# Climate Model Downscaling in Central Asia: A Dynamical and a Neural Network Approach

Bijan Fallah[1,2], Masoud Rostami[2,6], Emmanuele Russo[3], Paula Harder[4], Christoph Menz[2], Peter Hoffmann[2], Iulii Didovets[2], and Fred F. Hattermann[2,5]

[1]German Climate Computing Center (DKRZ), Hamburg, Germany.

[2]Potsdam Institute for Climate Impact Research (PIK), P.O. Box 601203, 14412 Potsdam, Germany

[3]ETH Zürich, Department of Environmental Systems Science, Universitätstrasse 16, 8092 Zürich, Switzerland

[4]Mila Quebec AI Institute, Montreal, Canada

[5]Eberswalde University for Sustainable Development (HNEE), Germany

[6]Laboratoire de Météorologie Dynamique (LMD), Sorbonne University (SU), Ecole Normale Supérieure (ENS), Paris, France

**Correspondence:** Bijan Fallah (fallah@pik-potsdam.de, fallah@dkrz.de)

**Abstract.** High-resolution climate projections are essential for estimating future climate change impacts. Statistical and dynamical downscaling methods, or a hybrid of both, are commonly employed to generate input datasets for impact modelling. In this study, we employ COSMO-CLM (CCLM) version 6.0, a regional climate model, to explore the benefits of dynamically downscaling a general circulation model (GCM) from CMIP6, focusing on climate change projections for Central Asia (CA). The CCLM, at 0.22° horizontal resolution, is driven by the MPI-ESM1-2-HR GCM (at 1° spatial resolution) for the historical period of 1985-2014 and the projection period of 2019-2100, under three shared socioeconomic pathways (SSPs): SSP1-2.6, SSP3-7.0, and SSP5-8.5 scenarios. Using the Climate Hazards Group InfraRed Precipitation with Station data (CHIRPS) gridded observation dataset as a reference, we evaluate the performance of CCLM driven by ERA-Interim reanalysis over the historical period. The added value of CCLM, compared to its driving GCM, is significant (more than 5 mm/day) over mountainous areas in CA, which are at higher risk of extreme precipitation events. Additionally, we employ CCLM to refine future climate projections. We present high-resolution maps of heavy precipitation changes based on CCLM and compare them with the CMIP6 GCM ensemble. Our analysis indicates a significant increase in the intensity and frequency of heavy precipitation events over CA areas already at risk of extreme climatic events by the end of the century. Finally, we train a convolutional neural network (CNN) to map a GCM simulation to its dynamically downscaled CCLM counterpart. The CNN successfully emulates the GCM-CCLM model chain over large CA areas, demonstrating added value when applied to a new GCM-CCLM

model chain. The scientific community interested in downscaling CMIP6 models could use our downscaling data, and the CNN architecture offers an alternative to traditional dynamical and statistical methods.

## 1 Introduction

The increasing global mean temperature due to anthropogenic greenhouse gas emissions presents a significant challenge for society, requiring the assessment and prediction of future impacts on human health, natural ecosystems, and economies across different regions of the world (Allan et al., 2021). Researchers conducting regional studies on vulnerability, impacts, and adaptation typically achieve reliable high-resolution climate projections through dynamical downscaling via RCMs (Rummukainen, 2010; Feser et al., 2011), statistical techniques (Maraun and Widmann, 2018; Fowler et al., 2007), or a hybrid of both approaches (Maraun et al., 2015; Meredith et al., 2018; Laflamme et al., 2016).

CA, recognised as one of the most vulnerable regions to climate change impacts, heavily depends on water resources from glaciers and rivers that are shrinking due to rising temperatures and decreasing precipitation (Reyer et al., 2017; Fallah et al., 2023; Didovets et al., 2024; Fallah and Rostami, 2024). The area faces significant challenges to food security, characterised by declining crop yields and an increased occurrence of severe and frequent extreme weather events like floods and landslides. These conditions damage infrastructure, livelihoods, and agriculture, resulting in population displacement and migration (Allan et al., 2021; Reyer et al., 2017).

Significant uncertainties inherent in the existing detailed observational and reanalysis datasets impede the development of high-resolution climate projections in CA (Fallah et al., 2016a). One option to complement these datasets is to use dynamical downscaling with RCMs. CMIP6 provides a framework for coordinated climate model experiments, enhancing our understanding of past, present, and future climate changes. Dynamical downscaling of CMIP6 models for the CA region is vital for accurately simulating extreme convective precipitation events, which are influenced by the orography of the region (Lundquist et al., 2019; Ban et al., 2015; Wang et al., 2013; Frei et al., 2003; Russo et al., 2019), large-scale atmospheric circulation, and sea surface temperature anomalies in the Indian Ocean and the Pacific (Kendon et al., 2014; Demory et al., 2020; Xu et al., 2022). This method enhances the resolution of a driving GCM and produces a physically consistent regional state of the climate. Despite some systematic biases, dynamical downscaling consistently provides high-quality datasets that accurately describe the climatology of all climate variables in CA (Qiu et al., 2022).

Various international institutions have collaborated within the Coordinated Regional Climate Downscaling Experiment (CORDEX) to address these issues and improve the inter-comparability of RCMs. CORDEX aims to create a robust framework for producing climate projections at a regional scale that is suitable for impact evaluation and adaptation planning globally. This effort aligns with the timeline of the Intergovernmental Panel on Climate Change's Sixth Assessment Report (Kikstra et al., 2022). However, most CORDEX research focuses on highly industrialised countries (Allan et al., 2021; Taylor et al., 2012). Developing regions, including CA, bear the brunt of global warming's consequences, yet they have access to only a limited number of CORDEX model simulations (Naddaf, 2022). As of the latest update, no simulation driven by CMIP6 has been planned for CORDEX-CA (see https://wcrp-cordex.github.io/simulation-status/CMIP6_downscaling_plans.html, last visited on 17.04.2024).

Beyond dynamical methods, recent developments in machine learning, including CNNs, offer promising avenues for statistical downscaling (Harder et al., 2023; Rampal et al., 2024). CNNs have proven effective in numerous earth science disciplines besides downscaling, such as classification (Gardoll and Boucher, 2022), segmentation (Galea et al., 2024), and prediction (Watson-Parris et al., 2022) thanks to their capacity to extract features from spatial data and identify nonlinear relationships between inputs and outputs. CNNs can recognise and encode spatial hierarchies in data (Zhu et al., 2017), making them exceptionally suitable for analysing geospatial data, a critical component in climate modelling. Unlike traditional statistical methods that often require manual selection and careful engineering of features, CNNs automatically learn the most predictive features directly from the data (Reichstein et al., 2019). They are generally more straightforward and efficient than traditional statistical downscaling methods for tasks aiming to predict or classify patterns distributed across spatial domains, such as temperature or precipitation patterns in climate models (Racah et al., 2017). CNNs are adept at maintaining spatial coherence in the output, which is critical in downscaling where preserving the geographical patterns of climate variables (like precipitation) is crucial (Kurth et al., 2018).

Researchers classify CNNs into two categories based on their last layer: 1) constrained and 2) unconstrained. Constrained CNNs integrate physical laws directly into the training process, such as mass, energy, or momentum conservation. This integration is achieved by modifying the loss function or the network's architecture to ensure compliance with these laws. In contrast, unconstrained CNNs do not explicitly incorporate physical laws or constraints. Instead, they rely solely on learning from the input data, generating output predictions based on the patterns detected in the data.

This study explores unconstrained and constrained CNN approaches to understand their effectiveness in downscaling and their performance when applied to GCMs not initially used for training.

The research questions guiding this study are:

- **Research Question 1:** How effectively can CMIP6 models be downscaled to enhance precipitation simulations for the CORDEX Central Asia region?
- **Research Question 2:** Can CNNs effectively downscale GCM outputs, and how do they perform when applied to GCMs that did not initially train them?

This article focuses on two main topics: 1) the added value of CCLM for representing precipitation over Central Asia, and 2) training a CCLM emulator using a CNN. We present data and methods in Section 2. Sections 3 and 4 introduce the results of dynamical and hybrid downscaling, respectively. Finally, we discuss the results and conclude in Section 5.

## 2 Data and Methods

The methodology employed in this study is illustrated in Figure 1. The following sections provide a detailed explanation of this methodology.

## 2.1 Employed Models and Experimental Setups

### 2.1.1 Regional Climate Model (RCM)

In this study, we conduct simulations using the CCLM regional climate model. Developed by the German Weather Service (DWD) and the German Climate Computing Center (Deutsches Klimarechenzentrum, DKRZ), CCLM originates from the COSMO numerical weather prediction model (Rockel and Geyer, 2008), which is widely utilised for short-term weather forecasting. Explicitly designed for regional climate simulation, CCLM enables researchers to investigate various aspects of the climate system, including temperature, precipitation, and extreme events. It has been extensively used to assess the impact of climate change across different regions such as Europe (Russo et al., 2021), Africa (Panitz et al., 2014; Dosio and Panitz, 2016), and Asia (Jacob et al., 2014; Kotlarski et al., 2014; Wang et al., 2013). Additionally, CCLM has been employed in climate projection studies to evaluate climate adaptation and mitigation strategies. The model has undergone thorough evaluation and validation (Fallah et al., 2016b; Russo et al., 2019; Kjellström et al., 2011), and its ability to generate realistic simulations of present climate conditions and variability has established it as one of the most widely used regional climate models in the scientific community (Sørland et al., 2021).

For our experiments, we utilised a model setup similar to the "optimal" configuration described by Russo et al. (2019). In their study, Russo et al. (2019) optimised the CCLM regional climate model for CA by adjusting albedo based on forest fraction ratios and soil conductivity to account for the soil's liquid water and ice proportions. These modifications significantly improved the model's climatological performance and the distribution of incoming radiation, leading to more accurate climate representations for the region. According to the CORDEX protocol, simulations are divided into two primary phases. The first phase, the evaluation run, involves a single model experiment over the period 1979-2014, using ERAInterim reanalysis data at a spatial resolution of T255 ($\sim 0.7°$). The second phase, the projection run, utilises boundary conditions from GCMs of the CMIP6 project for the period 1950-2100 under various SSPs. For this study, we selected the MPI-ESM1-2-HR GCM and considered SSP1-2.6, SSP3-7.0, and SSP5-8.5 scenarios. SSPs represent baseline scenarios that describe future pathways based on population growth, technological advancement, economic development, urbanisation, and investments in healthcare, education, land use, and energy (Riahi et al., 2017).

Historical data for this study are based on greenhouse gas levels, land use, and other climate forcings observed from 1850 to 2014. The Shared Socioeconomic Pathway (SSP) scenarios used in the projections are as follows:

- **SSP1-2.6** represents a "green" future, characterised by global efforts to protect resources, improve human well-being, and narrow income gaps. This scenario assumes low challenges to adaptation and low greenhouse gas emissions. Adaptation challenges in this context refer to the difficulties societies might face in adjusting to the impacts of climate change, including their susceptibility and the availability and effectiveness of mitigation technologies and strategies. Under SSP1-2.6, global cooperation and sustainable practices lead to advancements in technology and governance, significantly reducing vulnerability to climate change impacts. Societal structures are resilient, and resources are managed to minimise environmental stresses while maximising human well-being.

– **SSP3-7.0** depicts a future characterised by regional rivalry, where nationalism and regional conflicts dominate, global issues are neglected, and inequality increases. This scenario involves high challenges to adaptation and high greenhouse gas emissions.

– **SSP5-8.5** represents a future of fossil-fueled development with globally connected markets, rapid technological progress, and weak environmental policies. This scenario has low challenges to adaptation but results in very high greenhouse gas emissions.

For comparison and evaluation of our RCM simulations, we have selected two CORDEX-CA evaluation simulations from other models driven by ERAInterim at a 0.22° horizontal resolution: 1) **ERAInterim-RMIB-UGent-ALARO-0** (Giot et al., 2016) and 2) **ERAInterim-GERICS-REMO2015** (Jacob and Podzun, 1997; Fotso-Nguemo et al., 2017).

#### 2.1.2 CNNs

In this study, we develop a CNN-based emulator for the CCLM driven by the MPI-ESM1-2-HR GCM. This CNN utilises outputs from the GCM, covering both the historical period from 1985 to 2014 and future scenarios spanning 2019 to 2100, as inputs to model the responses of the CCLM, which serves as the target. Given the low annual precipitation and significant spatio-temporal variability in many regions of CA, a comprehensive dataset that includes various precipitation patterns from both GCMs and RCMs is essential for effectively training the CNN to map from GCM to RCM outputs. To enhance model training, we have augmented our dataset with ERA-Interim reanalysis data and corresponding CCLM simulations driven by it (ERAInterim-CCLM) (see Fig. 1).

We train our CNN model based on the architecture proposed by Harder et al. (2023), which incorporates physical constraints to ensure mass conservation and energy balance. The model architecture features:

– Conv (Convolutional Layer): These layers help extract various levels of features from low-resolution images, such as edges, textures, and other relevant image details.

– ReLU (Rectified Linear Activation Unit): This nonlinear activation function introduces non-linearity and returns the input unchanged if it is positive; otherwise, it returns zero. This function enables the network to learn complex patterns efficiently.

– TransConv (Transposed Convolutional Layer): This layer is crucial for downscaling. It increases the spatial dimensions of the feature maps, performing a sort of learned interpolation. This allows the model to add details to the downscaled images based on the features extracted and processed in the earlier layers.

– ResBlock (Residual Block): These blocks allow the model to refine the initial lower-resolution predictions, which are downscaled (interpolated outputs) to a higher resolution. They enhance the model's ability to add fine details and textures (high-frequency information), improving the perceptual quality and sharpness of the images at the increased resolution."

In the context of deep learning for climate modelling, the "perfect model" approach involves starting with high-resolution data and intentionally upscaling it to a lower resolution. The machine learning model is subsequently trained to reproduce the high-resolution data while receiving this artificial low-resolution input. The aim is to simulate a scenario where the "truth" (the original high-resolution data) is known and then to recover this high-resolution from the artificially upscaled data. This approach teaches the model the desired mapping from low to high resolution, enabling the model to effectively learn how to downscale or enhance resolution while minimising the loss of critical information. It is a controlled experiment that helps refine the model's capabilities.

The "imperfect model" approach, on the other hand, acknowledges that both the low-resolution (GCM output) and the high-resolution (RCM output) datasets have their inherent errors and limitations. In this scenario, we do not have a single source of truth but rather two separate sets of data:

- Low-resolution data: may capture global or large-scale phenomena but miss regional details (Xu et al., 2021; Chokkavarapu and Mandla, 2019).
- High-resolution data: provides detailed regional information but may still have errors or not perfectly reflect reality due to limitations in data collection, model configuration, or computational constraints (Muttaqien et al., 2021).

In this setup, CNN's challenge is learning to map between two independently imperfect datasets. The CNN is trained to predict high-resolution details from low-resolution inputs as accurately as possible despite the absence of perfect ground truth. This process involves understanding and modelling the uncertainties and biases inherent in both datasets.

Prior to training, the dataset was randomly shuffled at the pair level to ensure that each GCM input and its corresponding RCM output remained together, preserving the intrinsic relationships between the coarse and fine-resolution data. This approach prevents temporal or spatial autocorrelation from biasing the training process. It also improves the model's generalisation and performance by exposing it to various conditions. For the dataset distribution, 68,141 days (60%) of RCM simulation data were used for training, 22,714 days (20%) for validation, and 22,714 days (20%) for testing. The low-resolution (GCM) dataset consists of $30 \times 60$ grid points, and the high-resolution (RCM) dataset comprises $120 \times 240$ grid points over latitudes and longitudes, respectively, resulting in a downscaling factor (N) of 4.

#### 2.1.3 Constraint layers

We test the CNN with three different constraining methods in the last CNN layer (Harder et al., 2023): 1) soft constraining (SCL), 2) hard constraining (HCL) and 3) without constraining (NoCL). In the following, we explain briefly the three different constraining methodologies. The setup of constraining is as follows: consider a factor N for downscaling in all linear directions and let $n := N^2$ and $y_i$, $i = 1,...,n$ be the high-resolution patch values that correspond to low-resolution pixel $x$. The mass conservation law has the following form:

$$\frac{1}{n}\sum_{i=1}^{n} y_i = x. \tag{1}$$

**Hard constraining:** uses the SoftMax, which constrains quantities like water content by enforcing the output to be non-negative. The simplest way to ensure mass conservation would be to scale all small-scale values within a given large-scale grid cell with the ratio of the large-scale value and the sum of the small-scale values. However, Harder et al. (2022) demonstrated that employing the SoftMax constraints layer gives better results. The exponential ensures positive predictions and leads to more variance between subpixels in the super-resolved prediction. The multiplicative rescaling struggles when the sum of the small-scale values gets close to zero. Therefore, the SoftMax operator is used on the intermediate outputs of the CNN before the constraining layer ($\tilde{y}_i$) and multiplies it by the corresponding input pixel value $x$:

$$y_i = \exp(\tilde{y}_j) \cdot \frac{x}{\frac{1}{n}\sum_{i=1}^{n}\exp(\tilde{y}_i)}. \tag{2}$$

$y_i$ is the final output after applying the constraints. We have used the mean absolute error (MAE) as the loss function (Eq. 5).

**Soft constraining:** This is done by adding a regularisation term to the loss function. The MAE loss is then extended with an additional constraint violation (CV) loss term:

$$\text{Loss} = (1-\alpha) \cdot MAE + \alpha \cdot \text{CV}, \tag{3}$$

Where CV is the mean-squared error over all constraint violations between an input pixel x and the super-pixel (high-resolution grid-cell) $y_i$:

$$\text{CV} = \text{MSE}(\frac{1}{n}\sum_{i=1}^{n} y_i, x) \tag{4}$$

We use $\alpha = 0.99$.

**Without constraining**: In this setup, we remove the constraining layer after the last convolutional layer in the CNN.

We use 160 epochs, using a batch size of 64 and a learning rate of 0.001 for training with HCL and NoCL and 0.00001 for SCL. Training takes 15 hours on an NVIDIA Corporation Graphics Ampere 104 [GeForce Ray Tracing Texel eXtreme (RTX) 3060 Ti-Lite Hash Rate] graphics processing unit (GPU). We use the same model setup as in Harder et al. (2023).

It is important to note that the MAE can serve both as a loss function and an evaluation metric. As a loss function, it is used during training to optimise the neural network's parameters. Conversely, when used as an evaluation metric, it is calculated on the validation or test data sets to assess the model's performance using an independent dataset. Despite their different applications, MAE is suitable for both roles.

## 2.2 Evaluation and testing

According to Ciarlo et al. (2021), the choice of observational data can significantly influence the perceived added value of an RCM, particularly in detecting extreme events, where poor-quality data might misleadingly suggest improved model performance. They recommend using observations with spatiotemporal resolutions comparable to the model's for enhanced accuracy. In line with this, we use CHIRPS as our gridded observation to assess the added value of the CCLM driven by the GCM. CHIRPS provides a resolution of 0.05°, covers latitudes from 50°S to 50°N, and offers independent observations derived from satellite information and station data. This contrasts with reanalysis data, which depend on climate model simulations (Funk et al., 2015). We allocate 20$\%$ of the CCLM simulation data as the target to evaluate our CNN emulator instead of using CHIRPS directly. We measure the added value of the CNN by comparing the MAE of both the CNN outputs and the interpolated GCM outputs against the target CCLM output. This comparison assesses whether the CNN outperforms simple interpolation. The selected GCM and observational data are interpolated onto the RCM grid using the distance-weighted average method. Ciarlo et al. (2021) previously noted that such interpolation might create unrealistic values, as it does not account for the physical processes and could introduce artefacts depending on the interpolation method, the spatial distribution of data points, and the resolution ratio. Therefore, we use simple interpolation as a baseline, recognising its limitations in preserving the statistical properties of precipitation, which does not follow a normal distribution. Following (Hodson, 2022), we apply the MAE to quantify the biases in emulated and dynamically downscaled precipitation ($F$) against observations ($O$):

$$\mathrm{MAE} = \frac{1}{T}\sum_{t=1}^{T}|F_t - O_t| \tag{5}$$

Where $T$ represents the number of time steps over 30 years of daily data. We define added value (AV) as the reduction in MAE achieved by the downscaling relative to the driving GCM:

$$\mathrm{AV} = \mathrm{MAE}_{GCM} - \mathrm{MAE}_{CCLM} \tag{6}$$

Where $\mathrm{MAE}_{GCM}$ and $\mathrm{MAE}_{CCLM}$ are the differences between interpolated GCM and RCM with respect to the reference dataset.

As an additional metric, we also use the climatological bias, i.e., the difference between the model and observations:

$$\mathrm{BIAS} = \mathrm{PR}_{MODEL} - \mathrm{PR}_{OBS} \tag{7}$$

## 3 Results

Figure 3.a illustrates the topography of the CORDEX-CA simulation domain. Figure 1.b displays the mean daily precipitation, averaged over the years 1985-2014 (mm/day), derived from CHIRPS data. The regions with the highest precipitation are the mountainous areas of CA, particularly notable in the Asian summer monsoon region north of India and along the Himalayas in the southeastern part of the domain, where precipitation values are pronounced. Figure 3.c depicts the distribution of WorldClim weather stations (Fick and Hijmans, 2017) across CA, serving as a proxy for the density of station data used in the CHIRPS dataset. Observational data are sparsely distributed in East China, especially over the Tibetan Plateau. Consequently, data-model comparisons are considered unreliable in this region (Randall et al., 2007; Cui et al., 2021; Yan et al., 2020; Russo et al., 2019).

### 3.1 Added value of CCLM driven by ERAInterim

To characterise the overall performance of the CCLM model across time and space, Figures 4 and 5 present maps displaying annual, winter (DJF), and summer (JJA) MAE and mean biases. These biases in precipitation are calculated between the interpolated ERAInterim data and CCLM outputs driven by ERAInterim for the period 1985-2014, in comparison to CHIRPS (see Eq. 5 and Eq. 6). Figures 4.a-c illustrate the MAE for ERAInterim for annual, winter, and summer averages. The added value of the CCLM RCM compared to the interpolated ERAInterim is depicted in Figures 4.d-f. During the Asian summer monsoon, CCLM's MAE is high over the south and southeast of the domain (regions in magenta), whereas it is generally lower during winter. CCLM shows an MAE reduction in the mountainous areas of Afghanistan, Kyrgyzstan, and Tajikistan, as well as an increase near the domain's southern boundaries throughout the year and in the south and southeast during summer.

The AVs of GERICS-REMO2015 and RMIB-UGent-ALARO-0 driven by ERAInterim are presented in Figures 4.g-l, using CHIRPS as the observational dataset. The added value of RCM is most pronounced in areas with complex topography, especially during summer, across all three RCMs (Figs.4.d-l). Areas where the RCM has a smaller MAE than the reanalysis in comparison to observations are found over Tajikistan, Kyrgyzstan, northern Afghanistan, and part of the Himalayas—regions that are crucial water sources for former Soviet Union countries. Nevertheless, precipitation during the colder seasons may be more critical for water availability. The annual AV patterns still show positive values in these regions (Figure 4.d,g, and j). Across the entire domain, all three RCMs significantly reduce the large and local-scale bias of ERAInterim, especially in complex topographies. The nested RCMs exhibit similar MAE values near their lateral boundaries, relative to their driving model (Figure 4, *a*,*b*,*c*). Thus, negative AV quantities may result from boundary effects, particularly near the eastern and southeastern boundaries where monsoonal precipitation dominates. GERICS-REMO2015 displays pronounced negative added values annually and during winter above Tibet.

Additionally, model climatology biases are displayed in Figures 5. Once again, these biases are noticeable in the lower right corner of the domain during JJA and across the southern Tibetan Plateau throughout the year.

#### 3.1.1 Extreme precipitation patterns in CCLM and CMIP6 GCMs

Given that the CCLM simulation has demonstrated added value for precipitation over the mountainous regions of CA, we explore climate change signals in its high-resolution output. These high-resolution maps may inherit biases from the GCM-RCM selection and could vary under different anthropogenic forcings. We assume that many model biases are consistent across different time slices and, therefore, can be removed when calculating changes between the historical period (1985-2014) and future periods (2070-2099).

We present climate change trends in CCLM and the CMIP6 GCMs ensemble statistics (ensemble mean and standard deviation). We analysed 31, 33, and 38 models for SSP126, SSP370, and SSP585 scenarios, respectively, with a total of 158, 185, and 242 simulations (see Supplementary materials for the list of models used). We calculate statistics over each model's members to ensure equal weighting for individual models before building the final statistics. We have selected the yearly 99th percentile of daily precipitation (PR99), which accounts for the three days with the highest precipitation each year. Additionally, we chose the number of very heavy precipitation days during the period (ECA-RX20mm) as another index, which is commonly used in climate research to assess the impacts of heavy precipitation events on water resources, agriculture, and natural ecosystems (Klok and Klein Tank, 2008).

Figure 6 shows the changes in averaged PR99 at the end of the century (2070-2099) compared to the historical period (1985-2014) for CCLM (a,d,g) and CMIP6 GCMs (b,e,h) under different scenarios. The large-scale patterns remain consistent across all three scenarios, intensifying with increased anthropogenic influence. The standard deviation of the models' ensemble is depicted in Figures 6.c,f, i. Our analysis indicates that the Himalayas, particularly Nepal, North India, and Bhutan, exhibit the highest uncertainty among the GCMs in all scenarios. Except for this region and the eastern boundary of the domain, the standard deviation remains below 3 mm/day. Under the SSP585 and SSP370 scenarios, regions including Northwest India, North Pakistan, North and Southwest Iran, and the South and Southeast of the Black Sea are projected to experience increases in PR99 values exceeding 9 mm/day. A reduction in PR99 is detected in the eastern Mediterranean, specifically in Jordan, Syria, and southern Turkey. Similar patterns are observed in the CMIP6 ensemble mean, but due to averaging, the ensemble mean patterns are approximately $\pm$5 mm/day over these areas. Under the SSP126 scenario, which is aligned with the 2°C warming target, the previously observed increases in precipitation exceeding $\pm$9 mm/day for CCLM and $\pm$5 mm/day for GCMs are no longer evident. In CA, areas such as Kyrgyzstan, Tajikistan, northern Pakistan, and southwestern Iran are particularly vulnerable to rainfall-induced hazards, including landslides (Wang et al., 2021; Kirschbaum et al., 2010) and floods (e.g., the Pakistan floods of 2010 and 2022).

Figures 7.1, d, and g illustrate the ECA-RX20mm values for CCLM at the end of the century across three scenarios. The observed patterns align with those in Figure 6, underscoring an increase in the frequency of very heavy precipitation days, particularly marked over the Tibetan Plateau, as anthropogenic influences intensify. Similarly, Figures 7.b, e, and h reveal that the CMIP6 GCM ensemble mirrors the behaviour observed in CCLM. However, the ensemble standard deviations for ECA-RX20mm values rise over Tajikistan and Kyrgyzstan, as shown in Figures 7.c, f, and i. The growing frequency and intensity of extreme precipitation events over the elevated regions of Central Asia, driven by anthropogenic factors, are a cause for concern (Fallah et al., 2023). This CCLM simulation enhances our understanding of how dynamical downscaling's sensitivity to different levels of anthropogenic forcing can vary locally.

## 4 CCLM emulator using a CNN

We have demonstrated that dynamical downscaling adds significant value in capturing local climate change effects, particularly over areas influenced by complex topography. In this study, we create a CCLM emulator for precipitation over CA. As previously explained, a CNN trained on our GCM-RCM chain could serve as a fast, cost-effective downscaling method, though its efficacy needs to be rigorously assessed.

We aim to establish that this emulator outperforms simple interpolation, particularly in areas experiencing extreme precipitation. We aim to show that the CCLM emulator can replicate CCLM-like precipitation patterns when driven by the parent GCM.

Focusing on the CA domain, which encompasses the former Soviet Union countries (Kazakhstan, Kyrgyzstan, Tajikistan, Turkmenistan, and Uzbekistan), we exclude the broader CORDEX-CA domain shown in Figure 3. This domain is the region of interest in the Green Central Asia project https://www.greencentralasia.org/en, which the German Foreign Office finances. Figure 8.a illustrates the MAE of the interpolated MPI-ESM1-2-HR, using the CCLM output as the 'true' precipitation. CCLM generates distinct precipitation patterns, particularly in areas with complex topography. Assuming CCLM as the ground truth, we examine whether the CNN can replicate these outputs using the GCM as input. To assess the emulator's effectiveness, we present added value maps (relative to the parent GCM) in Figures 8.b-d. A comparison of MAE reduction maps reveals that the unconstrained CNN demonstrates significant skill over elevated regions of CA, whereas constrained runs show less noticeable pattern changes. For instance, the HCL and SCL emulators generate closely mingled negative and positive added values across elevated areas, while NoCL consistently exhibits positive values across the domain. Several artefacts in the MAE reduction maps of constrained models, particularly over northern India, reflect the shape of the GCM grid. We also produce boxplots of daily precipitation for the CA domain to explore distribution improvements (Figure 9). Correlation coefficients between time-series averages of precipitation across the domain and CCLM are presented in Figure 9 (values in parentheses). Among daily averages, NoCL achieves the best performance (highest correlation coefficient), although it records fewer outliers than CCLM and other model simulations. The distribution is concentrated around the median, exhibiting the narrowest interquartile range. The distribution profiles of both constrained models (HCL, SCL) resemble those of the interpolated GCM, expected since the constraints maintain mass consistency within corresponding grid boxes (Equation 1).

### 4.1 Applying the CNN to a different GCM

We evaluate the emulator's generalisation ability, i.e., its capacity to generate reliable predictions on new datasets. We conduct a new 15-year dynamical simulation using CCLM, driven by the EC-Earth3-Veg (Döscher et al., 2022) GCM under the SSP370 scenario from 2019 to 2033. This data serves as input to our CCLM emulator, which was previously trained to emulate CCLM outputs using MPI-ESMI-2 HR. We now use the emulator to reconstruct the local features of CCLM driven by EC-Earth3-Veg.

Figure 10.a presents the MAE of the interpolated EC-Earth3-Veg with respect to the dynamical downscaling with CCLM. Remarkably, the MAE pattern of EC-Earth3-Veg closely mirrors that of MPI-ESM1-2-HR (Figure 8.a). However, the NoCL emulator does not uniformly show positive error reduction across the domain (Figure 10.b). We chose NoCL for its superior performance among the three CNNs. The emulator attempts to establish relationships between MPI-ESM1-2-HR and CCLM, which may be specific to these models and might not necessarily apply to the new EC-Earth3-Veg and CCLM configuration. As demonstrated previously, the RCM state depends on the state of its driving GCM. CCLM is driven at the lateral boundaries by the GCM values for state variables (temperature, pressure, wind speed, etc.) and not by precipitation, which is the CNN's input. The precipitation inputs from the two GCMs carry different biases, complicating the transfer of mapping from MPI-ESM1-2-HR-driven CCLM outputs to those driven by EC-Earth3-Veg.

Despite these challenges, the CNN model demonstrates added values exceeding 1 mm/day in regions such as the Alborz Mountains and the southern Caspian Sea in northern Iran (highlighted in black rectangles in Figures 10.a and b) and parts of Tajikistan and Kyrgyzstan. Exploration of the daily precipitation distribution field-mean indicates that the CNN's median value and outliers are lower than those of the EC-Earth3-Veg and CCLM simulations (Figure 10.c). The day-to-day correlation has improved, although all models were trained on a shuffled dataset that ignored the memory in the time series. The trained NoCL model was provided with unshuffled EC-EARTH3-Veg data for new predictions, increasing the correlation coefficient from 0.815 (EC-Earth3-Veg) to 0.844 (NoCL). Over the highlighted area in Figure 10.b, where the NoCL model reduces MAE, the distribution of precipitation converges towards that of CCLM, encompassing the region with the highest rainfall in Iran, vital for a large portion of the population, including Tehran. Only the outliers larger than 20 mm/day are not reconstructed by NoCL.

As a further test of generalisation, we intentionally excluded the SSP370 scenario from the training process. This allowed us to apply the model to a specific simulation and assess its ability to handle an unknown forcing. Figure 11 demonstrates the AV of the CNN emulator for SSP370 in comparison to the dynamical downscaling with CCLM, revealing that the AV pattern is strikingly similar to that shown in Figure 8.d. This confirms that the CNN can learn and reproduce patterns under different forcing scenarios it was not explicitly trained on, as demonstrated by its performance with the SSP370 scenario.

## 5 Discussion and conclusions

Regional climate change impact assessments require high-resolution climate projections. The main strategies to produce such datasets are statistical and dynamical downscaling, as well as a hybrid of the two methods. Statistical downscaling often struggles to account for the dynamic influences of complex landscapes, including topography and varying surface parameters such as vegetation, soil types, and water bodies like lakes, which may affect the accuracy of statistical relationships (Li et al., 2022). For statistical downscaling methods applied to precipitation, observations need to contain detailed information about precipitation distribution in areas with complex topography (Lundquist et al., 2019).

Conversely, dynamical downscaling requires massive computational time and data storage. For example, a 30-year CCLM simulation driven by ERAInterim took roughly one week to complete using 216 processors of the HLRE-4 Levante computer at the German Climate Computing Center (DKRZ). Additionally, the added value of RCMs is still debated, as they are highly dependent on the driving GCMs.

In this study, we contributed to the dynamic downscaling efforts over the CORDEX-CA domain, taking a small step towards creating an RCM ensemble for CA. A single RCM simulation helps identify model biases and uncertainties that need to be addressed in future model improvements. It is essential to note that a single model run for CMIP6, instead of an RCM ensemble, may not provide a comprehensive understanding of potential climate change impacts on a region. Therefore, it is recommended that researchers conduct multiple simulations with different initial and boundary conditions and model configurations to account for the uncertainty associated with climate projections.

In the first part of the study, we demonstrated the added value of RCMs (using the CCLM model) over GCMs for CA in representing precipitation. Our CCLM run showed added value with respect to its driving GCM, comparable to the range of values obtained for other RCMs applied to the CORDEX-CA domain over the evaluation period. It also reproduced extreme precipitation patterns similar to the CMIP6 ensemble mean projections for the end of the century. Both the CCLM and CMIP6 ensembles indicated an increased risk (in terms of intensity and frequency) of heavy precipitation events in vulnerable regions of CA due to various human activities.

Our study evaluated the downscaling skill using high-resolution observations, a crucial step for accurately capturing localised climate phenomena. This evaluation was essential before further study steps and regional adaptation strategies could be implemented. However, as Volosciuk et al. (2017) noted, examining downscaling outputs at coarser resolutions can be equally informative. Their work emphasises that downscaling methods can introduce or fail to correct biases that differ significantly across spatial scales. By evaluating on a coarser grid, it is possible to distinguish between the inherent biases of the model and those introduced by the downscaling process. This distinction is crucial for understanding the limitations and strengths of downscaling methods in representing climatic variables across different scales.

We showed that a single GCM-RCM model chain could be used to train a climate emulator based on a CNN model. It learned nonlinear and physical relationships between the coarse and fine-resolution datasets, addressing the issue of spatial intermittency—where data points are unevenly distributed or missing across space—common in some statistical downscaling approaches (Harder et al., 2023). However, we also demonstrated that the CNN model had limitations when generalising, as it did not achieve a robust error-reduction pattern when given a different GCM as input. The learning process strongly depended on the GCM/CCLM relationships. More importantly, an RCM is usually forced to follow its driving GCM and can only produce extra information on a local scale. The presented CNN could be applied to other experiments of the same GCM, such as using the trained emulator for paleo-climate experiments or downscaling volcanic forcing experiments. This would aid the paleo-climate community in conducting proxy-model comparisons at local scales. However, previous studies have shown that the CNN suffered from the same generalisation problem when applied to a new GCM, and such applications must be tested (Jouvet and Cordonnier, 2023).

We deliberately excluded the SSP370 scenario from the training dataset to evaluate the model's generalisation capabilities for other scenarios of the same GCM. This strategy allowed us to assess whether the model could effectively infer and replicate patterns from untrained scenarios. Remarkably, the model's output for the SSP370 scenario exhibited an AV pattern mirrored the dynamical downscaling results of the CCLM driven by the same SSP370 scenario. This alignment strongly supported the notion that our CNN emulator could learn from its training data and generalise to new, unseen conditions. The similarity in AV patterns between the model output and the CCLM simulation underscored the robustness and adaptability of our model, affirming its potential for broader applications in climate modelling.

This work was an initial step in demonstrating the potential of such a hybrid approach. We encourage the community to explore different model structures and parameter combinations for further improvement. For example, our initial setups showed that a physically constrained CNN setup that applies a linear transformation to ensure mass or energy conservation between the low and high-resolution images did not successfully downscale precipitation. The original dataset might not satisfy the constraints, leading to suboptimal results. In contrast, with a higher degree of freedom, the unconstrained CNN produced patterns closer to the target RCM. Future studies could test alternative machine learning models, such as generative adversarial networks (GANs), which can generate more high-frequency patterns and improve the downscaled output. Additionally, incorporating more information into the CNN by adding characteristics like surface height, vegetation, land cover, and land use as new channels within the input layer could enhance model performance.

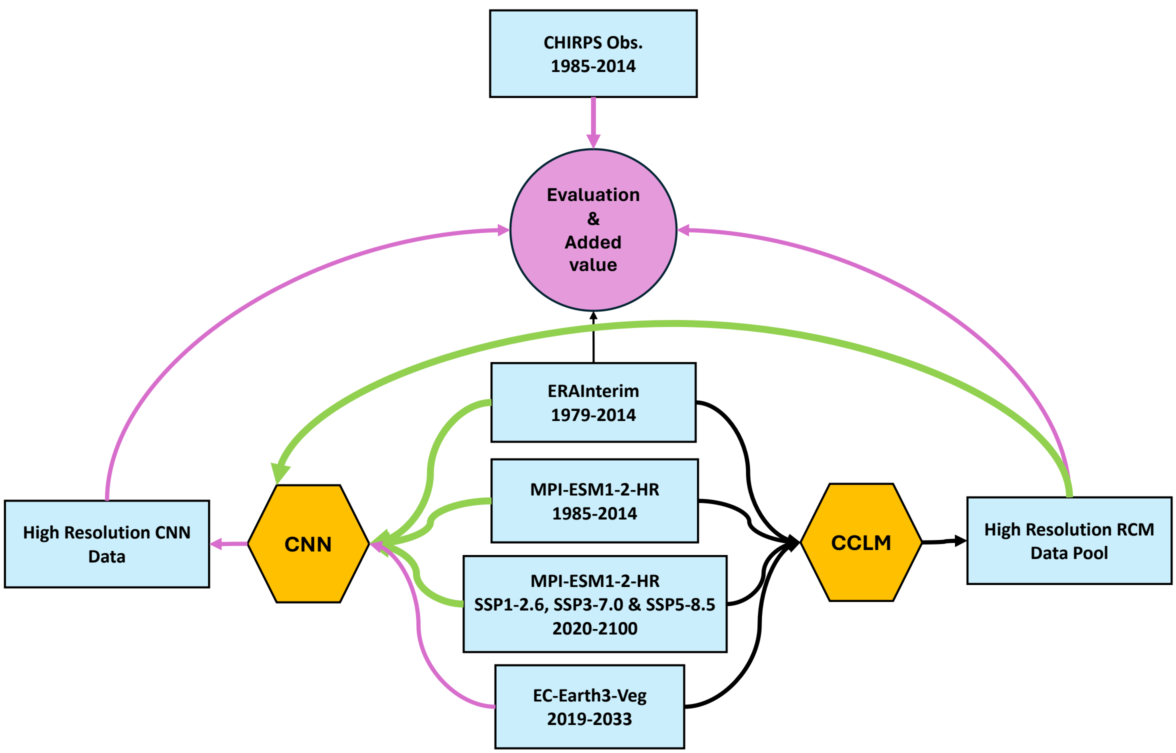



**Figure 1.** Schematic of the methodology used in this study. Green arrows show the data flow used for training the CNN and magenta for evaluation and calculation of the added values. Datasets are shown by rectangular, downscaling models by hexagonal and evaluation analysis by circle.

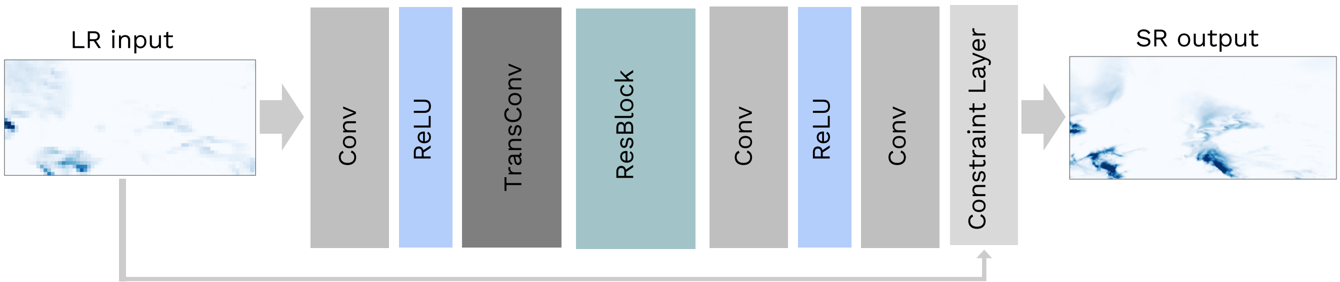


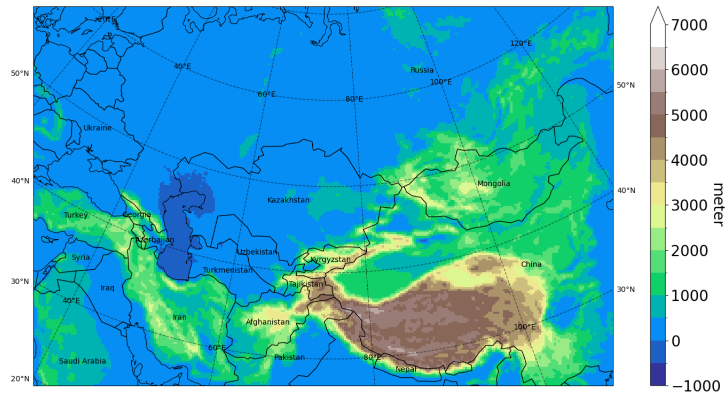

**Figure 2.** Schematic of the CNN architecture for 2 times upsampling with the constraints layer. The inputs are low-resolution (LR) images of size 30×60 and the output is a super-resolution (SR) image of size 60×120. This figure is modified from (Harder et al., 2023).

*Code availability.* The code for "Physics-Constrained Deep Learning for Climate Downscaling" is available on Zenodo at the following DOI: https://zenodo.org/record/8150694. This repository includes the input and output data, trained models, a snapshot of the code used in the deep-learning downscaling process, CCLM model setups for all Regional Climate Model (RCM) simulations conducted, and a list of CMIP6 models used for comparative analysis. Additionally, a Jupyter notebook for executing a test case of the "Physics-Constrained Deep Learning for Climate Downscaling" is available at Zenodo with the following DOI: https://zenodo.org/record/10417111.

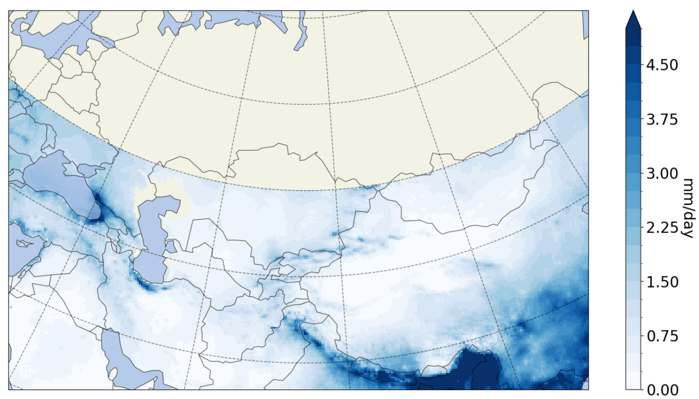
## Appendix A: CNN runs

We used the following commands for training the CNN model based on the Harder et al. (2023):

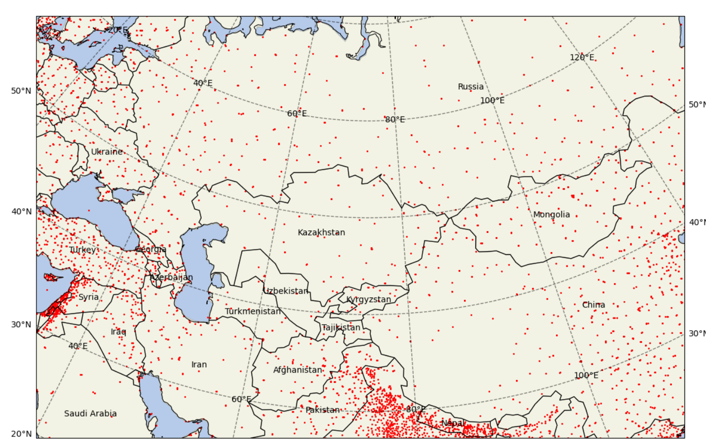
```
# for the run with soft constraining run, with a factor of alpha 0.99 :

$ python main.py --dataset dataset --model cnn --model_id
twc_cnn_soft_constraints_epochs_160_lr_0.00001_alpha_0.99
--constraints soft --loss mass_constraints --alpha 0.99
 --epochs 160 --batch_size 64 --lr 0.00001

# for the run with softmax constraining or hard constraining:
$ python main.py --dataset dataset --model cnn --model_id
twc_cnn_softmaxconstraints_epochs_200_batch_size_64_lr_0.001
--constraints softmax --lr 0.001 --epochs 160 --batch_size 64 --loss mae
```




(a)

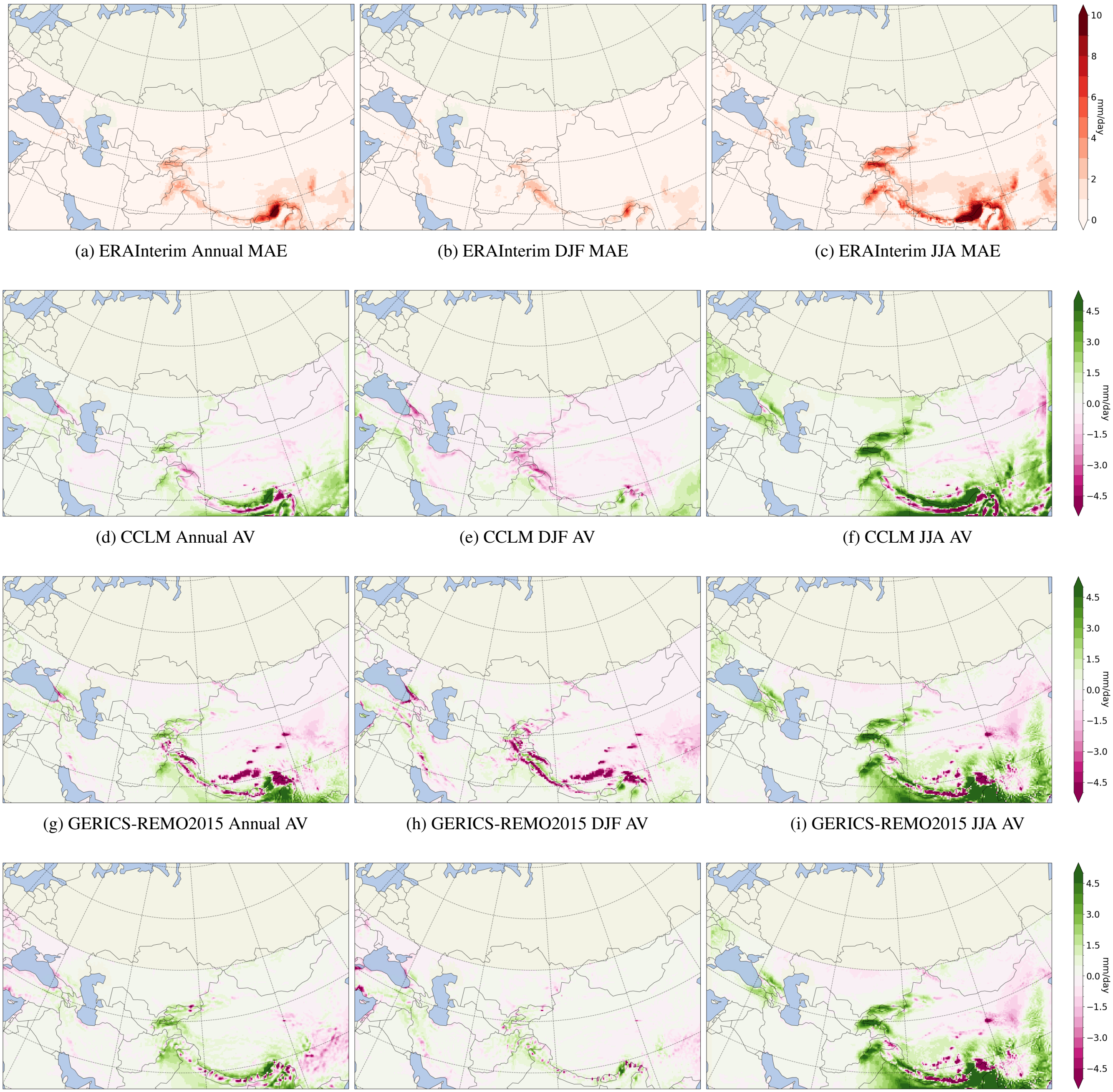



(b)




(c)

**Figure 3.** a) CCLM simulation domain over Central Asia and the topography (m), (b) CHIRPS climatology for 1985-2014 (average of daily values over all years in mm/day), and (c) WorldClim's weather stations (red dots).



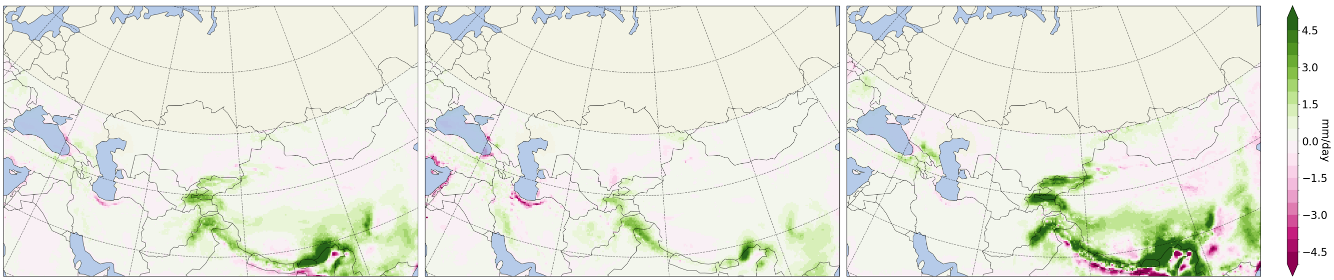

(a) ERAInterim Annual MAE (b) ERAInterim DJF MAE (c) ERAInterim JJA MAE

(d) CCLM Annual AV (e) CCLM DJF AV (f) CCLM JJA AV

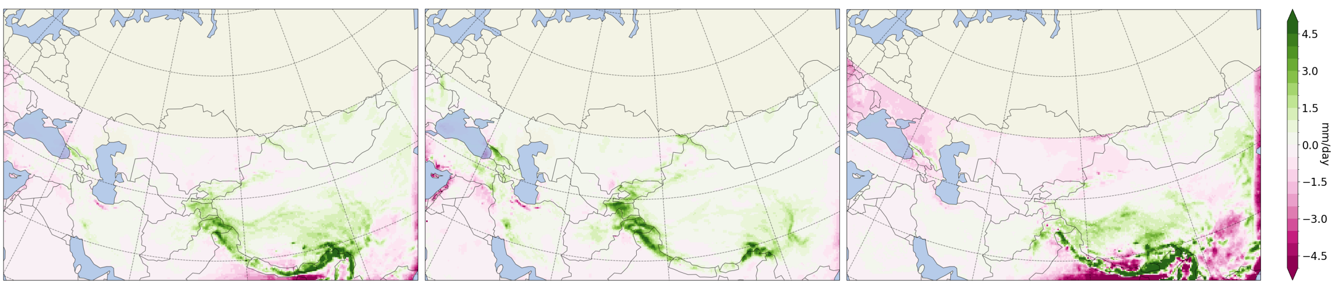
(g) GERICS-REMO2015 Annual AV (h) GERICS-REMO2015 DJF AV (i) GERICS-REMO2015 JJA AV

(j) RMIB-UGent-ALARO-0 Annual AV (k) RMIB-UGent-ALARO-0 DJF AV (l) RMIB-UGent-ALARO-0 JJA AV

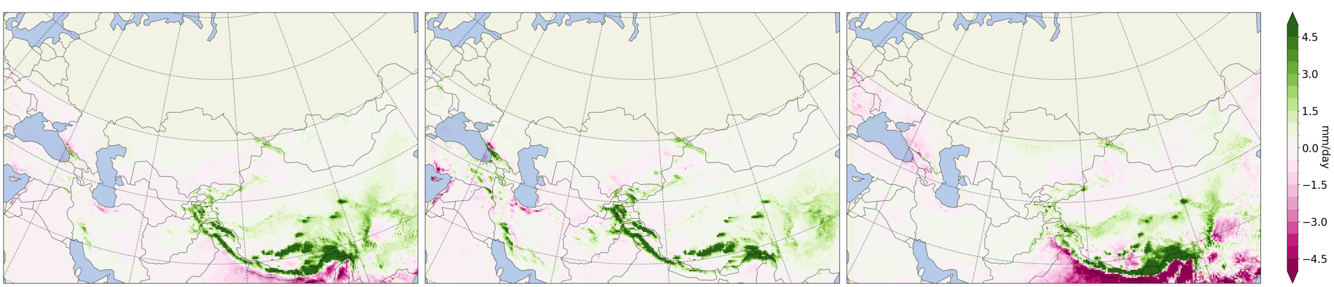
**Figure 4.** Mean absolute error (MAE) of daily precipitation (mm/day) from ERAInterim, as well as, added value (AV) as measured by MAE differences between ERAInterim and RCMs ($\mathrm{MAE}_{\mathrm{ERAInterim}} - \mathrm{MAE}_{\mathrm{RCM}}$) in mm/day for annual (a,d,j,i), December, January, February (b,e,h,k) and June, July, August (c,f,i,l). CHIRPS is used as observation. All datasets are interpolated to the CCLM grid.



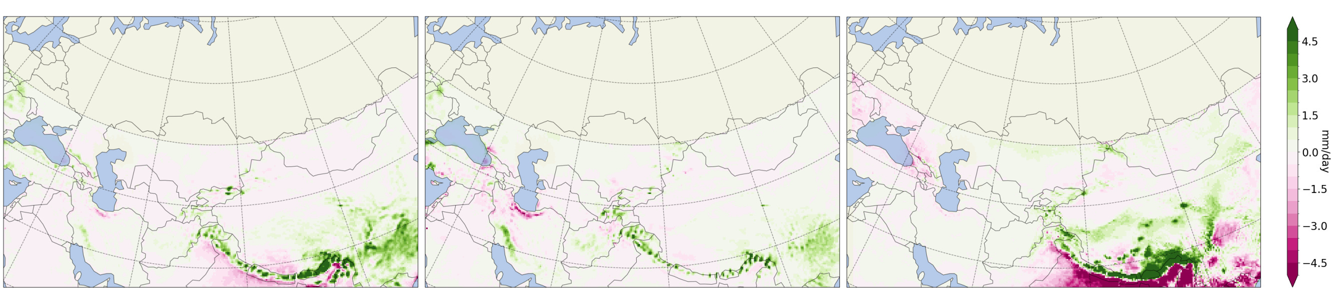

(a) ERAInterim Annual Bias (b) ERAInterim DJF Bias (c) ERAInterim JJA Bias




(d) CCLM Annual Bias (e) CCLM DJF Bias (f) CCLM JJA Bias

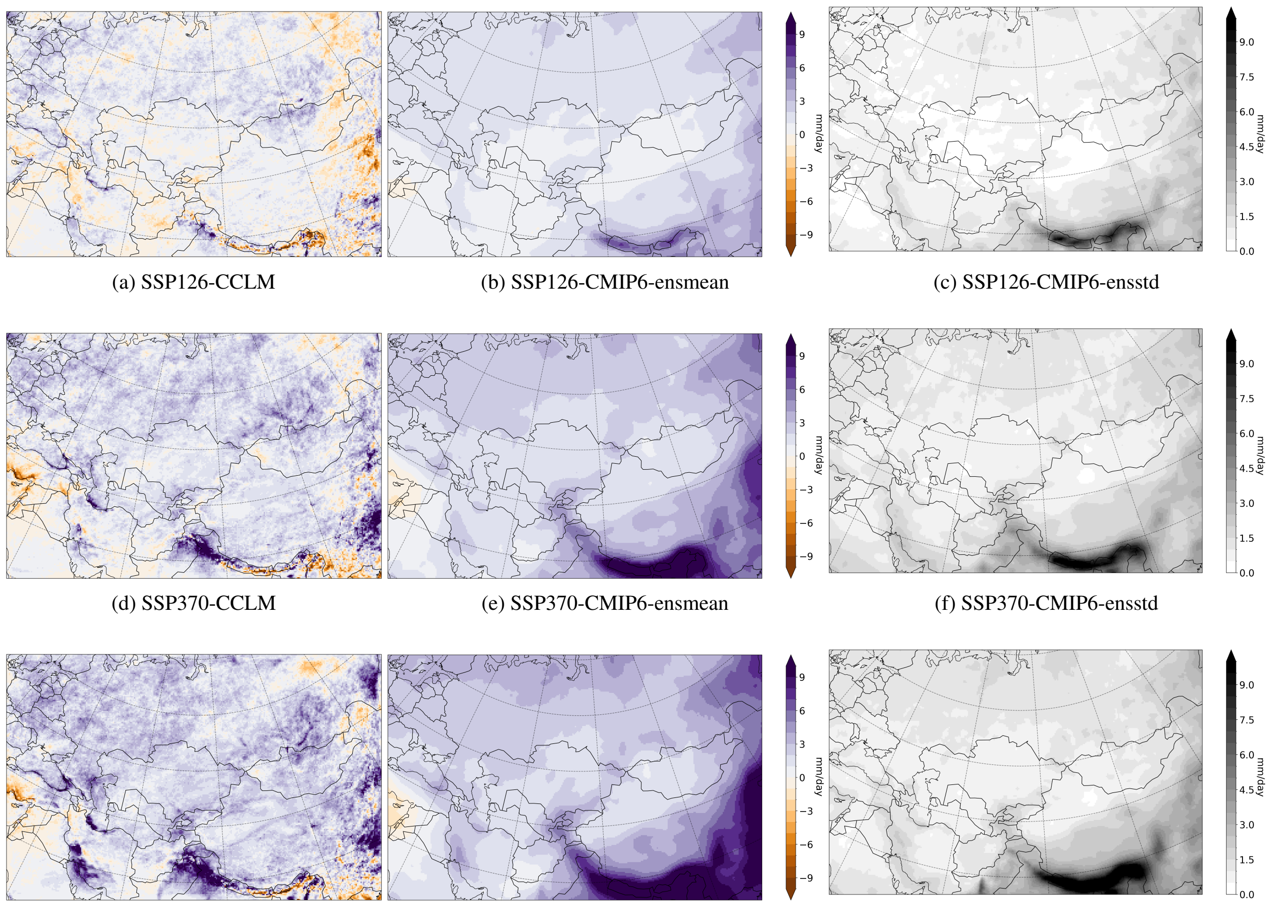



(g) GERICS-REMO2015 Annual Bias (h) GERICS-REMO2015 DJF Bias (i) GERICS-REMO2015 JJA Bias




(j) RMIB-UGent-ALARO-0 Annual Bias (k) RMIB-UGent-ALARO-0 DJF Bias (l) RMIB-UGent-ALARO-0 JJA Bias

**Figure 5.** Bias of climatological precipitation (mm/day) from ERAInterim, as well as, ERAInterim-driven RCMs ($PR_{\text{ERAInterim-CCLM}} - PR_{\text{OBS}}$) in mm/day for annual (a,d,j,i), December, January, February (b,e,h,k) and June, July, August (c,f,i,l). CHIRPS is used as observation.




(a) SSP126-CCLM (b) SSP126-CMIP6-ensmean (c) SSP126-CMIP6-ensstd

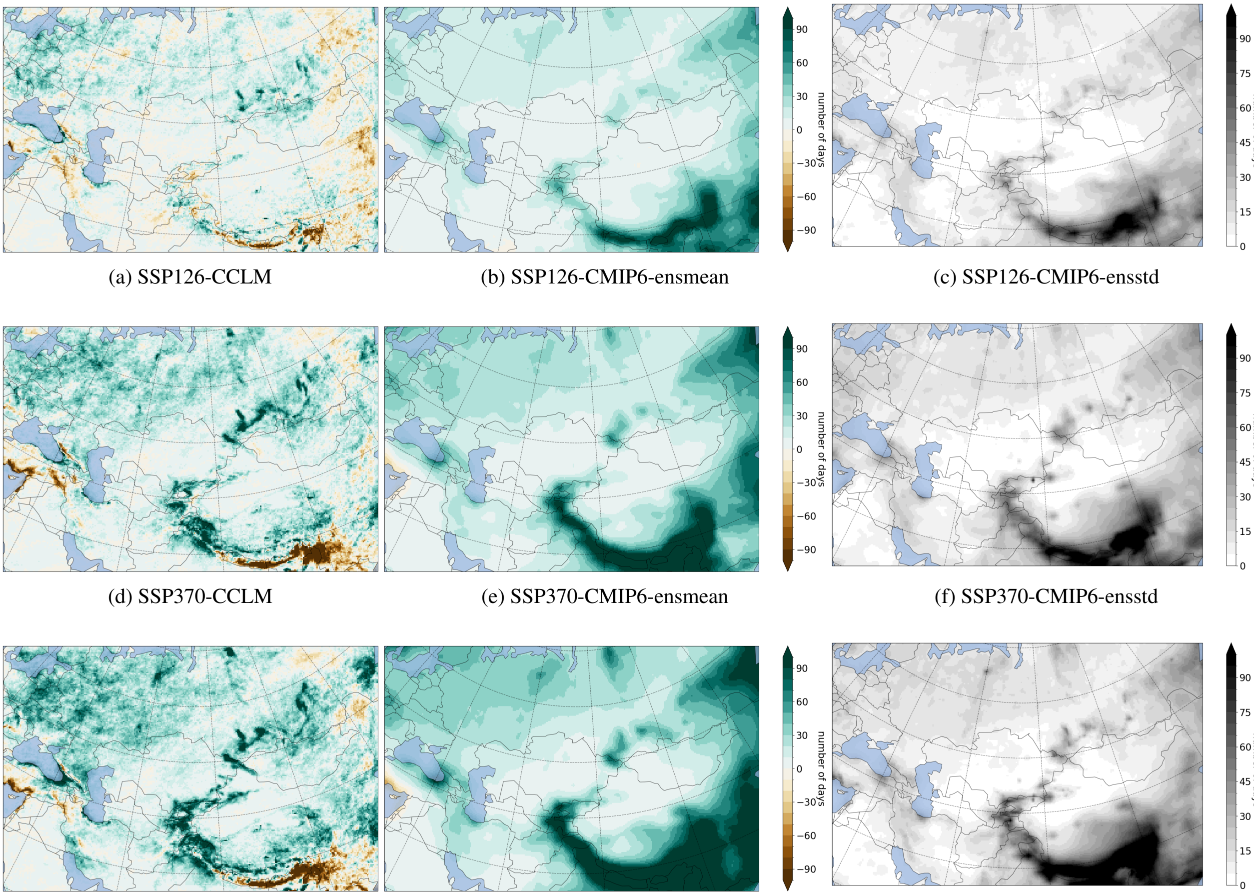
(d) SSP370-CCLM (e) SSP370-CMIP6-ensmean (f) SSP370-CMIP6-ensstd

(g) SSP585-CCLM (h) SSP585-CMIP6-ensmean (i) SSP585-CMIP6-ensstd

**Figure 6.** Changes in averaged yearly $99^{th}$ percentile (3 days per year) of total precipitation (mm/day) with respect to 1985-2014 references for a,b) SSP126, d,e) SSP370 and g,h) SSP585 at the end of the century (2070-2099) from CCLM and CMIP6 GCMs' ensemble mean. The ensemble's standard deviations are shown in c,f and i.

```
# for the standard CNN run without constraining:
$ python main.py --dataset dataset --model cnn --model_id
twc_cnn_noneconstraints_epochs_160_batch_size_64_lr_0.001
--constraints none  --lr 0.001 --epochs 160 --batch_size 64 --loss mae
```

Note that the datasets and codes are available at Zenodo (DOI: https://zenodo.org/records/10417111) with comprehensive details utilized in the paper.




(a) SSP126-CCLM (b) SSP126-CMIP6-ensmean (c) SSP126-CMIP6-ensstd

(d) SSP370-CCLM (e) SSP370-CMIP6-ensmean (f) SSP370-CMIP6-ensstd

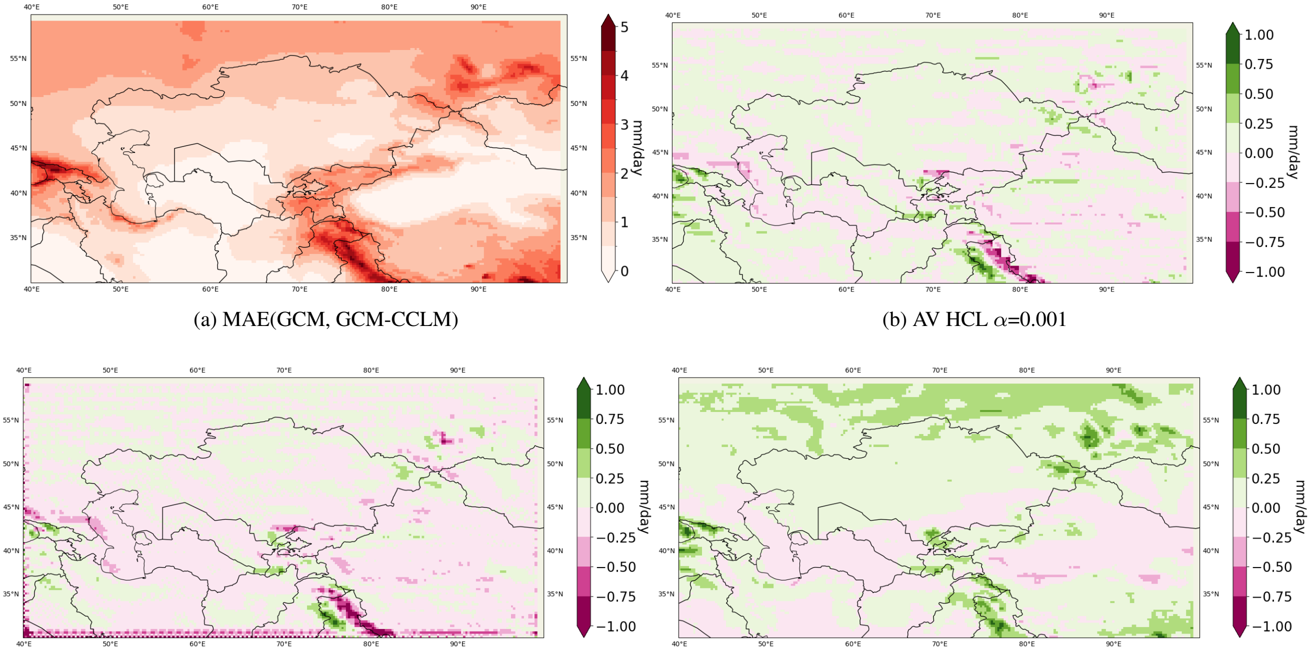
(g) SSP585-CCLM (h) SSP585-CMIP6-ensmean (i) SSP585-CMIP6-ensstd

**Figure 7.** Changes in number of days with precipitation more than 20 mm in the period with respect to 1985-2014 references for a,b) SSP126, d,e) SSP370 and g,h) SSP585 at the end of the century (2070-2099) from CCLM and CMIP6 GCMs' ensemble mean. The ensemble's standard deviations are shown in c,f and i.

*Author contributions.* BF conducted the dynamical and statistical downscaling with assistance from ER and PH, respectively. ER provided the setup for the CCLM simulations. PH provided the deep learning model code and setup. All authors contributed to the analysis of the results and the writing of the manuscript.

*Competing interests.* The authors declare that they have no competing interests.

*Acknowledgements.* BF thanks the German Climate Computing Center (DKRZ) for its support in using supercomputer data and resources. The German Foreign Office funded BF and ID via the Green Central Asia project (http://greencentralasia.org/en, last access: 4 July 2023).




(a) MAE(GCM, GCM-CCLM)

(b) AV HCL $\alpha$=0.001

(c) AV SCL $\alpha$=0.99

(d) AV NoCL

**Figure 8.** a) MAE (MPI-ESM1-2-HR,CCLM). MPI-ESM1-2-HR is remapped bilinearly to the $0.25 \times 0.25$ grid. b-d) Added Value (AV) or MAE(MPI-ESM1-2-HR,CCLM) - MAE(CNN,CCLM) for different constraining method.

The DKRZ and PIK provided the computational resources. The authors gratefully acknowledge the German Federal Ministry of Education and Research and the Land Brandenburg for supporting this project by providing resources on the high performance computer system at the Potsdam Institute for Climate Impact Research. BF thanks the CCLM community for providing the model code and the pre-processing code to convert the GCM to CCLM input files. BF is supported by the Coming Decade project at DKRZ.

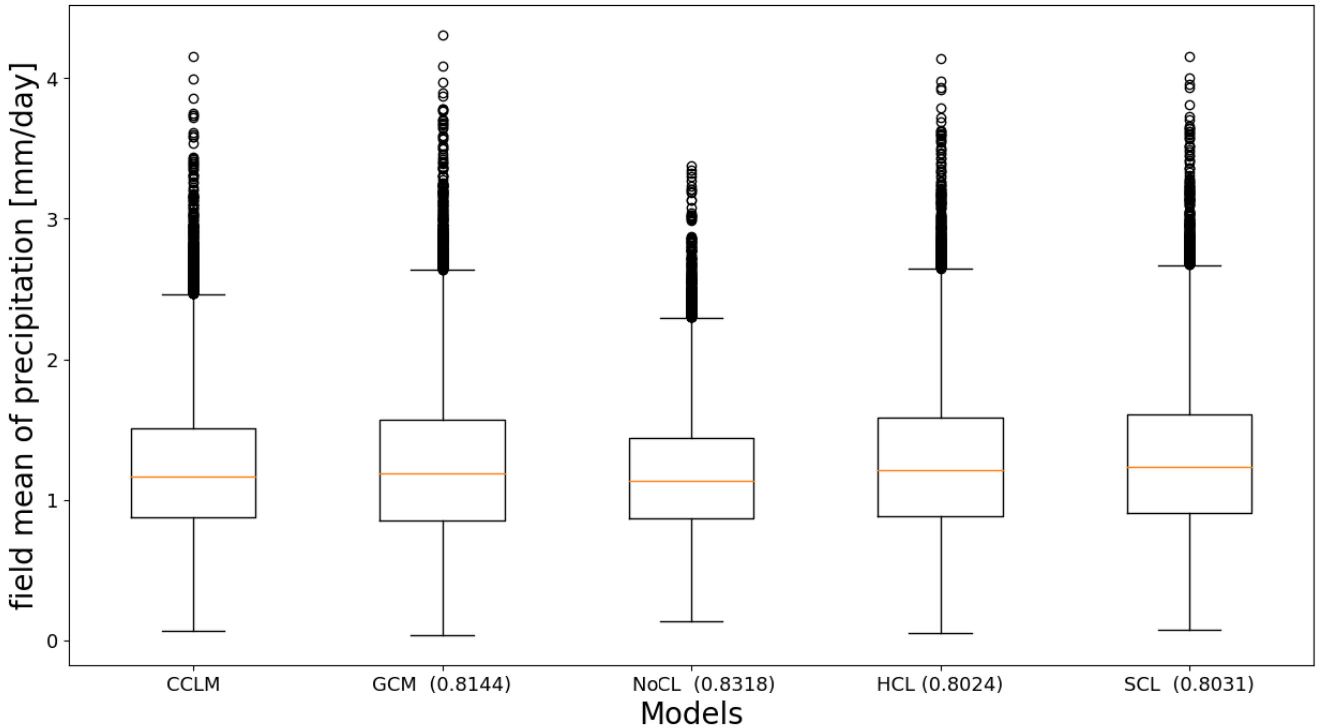



**Figure 9.** Boxplot of averaged daily precipitation over the Central Asian domain (shown in Figure 7) for different models and test dataset (22714 days or 62.2 years). Numbers in the parenthesis indicate the correlation coefficients between each model and the CCLM simulation.

## References


Allan, R. P., Hawkins, E., Bellouin, N., and Collins, B.: IPCC, 2021: summary for Policymakers, IPCC, 2021.

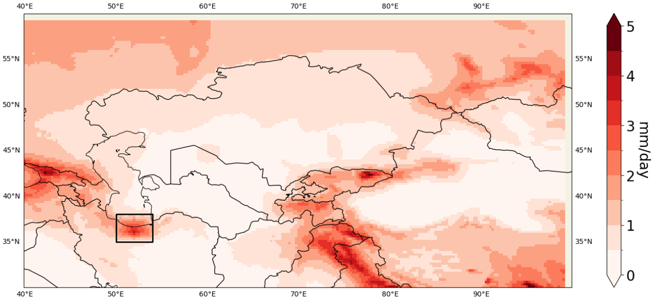
Ban, N., Schmidli, J., and Schär, C.: Heavy precipitation in a changing climate: Does short-term summer precipitation increase faster?, Geophysical Research Letters, 42, 1165–1172, 2015.

Chokkavarapu, N. and Mandla, V. R.: Comparative study of GCMs, RCMs, downscaling and hydrological models: a review toward future climate change impact estimation, SN Applied Sciences, 1, 1698, 2019.

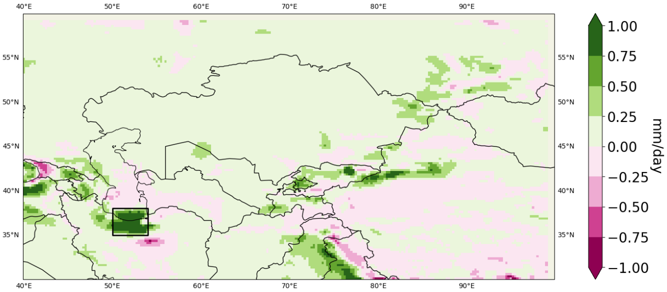
Ciarlo, J. M., Coppola, E., Fantini, A., Giorgi, F., Gao, X., Tong, Y., Glazer, R. H., Torres Alavez, J. A., Sines, T., Pichelli, E., et al.: A new spatially distributed added value index for regional climate models: the EURO-CORDEX and the CORDEX-CORE highest resolution ensembles, Climate Dynamics, 57, 1403–1424, 2021.

Cui, T., Li, C., and Tian, F.: Evaluation of temperature and precipitation simulations in CMIP6 models over the Tibetan Plateau, Earth and Space Science, 8, e2020EA001 620, 2021.

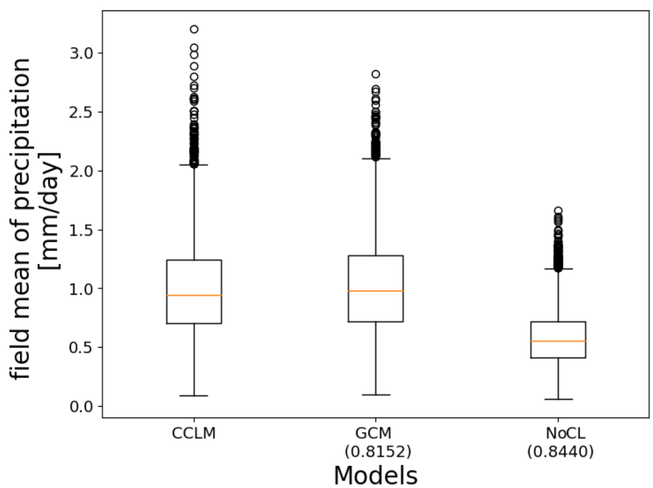
Demory, M.-E., Berthou, S., Fernández, J., Sørland, S. L., Brogli, R., Roberts, M. J., Beyerle, U., Seddon, J., Haarsma, R., Schär, C., et al.: European daily precipitation according to EURO-CORDEX regional climate models (RCMs) and high-resolution global climate models (GCMs) from the High-Resolution Model Intercomparison Project (HighResMIP), Geoscientific Model Development, 13, 5485–5506, 2020.



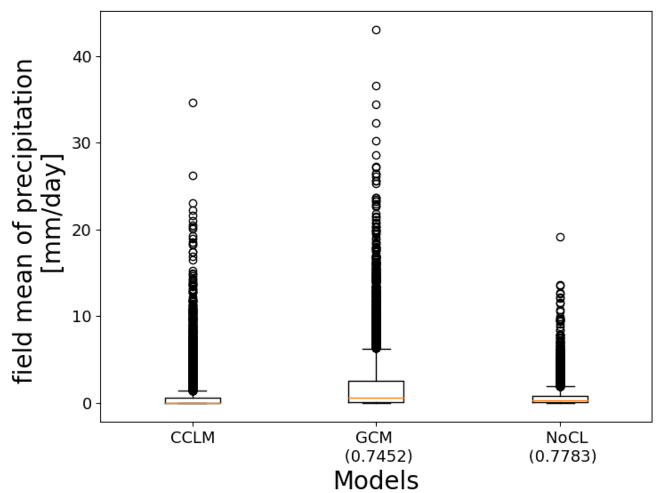

(a) MAE(EC-Earth3-Veg, CCLM)




(b) AV NoCL




(c) CA




(d) Box

**Figure 10.** a) MAE of GCM (EC-Earth3-Veg) vs CCLM run. GCM is remapped bilinearly to the 0.25×0.25 grid. b) Added value (AV) or MAE reduction (MAE(EC-Earth3-Veg,CCLM) - MAE(CNN,CCLM) for unconstrained method. c) and d) boxplots of averaged daily precipitation over the CA domain and the black box shown in a and b over North of Iran. Numbers in the parenthesis indicate the correlation coefficients of each model with respect to CCLM.

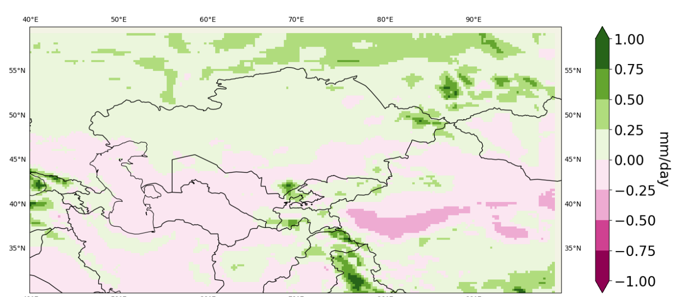



**Figure 11.** Added value (AV) or MAE reduction (MAE(EC-MPI-ESM1-2HR, CCLM)—MAE(CNN, CCLM)) for an unconstrained method that was not trained but applied to the SSP370 scenario.

Didovets, I., Krysanova, V., Nurbatsina, A., Fallah, B., Krylova, V., Saparova, A., Niyazov, J., Kalashnikova, O., and Hattermann, F. F.: Attribution of current trends in streamflow to climate change for 12 Central Asian catchments, Climatic Change, 177, 16, 2024.

Döscher, R., Acosta, M., Alessandri, A., Anthoni, P., Arneth, A., Arsouze, T., et al.: The EC-Earth3 Earth system model for the Coupled Model Intercomparison Project 6. Geosci. Model Dev. 15, 2973–3020, 2022.

Dosio, A. and Panitz, H.-J.: Climate change projections for CORDEX-Africa with COSMO-CLM regional climate model and differences with the driving global climate models, Climate Dynamics, 46, 1599–1625, 2016.

Fallah, B. and Rostami, M.: Exploring the impact of the recent global warming on extreme weather events in Central Asia using the counterfactual climate data ATTRICI v1. 1, Climatic Change, preprint, https://doi.org/10.21203/rs.3.rs-2106031/v1, 2024.

Fallah, B., Saberi, A. A., and Sodoudi, S.: Emergence of global scaling behaviour in the coupled Earth-atmosphere interaction, Scientific Reports, 6, 34 005, 2016a.

Fallah, B., Sodoudi, S., and Cubasch, U.: Westerly jet stream and past millennium climate change in Arid Central Asia simulated by COSMO-CLM model, Theoretical and Applied Climatology, 124, 1079–1088, 2016b.

Fallah, B., Russo, E., Menz, C., Hoffmann, P., Didovets, I., and Hattermann, F. F.: Anthropogenic influence on extreme temperature and precipitation in Central Asia, Scientific Reports, 13, 6854, 2023.

Feser, F., Rockel, B., von Storch, H., Winterfeldt, J., and Zahn, M.: Regional climate models add value to global model data: a review and selected examples, Bulletin of the American Meteorological Society, 92, 1181–1192, 2011.

Fick, S. E. and Hijmans, R. J.: WorldClim 2: new 1-km spatial resolution climate surfaces for global land areas, International journal of climatology, 37, 4302–4315, 2017.

Fotso-Nguemo, T. C., Vondou, D. A., Pokam, W. M., Djomou, Z. Y., Diallo, I., Haensler, A., Tchotchou, L. A. D., Kamsu-Tamo, P. H., Gaye, A. T., and Tchawoua, C.: On the added value of the regional climate model REMO in the assessment of climate change signal over Central Africa, Climate Dynamics, 49, 3813–3838, 2017.

Fowler, H. J., Blenkinsop, S., and Tebaldi, C.: Linking climate change modelling to impacts studies: recent advances in downscaling techniques for hydrological modelling, International Journal of Climatology: A Journal of the Royal Meteorological Society, 27, 1547–1578, 2007.

Frei, C., Christensen, J. H., Déqué, M., Jacob, D., Jones, R. G., and Vidale, P. L.: Daily precipitation statistics in regional climate models: Evaluation and intercomparison for the European Alps, Journal of Geophysical Research: Atmospheres, 108, 2003.

Funk, C., Peterson, P., Landsfeld, M., Pedreros, D., Verdin, J., Shukla, S., Husak, G., Rowland, J., Harrison, L., Hoell, A., et al.: The climate hazards infrared precipitation with stations—a new environmental record for monitoring extremes, Scientific data, 2, 1–21, 2015.

Galea, D., Ma, H.-Y., Wu, W.-Y., and Kobayashi, D.: Deep Learning Image Segmentation for Atmospheric Rivers, Artificial Intelligence for the Earth Systems, 3, 230 048, 2024.

Gardoll, S. and Boucher, O.: Classification of tropical cyclone containing images using a convolutional neural network: Performance and sensitivity to the learning dataset, Geoscientific Model Development, 15, 7051–7073, 2022.

Giot, O., Termonia, P., Degrauwe, D., De Troch, R., Caluwaerts, S., Smet, G., Berckmans, J., Deckmyn, A., De Cruz, L., De Meutter, P., et al.: Validation of the ALARO-0 model within the EURO-CORDEX framework, Geoscientific Model Development, 9, 1143–1152, 2016.

Harder, P., Yang, Q., Ramesh, V., Sattigeri, P., Hernandez-Garcia, A., Watson, C., Szwarcman, D., and Rolnick, D.: Generating physically-consistent high-resolution climate data with hard-constrained neural networks, 2022.

Harder, P., Hernandez-Garcia, A., Ramesh, V., Yang, Q., Sattegeri, P., Szwarcman, D., Watson, C., and Rolnick, D.: Hard-Constrained Deep Learning for Climate Downscaling, Journal of Machine Learning Research, 24, 1–40, 2023.

Hodson, T. O.: Root-mean-square error (RMSE) or mean absolute error (MAE): When to use them or not, Geoscientific Model Development, 15, 5481–5487, 2022.

Jacob, D. and Podzun, R.: Sensitivity studies with the regional climate model REMO, Meteorology and atmospheric physics, 63, 119–129, 1997.

Jacob, D., Petersen, J., Eggert, B., Alias, A., Christensen, O. B., Bouwer, L. M., Braun, A., Colette, A., Déqué, M., Georgievski, G., et al.: EURO-CORDEX: New high-resolution climate change projections for European impact research, Regional Environmental Change, 14, 563–578, 2014.

Jouvet, G. and Cordonnier, G.: Ice-flow model emulator based on physics-informed deep learning, Journal of Glaciology, pp. 1–15, 2023.

Kendon, E., Roberts, N., Fowler, H., Roberts, M., Chan, S., and Senior, C.: Heavier summer downpours with climate change revealed by weather forecast resolution model. Nat. Climate Change, 4, 570–576, 2014.

Kikstra, J. S., Nicholls, Z. R., Smith, C. J., Lewis, J., Lamboll, R. D., Byers, E., Sandstad, M., Meinshausen, M., Gidden, M. J., Rogelj, J., et al.: The IPCC Sixth Assessment Report WGIII climate assessment of mitigation pathways: from emissions to global temperatures, Geoscientific Model Development, 15, 9075–9109, 2022.

Kirschbaum, D. B., Adler, R., Hong, Y., Hill, S., and Lerner-Lam, A.: A global landslide catalog for hazard applications: method, results, and limitations, Natural Hazards, 52, 561–575, 2010.

Kjellström, E., Nikulin, G., Hansson, U., Strandberg, G., Ullerstig, A., Willén, U., and Wyser, K.: 21st century changes in the European climate: uncertainties derived from an ensemble of regional climate model simulations, Tellus A: Dynamic Meteorology and Oceanography, 63, 24–40, 2011.

Klok, E. and Klein Tank, A.: Updated and extended European dataset of daily weather observations, International Journal of Climatology, 28, 2081–2095, 2008.

Kotlarski, S., Keuler, K., Christensen, O. B., Colette, A., Déqué, M., Gobiet, A., Goergen, K., Jacob, D., Lüthi, D., Van Meijgaard, E., et al.: Regional climate modeling on European scales: a joint standard evaluation of the EURO-CORDEX RCM ensemble, Geoscientific Model Development, 7, 1297–1333, 2014.

Kurth, T., Treichler, S., Romero, J., Mudigonda, M., Luehr, N., Phillips, E., Mahesh, A., Matheson, M., Deslippe, J., Fatica, M., et al.: Exascale deep learning for climate analytics, in: SC18: International conference for high performance computing, networking, storage and analysis, pp. 649–660, IEEE, 2018.

Laflamme, E. M., Linder, E., and Pan, Y.: Statistical downscaling of regional climate model output to achieve projections of precipitation extremes, Weather and climate extremes, 12, 15–23, 2016.

Li, L., Bisht, G., and Leung, L. R.: Spatial heterogeneity effects on land surface modeling of water and energy partitioning, Geoscientific Model Development, 15, 5489–5510, 2022.

Lundquist, J., Hughes, M., Gutmann, E., and Kapnick, S.: Our skill in modeling mountain rain and snow is bypassing the skill of our observational networks, Bulletin of the American Meteorological Society, 100, 2473–2490, 2019.

Maraun, D. and Widmann, M.: Statistical downscaling and bias correction for climate research, Cambridge University Press, 2018.

Maraun, D., Widmann, M., Gutiérrez, J. M., Kotlarski, S., Chandler, R. E., Hertig, E., Wibig, J., Huth, R., and Wilcke, R. A.: VALUE: A framework to validate downscaling approaches for climate change studies, Earth's Future, 3, 1–14, 2015.

Meredith, E. P., Rust, H. W., and Ulbrich, U.: A classification algorithm for selective dynamical downscaling of precipitation extremes, Hydrology and Earth System Sciences, 22, 4183–4200, 2018.

Muttaqien, F. H., Rahadianti, L., and Latifah, A. L.: Downscaling for Climate Data in Indonesia Using Image-to-Image Translation Approach, in: 2021 International Conference on Advanced Computer Science and Information Systems (ICACSIS), pp. 1–8, IEEE, 2021.

Naddaf, M.: Climate change is costing trillions-and low-income countries are paying the price., Nature, 2022.

Panitz, H.-J., Dosio, A., Büchner, M., Lüthi, D., and Keuler, K.: COSMO-CLM (CCLM) climate simulations over CORDEX-Africa domain: analysis of the ERA-Interim driven simulations at 0.44 and 0.22 resolution, Climate dynamics, 42, 3015–3038, 2014.

Qiu, Y., Feng, J., Yan, Z., and Wang, J.: HCPD-CA: high-resolution climate projection dataset in central Asia, Earth System Science Data, 14, 2195–2208, 2022.

Racah, E., Beckham, C., Maharaj, T., Ebrahimi Kahou, S., Prabhat, M., and Pal, C.: Extremeweather: A large-scale climate dataset for semi-supervised detection, localization, and understanding of extreme weather events, Advances in neural information processing systems, 30, 2017.

Rampal, N., Hobeichi, S., Gibson, P. B., Baño-Medina, J., Abramowitz, G., Beucler, T., González-Abad, J., Chapman, W., Harder, P., and Gutiérrez, J. M.: Enhancing Regional Climate Downscaling through Advances in Machine Learning, Artificial Intelligence for the Earth Systems, 3, 230 066, 2024.

Randall, D. A., Wood, R. A., Bony, S., Colman, R., Fichefet, T., Fyfe, J., Kattsov, V., Pitman, A., Shukla, J., Srinivasan, J., et al.: Climate models and their evaluation, in: Climate change 2007: The physical science basis. Contribution of Working Group I to the Fourth Assessment Report of the IPCC (FAR), pp. 589–662, Cambridge University Press, 2007.

Reichstein, M., Camps-Valls, G., Stevens, B., Jung, M., Denzler, J., Carvalhais, N., and Prabhat, f.: Deep learning and process understanding for data-driven Earth system science, Nature, 566, 195–204, 2019.

Reyer, C. P., Otto, I. M., Adams, S., Albrecht, T., Baarsch, F., Cartsburg, M., Coumou, D., Eden, A., Ludi, E., Marcus, R., et al.: Climate change impacts in Central Asia and their implications for development, Regional Environmental Change, 17, 1639–1650, 2017.

Riahi, K., Van Vuuren, D. P., Kriegler, E., Edmonds, J., O'neill, B. C., Fujimori, S., Bauer, N., Calvin, K., Dellink, R., Fricko, O., et al.: The shared socioeconomic pathways and their energy, land use, and greenhouse gas emissions implications: an overview, Global environmental change, 42, 153–168, 2017.

Rockel, B. and Geyer, B.: The performance of the regional climate model CLM in different climate regions, as simulated in a transient climate change experiment, Climate Dynamics, 31, 713–728, 2008.

Rummukainen, M.: State-of-the-art with regional climate models, Wiley Interdisciplinary Reviews: Climate Change, 1, 82–96, 2010.

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, Geoscientific Model Development, 12, 5229–5249, 2019.

Russo, E., Fallah, B., Ludwig, P., Karremann, M., and Raible, C. C.: The long-standing dilemma of European summer temperatures at the mid-Holocene and other considerations on learning from the past for the future using a regional climate model, Climate of the Past Discussions, 2021, 1–27, 2021.

Sørland, S. L., Brogli, R., Pothapakula, P. K., Russo, E., Van de Walle, J., Ahrens, B., Anders, I., Bucchignani, E., Davin, E. L., Demory, M.-E., et al.: COSMO-CLM regional climate simulations in the Coordinated Regional Climate Downscaling Experiment (CORDEX) framework: a review, Geoscientific Model Development, 14, 5125–5154, 2021.

Taylor, K. E., Stouffer, R. J., and Meehl, G. A.: An overview of CMIP5 and the experiment design, Bulletin of the American meteorological Society, 93, 485–498, 2012.

Volosciuk, C., Maraun, D., Vrac, M., and Widmann, M.: A combined statistical bias correction and stochastic downscaling method for precipitation, Hydrology and Earth System Sciences, 21, 1693–1719, 2017.

Wang, D., Menz, C., Simon, T., Simmer, C., and Ohlwein, C.: Regional dynamical downscaling with CCLM over East Asia, Meteorology and Atmospheric Physics, 121, 39–53, 2013.

Wang, X., Otto, M., and Scherer, D.: Atmospheric triggering conditions and climatic disposition of landslides in Kyrgyzstan and Tajikistan at the beginning of the 21st century, Natural Hazards and Earth System Sciences, 21, 2125–2144, 2021.

Watson-Parris, D., Rao, Y., Olivié, D., Seland, Ø., Nowack, P., Camps-Valls, G., Stier, P., Bouabid, S., Dewey, M., Fons, E., et al.: ClimateBench v1. 0: A benchmark for data-driven climate projections, Journal of Advances in Modeling Earth Systems, 14, e2021MS002 954, 2022.

Xu, P., Wang, L., and Ming, J.: Central Asian precipitation extremes affected by an intraseasonal planetary wave pattern, Journal of Climate, 35, 2603–2616, 2022.

Xu, Z., Han, Y., Tam, C.-Y., Yang, Z.-L., and Fu, C.: Bias-corrected CMIP6 global dataset for dynamical downscaling of the historical and future climate (1979–2100), Scientific Data, 8, 293, 2021.

Yan, Y., You, Q., Wu, F., Pepin, N., and Kang, S.: Surface mean temperature from the observational stations and multiple reanalyses over the Tibetan Plateau, Climate Dynamics, 55, 2405–2419, 2020.

Zhu, X. X., Tuia, D., Mou, L., Xia, G.-S., Zhang, L., Xu, F., and Fraundorfer, F.: Deep learning in remote sensing: A comprehensive review and list of resources, IEEE geoscience and remote sensing magazine, 5, 8–36, 2017.