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. ~~To estimate~~ High-resolution climate projections are essential for estimating future climate change impacts~~, usually high-resolution climate projections are necessary~~. Statistical and dynamical downscaling methods, or a hybrid of both~~methods are mostly used to produce~~ , are commonly employed to generate input datasets for impact ~~modelers~~modelling. In this study, we ~~use the regional climate model (RCM)~~ employ COSMO-CLM (CCLM) version 6.0~~to identify the added value~~ , a regional climate model, to explore the benefits of dynamically downscaling a general circulation model (GCM) from ~~the sixth phase of the Coupled Model Inter-comparison Project (~~CMIP6~~) and its~~ , focusing on climate change projections ~~' signal over~~ for Central Asia (CA). ~~We use~~ The CCLM, at 0.22° horizontal resolution, is driven by the MPI-ESM1-2-HR GCM (at 1° spatial resolution) ~~to drive the CCLM (at 0.22° horizontal resolution)~~ for the historical period of 1985-2014 and the projection period of 2019-2100~~under three different~~ , 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 ~~ERAInterim~~ ERA-Interim reanalysis over the historical period. ~~CCLM's added value~~ The added value of CCLM, compared to its driving GCM, is significant (more than ~~5mm~~5 mm/day) over ~~CA mountainous areas~~ mountainous areas in CA, which are at higher risk of extreme precipitation events. Additionally, we employ ~~the~~ CCLM to refine future climate projections. We present high-resolution maps of heavy precipitation changes based on CCLM and compare them with the CMIP6 ~~GCMs~~ GCM ensemble. Our analysis ~~shows~~ indicates a significant increase in

heavy precipitation the intensity and frequency of heavy precipitation events over CA areas that are already at risk of extreme climatic events in the present dayby the end of the century. Finally, we train a convolutional neural network (CNN) to map a GCM simulation to its dynamically downscaled CCLM . We show that the CNN could emulate counterpart. The CNN successfully emulates the GCM-CCLM model chain over large CA areas. This emulator has added values , demonstrating added value when applied to a new GCM-CCLM model chain. Scientific communities The scientific community interested in downscaling CMIP6 models could use our downscaling data. The CNN architecture can be applied as , 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). Regional Researchers conducting regional studies on vulnerability, impacts, and adaptation necessitate typically achieve reliable high-resolution climate projections , which are typically achieved through dynamical downscaling via Regional Climate Models (RCMs) RCMs (Rummukainen, 2010; Feser et al., 2011), statistical techniques (Maraun and Widmann, 2018; Fowler et al., 2007)(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)(Maraun et al., 2015; Meredith et al., 2018; Laflamm
.

Central Asia (CA), recognized CA, recognised as one of the most vulnerable regions to climate change impacts, is heavily dependent 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)(Reyer et al., 2017; Fallah et
. The area faces significant challenges to food security, characterized 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).

Despite these critical concerns, the development of high-resolution climate projections in CA is impeded by the significant Significant uncertainties inherent in the existing high-resolution detailed observational and reanalysis datasets (Fallah et al., 2016a) . 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). Dynamical downscaling This method enhances the resolution of a driving GCM and produces a robust, physically consistent regional state of the climate. High-resolution atmospheric models have been shown to have better skills over complex topographies in estimating variables like precipitation than in situ observations, satellite-derived

and radar datasets (Lundquist et al., 2019). Many studies confirm that RCMs can better represent small-scale atmospheric features, especially for precipitation over complex topographies (Ban et al., 2015; Wang et al., 2013; Frei et al., 2003). This method is often preferred over statistical downscaling approaches, which assume that present statistical relationships will hold in the future (Hess et al., 2022). However, RCMs are computationally demanding and inherit a 'cascade of uncertainty', meaning that the uncertainties in the models will expand from one step or chain to another, highly affecting RCM outcomes and must be considered prior to performing climate projections (Mitchell and Hulme, 1999; Sørland et al., 2018). Despite these considerations, the added value of RCMs concerning their driving GCM is constantly debated in the community and is highly dependent on the driving GCM (Jacob et al., 2012; Lenz et al., 2017; Fotso-Nguemo et al., 2017; Di Luca et al., 2012, 2015). An RCM is tuned to perform over the target local region. However, a GCM is tuned to represent energy and water balance globally (Sørland et al., 2018). 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 models' inter-comparability of RCMs. CORDEX aims to create a better robust framework for producing climate projections at a regional scale that is suitable for impact evaluation and adaptation planning globally, aligned . 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 industrialized industrialised countries (Allan et al., 2021; Taylor et al., 2012). No simulation (except this study) driven by the CMIP6 model simulations has been planned so far for CORDEX-CA (see , last visited on 17.04.2024). Sadly, developing countries, Developing regions, including CA, bear the brunt of global warming's consequences, with yet they have access to only a limited number of CORDEX model simulations available for this region (Naddaf, 2022). The dynamical downscaling in CA can provide detailed insights into regional climate phenomena often not captured by coarser-resolution global models (Russo et al., 2019). Climate projections might be sensible to different parameter settings, emphasizing the need for careful calibration and validation of regional models. Dynamical downscaling's added value lies in its ability to tailor climate projections more closely to regional specifics, thereby improving the utility of climate data for regional climate change impact assessments (Russo et al., 2020). 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)(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 CNNsas the most popular choice, offer promising and potentially transformative 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 non-linear nonlinear relationships between inputs and outputs.

CNNs can recognize CNNs can recognise and encode spatial hierarchies in data (Zhu et al., 2017), making them exceptionally suitable for analysing geospatial data, which is fundamental a critical component in climate modelling. Unlike traditional

statistical methods that often require manual selection and careful engineering of features, ~~CNN automatically learns~~ CNNs automatically learn the most predictive features directly from the data (Reichstein et al., 2019). ~~CNNs can model complex non-linear relationships between input data and outputs, often present in climate data due to intricate interactions in weather systems. CNNs~~ They are generally more straightforward and efficient ~~for tasks that aim~~ than traditional statistical downscaling methods for tasks aiming to predict or classify ~~based on~~ patterns distributed across ~~the spatial domain~~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 ~~constraints or~~ laws directly into the training process~~. The constraining is done by changing~~ , such as mass, energy, or momentum conservation. This integration is achieved by modifying the loss function or the network's architecture to ~~enforce compliance with physical laws(i.e., conservation of mass, energy, or momentum). Unconstrained CNNs operate without explicitly incorporating~~ ensure compliance with these laws. In contrast, unconstrained CNNs do not explicitly incorporate physical laws or constraints~~into the network's architecture or loss functions. They focus~~ . Instead, they rely solely on learning from the input data~~to the~~ , generating output predictions based on the ~~data-driven patterns they detect.~~ patterns detected in the data.

This study explores unconstrained and constrained CNN approaches to understand their effectiveness in downscaling and ~~how they perform~~ their performance when applied to GCMs ~~on which they were not initially trained~~not initially used for training.

The research questions guiding this study are:

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

~~The manuscript will focus on three~~ This article focuses on two main topics: ~~1-added~~ 1) the added value of CCLM for ~~the representation of precipitation over CA, 2-dynamical downscaling signal of CCLM for heavy precipitation and 3-training~~ representing precipitation over Central Asia, and 2) training a CCLM emulator using a CNN. We present data and methods in ~~section~~ Section 2. ~~The~~ Sections 3 and 4 introduce the results of dynamical and hybrid downscaling~~are introduced in section 3 and 4~~, respectively. Finally, we discuss the results and ~~draw conclusions in section~~ 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. ~~In the following we will explain it in more details.~~

2.1 Employed Models and Experimental Setups

2.1.1 Regional Climate Model (RCM)

In ~~our~~ this study, we conduct ~~a series of simulations with the COnsortium for Small scale Modelling in CLimate Mode (CCLM) RCM. CCLM is a~~ simulations using the CCLM regional climate model~~developed~~. 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), ~~widely used~~ which is widely utilised for short-term weather forecasting. ~~The original core of COSMO-CLM or CCLM, was called Local MOdel (LM), developed by DWD for weather forecasting. The adopted LM version for climate purposes formed the CCLM (Böhm et al., 2003). CCLM is designed to simulate the regional climate at high spatial resolution, allowing researchers to study~~ Explicitly designed for regional climate simulation, CCLM enables researchers to investigate various aspects of the climate system, ~~such as~~ including temperature, precipitation, and extreme events. ~~CCLM has been utilized in numerous studies to evaluate~~ It has been extensively used to assess the impact of climate change ~~on various regions , including~~ 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). ~~It has also been used for~~ Additionally, CCLM has been employed in climate projection studies ~~and to assess the effectiveness of~~ to evaluate climate adaptation and mitigation strategies. The model has ~~been thoroughly evaluated and validated (Fallah et al., 2016b; Russo et al., 2019; Kjellström et al., 2011). Its ability to produce~~ 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 ~~the current climate and its variability has made it~~ 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 ~~used a similar model set-up as~~ utilised a model setup similar to the "optimal" ~~set-up provided in the studyof Russo et al. (2019) . The CORDEX protocolrequires a set of simulations that can be~~ 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 ~~main groups~~primary phases. The first ~~one, referred to as~~ phase, the evaluation run, ~~consists of~~ involves a single model experiment ~~performed~~ over the period 1979-2014, using ERAInterim reanalysis data at a spatial resolution of T255 ($\sim 0.7°$)~~as the driving data. In the second stream (projection), the models must runwith~~ . The second phase, the projection run, utilises boundary conditions from GCMs of the CMIP6 project for the period 1950-2100 under ~~different SSPs(here, we have chosen a single GCM:~~ various SSPs. For this study, we selected the MPI-ESM1-2-HR ~~and SSP126, SSP370 and SSP585 scenarios). SSPs are baseline scenarios describing the future development pathways depending on population , technology and economic~~ 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, ~~urbanization, investment in healthcareand~~ technological advancement, economic development, urbanisation, and investments in healthcare, education, land use, and energy (Riahi et al., 2017).

~~We have chosen the two available CORDEX-CA evaluation simulations from other models, driven by ERAInterim at 0.22° horizontal resolution, for comparison/evaluation of our RCM simulations, which are driven by ERAInterim for the evaluation period. The two simulations are 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~~

~~We create an emulator of CCLM using a CNN. We use the output of the CCLM Version 6.0 RCM, which is driven by the MPI-ESM1-2-HR GCM under four different scenarios (for 2019-2100). Historical is based on the data of~~ Historical data for this study are based on greenhouse gas levels, land use, and other climate forcings observed from 1850 to ~~2014 that were observed. SSP126~~ 2014. The Shared Socioeconomic Pathway (SSP) scenarios used in the projections are as follows:

- **SSP1-2.6** represents a "green" future~~where global resources are protected,~~ , characterised by global efforts to protect resources, improve human well-being~~is improved, and income gaps are narrowed~~, and narrow income gaps. This scenario ~~has~~ assumes low challenges to adaptation and low greenhouse gas emissions. ~~Challenges to adaptation~~ Adaptation challenges in this context refer to the ~~degree of difficulty that~~ difficulties societies might face in adjusting to the ~~environmental, economic, and social~~ impacts of climate change~~. Specifically, this term refers to a society's fundamental~~ , including their susceptibility and the ~~accessibility and efficacy of technologies and approaches designed to lessen the impacts of climate change. The adaptation challenges are minimal in the SSP126 scenario, which envisions a sustainable future. This implies that, under this scenario,~~ availability and effectiveness of mitigation technologies and strategies. Under SSP1-2.6, global cooperation and sustainable practices lead to advancements in technology and governance~~that significantly reduce~~ , significantly reducing vulnerability to climate change impacts. ~~Additionally, societal~~ Societal structures are resilient, and resources are managed to minimise environmental stresses ~~and maximise~~ while maximising human well-being. ~~SSP370 depicts a regional rivalry future~~

- **SSP3-7.0** depicts a future characterised by regional rivalry, where nationalism and regional conflicts ~~prevail~~dominate, global issues are ~~ignored~~neglected, and inequality ~~is increasing~~increases. This scenario ~~has~~ involves high challenges to adaptation and high greenhouse gas emissions. ~~SSP585 portrays a~~

- **SSP5-8.5** represents a future of fossil-fueled development ~~future where global markets are connected~~ with globally connected markets, rapid technological progress, ~~technological progress is fast, but environmental policiesare weak~~and weak environmental policies. This scenario has low challenges to adaptation ~~and~~ but results in very high greenhouse gas emissions. ~~As an additional dataset, we merge the~~

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 ~~and CCLM simulation~~ data and corresponding CCLM simulations driven by it (ERAInterim-CCLM) ~~to our previous data pool of GCM and RCM~~ (see Fig. 1). ~~We then~~

We train our CNN model based on the architecture proposed by Harder et al. (2023), which ~~can incorporate~~ incorporates physical constraints to ensure mass conservation and energy balance. ~~We evaluate our model in the CA domain. We have to note that we use not the whole GCM domain as input for the CNN but only the domain covering the CA (Fig. 3).~~ 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'~~ "perfect model" approach involves starting with high-resolution data ~~, which is considered accurate or nearly perfect, and intentionally degrading~~ 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'~~ "truth" (the original high-resolution data) is known ~~,~~ and then to recover this high-resolution from the artificially ~~degraded datausing deep learning techniques~~upscaled

data. This approach ~~is a crucial part of training, as it~~ teaches the model the desired mapping from low to high resolution, enabling the model to effectively learn how to ~~upscale~~ downscale or enhance resolution while ~~minimizing~~ minimising the loss of critical information. It ~~'s~~ 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: ~~This~~ may capture global or large-scale phenomena but miss regional details (Xu et al., 2021; Chokkavarapu and Mandla, 2019).
- High-resolution data: ~~This~~ 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, ~~the challenge for deep learning is to learn a mapping between these~~ CNN's challenge is learning to map between two independently imperfect ~~data sets. With using the CNN we try to train a model that can~~ datasets. The CNN is trained to predict high-resolution details from low-resolution inputs as accurately as possible despite the absence of ~~a~~ perfect ground truth. This process involves understanding and ~~modeling~~ modelling the uncertainties and biases inherent in both datasets.

~~Many regions of CA receive low precipitation throughout the year and the spatio-temporal variability of precipitation is large. One needs a large dataset of GCM output and the corresponding RCM with various precipitation patterns for training a CNN to find an RCM emulator that captures the mapping from GCM to RCM.~~

~~First, the daily datasets are shuffled randomly. We then have used a total number of 68141~~ 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,~~22714~~ 714 days (20~~%) and 22714~~ %) for validation, and 22,714 days (20~~%) RCM simulation days for training, validation and testing, respectively~~%) for testing. The low-resolution (GCM) ~~and~~ dataset consists of 30 × 60 grid points, and the high-resolution (RCM) ~~datasets have 30 × 60 and~~ dataset comprises 120 × 240 grid points over latitudes and longitudes, respectively~~. Therefore, the~~ , resulting in a downscaling factor (N) ~~is 4 in this case. For a complete explanation of the CNN architecture, we refer to the work of Harder et al. (2023) and the corresponding Zenodo repository at (last visited on 21st of June 2023).~~ of 4.

~~Figure 2 shows the schematic of the standard CNN (without constraint layers) architecture used for two times up-sampling in this study. We briefly explain the steps shown in the schematic:~~

- ~~Conv (Convolutional Layer): Initially, these layers help in extracting various levels of features from the low-resolution images, such as edges, textures, and other relevant image details.~~

- ~~ReLU (Rectified Linear Activation Unit): This non-linear activation function is a key player in our model's learning process. It introduces non-linearity, outputting the input directly if it's positive; otherwise, it outputs zero. This intriguing function helps the network learn complex patterns efficiently.~~

- ~~TransConv (Transposed Convolutional Layer): This layer is crucial for the task of upscaling. It increases the spatial dimensions of the feature maps, performing a sort of learned interpolation. This reassures us about the model's ability to add details to the upscaled images based on the features extracted and processed in the earlier layers.~~

- ~~ResBlock (Residual Block): They allow the model to learn corrections (or residuals) to the primary interpolated outputs, refining the details and adding high-frequency information that enhances the perceptual quality of the upscaled images. Adding original input features (from earlier layers) to the output of several convolutional layers ensures that no critical information is lost during processing.~~

2.1.3 Constraint layers

We test the CNN with three different constraining methods in the last CNN layer ~~: 1-~~ (Harder et al., 2023): 1) soft constraining (SCL), ~~2-~~ 2) hard constraining (HCL) and ~~3-~~ 3) without constraining (NoCL). ~~For a detailed information on the settings used we refer to the work of Harder et al. (2023).~~ In the following, we explain briefly the three different constraining methodologies. The ~~set-up~~ setup of constraining is as ~~following~~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: ~~it~~ uses the SoftMax~~constraining, which is a constraining for~~ , which constrains quantities like water content ~~. It enforces~~ by enforcing the output to be non-negative. ~~For constraining the predicted quantities, we use a SoftMax operator~~ 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 ~~neural networks~~ CNN before the constraining layer ($\tilde{y}_i$) and ~~multiply it with~~ 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 ~~regularization~~ regularisation term to the loss function. The MAE loss is then extended with an additional constraint violation (CV) loss term~~to~~:

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

~~where~~ Where CV is the ~~constraint violation, which 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 ~~the~~ $\alpha = 0.99$~~in this study~~.

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

~~The constraint layers are applied at the end of the CNN architecture, and all satisfy the criteria that the resulting high-resolution patch conserves the values in low-resolution pixels. The performance of the different settings is assessed through the MAE.~~

~~We use the mean absolute error (MAE) as the loss function. We use~~ We use 160 epochs, ~~with~~ 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~~ Ti-Lite Hash Rate] graphics processing unit (GPU). We use the same model ~~set-up~~ setup as in Harder et al. (2023)~~, and the computational cost of the CNN is very high, therefore, we did not conduct any cross-validation in this study~~.

~~We must~~ It is important to note that the MAE can ~~be used as both~~ serve both as a loss function and an evaluation metric. ~~A loss function~~As a loss function, it is used during training to ~~optimize~~ optimise the neural network~~parameters, while~~ 's parameters. Conversely, when used as an evaluation metric, it is calculated on the validation or test data ~~set to evaluate the modelon~~ sets to assess the model's performance using an independent dataset. ~~Those are two different use cases, but both can use an MAE~~ 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 ~~significantly influences the added value calculation~~ can significantly influence the perceived added value of an RCM, ~~as well as the extreme eventsdetection. To reduce these issues, they recommended to use observations with a resolution~~ 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 ~~one of the model. Therefore, for assessing the added value of CCLM with respect to the driving GCM~~model's for enhanced accuracy. In line with this, we use ~~the Climate Hazards Group InfraRed Precipitation with Station data (CHIRPS~~

) CHIRPS as our gridded observation . CHIRPS has to assess the added value of the CCLM driven by the GCM. CHIRPS provides a resolution of 0.05°and covers the area between , covers latitudes from 50°S-50S to 50°N. CHIRPS is based on , and offers independent observations derived from satellite information and station data, and, in contrast to . This contrasts with reanalysis data, it is independent of which depend on climate model simulations . Therefore, CHIRPS could be an excellent alternative to similar but not identical coarse datasets like Global Precipitation Climatology Centre (GPCC) (Becker et al., 2013) for data-sparse regions with convective rainfall (Funk et al., 2015).

For testing the CNN methods, instead of CHIRPS, we use the corresponding CCLM simulation ((Funk et al., 2015). We allocate 20% of the data, as mentioned above) as our target . We calculate the metrics on the CNN and 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 with respect to against the target CCLM output.

This comparison assesses whether the CNN outperforms simple interpolation. The selected GCM , RCM and observational data are interpolated onto the RCM grid using the distance-weighted average method. Interpolation of the coarser grid to a higher resolution Ciarlo et al. (2021) previously noted that such interpolation might create unrealistic values. This issue was discussed in the work of Ciarlo et al. (2021). Usually, the interpolation , as it does not account for the physical processes and constraints that govern the original data, the statistical properties (like mean, variance and skewness) are not preserved, and it introduces artefacts and errors that depend on the choice of could introduce artefacts depending on the interpolation method, the spatial distribution of the data pointsdata points, and the resolution ratio. Therefore, dynamical/statistical downscaling is used to increase the resolution of the climate data, and we use simple interpolation as a baselinein our study.

Since precipitation, recognising its limitations in preserving the statistical properties of precipitation, which does not follow a normal distribution, following Hodson (2022), we use . Following (Hodson, 2022), we apply the MAE to explore the bias of the quantify the biases in emulated and dynamically downscaled precipitation (FF) against observations (OO):

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

where Where T is equal to represents the number of time steps . We quantify the over 30 years of daily data. We define added value (AV) as the ability of the downscaling approach to decrease the MAE of the driving GCMwhen calculated against the reference dataset (CHIRPS or target CCLM simulation), i.e. reduction in MAE achieved by the downscaling relative to the driving GCM:

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

where

Where MAE_{GCM} and MAE_{CCLM} are the differences ~~of~~ 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:

$$BIAS = PR_{MODEL} - PR_{OBS} \quad (7)$$

3 Results

Figure 3.a ~~shows~~ illustrates the topography of the CORDEX-CA simulation domain. Figure 1.b ~~presents~~ displays the mean daily precipitation~~averaged over all years~~ , averaged over the years 1985-2014 (mm/day)~~as~~ , derived from CHIRPS data~~for the period 1985-2014~~. The regions with the highest ~~values of~~ precipitation are the mountainous areas of CA~~. Additionally, also~~ , particularly notable in the Asian summer monsoon region north of India and along the Himalayas in the southeastern part of the domain~~present pronounced precipitation values~~ , where precipitation values are pronounced. Figure 3.c ~~shows~~ depicts the distribution of ~~the~~ WorldClim weather stations (Fick and Hijmans, 2017) ~~over CA, representing~~ across CA, serving as a proxy for the density of ~~the~~ station data used in the CHIRPS dataset. ~~Over~~ Observational data are sparsely distributed in East China, especially over the Tibetan Plateau~~, the observation data distribution is sparse. The~~ . Consequently, data-model ~~comparison is considered unreliable over~~ 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 characterize~~

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

~~Added values~~ The AVs of GERICS-REMO2015 and RMIB-UGent-ALARO-0 driven by ERAInterim are ~~shown in Figure~~ presented in Figures 4.g-l~~respectively. The CHIRPS dataset is again used~~ , using CHIRPS as the observational dataset~~O~~.

The added value of RCM is ~~the most pronounced over~~ most pronounced in areas with complex topography~~and~~ , especially during summer, ~~for all three RCMS considered~~ across all three RCMs (Figs.4.d-l). Areas where the RCM has a smaller MAE than the reanalysis ~~with respect~~ in comparison to observations are ~~located~~ found over Tajikistan, Kyrgyzstan, ~~North of Afghanistan~~northern Afghanistan, and part of the Himalayas~~. Mountain areas of Tajikistan and Kyrgyzstan are the main source of water for the~~ —regions that are crucial water sources for former Soviet Union countries. ~~However~~Nevertheless, precipitation during the colder seasons ~~might be of more importance~~ may be more critical for water availability. The annual AV patterns still show positive values ~~over those areas~~ in these regions (Figure 4.d,g, and j). ~~Considering the whole~~ Across the entire domain, all three RCMs ~~sensibly~~ significantly reduce the large and local-scale bias of ERAInterim~~against CHIRPS, especially for~~ , especially in complex topographies. The nested RCMs ~~show similar values of~~ exhibit similar MAE values near their lateral boundaries, ~~with respect~~ relative to their driving model (Figure 4, *a*,*b*,*c*). ~~Therefore~~Thus, negative AV quantities ~~might originate from the boundary effect, especially near the east~~ may result from boundary effects, particularly near the eastern and southeastern boundaries ~~, where the monsoonal precipitation is dominant~~where monsoonal precipitation dominates. GERICS-REMO2015 ~~shows~~ displays pronounced negative added values ~~for annual and~~ annually and during winter above Tibet.

~~As an additional check, we also show the bias in the climatologies of models in figures~~ Additionally, model climatology biases are displayed in Figures 5. Once again~~the biases are pronounced on the right bottom~~ , these biases are noticeable in the lower right corner of the domain during ~~the JJA and south~~ 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 ~~shown some~~ demonstrated added value for precipitation over ~~mountainous areas~~ the mountainous regions of CA, we explore climate change signals in its high-resolution output. ~~The resulting~~ These high-resolution maps ~~might have biases inherited~~ may inherit biases from the GCM-RCM selection and could vary under different anthropogenic forcings. We assume that many model biases ~~remain conserved among the~~ are consistent across different time slices and, therefore, can be removed when calculating ~~the~~ changes between the historical period (1985-2014) and future periods (2070-2099).

We present ~~the resulting~~ climate change trends in CCLM and the CMIP6 GCMs ensemble statistics (ensemble mean and standard deviation). We ~~analyzed~~ analysed 31, 33, and 38 models for SSP126, SSP370, and SSP585 scenarios, respectively, with a total ~~number of simulations of~~ of 158, 185, and 242 ~~, respectively~~ simulations (see Supplementary materials for the list of models used~~in this study). To give the same weight to individual models, we first calculate the statistics over all the members of each modeland then build~~). We calculate statistics over each model's members to ensure equal weighting for individual models before building the final statistics. We have ~~chosen the yearly 99^{th}~~ selected the yearly 99th percentile of daily precipitation (PR99~~hereafter~~), which ~~considers~~ accounts for the three days ~~of the year~~ with the highest precipitation ~~. We also~~ each year. Additionally, we chose the number of very heavy precipitation days ~~in~~ during the period (~~ECA_RX20mm) as a different index, one of several precipitation-related indices used to monitor and analyze climate variability and change. For example, this index is often~~ ECA-RX20mm) as another index, which is commonly used in climate research to assess the impacts of ~~very~~ heavy precipitation events on water resources, agriculture, and natural ecosystems (Klok and Klein Tank, 2008). ~~Figure 6 presents~~

Figure 6 shows the changes in averaged PR99 at the end of the century (2070-2099) ~~with respect~~ compared to the historical period (1985-2014) for CCLM (a,d~~and ~~,g) and CMIP6 GCMs (b,e~~and ~~,h) under different scenarios. The ~~downscaling signals indicate that those characteristics depend on the scenario and time period. The~~ large-scale patterns remain ~~the same among all three selected scenarioswith intensification when the anthropogenic influenceincreases~~consistent across all three scenarios, intensifying with increased anthropogenic influence. The standard deviation of the models' ensemble is ~~shown~~ depicted in Figures 6.c,f~~and i. According to our analysis ~~, ~~~~, i. Our analysis indicates that the Himalayas, ~~especially~~ particularly Nepal, North India, and Bhutan, ~~have~~ exhibit the highest uncertainty among the GCMs ~~and~~ in all scenarios. Except for this ~~area~~ region and the eastern boundary of the domain, the standard deviation remains ~~under~~ below 3 mm/day. Under the ~~pessimistic~~ SSP585 and ~~the regional rivalry~~ SSP370 scenarios, ~~areas with more than 9 mm/day increase in PR99 for CCLM over~~ regions including Northwest India, North Pakistan, North ~~Iran, Southwest of Iranexist and~~ and Southwest Iran, and the South and Southeast of ~~Black Sea~~the Black Sea are projected to experience increases in PR99 values exceeding 9 mm/day. A reduction ~~pattern is detected East of the Mediterranean Sea~~ in PR99 is detected in the eastern Mediterranean, specifically in Jordan, Syria, and ~~South of~~ southern Turkey. Similar patterns are ~~also~~ observed in the CMIP6 ensemble mean~~. However, due to the ~~~~, but due to averaging, the ~~GCMs'~~ ensemble mean patterns are ~~around ±~~ approximately ±5 mm/day over ~~those~~ these areas. Under the SSP126 scenario, which ~~agrees~~ is aligned with the 2°C warming target, the ~~increasing patterns of more than ±~~ previously observed increases in precipitation exceeding ±9 mm/day for CCLM and ~~±~~ ±5 mm/day for GCMs ~~disappeared~~are no longer evident. In CA, areas ~~of increased PR99 over~~ such as Kyrgyzstan, Tajikistan, ~~North of Pakistan~~northern Pakistan, and southwestern Iran are particularly vulnerable to rainfall-induced hazards, including landslides (Wang et al., 2021; Kirschbaum et al., 2010) and ~~Southwest Iran are regions with a considerable risk of rainfall-triggered events like landslides (Wang et al., 2021; Kirschbaum et al., 2010) and~~ floods (~~for example,~~ e.g., the Pakistan floods of 2010 and 2022).

Figures 7.1, d~~and g show the ECA_RX20mm~~ , and g illustrate the ECA-RX20mm values for CCLM ~~for the three scenarios~~ at the end of the century ~~. The patterns are like those shown~~ across three scenarios. The observed patterns align with those in Figure 6, ~~indicating that the number (frequency)~~ underscoring an increase in the frequency of very heavy precipitation days~~also increases with an enhanced anthropogenic influence, particularly~~ , particularly marked over the Tibetan Plateau~~. From~~ , as anthropogenic influences intensify. Similarly, Figures 7.b, e~~and h, we conclude~~ , and h reveal that the CMIP6 GCM ensemble ~~also presents a very similar behavior to CCLM. The~~ mirrors the behaviour observed in CCLM. However, the ensemble standard deviations ~~, however, increase~~ for ECA-RX20mm values rise over Tajikistan and Kyrgyzstan~~for ECA_RX20mm values (~~, as shown in Figures 7.c, f~~and i). The increased~~ , and i. The growing frequency and intensity of extreme precipitation ~~over elevated areas of CA due to anthropogenic forcing is alerting (Fallah et al., 2023). The presented CCLM simulation contributes to study the sensitivity of dynamical downscaling~~ 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 ~~at the local scale. This information might be of interest for the scientific community working on the impact of climate change in CA.~~ can vary locally.

4 CCLM emulator using a CNN

We have ~~shown that the dynamical downscaling added value to explore the local effects of climate change during the historical period, especially over areas with enhanced topographical forcings. Here~~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 ~~an emulator of CCLM~~ a CCLM emulator for precipitation over CA. As ~~explained previously~~ previously explained, a CNN ~~could be~~ trained on our GCM-RCM chain ~~and be applied~~ could serve as a fast~~and computationally cheap downscaling method. However, the skill of such a model must be explored and verified~~, cost-effective downscaling method, though its efficacy needs to be rigorously assessed.

~~Here we want to demonstrate that the emulator is better at downscaling than a~~ We aim to establish that this emulator outperforms simple interpolation, ~~especially for areas receiving extreme precipitationvalues. More specifically, our goal is~~ particularly in areas experiencing extreme precipitation. We aim to show that the CCLM emulator can ~~produce~~ replicate CCLM-like ~~patterns when fed~~ precipitation patterns when driven by the parent GCM.

~~For the CNN approach, we focus~~ Focusing on the CA domain~~covering only~~ , which encompasses the former Soviet Union countries (Kazakhstan, Kyrgyzstan, Tajikistan, Turkmenistan, and Uzbekistan)~~and not the~~ , we exclude the broader CORDEX-CA domain ~~previously~~ shown in Figure 3. This domain is the region of interest in the Green Central Asia project https://www.greencentralasia.org/en, which ~~is financed by~~ the German Foreign ~~office~~Office finances. Figure 8.a ~~shows~~ illustrates the MAE of the interpolated MPI-ESM1-2-HR, using the CCLM ~~driven by it from the test dataset as the "true" precipitation. As can be seen, CCLM produces different precipitation values compared to its driving GCM, especially over regions~~ output as the 'true' precipitation. CCLM generates distinct precipitation patterns, particularly in areas with complex topography. ~~Here, we assume that the CCLM is~~ Assuming CCLM as the ground truth~~and check if~~ , we examine whether the CNN can ~~produce it~~ replicate these outputs using the GCM as input~~data. To evaluate the performance of the emulator, we show the maps of added value~~ . To assess the emulator's effectiveness, we present added value maps (relative to the parent GCM) in Figures 8.b-d. ~~Comparison~~ A comparison of MAE reduction maps ~~shows~~ reveals that the unconstrained CNN ~~produces significant skills~~ demonstrates significant skill over elevated regions of CA~~and the constrained runs do not present considrable patterns of~~ , whereas constrained runs show less noticeable pattern changes. For ~~example, there are areas of~~ instance, the HCL and SCL emulators generate closely mingled negative and positive added values ~~remarkably close together over elevated areasof CA created by HCL and SCL emulators. NoCL, in contrast, shows systematic positive values over large parts of~~ across elevated areas, while NoCL consistently exhibits positive values across the domain. ~~There are several artifacts~~ Several artefacts in the MAE reduction maps of constrained models, ~~especially over North of India, which represent the GCM grid shape. We produce the~~ particularly over northern India, reflect the shape of the GCM grid. We also produce boxplots of daily precipitation ~~over~~ for the CA domain ~~covering the former soviet union to explore the improvement in the distributions~~ to explore distribution improvements (Figure 9). ~~The correlation coefficients between the~~ Correlation coefficients between time-series ~~of average precipitation over the domain with respect to CCLM are also~~ averages of precipitation across the domain and CCLM are presented in Figure 9 (values in ~~the~~ parentheses). ~~For the~~ Among daily averages, NoCL ~~presents~~ achieves the best performance

(highest correlation coefficient). However, the values of outliers are smaller than the ones from CCLM and all , although it records fewer outliers than CCLM and other model simulations. The distribution is more condensed concentrated around the median(smallest interquartile range), exhibiting the narrowest interquartile range. The distribution profiles of both constrained models (HCL, SCL) is like resemble those of the interpolated GCMone. This was expected, since the constraining conserves the mass of high-resolution grid-boxes within the corresponding low-resolution grid-box , 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 generalization generalisation ability, i.e.the ability to create , its capacity to generate reliable predictions on a new data setnew datasets. We conduct a new 15-year dynamical simulation with CCLMusing CCLM, driven by the EC-Earth3-Veg (Döscher et al., 2022) GCM under ssp370 the SSP370 scenario from 2019 to 2033. We use this data This data serves as input to our CCLM emulator, which was previously trained to emulate CCLM outputs using MPI-ESMI-2 HRas input GCM. 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, i.e., the CCLM simulation driven by EC-Earth3-Veg. The . Remarkably, the MAE pattern of EC-Earth3-Veg is remarkably like the one from closely mirrors that of MPI-ESM1-2-HR (Figure 8.a). However, the CCLM emulator based on the NoCL CNN model does not NoCL emulator does not uniformly show positive error reduction everywhere in across the domain (Figure 10.b). We chose the NoCL CNN because it showed the best NoCL for its superior performance among the constrained ones. Training the CNN on the three CNNs. The emulator attempts to establish relationships between MPI-ESM1-2-HR /CCLM might have ignored learning processes which overcome considerable biases in the driving GCM. The CCLM emulator tries to find relations between the MPI-ESM1-2-HR and CCLM, which might may be specific to these two models and there is no guarantee that those relationships also models and might not necessarily apply to the new EC-Earth3-Veg and CCLM driven by EC-Earth3-Veg. This new GCM-RCM chain is extremely sensitive to the characteristics of the EC-Earth3-Veg model because, as we showed configuration. As demonstrated previously, the RCM state follows depends on the state of its driving GCM. We note that CCLM is driven at the lateral boundaries by the GCM values for the state variables of CCLM state variables (temperature, pressure, wind speed, etc.) . Precipitation is not used for driving the RCM. The CNNinput is the GCM precipitation , which has different biases in the two GCM, and therefore the mapping from the MPI-ESM1-2-HR-precipitation to the CCLM precipitation cannot be successfully transferred to 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.

Knowing these limitationsDespite these challenges, the CNN model shows added values of more than demonstrates added values exceeding 1 mm/day over in regions such as the Alborz Mountains and South of the the southern Caspian Sea in the North of Iran (black rectangular northern Iran (highlighted in black rectangles in Figures 10.a and b) and some parts of Tajikistan and Kyrgyzstan. Exploring the field mean of Exploration of the daily precipitation distribution field-mean indicates that the CNN's median value and the outliers are lower than both those of the EC-Earth3-Veg and CCLM simulations (Figure

10.c). ~~Only the~~ The day-to-day correlation ~~is being improved. As mentioned before, all model~~ has improved, although all models were trained on ~~the shuffled dataset and~~ a shuffled dataset that ignored the memory in the time series~~but here the~~ . The trained NoCL model was ~~given~~ provided with unshuffled EC-EARTH3-Veg ~~to make new predictions. The correlation coefficient increases using the NoCL model~~ data for new predictions, increasing the correlation coefficient from 0.815 (EC-Earth3-Veg) to 0.844 (NoCL). Over the ~~black rectangular box~~ highlighted area in Figure 10.b, ~~the region where the NonCL model reduces the MAE, i.e.,~~ where the NoCL model reduces MAE, the distribution of precipitation converges ~~to the one from CCLM(Figure 10.d) and receives the highest amount of precipitation in Iranand supplies water~~ towards that of CCLM, encompassing the region with the highest rainfall in Iran, vital for a large portion of ~~population in the country, including the capital city Tehranwith a population of over 10 million people~~the population, including Tehran. Only the outliers larger than 20 mm/day are not reconstructed by ~~the~~ NoCL.

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

5 Discussion and conclusions

Regional climate change impact assessments require ~~high resolution~~ 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 ~~(SD) usually has limited capability to consider~~ often struggles to account for the dynamic influences of ~~the complex topography. The large-scale domain does not reflect the spatial diversity and variation of the local climate and the topography~~complex landscapes, including topography and varying surface parameters such as vegetation, soil types, and water bodies like lakes, which may affect the accuracy of ~~the~~ statistical relationships (Li et al., 2022). For ~~SD~~ statistical downscaling methods applied to precipitation, ~~the~~ observations need to contain detailed information about ~~the~~ precipitation distribution in areas with complex topography (Lundquist et al., 2019). ~~On the other hand~~

Conversely, dynamical downscaling requires massive computational time and data storage~~space. A~~ . For example, a 30-year CCLM simulation driven by ERAInterim took roughly one week to ~~finish~~ 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~~since~~ , as they are highly dependent on the driving GCMs.

In this study, we ~~contribute to the few dynamical~~ contributed to the dynamic downscaling efforts over the CORDEX-CA domain, taking a small step towards creating an RCM ensemble ~~creation~~ for CA. A single RCM simulation ~~can help~~ helps identify model biases and uncertainties that need to be addressed in future model improvements. It is essential to note that ~~relying~~

~~solely on~~ a single model run for CMIP6~~instead,~~ , instead of an RCM ensemble, may not provide ~~any~~ a comprehensive understanding of ~~the~~ potential climate change impacts on a region. Therefore, it is recommended that researchers conduct multiple simulations with different initial and boundary conditions and ~~different~~ model configurations to account for the uncertainty associated with climate projections.

In ~~a~~ the first part of the study~~we demonstrate~~ , we demonstrated the added value of RCMs (~~here we chose to use the COSMO-CLM/~~using the CCLM model) over GCMs for CA in ~~the representation of~~ representing precipitation. Our CCLM run ~~shows~~ 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 ~~reproduces extreme precipitation changing patterns like~~ reproduced extreme precipitation patterns similar to the CMIP6 ensemble mean ~~at~~ projections for the end of the century. Both the CCLM and CMIP6 ~~ensemble present elevated risk (frequency and intensity~~ ensembles indicated an increased risk (in terms of intensity and frequency) of heavy precipitation events ~~over vulnerable areas~~ in vulnerable regions of CA due to ~~different anthropogenic influences~~various human activities.

Our study evaluated the downscaling skill ~~primarily using higher resolution observations, which are critical for capturing localized climate phenomenarelevant to~~ 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 ~~emphasizes~~ 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.

~~Additionally, acknowledging the computational and memory constraints of running an RCM at high resolution, here we also show~~ We showed that a single GCM-RCM model chain ~~can~~ could be used to train a climate emulator based on a CNN model. It ~~can learn some~~ learned nonlinear and physical relationships between the coarse and fine-resolution datasets~~. This can overcome the problem~~ , addressing the issue of spatial intermittency~~seen~~ —where data points are unevenly distributed or missing across space—common in some statistical downscaling approaches (Harder et al., 2023). However, we ~~have also shown~~ also demonstrated that the CNN model ~~has limitations , namely when generalizing~~had limitations when generalising, as it did not achieve a robust error-reduction pattern when given a different GCM as input. The learning process ~~depends strongly~~ strongly depended on the GCM/CCLM relationships. More importantly, an RCM is usually forced to follow its driving GCM and ~~only on local scales can~~ can only produce extra information ~~. An application of the presented CCN could be to apply it for~~ on a local scale. The presented CNN could be applied to other experiments of the same GCM~~: One could use~~ , such as using the trained emulator for paleo-climate ~~experiment of the parent GCM to create more than 10,000 years of downscaled simulation. One can also downscale the~~ experiments or downscaling volcanic forcing experiments~~using the trained emulator. This will~~ . This would aid the paleo-climate community in conducting proxy-model comparisons at local scales. However, previous studies have shown that the CNN ~~suffers~~ suffered from the same generalisation problem ~~as~~ when applied to a new GCM, and such applications must be tested (Jouvet and Cordonnier, 2023).

~~In an effort to evaluate the model's generalization capabilities, we~~ 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 ~~was implemented~~ 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 ~~exhibits~~ exhibited an AV pattern ~~that closely mirrors~~ mirrored the dynamical downscaling results ~~obtained with the CCLM ,~~ of the CCLM driven by the same SSP370 scenario. This alignment strongly ~~supports~~ supported the notion that our CNN emulator ~~is not only capable of learning~~ could learn from its training data ~~but also proficient in generalizing~~ and generalise to new, unseen conditions. The similarity in AV patterns between the model output and the CCLM simulation ~~underscores~~ underscored the robustness and adaptability of our model, affirming its potential for broader ~~applicative scenarios~~ applications in climate modelling.

~~We note that this work is only a step to demonstrate~~ This work was an initial step in demonstrating the potential of such a hybrid approach~~, and we~~ . We encourage the community to explore different model structures and parameter combinations for further improvement. For example, our ~~few model set-ups showed that using~~ initial setups showed that a physically constrained CNN ~~set-up,~~ setup that applies a linear transformation to ~~the high-resolution image to ensure that the total~~ ensure mass or energy ~~is conserved~~ conservation between the low and high-resolution images ~~,~~ did not successfully downscale ~~the~~ precipitation. The ~~constraints might not be satisfied in the original dataset and therefore the constrained model set-up did not lead to better~~ 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. ~~Alternative~~ Future studies could test alternative machine learning models, such as generative adversarial networks (GANs), which can generate more high-frequency patterns ~~, might~~ and improve the downscaled ~~pattern, and could be tested in future studies. An additional set-up might be to provide more information to~~ output. Additionally, incorporating more information into the CNN by adding characteristics like surface height, vegetation, ~~land-cover, land-use, etc.~~ land cover, and land use as new channels within the input layer could enhance model performance.

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
```

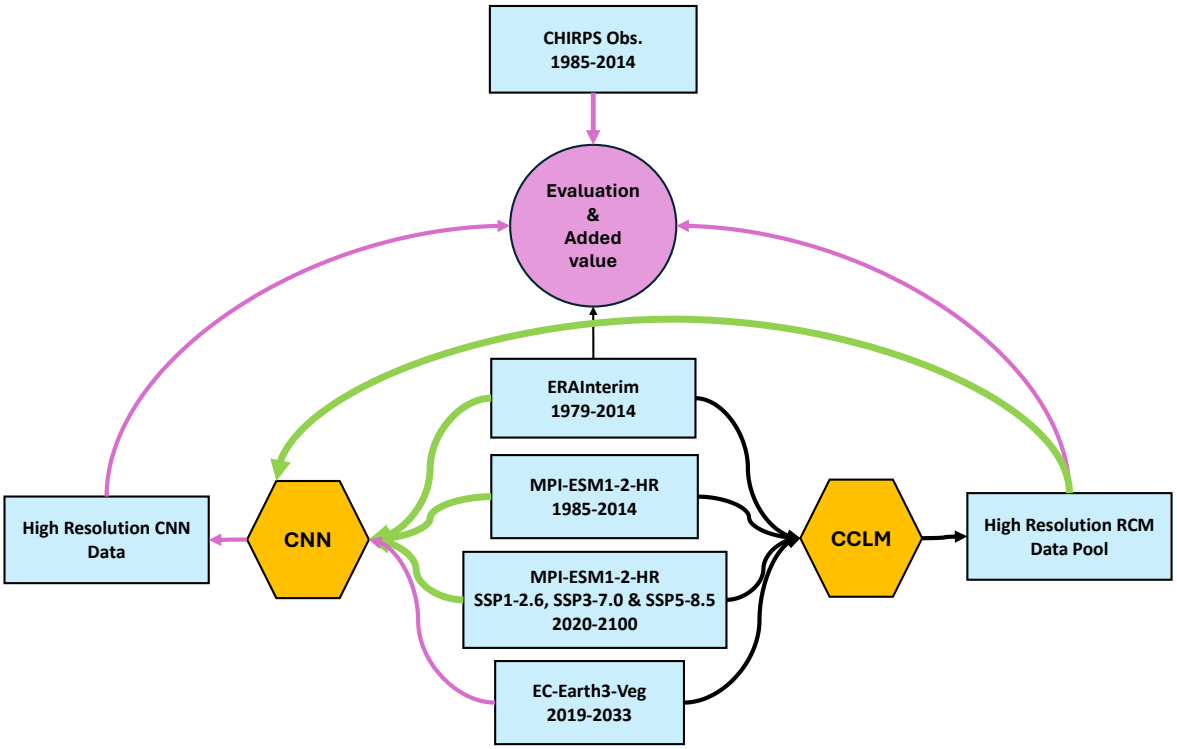


Figure 1. Schematic of the methodology used in this study. Green arrows show the data flow used for training the CNN and magenta for evaluation and calculation of the added values. Datasets are shown by rectangular, downscaling models by hexagonal and evaluation analysis by circle.

```
    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

    # 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
```

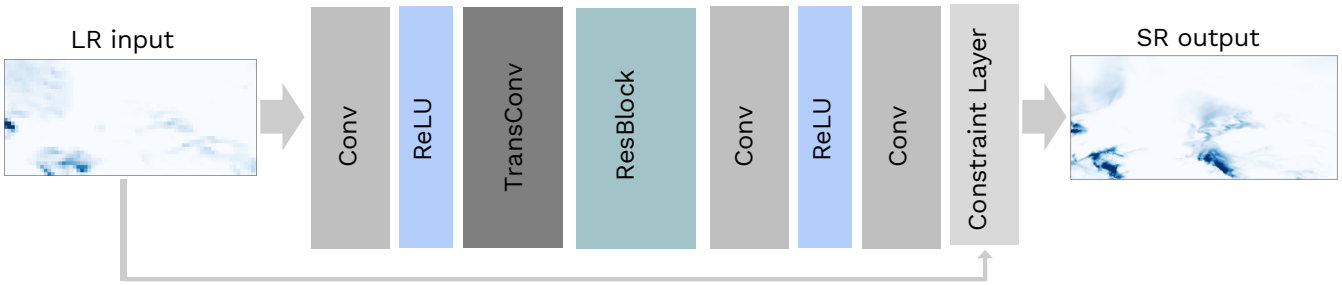


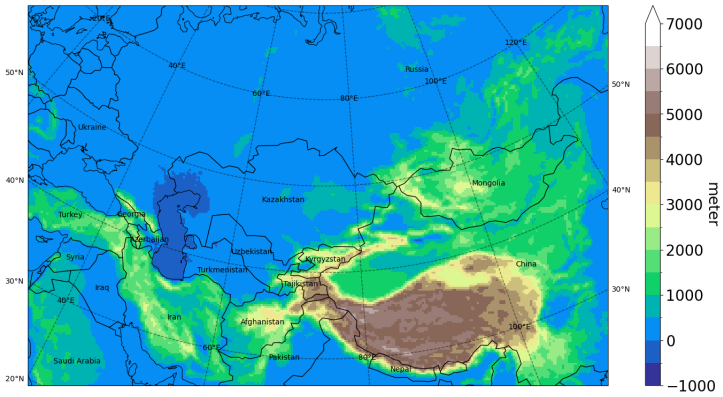
Figure 2. Schematic of the CNN architecture for 2 times upsampling with the constraints layer. The inputs are low-resolution (LR) images of size 30×60 and the output is a super-resolution (SR) image of size 60×120. This figure is modified from (Harder et al., 2023).

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

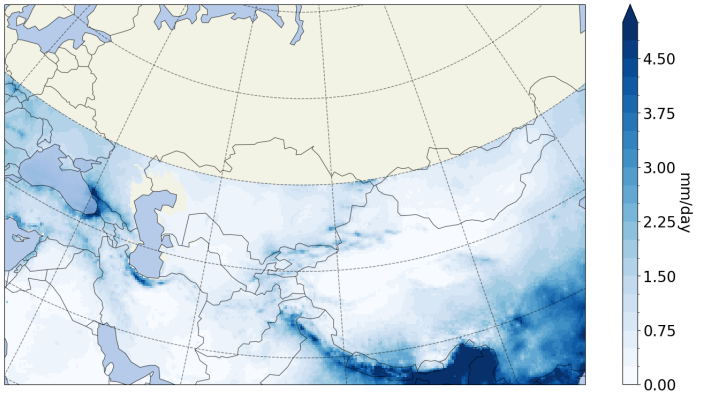
Author contributions. BF conducted the dynamical and statistical downscaling with assistance from ER and PH, respectively. ER provided the setup for the CCLM simulations. PH provided the deep learning model code and setup. All authors contributed to the analysis of the results and the writing of the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

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

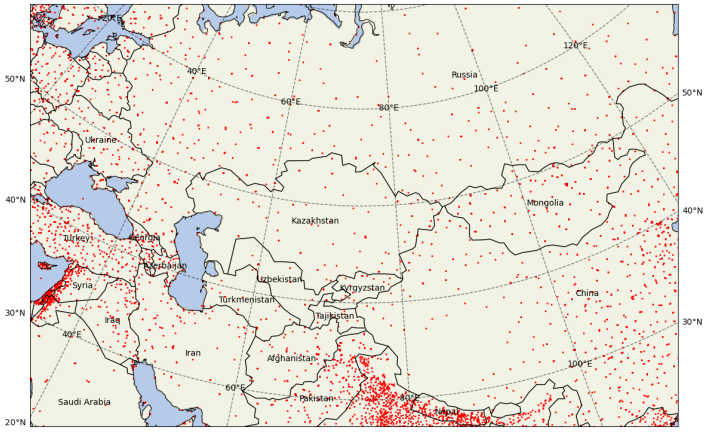
Acknowledgements. BF thanks the German Climate Computing Center (DKRZ) for its support in using supercomputer data and resources. The German Foreign Office ~~funds~~ funded BF and ID via the Green Central Asia project (http://greencentralasia.org/en, last access: 4 July 2023). The DKRZ and PIK provided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



(a)

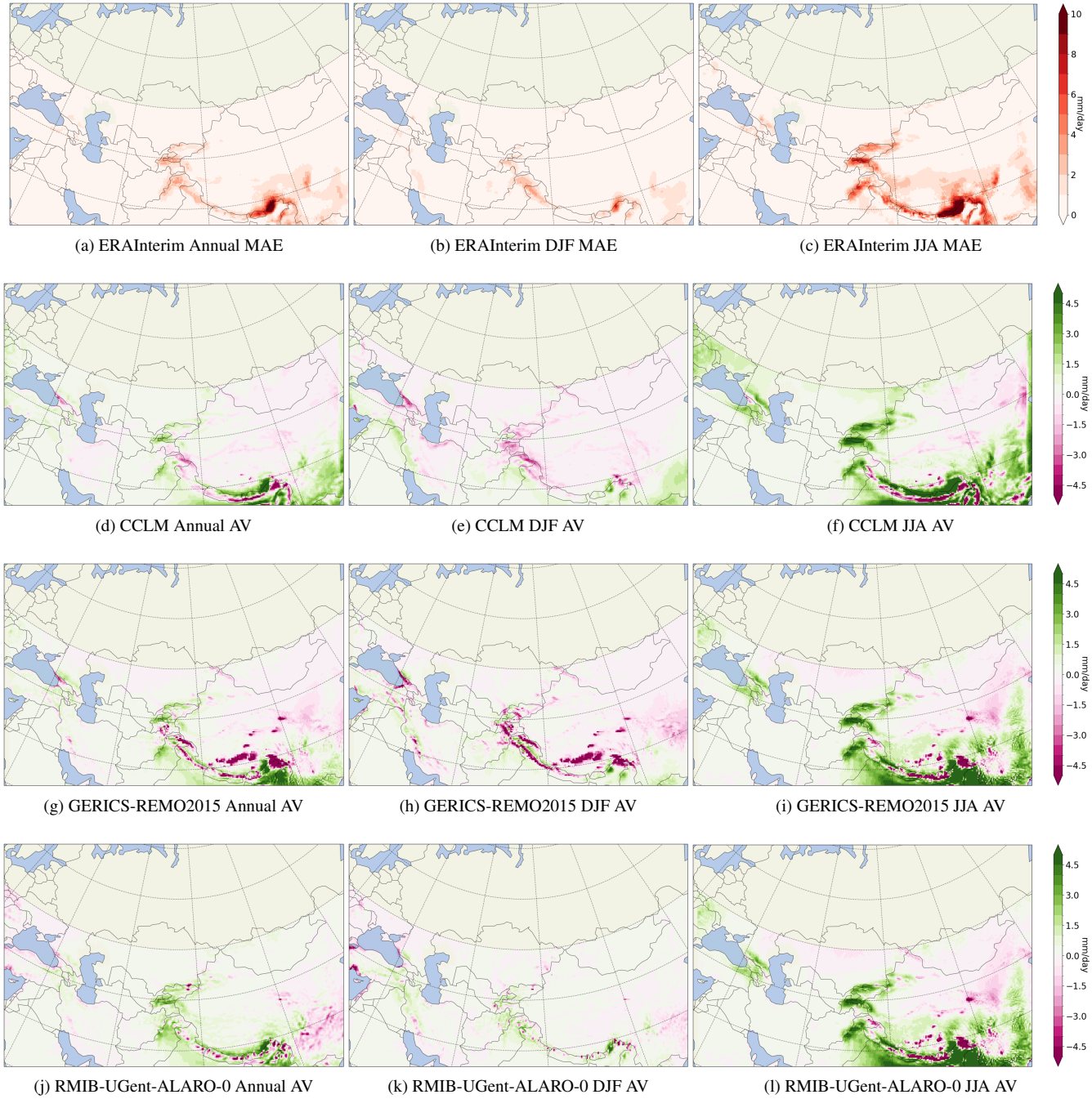


(b)



(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).



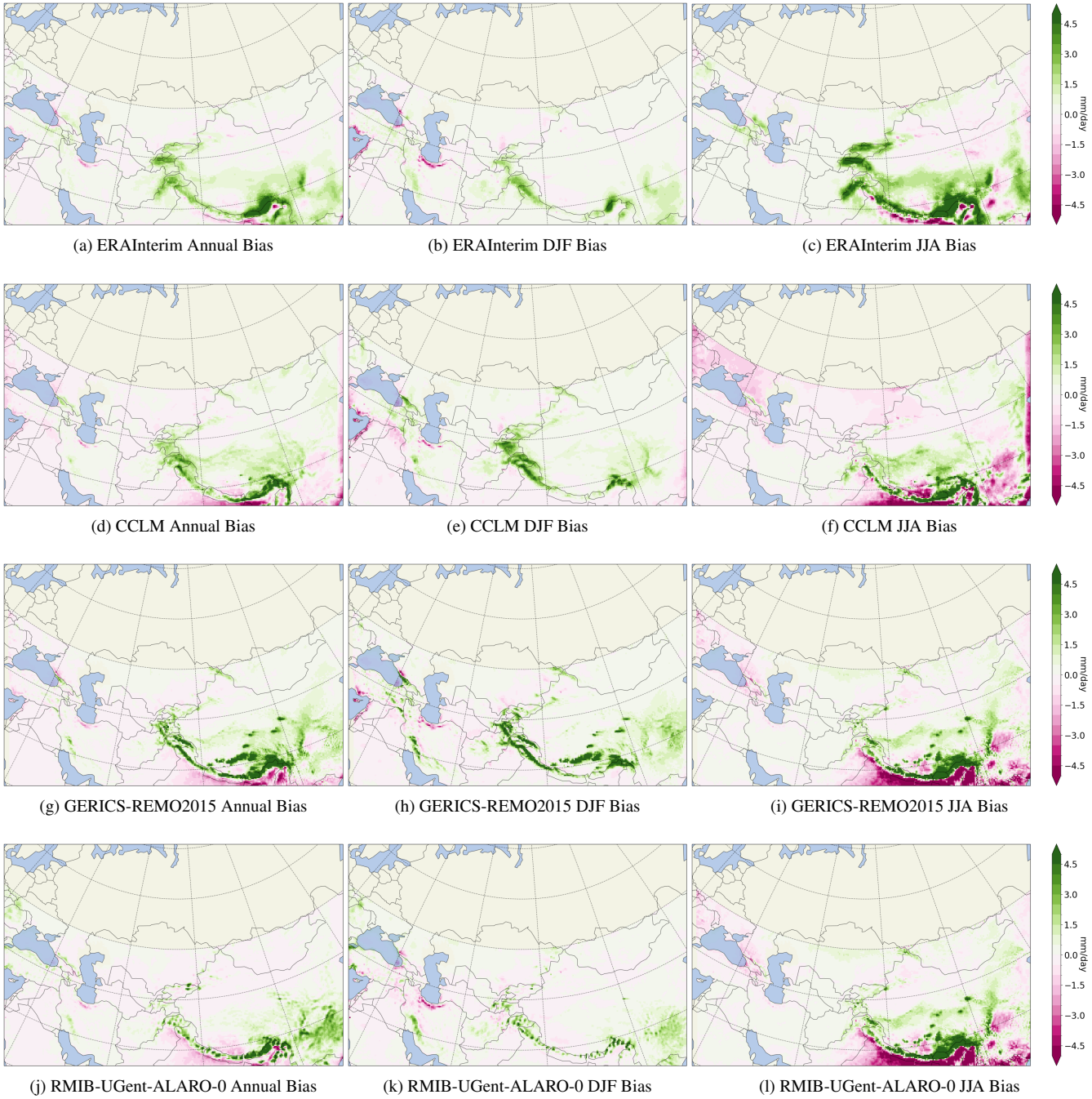
(a) ERAInterim Annual MAE (b) ERAInterim DJF MAE (c) ERAInterim JJA MAE

(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 ~~average~~ 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 ($MAE_{ERAInterim} - MAE_{RCM}$) 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 ~~daatasets~~ datasets are interpolated to the CCLM grid.



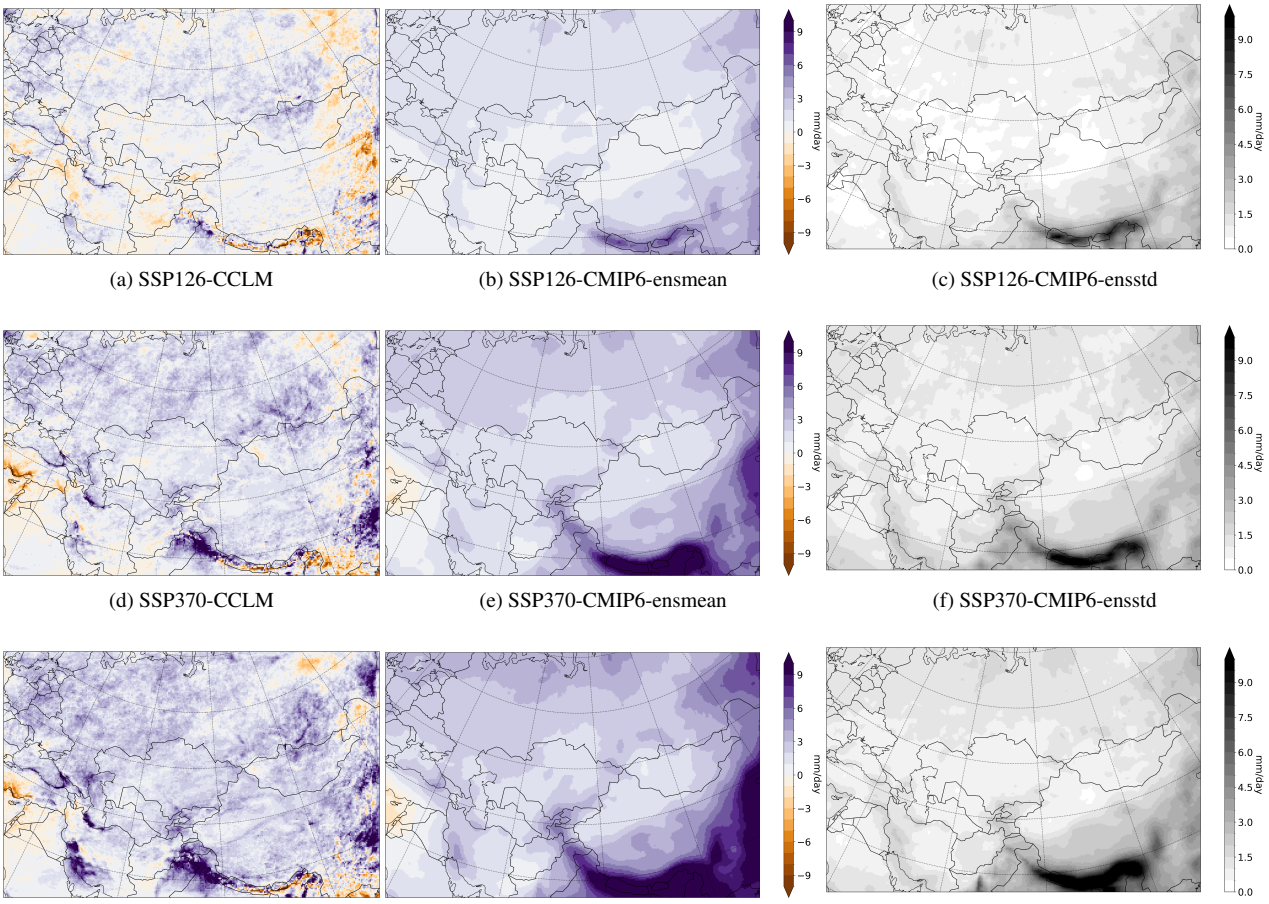
(a) ERAInterim Annual Bias (b) ERAInterim DJF Bias (c) ERAInterim JJA Bias

(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 ($PR_{ERAInterim\text{-}CCLM} - PR_{OBS}$) 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

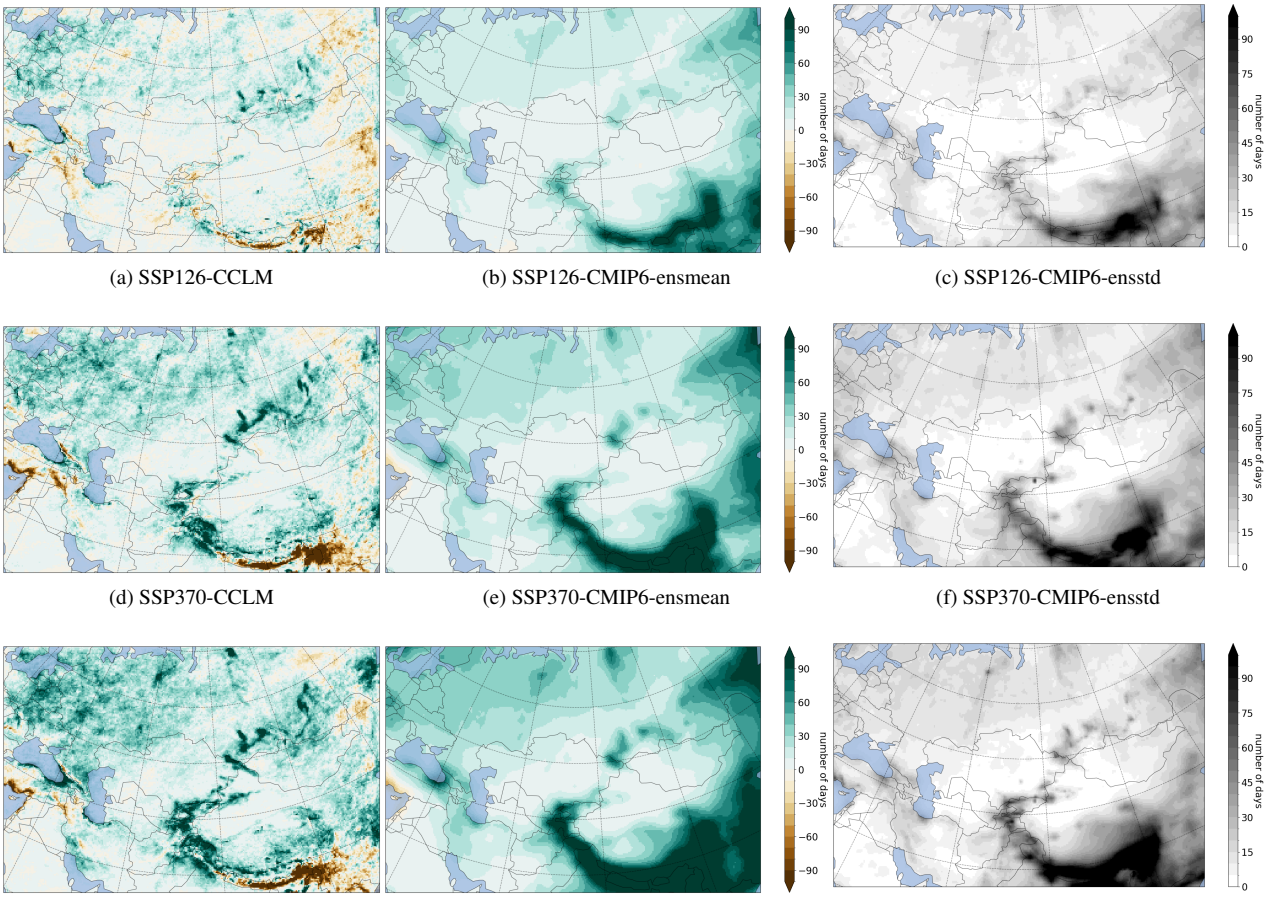
Figure 6. Changes in averaged yearly 99^{th} percentile (3 days per year) of total precipitation (mm/day) with respect to 1985-2014 references for a,b) SSP126, d,e) SSP370 and g,h) SSP585 at the end of the century (2070-2099) from CCLM and CMIP6 GCMs' ensemble mean. The ensemble's standard deviations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

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.

Becker, A., Finger, P., Meyer-Christoffer, A., Rudolf, B., Schamm, K., Schneider, U., and Ziese, M.: A description of the global land-surface precipitation data products of the Global Precipitation Climatology Centre with sample applications including centennial (trend) analysis from 1901–present, Earth System Science Data, 5, 71–99, 2013.



(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

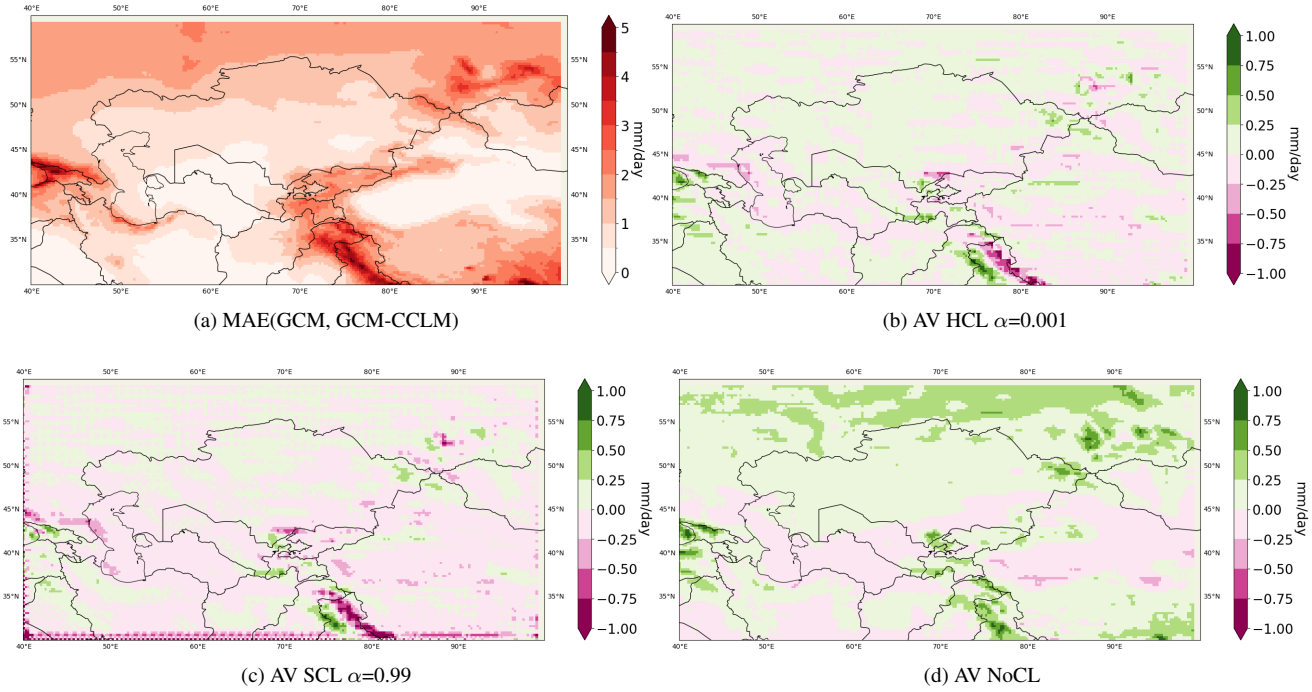
Figure 7. Changes in number of days with precipitation more than 20 mm in the period with respect to 1985-2014 references for a,b) SSP126, d,e) SSP370 and g,h) SSP585 at the end of the century (2070-2099) from CCLM and CMIP6 GCMs' ensemble mean. The ensemble's standard deviations are shown in c,f and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

Böhm, U., Gerstengarbe, F.-W., Hauffe, D., Kücken, M., Österle, H., and Werner, P. C.: Dynamic regional climate modeling and sensitivity experiments for the northeast of Brazil, Global Change and Regional Impacts: Water Availability and Vulnerability of Ecosystems and Society in the Semiarid Northeast of Brazil, pp. 153–170, 2003.

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.



(a) MAE(GCM, GCM-CCLM)

(b) AV HCL α=0.001

(c) AV SCL α=0.99

(d) AV NoCL

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

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.

Di Luca, A., de Elía, R., and Laprise, R.: Potential for added value in precipitation simulated by high-resolution nested regional climate models and observations, Climate dynamics, 38, 1229–1247, 2012.

Di Luca, A., de Elía, R., and Laprise, R.: Challenges in the quest for added value of regional climate dynamical downscaling, Current Climate Change Reports, 1, 10–21, 2015.

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.

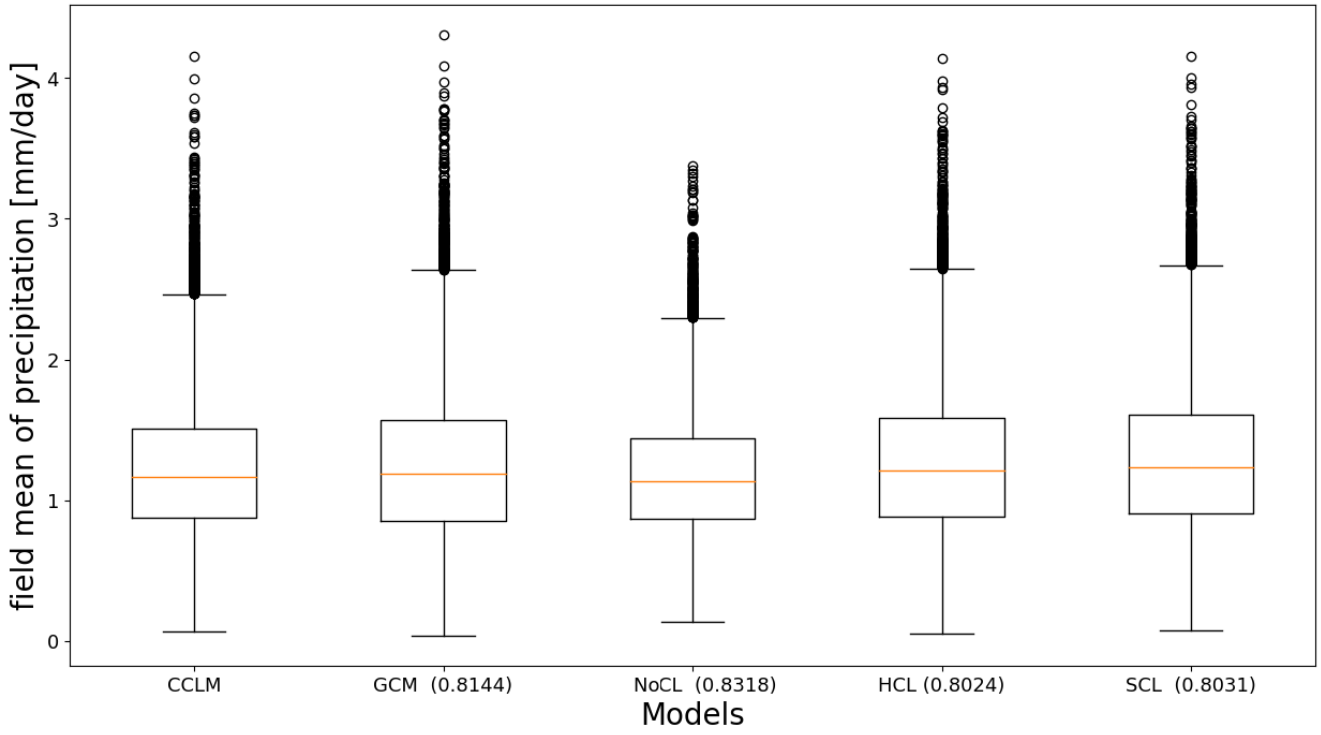


Figure 9. Boxplot of averaged daily precipitation over the Central Asian domain (shown in Figure 7) for different models and test dataset (22714 days or 62.2 years). Numbers in the parenthesis indicate the correlation coefficients between each model and the CCLM simulation.

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.

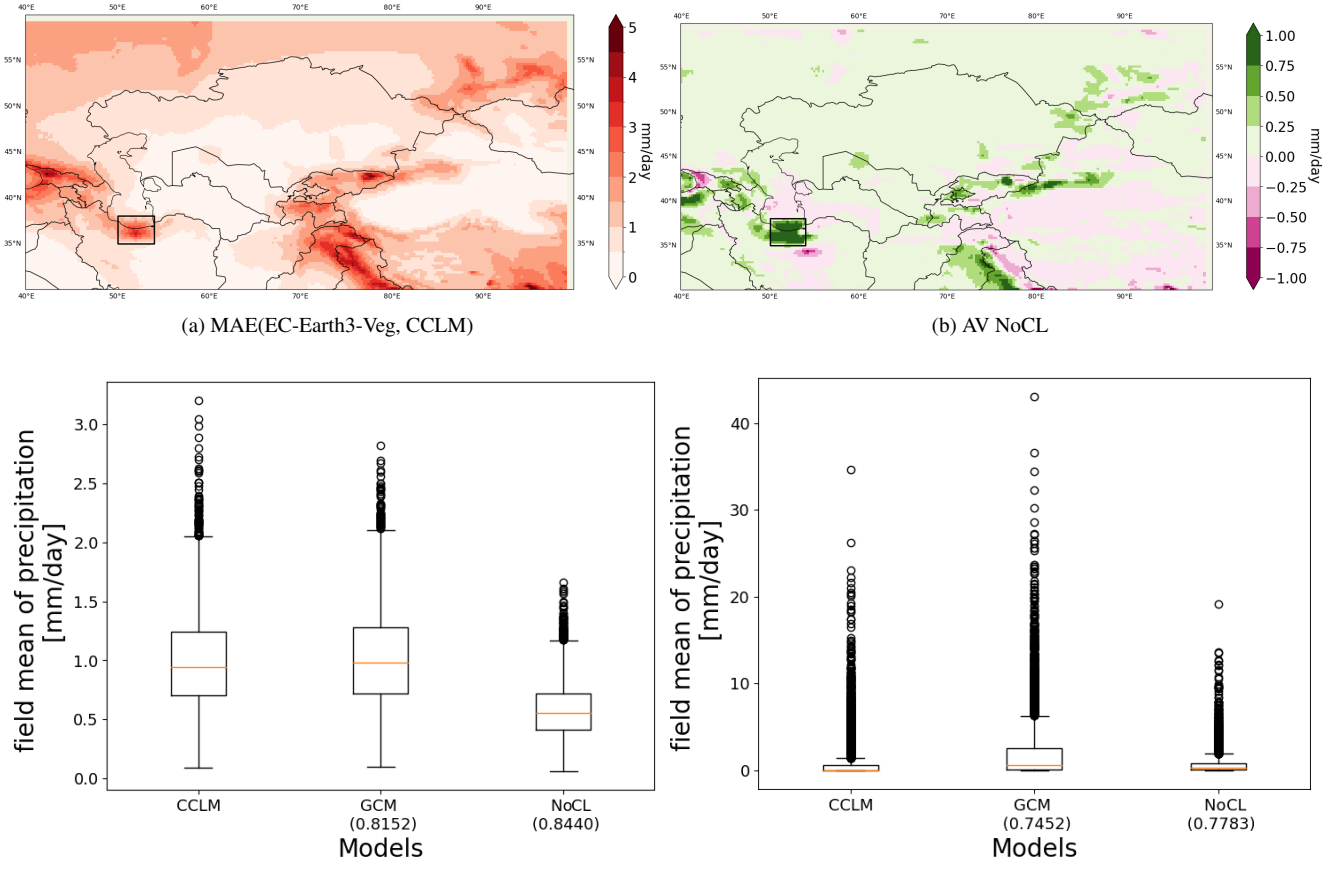
Fallah, B., Russo, E., Menz, C., Hoffmann, P., Didovets, I., and Hattermann, F. F.: Anthropogenic influence on extreme temperature and precipitation in Central Asia, Scientific Reports, 13, 6854, 2023.

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.



(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.

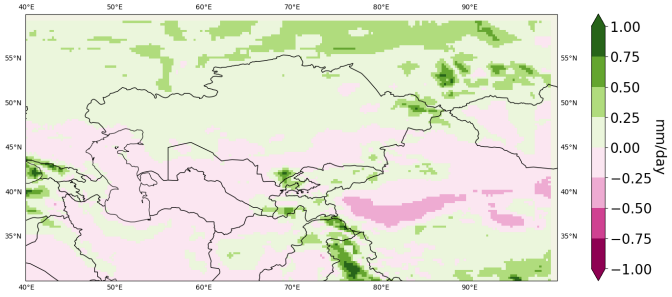


Figure 11. Added value (AV) or MAE reduction (MAE(EC-MPI-ESM1-2HR, CCLM)—MAE(CNN, CCLM)) for an unconstrained method that was not trained but applied to the SSP370 scenario.

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.

Hess, P., Drüke, M., Petri, S., Strnad, F. M., and Boers, N.: Physically constrained generative adversarial networks for improving precipitation fields from Earth system models, Nature Machine Intelligence, 4, 828–839, 2022.

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., Elizalde, A., Haensler, A., Hagemann, S., Kumar, P., Podzun, R., Rechid, D., Remedio, A. R., Saeed, F., Sieck, K., et al.: Assessing the transferability of the regional climate model REMO to different coordinated regional climate downscaling experiment (CORDEX) regions, Atmosphere, 3, 181–199, 2012.

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.

Lenz, C.-J., Früh, B., and Adalatpanah, F. D.: Is there potential added value in COSMO–CLM forced by ERA reanalysis data?, Climate Dynamics, 49, 4061–4074, 2017.

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.

Mitchell, T. D. and Hulme, M.: Predicting regional climate change: living with uncertainty, Progress in Physical Geography, 23, 57–78, 1999.

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., Sørland, S. L., Kirchner, I., Schaap, M., Raible, C. C., and Cubasch, U.: Exploring the parameter space of the COSMO-CLM v5. 0 regional climate model for the Central Asia CORDEX domain, Geoscientific Model Development, 13, 5779–5797, 2020.

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., Schär, C., Lüthi, D., and Kjellström, E.: Bias patterns and climate change signals in GCM-RCM model chains, Environmental Research Letters, 13, 074 017, 2018.

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.