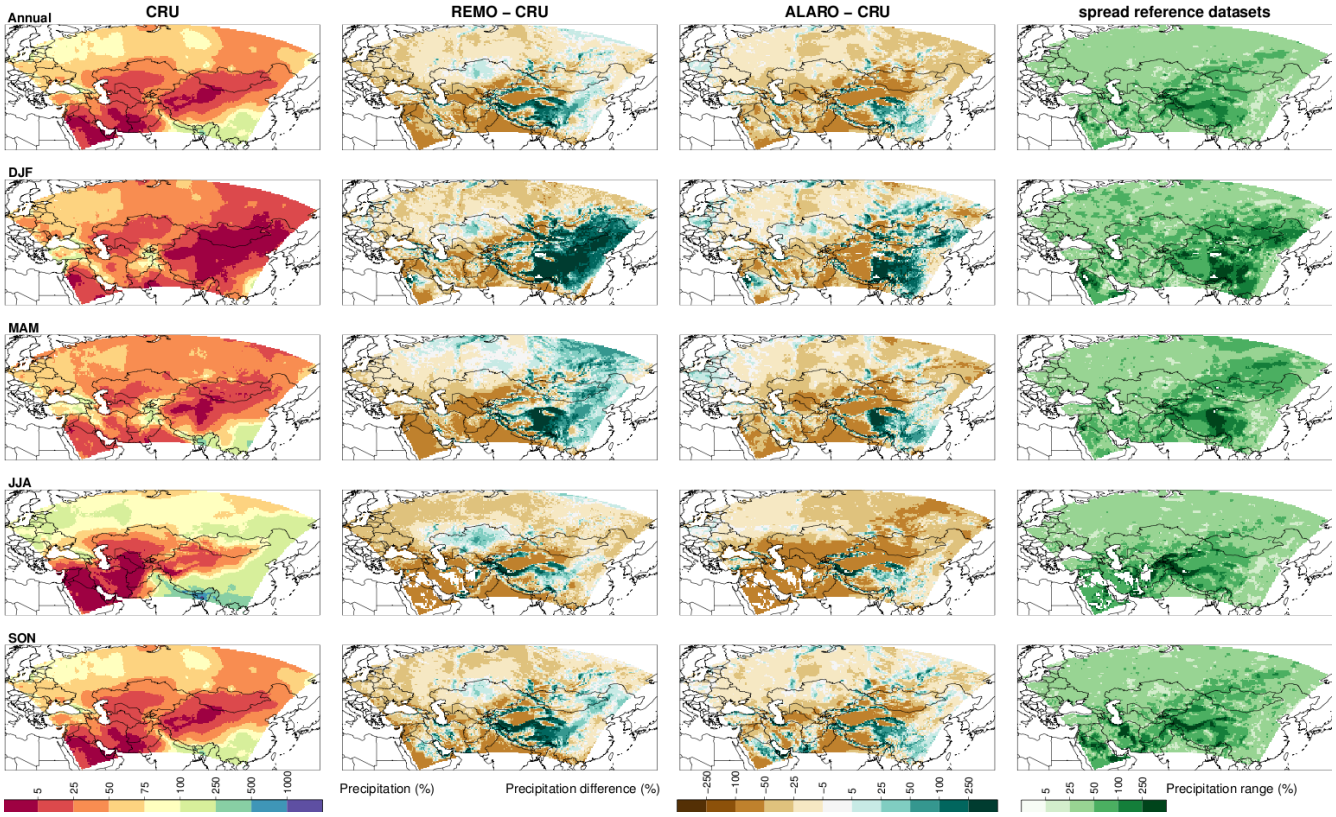

Figure 9: Left: mean monthly precipitation amounts (mm month-1) over the CAS-CORDEX domain based on the observational CRU dataset for the 1980-2017 period on annual level and for winter (DJF), spring (MAM), summer (JJA) and autumn (SON). In the middle: relative difference between the average annual and seasonal CRU precipitation and the precipitation simulated by REMO and ALARO-0 (%). Right: the range in precipitation (%) between the different reference datasets (CRU, MW, GPCC and ERA-Interim).

Next to these biases in the monsoon region, both models show dry biases over the Tarim basin and the south-western part of the domain during spring and summer. The Taklamakan and Arabian deserts are located here, which are already dry regions in the CRU dataset (Fig. 9). The absolute biases over this region are less pronounced in Fig. S2. 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 (Fig. S2).

From Fig. 10 can be deduced that ALARO-0 is better than REMO in capturing the annual and seasonal variations in precipitation since the RSVs are closer to 1. Additionally, ALARO-0 better captures the spatial patterns since the correlations are larger than those for REMO. The dry biases for ALARO-0 in Table 5 are thus caused by the simulation of systematically

less precipitation compared to CRU over most parts of the domain (Fig. 9 and 11). Both RCMs show the largest error in normalized RSV during spring. This too large spatial variation is due to an overestimation of the precipitation in the wettest region combined with an underestimation in the driest region of the CAS-CORDEX domain (Fig. 9). During summer, both RCMs underestimate the variability in precipitation (Fig. 10).

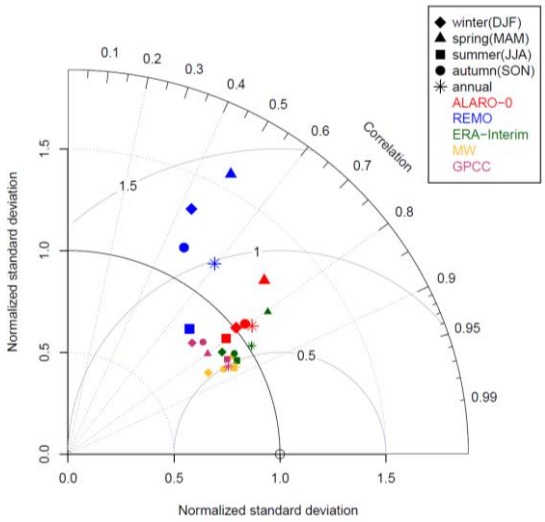

Figure 10: Normalized Taylor diagram representing the model performance of precipitation for seasonal and annual means for both RCMs (ALARO-0 and REMO), gridded observational datasets (MW, GPCC) and the ERA-Interim reanalysis data with respect to CRU.

3.3.2 Annual cycles over subdomains

The annual cycles over the subdomains show that ALARO-0 and REMO indeed mostly underestimate the precipitation values of CRU in the different subdomains, but for East Europe and the Tibetan Plateau the precipitation amounts are higher than those of MW and GPCC and are thus within the range of observational spread (Fig. 11). ALARO-0 does underestimate the precipitation slightly in May and June over West Siberia and in June and July over East Siberia. For the West Central Asian subdomain, both RCMs underestimate the precipitation in spring and summer. REMO overestimates the precipitation slightly over the East Siberian subdomain in spring. Additionally, it is seen that REMO is unable to simulate the annual cycle of precipitation correctly over the subdomain of the Tibetan Plateau (Fig. 11). The precipitation rates are too high, except during the summer when the Asian Monsoon takes place.

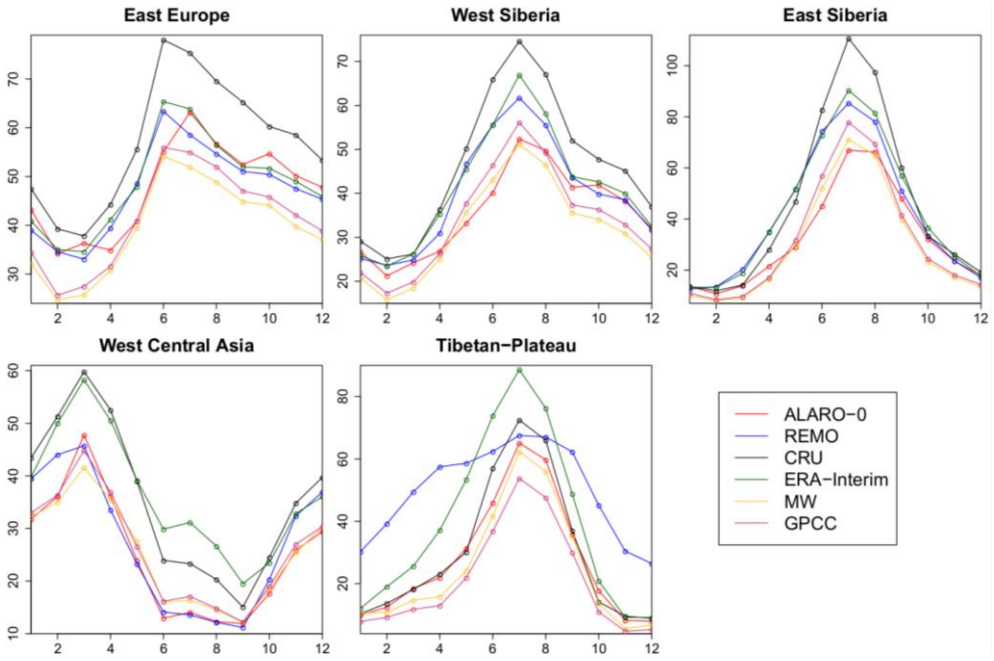

Figure 11: Annual cycles of precipitation (mm/month) for both RCMs (ALARO-0 and REMO) compared to the ERA-Interim reanalysis, MW, GPCC and CRU observational data over five subdomains.

4 Discussion

4.1 Temperature

4.1.1 CAS-CORDEX domain

When comparing the above results of temperature with the other reference datasets (Fig. 3), the normalized standard deviation of ERA-Interim and MW differs less from CRU than the RCMs during spring and summer. 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 well, even though they were slightly deviating from the spatial temperature patterns in the CRU data. The latter can also be explained by the spread of the reference datasets in Fig. 2: larger biases between the RCMs and CRU are especially located in regions where the spread between the different reference datasets is high, which means that there is a large observational uncertainty at those locations. Figure 3 shows that the larger RSVs of the RCMs during summer are partly 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 observed spatial patterns are less reliable during summer since the two other reference datasets both show a lower spatial correlation with CRU during summer compared to the other seasons. The lower performance of the RCMs during summer can thus partly be explained by the observational uncertainty in

435 spatial variation of temperatures. This is more pronounced for the summer season since the spatial variation in temperature is lower during this season. Ozturk et al. (2016) reported a lower spatial correlation during summer with 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° horizontal resolution when compared to the results of Ozturk et al. (2016). Zhu et al. (2020) obtained spatial correlations that are slightly lower than those obtained for ALARO-0 and REMO. They obtained a slightly larger spatial
440 variation during winter and similar results for the spatial variation in summer and on annual level. Although, it should be mentioned that their domain is smaller than the CAS-CORDEX domain and they used a different observational dataset which makes comparison difficult.

Larger differences between temperatures of the reference datasets in the region of the Tibetan Plateau (Fig. 2) were also observed by Ozturk et al. (2012 and 2016) and Russo et al. (2019) and this is due to the fact that gridded data is based on
445 measurements of meteorological stations in the valleys (New et al., 1999). This is the case for the gridded observational data of CRU and MW (Fig. 4). The gridded observations are thus less reliable over the Himalayas and Tibetan Plateau and cause a bias of the RCMs within the range of observational uncertainty. Further, the amplification of the biases over the mountainous regions for the RCMs can be attributed to the used assumption of the lapse rate of 0.0064 K m-1 for the elevation correction (Kotlarski et al., 2014).

450 When comparing the mean spatial biases for the 1980-2017 period (Table 2), then it is seen that the differences between the observational datasets are smaller than the differences between the RCMs and CRU, except for autumn for both RCMs and for REMO on the annual level. Additionally, Fig. 2 and 4 show that for most parts of the domain the mean temperatures of ALARO-0 and REMO are lying within the range of spread between the reference datasets during autumn. From this we conclude that both RCMs simulate temperatures in autumn within the range of observational uncertainty. During winter, spring
455 and summer none of the RCMs are able to reproduce temperatures that can be completely explained by the observational uncertainty (Fig. 2 and Table 2).

In the following subsection the temperature biases over snow covered areas during winter and spring are explained. For summer temperatures, Russo et al. (2019) found with COSMO-CLM 5.0 a spatial pattern with a cold temperature bias in the north and warm biases in the southern part of the domain except for some locations on the Tibetan Plateau, which is similar to ALARO-
460 0. In general both ALARO-0 and REMO produced biases within a similar order of magnitude as was obtained with other RCMs over the CAS-CORDEX region (Russo et al., 2019) and Central Asian subdomains (Wang et al., 2020; Zhu et al., 2020).

4.1.2 Spring and winter biases in northern subdomains

The cold bias for REMO during winter over the East European subdomain is likely due to the surface treatment of the model
465 when there is snow (Pietikäinen et al., 2018). Pietikäinen et al. (2018) already reported that the thermodynamics of the snow layer plays an important role in the cold bias that appears over East Europe during the months when snow cover is present.

Although this cold bias occurs in the north-west, both models are producing on average temperatures that are too warm in winter and too cold in spring (Table 2).

New et al. (1999) mentioned that CRU contains colder temperatures in winter over Russia. The range between the different reference datasets is larger for the East Siberian subdomain, indicating that there is a larger uncertainty for this subdomain during winter. This observational uncertainty could explain the warm biases for both RCMs over the mountain ranges Altai, Yablonovy and Stanovoy since the spread between the reference datasets is larger than the obtained biases (Fig. 2). However, this is not the main reason for the warm bias over Russia since the spread between the reference datasets is smaller than the obtained biases.

Moreover, during winter the RCMs simulate warm biases in different regions, while in spring they both show a cold bias over the north (Fig. 2 and 4). Compared to the northern part in the CAS-CORDEX region, a similar warm bias during winter was found over Scandinavia in the EURO-CORDEX runs with ALARO-0 (Giot et al., 2016). Both regions have a similar climate which suggests that similar physical processes might be at the basis of these biases. The warm bias during winter and cold bias during spring in the north-eastern part of the domain are not due to a shift in the annual cycle in the northern part of the domain, although there is a delay in warming temperatures during spring.

A limited warm bias arises in the north during autumn, when the first snow cover appears over this region. This bias increases when the snow covered region expands. ALARO-0 seems to underestimate cooling above snow cover during stable conditions (Fig. 4). Mašek (2017) linked too warm temperatures above snow to the used single layer snow scheme (Douville et al., 1995). REMO is using a multi-layer snow scheme and does not encounter this problem.

A similar strong warm bias in the north, as found for ALARO-0 in winter, was also found by Ozturk et al. (2012) and Russo et al. (2019) for the RegCM and COSMO-CLM 5.0 models, respectively. Ozturk et al. (2012) related this warm bias to shortcomings in the simulation of snow, whereas Russo et al. (2019) found that changes in the snow scheme did not affect the simulation results significantly and did not reduce the warm bias in the north-east during winter. This shows that a complexer multi-layer snow scheme might not be enough to solve the warm bias for ALARO-0 during winter. Therefore, further investigation should be done to see whether the warm bias in winter over the northern part of the domain is due to the inability of the current snow scheme to reproduce the heat conductivity of snow.

In spring, the warm temperature bias of the ALARO-0 simulation over the northern subdomain evolves into a significant cold bias. This remarkable evolution is probably related to another issue related to the snow scheme as we find a delay in the springtime melting of the snowpack (not shown). Additionally, ALARO-0 simulates too high pressure values over the northern area (not shown). Further research is needed to clarify whether this overestimation of the Siberian High in the ALARO-0 simulations is coupled to the difficulties with the snow cover.

4.2 Diurnal temperature range

Spatially averaged biases are warmer for the minimum temperature and colder for the maximum temperature, when compared to those of the mean temperature (Tables 2, 3 and 4). 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 and colder for the maximum temperature than it was the case for the mean temperature. This causes a stronger warm bias in winter for the minimum temperature and a stronger cold bias for maximum temperature in spring, which is especially visible in the northern part of the domain for the ALARO-0 model (Fig. 2, 5 and 7). Moreover, the cold bias in the north during spring for the ALARO-0 model is weaker for the minimum temperature than for the mean temperature. The REMO model shows warmer biases over Mongolia during winter and spring for minimum temperature and colder biases in maximum temperature in the north during spring when compared to the mean temperature.

Although the magnitude of the biases is different for mean, minimum and maximum temperature, similar spatial patterns are found in the biases of both RCMs over the different seasons and for the annual mean (Fig. 2, 5 and 7). This means that these variables are spatially highly correlated with each other in both models and observations. When comparing the metrics in Fig. 6 and 8 to those of the mean temperature (Fig. 3), then it is seen that the metrics of mean, minimum and maximum temperature are similar for both RCMs during the different seasons. However, both RCMs overestimate the normalized RSV of the maximum temperature for all seasons (Fig. 8), which differs from the mean temperature where ALARO-0 and REMO underestimated the normalized RSV during winter (Fig. 3). This indicates that there is a slightly larger spatial variation in winter maximum temperatures simulated by the RCMs with respect to CRU, while for mean temperatures a smaller spatial variation is simulated. Additionally, both minimum and maximum temperatures have a similar temporal pattern as for the mean temperature, e.g. the smallest range in mean, minimum and maximum temperature biases is obtained in autumn for both RCMs (Fig. 4). Moreover, the underestimation of the minimum and maximum temperatures in spring is more pronounced for ALARO-0 than for REMO (Fig. 4).

The RCMs underestimate the diurnal range, which is similar to the findings over other regions (Laprise et al., 2003; Kyselý and Plavcová 2012). The underestimation of the diurnal range over the CAS-CORDEX domain was also observed by Russo et al. (2019) for the winter and summer seasons. Their RCM produced smaller diurnal ranges compared to different observational datasets and the comparison between the observational datasets pointed out that the observational uncertainty is high for the diurnal range in the north-eastern part of the domain, which makes it difficult to evaluate the diurnal range accurately over this area. In particular ALARO-0 shows a very small range in the diurnal cycle of temperatures due to very high minimum temperatures (Fig. 4) and this could be due to the inability of the model to simulate temperatures correctly over snow cover during stable conditions (Mašek, 2017).

The evaluation of temperature and its diurnal cycle shows that a bias adjustment is essential before the climate data is applied in impact modelling for some regions, e.g. Tibetan Plateau and East Siberia. The current research is done within the AFTER project (Kotova et al, 2018). Within this project different bias-adjustment techniques are applied to the set of climate simulations. This will enable impact modellers to optimally use our climate data in their models for crop production, biomass production, etc.

4.3 Precipitation

The precipitation of ALARO-0 and REMO is for the majority of the grid points situated within the spread of the different gridded datasets during the different seasons (Fig. 9). However, there are some subregions where the precipitation of ALARO-0 and/or REMO exceeds the observational spread for a specific season. For example, both RCMs show slightly lower precipitation amounts in summer over West Central Asia compared to the different reference datasets (Fig. 11). Ozturk et al. (2012) 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. They also obtained a dry bias in summer over the north-western and south-western part of the domain. Additionally, 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 in mountainous areas except for some parts of the Tibetan Plateau (Ozturk et al., 2012; Russo et al., 2019).

ALARO-0 and REMO produce smaller spatially averaged precipitation biases over the CAS-CORDEX region at a horizontal resolution of 0.22° than the RegCM4.3.5 model at a resolution of 0.50°, except during summer (Ozturk et al., 2016). 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 (Fig. 10). ALARO-0 obtains higher values for the spatial correlations and they are close to those of the other observational datasets.

The overestimation of precipitation by the RCMs over the Himalaya, Altay, Tian Shan and Kunlun Mountains on annual level is partly due to the fact that gridded observational datasets CRU, MW and GPCC 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). This explains as well why the gridded observational datasets show a drier environment than the ERA-Interim reanalysis dataset in the eastern part of the domain (East Siberia and Tibetan Plateau), particularly during spring (Fig. 11) (Hu et al., 2018). Moreover, this pronounced difference during spring between the observational gridded datasets on the one hand and the RCMs and ERA-Interim reanalysis data on the other hand explains why the scores with respect to CRU are worse during spring (Fig. 10). This difference between the observational and reanalysis datasets makes it difficult to draw sound conclusions over the south-eastern part of the domain during spring, when the monsoon takes place.

It is known that CRU data shows higher precipitation rates at most of the grid points in eastern Russia due to poor station coverage (New et al., 1999). 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. 10). When averaging over the complete domain, then the output of both ALARO-0 and REMO is within the range of the spread between the reference datasets for the different seasons (Table 5).

Table 5 and Fig. 11 show that CRU contains higher precipitation amounts compared to the two other observational datasets, MW and GPCC. This explains the systematic dry bias that was found for ALARO-0 during all seasons when compared to CRU (Table 5). The underestimation in precipitation by ALARO-0 during spring in the north-eastern part of the domain might

be related to the Siberian High that remains too strong during spring (not shown). REMO simulated wetter circumstances with respect to all reference datasets over East Siberia during spring and over the Tibetan Plateau during all seasons except for summer (Fig. 11). The wet bias over East Siberia during spring is in absolute values very low when compared to the subdomain Tibetan Plateau (Fig. 11 and S2). Russo et al. (2019) found a similar spatial pattern of a wet bias during winter (autumn and spring were not discussed) over the south-eastern region with their COSMO-CLM model as presented here for REMO (Fig. 9).

The precipitation amounts of REMO tend in the north more towards those of ERA-Interim (Fig. 11). The similarities between ERA-Interim and REMO for precipitation are probably due to the fact that both use a modified convection scheme that is based on Tiedtke (1989) (Table S1; www.ecmwf.int, consulted on 07/07/2020), while ALARO-0 uses the 3MT cloud microphysics scheme and shows a different behaviour. For example, the weak wet bias which was observed in the north-eastern 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 MW and GPCC data (Fig. 11). Additionally, REMO is not able to reproduce the annual cycle of precipitation over the Asian monsoon region. Remedio et al. (2019) found as well a shift in precipitation for REMO over the subtropical region where the Asian monsoon takes place with wetter winter and spring seasons and a drier summer season.

It can be concluded that REMO and ALARO-0 simulated precipitation for the different subregions and seasons mostly within the range of the observational spread, although it should be mentioned that the observational uncertainty is large. MW, GPCC 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 datasets on the RSV is larger for precipitation than for temperature. Moreover, both models are worse in simulating the spatial correlation of precipitation (Fig. 10) compared to the mean, minimum and maximum temperature (Fig. 3, 6 and 8). The lower accuracy of simulated precipitation is due to the fact that precipitation is less systematically affected by land cover and topography compared to temperature (Kotlarski et al., 2014). Furthermore, the uncertainty range and error in the observational products should be restricted in the future to improve the evaluation of precipitation (Russo et al., 2019).

5 Conclusion

The evaluation over the CAS-CORDEX domain of ALARO-0 and REMO, run at 0.22° resolution, showed that both RCMs reproduced in general realistic spatial patterns for temperature and precipitation. Both RCMs perform best during autumn, showing biases within the range of observational uncertainty for temperature and precipitation. Nevertheless, there are significant biases in several regions during several seasons e.g. a warm bias in the north during winter and a wet bias in spring over the Asian monsoon region. For ALARO-0 the northern part of the CAS-CORDEX domain is subject to significant positive temperature biases in winter, followed by large negative temperature biases in spring. This behaviour is probably linked to limitations of the used snow scheme. The evaluation of minimum and maximum temperatures showed that the RCMs

underestimate the daily temperature range. This illustrates the added value of taking more evaluation variables into account than only the commonly evaluated variables mean temperature and precipitation.

The values of spatial variation and pattern correlation for mean temperature of both RCMs correspond closely to the values obtained with other reference datasets. These metrics indicated a less good performance for precipitation data 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.

We conclude that REMO and ALARO-0 can be used to perform climate projections over Central Asia since they perform similarly to experiments with other models over the same domain. REMO is better than ALARO-0 in reproducing the seasonal mean temperatures over the entire domain except during autumn, while ALARO-0 is very good in estimating the precipitation. However, deficiencies described in this evaluation study should be kept in mind. Climate data produced by both RCMs can be used for impact studies if a suitable bias adjustment is applied for those subregions where the RCMs perform less well e.g. East Siberia and Tibetan Plateau.

Code availability

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: 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: 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., 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 - 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 EWI. The CORDEX-CORE REMO simulations were performed under the GERICS/HZG share at the German Climate Computing Centre (DKRZ).

We would like to thank Ján Mašek for his insights on the warm bias above snow cover.

Competing interests

The authors declare that they have no conflict of interest.

References

Akperov, M., Rinke, A., Mokhov, I.I., Matthes, H., Semenov, V.A., Adakudlu, M., Cassano, J., Christensen, J.H., Dembitskaya, M.A., Dethloff, K. and Fettweis, X.: Cyclone activity in the Arctic from an ensemble of regional climate models (Arctic CORDEX). Journal of Geophysical Research: Atmospheres, 123, 2537-2554, https://doi.org/10.1002/2017JD027703, 2018.

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.

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.

Collins, M., AchutaRao, K., Ashok, K., Bhandari, S., Mitra, A.K., Prakash, S., Srivastava, R. and Turner, A.: Observational challenges in evaluating climate models, Nature Climate Change, 3, 940–941. https://doi.org/10.1038/nclimate2012, 2013.

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 2019).

Cabos, W., Sein, D.V., Durán-Quesada, A., Liguori, G., Koldunov, N.V., Martínez-López, B., Alvarez, F., Sieck, K., Limareva, N. and Pinto, J.G.: Dynamical downscaling of historical climate over CORDEX Central America domain with a regionally coupled atmosphere–ocean model. Climate dynamics, 52, 4305-4328, https://doi.org/10.1007/s00382-018-4381-2, 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, 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 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 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.

Douville, H., Royer, J-F. and Mahfouf., J-F.: A new snow parameterization for the Meteo-France climate model, Climate Dynamics, 12, 21-35, 1995.

ECMWF: Atmospheric physics, https://www.ecmwf.int/en/research/modelling-and-prediction/atmospheric-physics, (accessed on 7 July 2020).

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-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.

Gibson, P.B., Waliser, D.E., Lee, H., Tian, B. and Massoud, E.: Climate model evaluation in the presence of observational uncertainty: precipitation indices over the Contiguous United States. Journal of Hydrometeorology, 20,1339-1357, 2019.

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 Environment and Resources, 40, 467–490, 2015.

Giorgi, 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.: 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.

Gómez-Navarro, J., Montávez, J., Jerez, S., Jiménez-Guerrero, P., and Zorita, E.: What is the role of the observational dataset in the evaluation and scoring of climate models?, Geophys. Res. Lett., 39, L24701, https://doi.org/10.1029/2012GL054206, 2012.

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.

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.

Haarsma, R. J., Roberts, M. J., Vidale, P. L., Senior, C. A., Bellucci, A., Bao, Q., Chang, P., Corti, S., Fučkar, N. S., Guemas, V., von Hardenberg, J., Hazeleger, W., Kodama, C., Koenigk, T., Leung, L. R., Lu, J., Luo, J.-J., Mao, J., Mizielinski, M. S., Mizuta, R., Nobre, P., Satoh, M., Scoccimarro, E., Semmler, T., Small, J., and von Storch, J.-S.: High Resolution Model Intercomparison Project (HighResMIP v1.0) for CMIP6, Geosci. Model Dev., 9, 4185-4208, doi:10.5194/gmd-9-4185-2016, 2016.

Hagemann, S.: An improved land surface parameter data set for global and regional climate models, Max Planck Institute for 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., Osborn, T.J., Jones, P.D. and Lister, D.H.: Version 4 of the CRU TS monthly high-resolution gridded multivariate climate dataset, Scientific Data, 7 (109), doi:10.1038/s41597-020-0453-3, 2020.

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.

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.

Iturbide, M., Gutiérrez, J. M., Alves, L. M., Bedia, J., Cimadevilla, E., Cofiño, A. S., Cerezo-Mota, R., Di Luca, A., Faria, S. H., Gorodetskaya, I., Hauser, M., Herrera, S., Hewitt, H. T., Hennessy, K. J., Jones, R. G., Krakovska, S., Manzanas, R., Marínez-Castro, D., Narisma, G. T., Nurhati, I. S., Pinto, I., Seneviratne, S. I., van den Hurk, B., and Vera, C. S.: An update of IPCC climate reference regions for subcontinental analysis of climate model data: Definition and aggregated datasets, Earth Syst. Sci. Data Discuss., https://doi.org/10.5194/essd-2019-258, in review, 2020.

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.

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., 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., 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.

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 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.

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 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.

Mašek, J.: Problem with screen level temperatures above snow in ISBA scheme, report RC LACE, 2017.

Morcrette, J.-J., Smith, L. and Fouquart, Y.: Pressure and temperature dependence of the absorption in longwave radiation 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.

Nikulin, G., Jones, C., Giorgi, F., Asrar, G., Büchner, M., Cerezo-Mota, R., Christensenf, O. B., Déqué, 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.

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., Türkeș, 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.

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.

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., 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 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.

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., 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-14-00176.1, 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 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, 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., 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, 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.

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, 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.

Tiedtke, M. (1989). A comprehensive mass flux scheme for cumulus parameterization in large-scale models. Mon. Wea. Rev., 117, 1779-1800.

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 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 temperature and rainfall in Vietnam, Vietnam Journal of Earth Sciences, 41, 374–387, doi:10.15625/0866-7187/41/4/14259, 2019.

Wang, Y., Feng, J., Luo, M., Wang, J. and Yuan, Q.: Uncertainties in simulating Central Asia: sensitivity to physical parameterizations using WRF, International Journal of Climatology, doi:10.1002/joc.6567, 2020.

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

Zhu, X., Wei, Z., Dong, W., Ji, Z., Wen, X., Zheng, Z., Yan, D. and Chen, D.: Dynamical downscaling simulation and projection for mean and extreme temperature and precipitation over central Asia, Climate Dynamics, 54, 3279-3306, doi:10.1007/s00382-020-05170-0, 2020.

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: 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 - 2018	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	0.50° or 0.25°	precipitation	monthly	1891 - 2016	global land mass (excluding Antarctica)
ERA-Interim	ERA-Interim	reanalysis data	0.70°	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 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

895

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

900

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

905