https://doi.org/10.5194/essd-2021-418
Preprint. Discussion started: 10 December 2021

Open Access Earth System Science Data Discussions

Deep-Learning-Based Harmonization and Super-Resolution of Near-Surface Air Temperature from CMIP6 Models (1850-2100)

Xikun Wei[1], Guojie Wang[1]*, Donghan Feng[1], Zheng Duan[2], Daniel Fiifi Tawia Hagan[1], Liangliang Tao[1], Lijuan Miao[1], Buda Su[1], Tong Jiang [1]

[1]School of Geographical Science, Nanjing University of Information Science and Technology (NUIST), Nanjing 210044, China

[2]Department of Physical Geography and Ecosystem Science, Lund University, 223 62 Lund, Sweden

Correspondence to: Guojie Wang (gwang@nuist.edu.cn)

Abstract. Future global temperature change would have significant effects on society and ecosystems. Earth system models (ESM) are the primary tools to explore the future climate change. However, ESMs still exist great uncertainty and often run at a coarse spatial resolution (The majority of ESMs at about 2 degree). Accurate temperature data at high spatial resolution are needed to improve our understanding of the temperature variation and for many applications. We innovatively apply the deep-learning(DL) method from the Super resolution (SR) in the computer vision to merge 31 ESMs data and the proposed method can perform data merge, bias-correction and spatial-downscaling simultaneously. The SR algorithms are designed to enhance image quality and outperform much better than the traditional methods. The CRU TS (Climate Research Unit gridded Time Series) is considered as reference data in the model training process. In order to find a suitable DL method for our work, we choose five SR methodologies made by different structures. Those models are compared based on multiple evaluation metrics (Mean square error(MSE), mean absolute error(MAE) and Pearson correlation coefficient(R)) and the optimal model is selected and used to merge the monthly historical data during 1850-1900 and monthly future scenarios data (SSP1-2.6, SSP2-4.5, SSP3-7.0, SSP5-8.5) during 2015-2100 at the high spatial resolution of 0.5 degree. Results showed that the merged data have considerably improved performance than any of the individual ESM data and the ensemble mean (EM) of all ESM data in terms of both spatial and temporal aspects. The MAE displays a great improvement and the spatial distribution of the MAE become larger and larger along the latitudes in north hemisphere, presenting like a 'tertiary class echelon' condition. The merged product also presents excellent performance when the observation data is smooth with few fluctuations in time series. Additionally, this work proves that the DL model can be transferred to deal with the data merge, bias-correction and spatial-downscaling successfully when enough training data are available. Data can be accessed at https://doi.org/10.5281/zenodo.5746632 (Wei et al., 2021).

https://doi.org/10.5194/essd-2021-418
Preprint. Discussion started: 10 December 2021

Open Access Earth System Science Data Discussions

1 Introduction

From the Sixth Assessment Report by the Intergovernmental Panel on Climate Change(IPCC), the global air temperature increased by 1.09°C during the period 2011-2020 compared to the pre-industrail period (1850-1900). Global warming causes a large number of effects on the nature environment and mankind in water cycle(such as the speed of tropical cyclone translation)(Yamaguchi et al., 2020), food safety (for example, continuing warming will result in an extra about 20% population suffer from hunger in 2050)(Hasegawa et al., 2021), extreme disaster(such as the future extreme snowfall will intensively increase especially in Asia and North American and the risk of extreme fire weather have an unexpected rise) (Quante et al., 2021;Touma et al., 2021)), social and economic development(the drought loss will arrive 65 billion per year in Europe, for example) (Naumann et al., 2021), diversity of species(the evolution of fish, for instances) (Avaria-Llautureo et al., 2021) and other fields. Besides it, the most direct influence is the decline of ice cover in the ocean. In past decades, the South Pole experience an unprecedented warming speed (more the three times than the whole earth) (Clem et al., 2020) and ice cover in Arctic gone through a decrease of average 12% every decade (Zhukov and Gushcha, 2020). The loss of mass ice will also bring an extra increase of temperature at a long time scale (Wunderling et al., 2020). If temperature have an extra increase 2C° will result in an 20cm rise of the global ocean (Jevrejeva et al., 2016). In this background, the society and scientific research groups pay a lot of attention to the global temperature changes (Choi et al., 2020). Hence, in order to improve our understanding of the spatial and temporal changes of future temperature, a more precise future temperature data at high spatial resolution is needed.

The meteorological traditional stations can provide long time series point data. Besides it, with the development of remote sensing technology, the number of satellites specially used for meteorological observation is increasing, which also brings a large number of meteorological satellite observation data. The satellite can provide observed raster data with relatively short time series, those data cannot predict the future climate change. The climate model and Earth system models(ESMs) that consider comprehensive physical and biology process are acted as major tools for projecting future climate variation. The projecting data of ESMs provide opportunity to find out climate changes and potential influence by those changes in the global and regional view (Zhukov and Gushcha, 2020;Yuan et al., 2021). For the aim of deal with extending range of scientific questions proposed by more and more research communities, the coupled model inter-comparison project(CMIP) is continue revising the organization of CMIP (Eyring et al., 2016). The latest one is Phase6 CMIP(CMIP6) that has a substantial augmentation of model numbers and range of experiment. Compared with the previous Phase5 CMIP(CMIP5), ESMs of CMIP6 achieve improvement in spatial resolution, making progress in physical parameterizations (In the expression of cloud, for instance), and including the extra Earth system processes (nutrient restriction on the terrestrial carbon cycle, for example) and components (such as ice cover) (Eyring et al., 2019). Because the effect of climate variation for the environment and society will depend not only on the nature process of Earth system affected by the changes of radiative forcing, but also on how mankind by changes in technology, economies, lifestyle and policy (Moss et al., 2010). New scenarios named shared

https://doi.org/10.5194/essd-2021-418
Preprint. Discussion started: 10 December 2021

socioeconomic pathway(SSPs) (van Vuuren et al., 2014;Riahi et al., 2016) have been developed based on the anticipated challenges to adaptation and mitigation applied in CMIP6(O'Neill et al., 2016). The SSPs include five alternative pathways for socioeconomic Development, including sustainable development(SSP1) (Riahi et al., 2016), middle-of-the-road development(SSP2) (Fricko et al., 2017), regional rivalry(SSP3) (Fujimori et al., 2017), A Road Divided(SSP4) (Calvin et al.,2017) and fossil-fueled development(SSP5) (Kriegler et al., 2017). Those new scenarios can help people understand climate and socioeconomic futures better.

There are many studies analyzing future climate changes based on the ESM data (Gillett et al., 2019;Parsons et al., 2020). In general, these studies highlight three main problems in using ESM data: (1) spatial downscaling of ESMs data to generate data at higher spatial resolution; (2) bias reduction or correction of ESMs; (3) merging of the multi-ESMs to obtain improved accuracy. The output of ECMs have certain errors and raw low spatial resolution on account of the limited computation resources and several simplified hypothesis and uncertainties in model structures and parameterizations (Phillips and Gleckler, 2006). Current ESMs are run at a coarse spatial resolution and can only provide outputs at the spatial resolution ranging from 0.5 to 3.75 degree with the majority at about 2 degrees, and these spatial resolution is not sufficient for elaborated studies (White and Toumi., 2013). Spatial downscaling methods are widely used to improve the resolution of the ESM (Shrestha et al., 2014;Baghanam et al., 2019). It can be divided into statistical and dynamic downscaling techniques (Kannan et al.,2013). Dynamical downscaling is mainly based on the high-resolution regional climate model (RCM) and use the forced lateral boundary from the ECM (Adachi and Tomita, 2020). As for statistic downscaling, it builds the relationship between the ECM and the local-scale meteorological data and then generate high-resolution data (Maraun and Widmann, 2018). Because the dynamical downscaling methods are complex and need high computation cost, statistical downscaling approach is more commonly used. The most popular techniques of statistical downscaling are regression-based approaches, for examples, multiple and generalized linear regression models (Das et al., 2019; Asong et al., 2016), due to their low requirement of computation resources and realize simply. Though the enhancement of the resolution, but the considerable bias still exists no matter in the statistic or dynamic downscaling ways (Miao et al., 2015).

For the aim to reduce the bias, many bias correction approaches are developed includes statistic characteristics bias correction (such as mean, variance and standard deviation) (Ho et al., 2012; Fang et al., 2015), probability distribution bias correction (For example, using the cumulative density function (CDF) to match the ESM and observation data) (Jakob Themeßl et al., 2011) and non-stationary bias correction methods (Miao et al., 2016). The bias correction methods can help the downscaling techniques to get a closer result to observation data (Fan et al., 2021). But the corrected and downscaling methods bring some uncertainties such as the loss of production from extreme disasters was underestimated by the bias-corrected and downscaled ECM data (Lafferty et al., 2021). In different application tasks, different bias correction methods get the best performance (For instance, two nonparametric-transformation approaches perform better than other 9 methods in precipitation bias correction in Norway and regression technique outperforms other 11 methodologies in terms of correlation in the temperature bias correction

https://doi.org/10.5194/essd-2021-418
Preprint. Discussion started: 10 December 2021

in Spain) (Gudmundsson et al,.2012; Gutiérrez et al,.2013). There are some limitations exist in current ECM bias correction approaches (Such as the ECM mean bias correction only consider the ECM mean bias and quantile-quantile correction brings an extra bias to the spatial gradient of variables) (Bruyère et al.,2014; Colette et al.,2012). In addition, the majority of bias correction methods are applied to individual model dataset, which introduces greater uncertainty in the view of projecting the future climate.

Considering different ESMs have different strengths which can be potentially combined to complement each other, many efforts have been made to develop methods to merge multiple ESMs to obtain improved outputs. In previous researches, because of the unstable of single ESM, the ensemble mean(EM) method is considered as a simple and effective way to merge ESMs (A. P. Weigel, 2008). The evaluation result of CMIP6 model show no single model performs best in all regions under different evaluation rules (Papalexiou et al., 2020). Because ESMs have different resolutions and showing different advantages in variant regions, EM method owns a better precise analyze the future temperature data regionally and globally (Fan et al., 2020;Lovino, 2021;You et al., 2021;Tang et al.,2021). The EM method is easy and effective and still remain improve space. In recent years, with a rapid increase amount of 'earth big data' such as earth observation and model simulation data, DL based methods play an increasingly important role in Earth Science because of its better capacity to process big data (Reichstein et al., 2019). Those new approaches have been successfully used in many fields in Earth Science(Yuan et al., 2020;Toms et al., 2020), such as data fusion and downscaling(coarse-resolution temperature and precipitation, for example)(Huang, 2020), Land cover mapping(Tong et al., 2020), information reconstruction(such as reconstruct miss data of the temperature) (Kadow et al., 2020;Barth et al., 2020; Tang et al., 2021), information predication(multi-year ENSO forecast and forecast of complicated tropical instability waves, for instances)(Ham et al., 2019;Zheng et al., 2020) and environmental parameter retrieval(such as PM2.5 and gas concentrations) (Zang, 2021;Tian et al., 2021). The output data from CMIP6 models will arrive 30 petabytes(PB) in the estimation (Stockhause and Lautenschlager, 2017).

The traditional methods need to solve these problem separately in three process and bring extra uncertainties to the future projecting climate data. New DL methods provide us with new opportunity for dealing with the questions when we use the ECM. We can transfer some mature DL algorithms from Super-Resolution(SR) field from the computer vision to solve these question together. Different from other traditional methods, we can input many low resolution temperature data and get one high resolution temperature data, it means that those models can learn the information from different ESMs and do the data merge, bias correction and data downscaling at the same time and reduce uncertainty. The first deep-based approach proposed (Yoon et al., 2015) named SRCNN performed considerably better the traditional SR methods. Then, many DL approaches have been applied to the SR fields and increasing efforts are being made to improve the performance of those algorithms (Shi et al., 2016;Seif and Androutsos, 2018;Blau et al., 2019;Chen et al., 2020). More and more ESMs temperature in different scenarios and historical from CMIP6 become available now. The increasing data quantity can provide enough training samples for the deep learning models, so we can design this work. In this work, we train five different DL networks based on the CRU TS

https://doi.org/10.5194/essd-2021-418
Preprint. Discussion started: 10 December 2021

temperature and 31 ESMs temperature over the time period of 1901 to 2014 then compare the five DL networks to choose the best model to merge the future scenarios land temperature of SSP1-2.6, SSP2-4.5, SSP3-7.0, SSP5-8.5 in the time period of 2015-2100.

2 Data and Methods

2.1 Model data and Observation data

2.1.1CMIP6 Model data

In this study, we collect global surface temperature from 31 CMIP6 ESMs that provides both historical and future data. The most commonly used is the first realization of each model in former studies (Norris, J et al.,2021), so we also use the first realization simulations (r1ilp1f1, except where unavailable). These models are summarized in Table 1. This data is open access, we download it from https://esgf-node.llnl.gov/projects/cmip6/. In CMIP5, only using the Representative Concentration Pathway(RCP) to simulate future climate, includes RCP2.6, RCP4.5, RCP6.0 and RCP8.5 that represent the radiative forcing will reach about 2.6 W/m^2, 4.5 W/m^2、6.0 W/m^2、8.5 W/m^2 in 2100. And in the CMIP6, three new RCPs (RCP1.9, RCP3.4, RCP7.0) have been proposed to meet more research meet. Five sets of experiments use in this study include one historical simulations in 1850-2014 and four scenarios (SSP1-2.6, SSP2-4.5, SSP3-7.0, SSP5-8.5) which are combined SSP and RCP are selected as future projections in the period of 2015 to 2100. And all model surface temperature data are acted as the input data for the model. The raw resolution of most ESMs are close to 2°×2°. In order to retain more information in the origin data, we resampled all ESMs data to 2°×2° resolution using bilinear interpolation, and the output of our model is one merged data at 0.5°×0.5° resolution. We split the data to 80% as the train dataset (1901-1992) and 20% as the validate dataset (1993-2014).

Table1 Summary of 31 different ESMs used in this study

Institution(country)	Model name	Resolution (lon × lat)	Used member
CSIRO(Australia)	ACCESS-CM2	1.875º×1.25º	r1i1p1f1
	ACCESS-ESM1-5	1.875°×1.241°	r1i1p1f1
AWI(Germany)	AWI-CM-1-1-MR	1.875°×1.875°	r1i1p1f1
BCC(China)	BCC-CSM2-MR	2.25°×2.25°	r1i1p1f1
CAMS(China)	CAMS-CSM1-0	1.125°×1.125°	r1i1p1f1
NCAR(USA)	CESM2-WACCM	2.5 °×1.875 °	r1i1p1f1
CMCC(Italy)	CMCC-CM2-SR5	1.25 °×0.9375 °	r1i1p1f1
	CNRM-CM6-1-HR	1.25°×0.9375°	r1i1p1f2

https://doi.org/10.5194/essd-2021-418
Preprint. Discussion started: 10 December 2021

CNRM-CERFACS(France)	CNRM-CM6-1	1.40625°×1.40625°	r1i1p1f2
	CNRM-ESM2-1		r1i1p1f2
CCCMA (Canada)	CanESM5-CanOE	2.8125°×2.8125°	r1i1p2f1
	CanESM5		r1i1p1f1
EC-Earth-Consortium(EU)	EC-Earth3-Veg-LR	0.703125°×0.703125°	r1i1p1f1
	EC-Earth3-Veg		r1i1p1f1
CAS(China)	FGOALS-f3-L	2.5°×2°	r1i1p1f1
	FGOALS-g3	2°×2.25°	r1i1p1f1
NOAA-GFDL(USA)	GFDL-ESM4	2.5°×2°	r1i1p1f1
NASA (USA)	GISS-E2-1-G	2.5°×2°	r1i1p1f2
INM(Russia)	INM-CM4-8	2°×1.5°	r1i1p1f1
	INM-CM5-0		r1i1p1f1
IPSL(France)	IPSL-CM6A-LR	2.5°×1.259°	r1i1p1f1
NIMS-KMA(Korea)	KACE-1-0-G	1.875º×1.25º	r1i1p1f1
UA(USA)	MCM-UA-1-0	3.75°×2.25°	r1i1p1f2
MIROC(Japan)	MIROC-ES2L	2.8125°×2.8125°	r1i1p1f2
	MIROC6	2.8125°×0.703125°	r1i1p1f1
MPI-M(Germany)	MPI-ESM1-2-HR	0.9375°×0.9375°	r1i1p1f1
	MPI-ESM1-2-LR	1.875°×1.875°	r1i1p1f1
MRI(Japan)	MRI-ESM2-0	1.125°×1.125°	r1i1p1f1
NCC(Norway)	NorESM2-LM	5°×3.75°	r1i1p1f1
	NorESM2-MM	2.5°×1.875°	r1i1p1f1
AS-RCEC(China)	TaiESM1	1.25°×0.9375°	r1i1p1f1

2.1.2Observation data

The CRU TS (Climate Research Unit gridded Time Series) is an extensively monthly land climate dataset includes temperature and precipitation(http://doi.org/10/gbr3nj) over global land except the Antarctica with a high spatial resolution at 0.5°×0.5° from 1901 to 2019. It was made by the analysis of more than 4000 independent weather station records and cover all the global land excepted Antarctica and has high credibility (Harris et al., 2020). From the birth of the first version of CRU TS data in 2000, it has been widely used in many research fields such as data evaluation (Fan et al., 2020), climate variability analysis (Wang et al., 2013) and independent climate mode (Renard and Tilman, 2019) due to the high quality. Because the observation data in the ocean region such as ERSST (Huang et al., 2017a) does not match the resolution and time-series with the CRU

https://doi.org/10.5194/essd-2021-418
Preprint. Discussion started: 10 December 2021

dataset, our study concentrates on the land region only. We use the 2m air temperature from the latest version of the CRU TS as the reference data to train our model.

2.2 Methods

In this study, we aimed to build a relationship with the historical CMIP6 ESMs temperature data and the observation data based on the DL models in the period 1901-2014. Then use the optimal model to predict the future data in four scenarios (SSP1-2.6, SSP2-4.5, SSP3-7.0, SSP5-8.5) and extend the historical data from 1850 to 1900. The flowchart of this study is presented in Figure1.

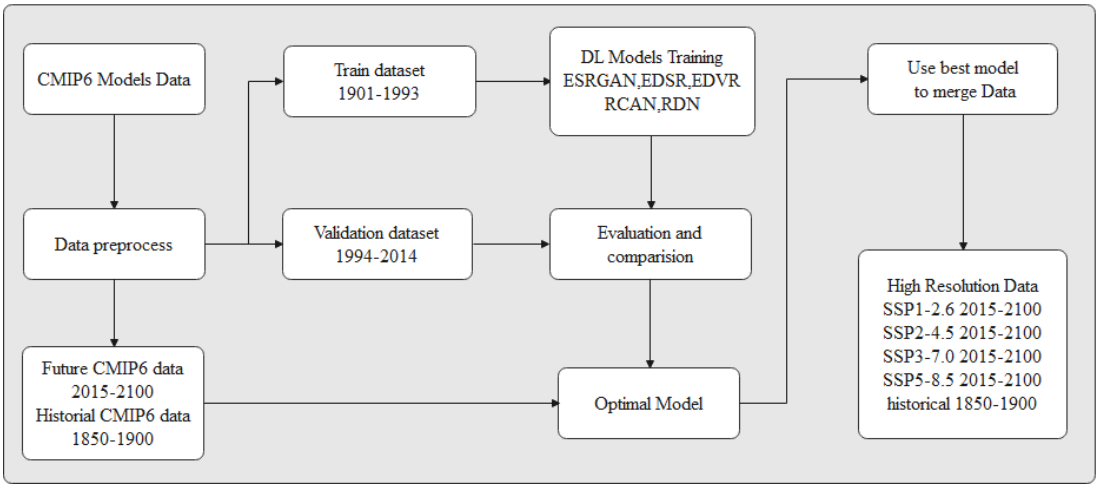
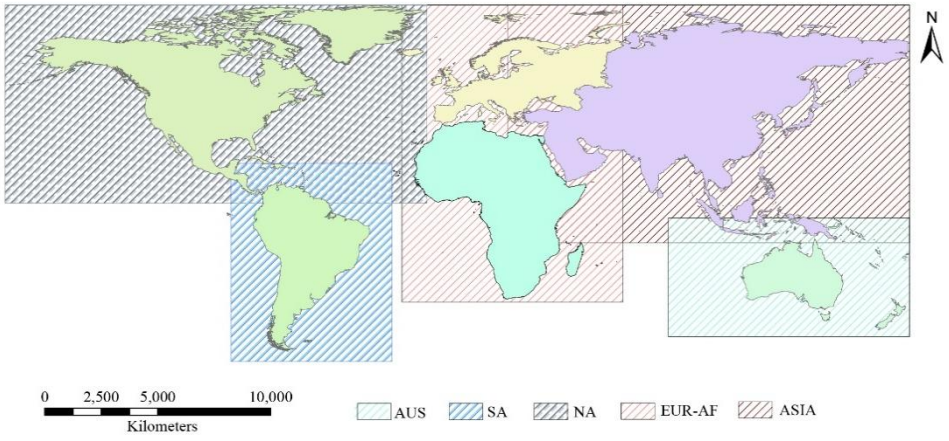


Figure1 Flowchart of this work. The DL models includes Enhanced Super-Resolution Generative Adversarial Networks (ESRGAN), Enhanced Deep Residual Networks for Single Image Super-Resolution(EDSR), Video Restoration with Enhanced Deformable Convolutional Networks(EDVR), Image Super-Resolution Using Very Deep Residual Channel Attention Networks(RCAN) and Residual Dense Network for Image Super-Resolution (RDN).



https://doi.org/10.5194/essd-2021-418
Preprint. Discussion started: 10 December 2021


Open Access Earth System Science Data Discussions

Figure2. Different sub-regions in our study. We divide the global land based on the continent. Australia (AUS), South American(SA), North American(NA), Europe and African(EUR-AF), and Asia.

2.2.1 Five different DL models

In order to find a suitable model in our work, wo choose different methods made by different classic DL structure includes Residual Network(ResNet), Attention mechanism, dense convolutional Network (DenseNet), Generative Adversarial Networks(GAN) and Deformable Convolution Networks (DCN). Five mature SR methods are designed for our work. Four networks are single image SR includes Enhanced Deep Residual Networks for Single Image Super-Resolution(EDSR), Residual Dense Network for Image Super-Resolution (RDN), Image Super-Resolution Using Very Deep Residual Channel Attention Networks(RCAN) and Enhanced Super-Resolution Generative Adversarial Networks (ESRGAN) and one for video super-resolution named Video Restoration with Enhanced Deformable Convolutional Networks(EDVR). Each method is briefly described below.

ResNet solve the limitation on network depth because of the gradient disappeared (He et al., 2016). It achieved great success and widely applied in different fields of computer vision include image super resolution named SRResNet (Ledig et al., 2017). On the basis of SRResNet, a better ResNet structural proposed (Wang et al., 2019a) called EDSR, the most obvious feature is removing the batch normalization layers in the network and we can see it in FigureS1(a). Besides increase of the performance, this new baseline model need few parameters and save about 40% computational resources.

DenseNet is a novel network can relive the phenomenon of the gradient vanish, the main character is that the input of each layer comes from the output of all previous layers and it means every layer has direct access to the final result, has a narrow internet structure and few parameters (Huang et al., 2017b). Because of most deep CNN based SR networks do not fully utilize hierarchical features from the initial low-resolution(LR) image, RDN with a creative residual dense network based on ResNet and Densenet showed in FigureS1(b) is put forward to deal with this problem (Xu et al., 2018).

Attention mechanism was an indispensable part of model in different regions of natural language processing (NLP), such as speech recognition, question answering and algorithm-learning (Vaswani et al., 2017). Attention function solve a problem that long distance information is very difficult to remember. After the new network structure transformer created mainly consisted of attention blocks (Vaswani et al., 2017), attention mechanism became popular in the other fields of computer vision includes image SR. Researchers invented a residual in residual(RIR) structure to form very deep network and add channel attention to residual block showed in FigureS1(c) (Zhang et al., 2018). RCAN can learn low and high frequency information and adaptive rescaling of channel characteristics by considering the interdependence between channels.

https://doi.org/10.5194/essd-2021-418
Preprint. Discussion started: 10 December 2021


Open Access Earth System Science Data Discussions

GAN is consisted of two models: a generative model G to capture the feature and a discriminative model D to estimate the result from the G, the task of D is to minimize the probability of D making an error (Goodfellow et al., 2014). In order to solve the problem that the finer texture detail is hard to recover at large upscaling factors, an image SR method is proposed based on GAN named SRGAN (Ledig et al., 2017). But the result of the SRGAN often accompanied with unpleasant artifacts, the ESRGAN is designed to improve the performance of this network (Wang et al., 2019c).

CNN has a fixed model geometric structure restrict the model geometric transformations. For the aim to enhance the model transformation, DCN is proposed on account of the thought enlarging the spatial sampling regions (Qi et al., 2017). EDVR apply DCN and Pyramid structure to design a Pyramid, Cascading and Deformable(PCD) network and the structure showed in FigureS2 (Wang et al., 2019b). Although this method is proposed to do video restoration, it also achieve a state-of-art performance in video SR. Different from the single image SR algorithm, the EDVR need a time series sequence input to learn spatial and temporal information and output the central frame image. As we all know, the climate data have a strong relationship in the time dimensional, so we also use the method to have an attempt.

2.2.2 Evaluation metrics

2.2.2.1Train and validation metrics

In order to evaluate different SR models fairly, we adopt the same loss function for different networks. In the training process the Mean square error(MSE) is used as the loss function, the equation is described as follows:

$$MSE = \frac{1}{N}\sum_{n=1}^{N}|Y_n - P_n|^2 \quad (1)$$

where, N is the number of samples, Y_n is the observation temperature and P_n is the output of model data. In the validation process, in order to evaluate the result more comprehensive, mean absolute error(MAE) and Pearson correlation coefficient(R) are applied to choose best model, in addition to the MSE. MAE and R are calculated as follows:

$$MAE = \frac{1}{N}\sum_{n=1}^{N}|Y_n - P_n| \quad (2)$$

$$R = \frac{\sum_{n=1}^{N}(Y_n-\bar{Y})(P_n-\bar{P})}{\sqrt{\sum_{n=1}^{N}(Y_n-\bar{Y})^2}\sqrt{\sum_{n=1}^{N}(P_n-\bar{P})^2}} \quad (3)$$

The parameter meaning is the same as formula (1).

Besides it, we constructed Taylor diagrams (Taylor 2001) to compare the observation, merged data and model data, which needs R, standard deviations of error and unbiased root-mean-square deviation (ubRMSD):

$$RMSD = \sqrt{\frac{1}{N}\sum_{n=1}^{N}(Y_n - P_n)^2} \quad (4)$$

$$BIAS = \frac{1}{N}\sum_{n=1}^{N}(Y_n - P_n) \quad (5)$$

$$ubRMSD = \sqrt{RMSD^2 - BIAS^2} \quad (6)$$

https://doi.org/10.5194/essd-2021-418
Preprint. Discussion started: 10 December 2021

2.2.2.2 The SPAtial-EFficiency metric (SPAEF)

Besides some common error metric for the evaluation, we also apply a novel spatial performance SPAEF that consisted of Pearson correlation coefficient, coefficient of variation and histogram intersection. Only one metric is not able to reflect the whole pattern information, three parts of SPAEF is proved to be individual, allowing them supplement each other in meaningful way (Koch et al., 2018).

$$\beta = \left(\frac{\sigma_{\text{sim}}}{\mu_{\text{sim}}}\right) / \left(\frac{\sigma_{\text{obs}}}{\mu_{\text{obs}}}\right) \quad (7)$$

$$\gamma = \frac{\sum_{j=1}^{n} min(K_j, L_j)}{\sum_{j=1}^{n} K_j} \quad (8)$$

$$SPAEF = 1 - \sqrt{(R-1)^2 + (\beta-1)^2 + (\gamma-1)^2} \quad (9)$$

where, β is the ratio of the coefficient of variation between the observation data and the model simulated data. σ_{sim} and μ_{sim} are standard deviation and mean value of model output separately; σ_{obs} and μ_{obs} are standard deviation and mean value of model output respectively. γ is the histogram intersection for the histogram K of the observation data and the histogram L of the output of model, each containing n bins. For the aim to comparison of different units variables and to guarantee bias insensitivity, the z score method is used to calculate γ .The closer this value is to one, the better the data quality.

3 Results and discussion

https://doi.org/10.5194/essd-2021-418
Preprint. Discussion started: 10 December 2021

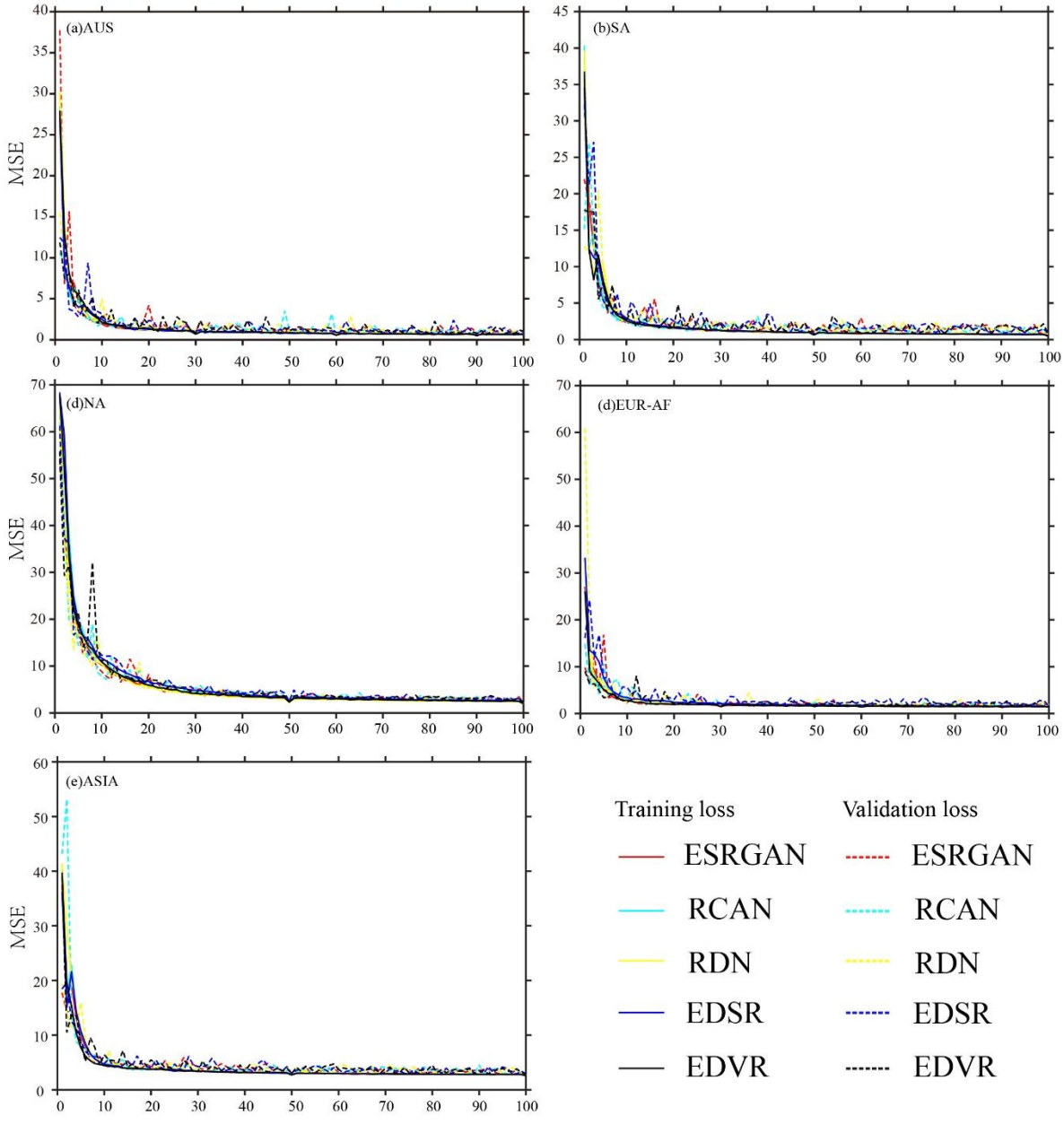


Figure3. The epoch loss change accompanying the train process in different models and different regions. The solid line represents loss in train dataset and the dotted line represents loss in validation dataset.

We divide global land to five regions (AUS, SA, NA, EUR-AF, ASIA) as showed in Figure2. Considering the boundary effects, we reserve duplicate areas between different zones. In the training process, we introduce the land-ocean mask to make the result not disturbed by ocean region. For the aim to choose a DL model most suitable for our work under equity principle, we

https://doi.org/10.5194/essd-2021-418
Preprint. Discussion started: 10 December 2021


Open Access Earth System Science Data Discussions

use the same parameter setting such as loss function(MSE), learning rate and train epochs to train different models. One epoch iterates over the training dataset about 1000 set once, we apply 100 epochs in our train. Figure3 depicts the loss in train and validation dataset variation along the epoch. A rule can be induced from each of the loss line pictures clearly, after the big drop of loss at the start of the training process and the loss line of train and validation is declined with fluctuation subsequently. Almost all model achieve convergence at about 50 epochs. We also use learning rate attenuation strategy based on the change of epoch to reach a better result. According to the loss performance of validation dataset, we save the best model of individual model in different areas.

The evaluation metrics calculated on the validation dataset of different model in various regions are showed in Table 2. It is noteworthy that three metrics (MSE, MAE and R) have little difference between various models. In terms of those metrics, the output of the EDSR models achieve the best performance in all regions except Region-1. The EDVR get the best scores from the MAE and MSE in AUS, is slightly higher more than the EDSR (second to the EDVR), but the R of EDSR is better than the EDVR. All regions have a very high values of R and reflecting a relatively large difference in MSE and MAE. In NA and ASIA, we can get a conclusion that the error is obvious bigger than the other regions based on the MSE and MAE. The SA region has the least error. The rest of regions reach a relative low error result compared with the NA and ASIA. Totally, the EDSR model outperforms the other DL approaches, we choose EDSR as the optimal method in the following work. In order to understand the spatial error distribution, we calculate the mean MAE by pixel of all model output in the validation dataset. Additionally, we also compute the mean spatial deviation by the EM data and compare with DL results to see the improvement brought by the DL model. In order to have an intuitive comparison, we use bilinear interpolation resample all the ESM model data to a 0.5°×0.5°same to the observation data and merged data.

Table2 Evaluation metrics in validation dataset for the five different models. The values in bold means the best performance.

Region	ESRGAN	RCAN	RDN	EDSR	EDVR
AUS	MSE:1.09	MSE: 1.10	MSE:1.07	MSE: 1.05	**MSE: 1.03**
	MAE:0.80	MAE:0.80	MAE:0.79	MAE:0.79	**MAE:0.77**
	R:0.9853	R:0.9855	R:0.9859	**R:0.9860**	R:0.9859
SA	MSE:0.77	MSE:0.72	MSE:0.74	**MSE:0.67**	MSE: 0.72
	MAE:0.66	MAE:0.63	MAE:0.64	**MAE:0.60**	MAE:0.64
	R:0.9937	R:0.9940	R:0.9945	**R:0.9947**	R:0.9938
NA	MSE:3.90	MSE:3.63	MSE:3.86	**MSE:3.36**	MSE: 3.84
	MAE:1.45	MAE:1.39	MAE:1.42	**MAE: 1.33**	MAE:1.42
	R:0.9936	R:0.9939	R:0.9937	**R:0.9943**	R:0.9935
	MSE: 2.56	MSE:2.45	MSE:2.55	**MSE:2.32**	MSE: 2.48

https://doi.org/10.5194/essd-2021-418
Preprint. Discussion started: 10 December 2021

EUR-AF	MAE:1.10	MAE:1.07	MAE:1.10	**MAE: 1.02**	MAE:1.07
	R:0.9922	R:0.9928	R:09923	**R:0.9933**	R:0.9926
	MSE: 4.73	MSE:4.41	MSE:4.43	**MSE:4.16**	MSE: 4.53
ASIA	MAE:1.55	MAE:1.47	MAE:1.47	**MAE:1.42**	MAE:1.47
	R:0.9917	R:0.9926	R:0.9924	**R:0.9930**	R:0.9920

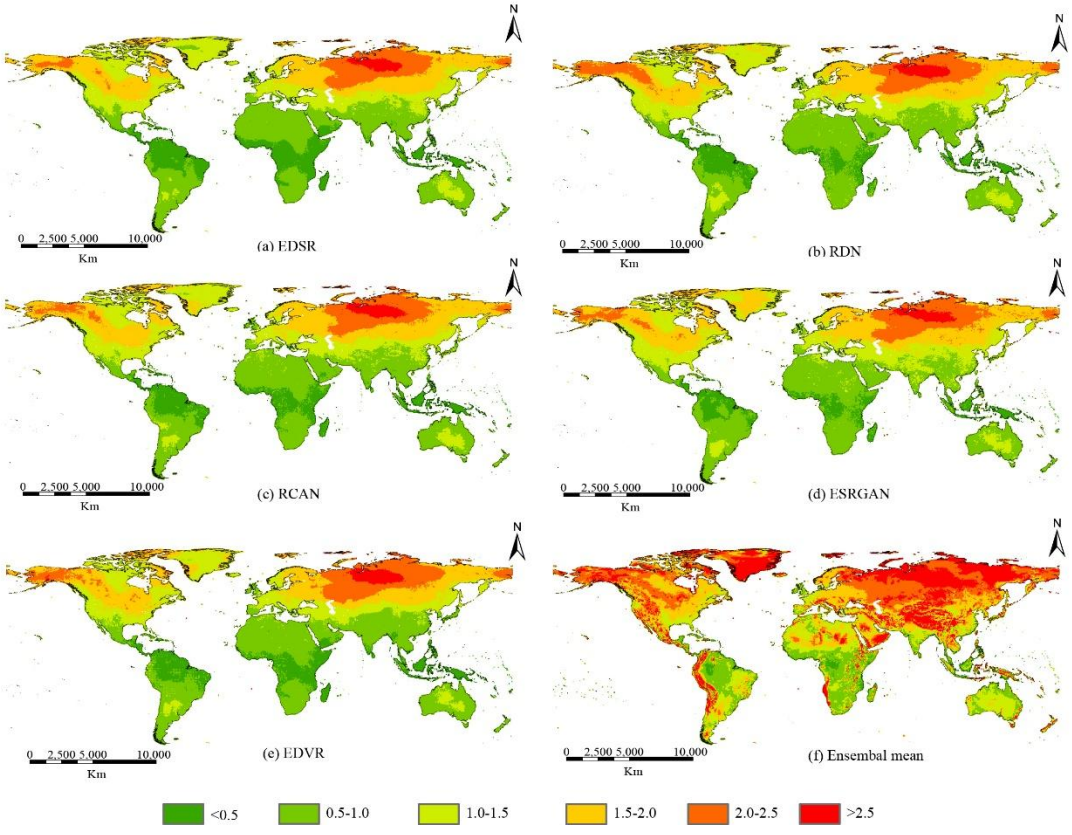


Figure4. Spatial distribution of MAE compared with observation data in validation dataset for five different DL models ((a) EDSR, (b) RDN, (c) RCAN, (d) ESRGAN, (e) EDVR and the (f) CMIP6 EM data).

Figure 4 shows the spatial distribution of the metric MAE for various methods in validation dataset. Figure4(a)-(e) show the results of DL models and Figure4(f) reveals the mean error of EM methodology. We can see the result of DL models have a significant improvement in almost every pixel compared with the EM result clearly. From the overall view, the performance of DL model decided by the input ESM data to a great extent. We can see from the EM result that the whole northern hemisphere including Asia, Europe and North American display high errors (MAE more than 2 in most regions), indicating that the raw model data is not accurate. In the south hemisphere, the situation is much better, only some sub-regions display high error

https://doi.org/10.5194/essd-2021-418
Preprint. Discussion started: 10 December 2021

(MAE exceeds 2) and the most area with a relatively low error (MAE less than 1.5). Remarkably, the DL model results have a similar spatial pattern of MAE, the majority districts own a favorable small error expect the upper northern latitude areas. An evident regularity can be summarized, the error becoming larger and larger along the latitudes in north hemisphere, presenting like a 'tertiary class echelon' condition. As we all know, Tibet plateau area is one of the polar regions in the word, is very hard to simulate. From our result, it is glad to see that the error in Tibet plateau area improves a lot. The same conclusion can be obtained that the EDSR is the optimal model in our task. Although DL methods own a similar spatial distribution of MAE, however we can notice the fewest high MAE region in the Figure5(a) (The EDSR model result) on closer inspection.

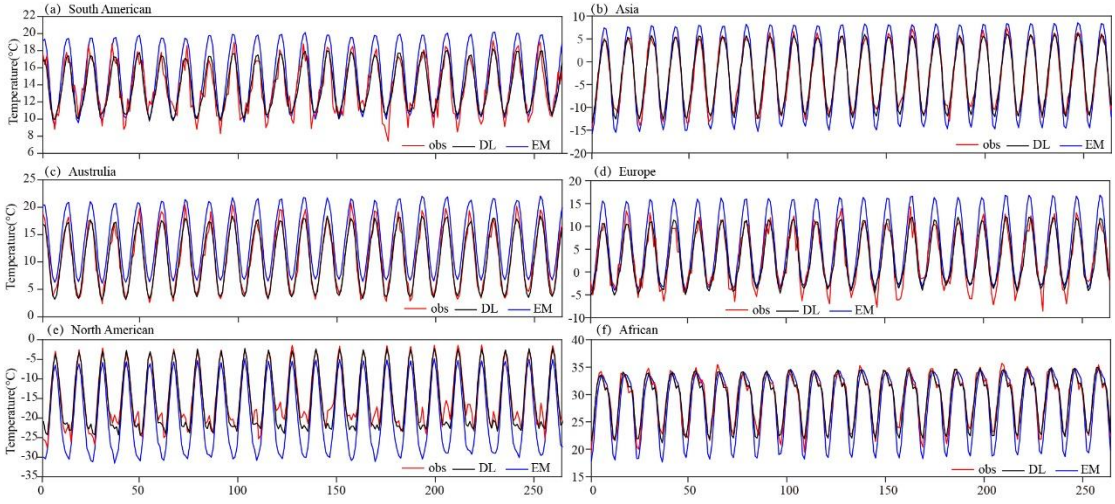


Figure5. Time series temperature of the observation, DL merged data and EM data in one pixel in validation dataset, (a)-(f) represent the pixel in different continent. The location of these pixels are (-77.75, -7.25), (87.75.31.25), (148.25, -35.25), (10.25,46.75), (-44.75,68.75) and (13.75, -19.25) corresponding to (a) to (f) respectively.

We randomly selected six pixels in individual continent to evaluate the performance of the time dimension. The time series showed in Figure5. It can be seen that the merged data has a remarkably better performance than the EM data compared to the observation data. Generally, the peak high and low temperature tends to have higher bias over the time series. Six pixel can be classed to three group based on the results. The Figure5(a) and Figure5(d) are in one group, the observation data have a lot of small fluctuation besides the cyclic variation, the merged data notwithstanding reduce the bias but not reconstruct the wave motion of the whole series. The second group consisted of Figure5(b) and Figure5(c), data of observation exhibits a relatively smooth period and the merged data has a good fit with the reference data. The last condition depicted by the Figure5(e) and Figure5(f), the observed data also have a smooth period but exist a fluctuation in place of the high and low in the cycle. It is surprised to find the merged data has learned this wave despite there are still considerable errors in those places. According to those results, we can get a consequence that the merged data will have an excellent result when the observation data with a few fluctuations. On the contrary, if the observation data appear a lot of wave, the performance of the merged data will decrease.

https://doi.org/10.5194/essd-2021-418
Preprint. Discussion started: 10 December 2021

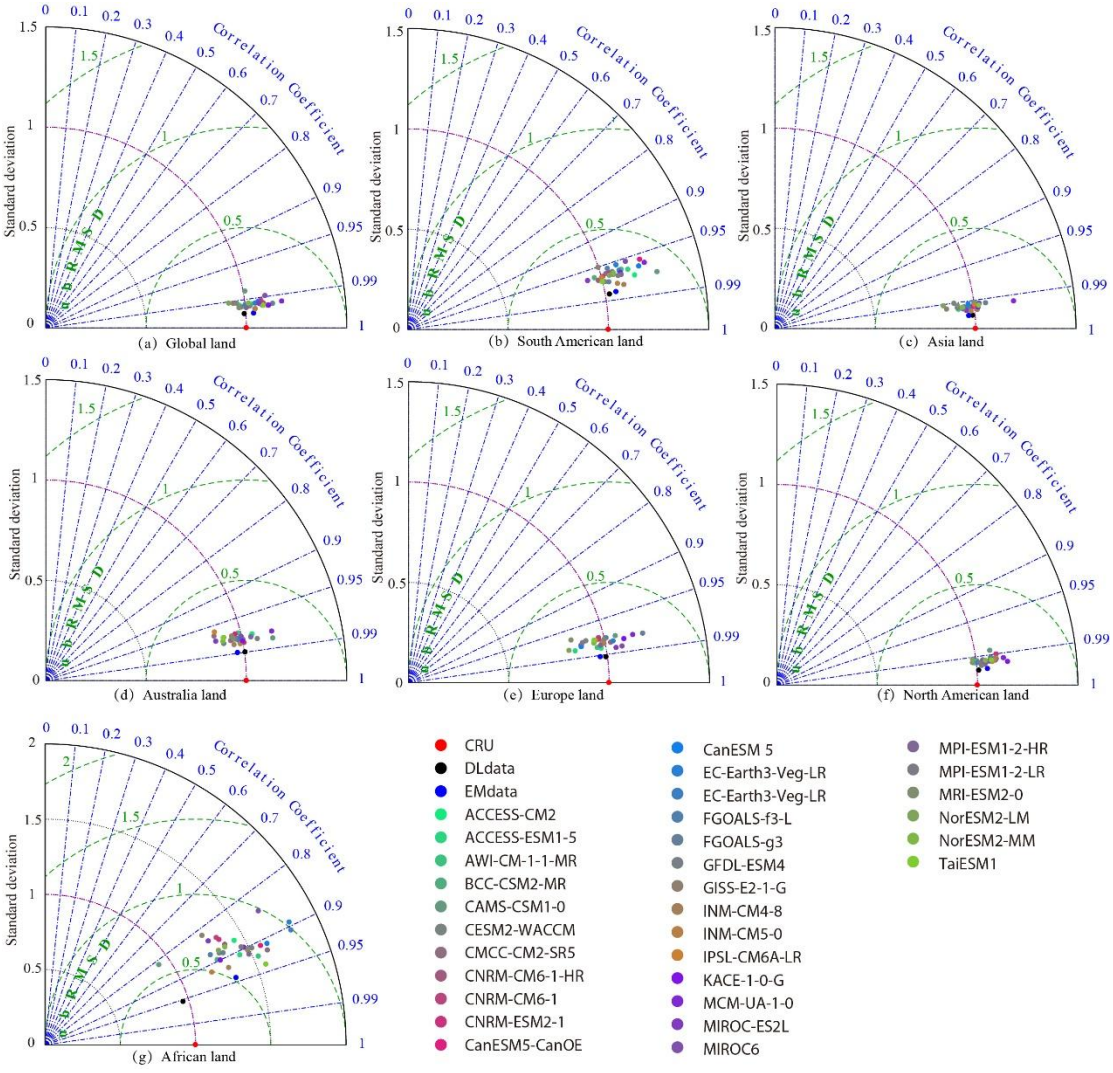


Figure 6 The monthly mean temperature of the Global and regions was used for Taylor diagrams to compare each of the ESMs, the observation, DL merged data, EM data in the period of validation dataset. The vertical coordinate is the standard deviation. Green concentric circles of dashed lines are ubRMSD. The angular coordinate is used to show the R.

We use the Taylor diagrams to farther evaluate the DL merged data, EM data and the observation. Figure6 depicts the Taylor diagrams for results of the monthly mean temperature during the period 1993-2014 in different data. It is not surprising to find that the merged data achieve an excellent stable performance (low bias and high R) in all regions and they are superior to any other data. From the overall view, all ESMs and EM data have a very high R exceeding 0.95, and the EM data display a relatively better performance than the individual model globally. On the continent scale, there are clear differences between individual ESMs. In Asia and North American showed in Figure6(c) and Figure6(f), the performance of individual ESM has

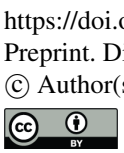
https://doi.org/10.5194/essd-2021-418
Preprint. Discussion started: 10 December 2021

no obvious variations and get a similar result like the global. And in other continents, the whole result of ESMs showing a trend of dispersion and the performance is relatively low. Only from the angle of the R, the EM data also achieve a splendid result second to the merged data. But from the view of the standard deviation and ubRMSD, the performance of the EM data shows a small but clear distances compared with the merged data, even not as effective as single ESM data in some areas. The worst result is observed in the African as shown in Figure6(g), all ESMs with a R less than 0.95 and high ubRMSD larger than 0.5, representing a more scattered distribution.

In order to achieve a better understanding of the DL merged data, the Taylor diagrams for annual climatology temperature, the result depicted the DL merged data display the best performance. Besides it, we also use the innovative metric SPAEF to compare DL data with individual ESM data and EM data. The SPAEF of the merged data, EM data and different ESMs data depicted in Figure7. It is not hard to find the scores of the merged data are ranked the first regardless of from the view of the global and different continent land. On the Contrary, the SPAEF of EM data and ESMs data have a diverse consequence in the global and region area. We can see from the Figure7(a) that the EM data also shows an excellent SPAEF scores only second to the merged data globally. But in the continent perspective, the performance of EM data is not stable, the similar situation also happened on the ESMs data, revealed by Figure7(c) and Figure7(g) the SPAEF of EM data in Asia and African is obviously lower than some single ESM result. It's worth noting that the range of SPAEF is not restrict in 0 to 1 like some other correlation metrics. Totally, global scale scores of all data are exceed 0.6, but there is a lot of negative values appear in some regional results include Asia, Australia, Europe and North American reflected in Figure7(c)-(f). Additionally, comparison of global and regional results, almost all data SPAEF performance experience a clear decrease from global to region. It is surprise to realize that three ESMs (KACE-1-0-G, MCM-UA-1-0 and MIROC-ES2L) data achieve almost negative SPAEF in Australia, Europe and North American, it mainly related to those model have a raw low resolution (less than 2°×2°) showed in Table 1.

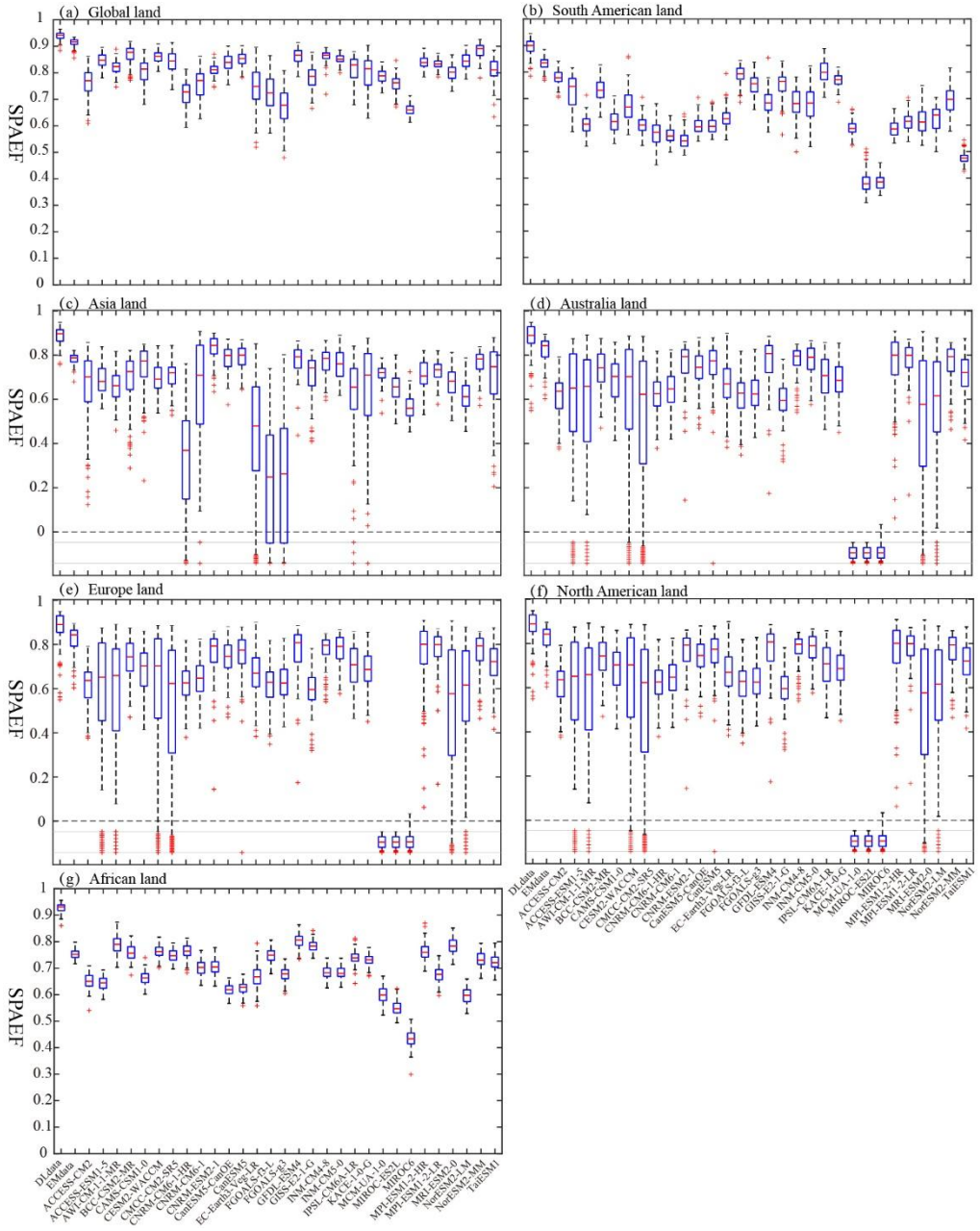
Based on the ESMs data, we use the best DL model not only merge the future temperature dataset and also merge the historical data from 1850 to 1900. From the Figure8(a), we can see anomalies of merged data and all CMIP6 models in the period of 1850 to 2100, using the data of observation data in 1995 to 2014 as the reference period. Obviously, the largest increase of temperature happened in the SSP-5.85, the up trends are also showed in SSP3-7.0, SSP2-4.5 and SSP1-2.6 circumstances. From the overall view, the EM data tend to lower than the observation data and the DL merged data have a closer result with the observed data globally. So, the merged data is slightly higher than the EM data regardless of in the historical and future, consistency of results can prove the stable ability of the DL model in our work. It is no doubt that we can reach a conclusion that the merged data is credible. In Figure8(b), we use EM data and DL data to calculate the temperature variation respectively in different periods. In order to match the reference period of 1995 to 2014, we select mean temperature of 1851-1870 and 2081-2100 to represent the historical temperature and future temperature. The result of DL data shows a smaller increase in historical and different scenarios than the EM data. This phenomenon reveals the EM data exist a high estimation of temperature increase. The regional result is showed in Figure S4 and Table S1, this situation is most obvious in the South

https://doi.org/10.5194/essd-2021-418
Preprint. Discussion started: 10 December 2021



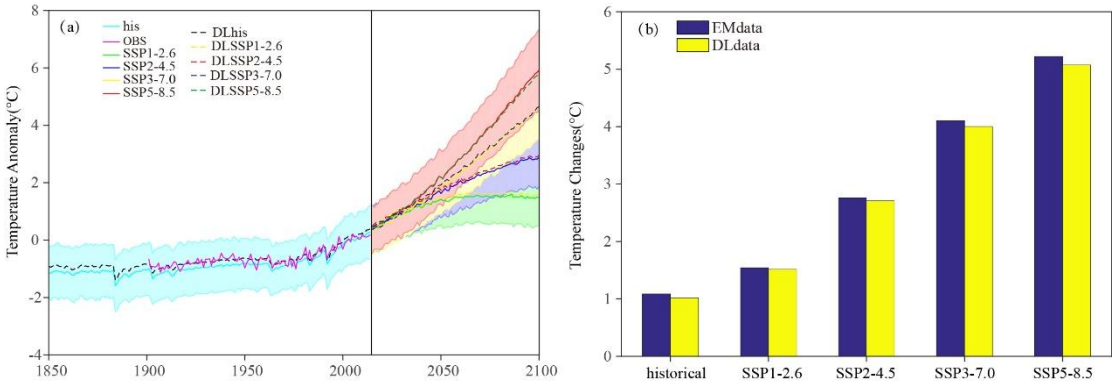
Open Access Earth System Science Data Discussions

American. Apart from it, the future four scenarios temperature decade trend presented in FigureS5, showing a clear upward tendency.



370 Figure7 SPAEF of the global land surface and each individual continent comparing each of the CMIP6 models 、the observations、DL merged data and EM data for the period 1901–2014.

https://doi.org/10.5194/essd-2021-418
Preprint. Discussion started: 10 December 2021



375 Figure8 Time series of Global land annual mean temperature anomalies of CMIP models, DL merged data and observation data over the global land surface and temperature change in historical and different scenarios calculated by EM data and DL data. The solid line in Figure8(a) are EM data of CMIP6 models and the shaded areas are the standard deviation.

From our result, it is not surprise to find that the DL merged data outperforms the individual ESM and EM data in all aspects 380 from the evaluation metrics. The performance of EM data is second to the DL merge data in the whole although it poorer than single model in regional scale. As we all know, the deep-learning methods are data-driving algorithms (Reichstein et al., 2019), it means the result is determined by the input data to a large extent. From our result and other studies (Fan et al., 2020), the errors of raw ECM data in north hemisphere are very high (MAE of most regions exceeds 2), the DL merged data have a greater improvement in north hemisphere except the high north hemisphere still exist some high error regions (MAE exceeds 385 than 1.5). It is worth noting that the bias in the Tibet plateau area is relatively low (MAE less than 1) under some raw model have a 5 degree under estimate in this area (Fan et al., 2020). The consequence prove the DL model do a good bias correction during the spatial downscaling and date merge.

At the pixel time series, the DL merged data cannot achieve a good consequence when the temperature change has a lot of 390 small wave in a big cycle and in the other condition it show a good performance. All ESM do not perform well in African than other continents from the time series evaluation (Figure6 (g)), mainly due to the limited availability of ground measurements for CRU in African (Collins, 2011). The merged data have a relatively low progress in time dimension compared with the spatial dimension. This phenomenon is related with the best DL model we selected is made by a single image DL approach not consider the information in time dimension.

395

In former studies, the DL methods based on simply basic structures were reported with good performance in the downscaling work (Huang, 2020). Instead of building a small and easy DL network, our work applied mature methods with relatively more complexity. In this study, for the aim to choose a suitable method for this work, 4 single image SR algorithm and 1 video SR

https://doi.org/10.5194/essd-2021-418
Preprint. Discussion started: 10 December 2021

algorithm have been used in our work to select a best model to merge the data. Actually, the result makes us a little surprise because the best model we selected according to the validation metrics is not own the outperformance than other methods in the computer vision task and not the newest algorithm. Besides it, in terms of the theory, video SR algorithm is thinking about the information both in space and time dimension (Chan et al., 2021), combined with the climate data is full of space and time relationship, this method should get the best result compared with the only consider spatial feature model. But the result from the EDVR not achieve the state of art of our work. This result depicts that the experience from the computer vision cannot fully applied into other fields and we need to do more comparative experiment when we consider specific applications.

Our work used DL methods to make high-resolution future temperature under four sceneries (SSP1-2.6, SSP2-4.5, SSP3-7.0, SSP5-8.5). However, there are some limitations of this work. Firstly, the observation data we use is a reanalysis gridded data made by information of meteorological stations, it also exists some errors and cannot be avoided, which affects the accuracy of this work. In the future work, the weather station data need to be considered with reanalysis data together to achieve a better result. Secondly, no suitable ocean observation data are available to match the CRU temperature data in the spatial resolution and time series simultaneously, so we only merge a high-resolution data in the land area. Additionally, the lack of SST also makes us unable to consider SST in the model via the training process. In the future study, we think taking the ocean into account will most likely improve the model outputs because the DL model can learn extra information between the ocean and land. Thirdly, climate data is full of temporal information, our work does not explicitly take advantages of the temporal information. Similarly, in some other application field (such as temperature reconstruction work) (Kadow et al., 2020), how to make full use of information from the time dimension is also a question. Therefore, how to make full use of the temporal information to improve the performance of the merged data is an interesting topic that is needed to be resolved in the future work.

4 Data availability

All data we used in this work can be downloaded from the links given in Section 2. A copy of our merged future global land high-resolution temperature product has been registered with Zenodo and is available at https://doi.org/10.5281/zenodo.5746632 (Wei et al., 2021).

5 Conclusions

In this study, we use the deep learning (DL) method to generate the temperature data for the global land (except Antartica) at higher spatial resolution (0.5 degree) based on 31 different CMIP6 ESMs. Our methods can perform bias correction, spatial downscaling and data merging simultaneously. Five different DL methods are evaluated and the optimal model is selected to generate the historical data during 1851-1900 and future scenarios simulation data (SSP1-2.6, SSP2-4.5, SSP3-7.0, SSP5-8.5)

from 2015 to 2100. The merged data have a remarkably better quality compared with the individual ESMs and ensemble of all ESM in terms of both spatial dimension and time dimension. The error metrics (like R and ubRMSD) also show that the merged data have better accuracy in global and region areas.

The merged data have a great improvement in the spatial error, besides the high latitude of Asia and North American, the most areas owns a low MAE (less than 1 degree). It means the merged data take full advantages of the raw CMIP6 model data. From the pixel time series, the DL merged data can present excellent performance when the observation is smooth with few fluctuations. From the total view, the merged data also make a progress in the time dimension. It is easy to find that the merged data have a better result in the spatial distribution than the time variation because the DL model is mainly pay more attention on the spatial feature.

Our work is one of the first studies that perform spatial downscaling, bias correction and date merging simultaneously. Our results demonstrate that the DL model from the SR in computer vision can be successfully transferred to deal with the data merge and data downscaling problems when enough training data are available. The generated merged data with improved accuracy have great potentials for many applications such as spatial-temporal change of future climate, how climate change influences on water resources and agriculture.

Author contributions. XW conducted the research, completed the original draft, and revised it. DF, DFTH, LM and LT contributed to data processing., ZD, JT and BS revised the draft. GW, the corresponding author, contributed to conceptual designing, reviewing of the manuscript, funding acquisition, and project administration. All coauthors reviewed the manuscript and contributed to the writing process.

Competing interests. The authors declare that they have no conflict of interest.

Acknowledgments. This work was supported by National Key Research and Development Program of China (2019YFC1510203), General project of national natural Science Foundation of China (41875094) and The Sino-German Cooperation Group Project of China(GZ1447).

References

Adachi, S.A., Tomita, H.: Methodology of the constraint condition in dynamical downscaling for regional climate evaluation: a review. J. Geophys. Res.: Atmosphere 125 (11), e2019JD032166. https://doi.org/10.1029/2019JD032166, 2020.

A. P. Weigel, M. A. L. and C. A.: Can multi-model combination really enhance the prediction skill of probabilistic ensemble forecasts?, , 1227(July), 496, doi:10.1002/qj, 2008.

Asong, Z.E., Khaliq, M.N., Wheater, H.S. :Projected changes in precipitation and temperature over the Canadian Prairie Provinces using the Generalized Linear Model statistical downscaling approach. J. Hydrol. 539, 429–446, 2016.

https://doi.org/10.5194/essd-2021-418
Preprint. Discussion started: 10 December 2021

Avaria-Llautureo, J., Venditti, C., Rivadeneira, M. M., Inostroza-Michael, O., Rivera, R. J., Hernández, C. E. and Canales-Aguirre, C. B.: Historical warming consistently decreased size, dispersal and speciation rate of fish, Nat. Clim. Chang., 11(9), 787–793, doi:10.1038/s41558-021-01123-5, 2021.

Baghanam, A. H., Nourani, V., Sheikhbabaei, A. and Seifi, A. J.: Statistical downscaling and projection of future temperature change for Tabriz city , Iran.[J]. IOP Conference Series Earth and Environmental Science, 491:012009, 2020.

Barth, A., Alvera-Azcárate, A., Licer, M. and Beckers, J. M.: DINCAE 1.0: A convolutional neural network with error estimates to reconstruct sea surface temperature satellite observations, Geosci. Model Dev., 13(3), 1609–1622, doi:10.5194/gmd-13-1609-2020, 2020.

Blau, Y., Mechrez, R., Timofte, R., Michaeli, T. and Zelnik-Manor, L.: The 2018 PIRM challenge on perceptual image super-resolution, in Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), vol. 11133 LNCS, pp. 334–355., 2019.

Bruyère, C. L., Done, J. M., Holland, G. J. & Fredrick, S.: Bias corrections of global models for regional climate simulations of high-impact weather. Clim. Dyn. 43, 1847–1856., 2014.

Calvin, K., Bond-Lamberty, B., Clarke, L., Edmonds, J., Eom, J., Hartin, C., Kim, S., Kyle, P., Link, R., Moss, R., McJeon, H., Patel, P., Smith, S., Waldhoff, S. and Wise, M.: The SSP4: A world of deepening inequality, Glob. Environ. Chang., 42, 284–296, doi:10.1016/j.gloenvcha.2016.06.010, 2017.

Chan, K. C. K., Wang, X., Yu, K., Dong, C. and Loy, C. C.: BasicVSR: The Search for Essential Components in Video Super-Resolution and Beyond, [online] Available from: http://arxiv.org/abs/2012.02181, 2021.

Chen, Y., Liu, S. and Wang, X.: Learning continuous image representation with local implicit image function, arXiv, 2020.

Choi, D., Gao, Z. and Jiang, W.: Attention to global warming, Rev. Financ. Stud., 33(3), 1112–1145,doi:10.1093/rfs/hhz086, 2020.

Clem, K. R., Fogt, R. L., Turner, J., Lintner, B. R., Marshall, G. J., Miller, J. R. and Renwick, J. A.: Record warming at the South Pole during the past three decades, Nat. Clim. Chang., 10(8), 762–770, doi:10.1038/s41558-020-0815-z, 2020.

Colette, A., Vautard, R. & Vrac, M.: Regional climate downscaling with prior statistical correction of the global climate forcing. Geophys. Res. Lett. 39, L13707 ,2012.

Collins, J. M.: Temperature variability over Africa, J. Clim., 24(14), 3649–3666, doi:10.1175/2011JCLI3753.1, 2011.

Das, L., Akhter, J.: How well are the downscaled CMIP5 models able to reproduce the monsoon precipitation over seven homogeneous zones of India? Int. J. Climatol. 39, 3323–3333,2019.

Eyring, V., Bony, S., Meehl, G. A., Senior, C. A., Stevens, B., Stouffer, R. J. and Taylor, K. E.: Overview of the Coupled Model Intercomparison Project Phase 6 (CMIP6) experimental design and organization, Geosci. Model Dev., 9(5), 1937–1958, doi:10.5194/gmd-9-1937-2016, 2016.

Eyring, V., Cox, P. M., Flato, G. M., Gleckler, P. J., Abramowitz, G., Caldwell, P., Collins, W. D., Gier, B. K., Hall, A. D., Hoffman, F. M., Hurtt, G. C., Jahn, A., Jones, C. D., Klein, S. A., Krasting, J. P., Kwiatkowski, L., Lorenz, R.,

https://doi.org/10.5194/essd-2021-418
Preprint. Discussion started: 10 December 2021

Maloney, E., Meehl, G. A., Pendergrass, A. G., Pincus, R., Ruane, A. C., Russell, J. L., Sanderson, B. M., Santer, B. D., Sherwood, S. C., Simpson, I. R., Stouffer, R. J. and Williamson, M. S.: Taking climate model evaluation to the next level, Nat. Clim. Chang., 9(2), 102–110, doi:10.1038/s41558-018-0355-y, 2019.

Fan, X., Duan, Q., Shen, C., Wu, Y. and Xing, C.: Global surface air temperatures in CMIP6: Historical performance and future changes, Environ. Res. Lett., 15(10), doi:10.1088/1748-9326/abb051, 2020.

Fan, X., Jiang, L. and Gou, J.: Statistical downscaling and projection of future temperatures across the Loess Plateau, China, Weather Clim. Extrem., 32, 100328, doi:10.1016/j.wace.2021.100328, 2021.

Fang, G. H., Yang, J., Chen, Y. N. and Zammit, C.: Comparing bias correction methods in downscaling meteorological variables for a hydrologic impact study in an arid area in China, Hydrol. Earth Syst. Sci., 19(6), 2547–2559, doi:10.5194/hess-19-2547-2015, 2015.

Fricko, O., Havlik, P., Rogelj, J., Klimont, Z., Gusti, M., Johnson, N., Kolp, P., Strubegger, M., Valin, H., Amann, M., Ermolieva, T., Forsell, N., Herrero, M., Heyes, C., Kindermann, G., Krey, V., McCollum, D. L., Obersteiner, M., Pachauri, S., Rao, S., Schmid, E., Schoepp, W. and Riahi, K.: The marker quantification of the Shared Socioeconomic Pathway 2: A middle-of-the-road scenario for the 21st century, Glob. Environ. Chang., 42(July), 251–267, doi:10.1016/j.gloenvcha.2016.06.004, 2017.

Fujimori, S., Hasegawa, T., Masui, T., Takahashi, K., Herran, D. S., Dai, H., Hijioka, Y. and Kainuma, M.: SSP3: AIM implementation of Shared Socioeconomic Pathways, Glob. Environ. Chang., 42, 268–283, doi:10.1016/j.gloenvcha.2016.06.009, 2017.

Gillett, N. P., Kirchmeier-young, M., Ribes, A., Shiogama, H., Hegerl, G. C., Knutti, R., Gastineau, G., John, J. G. and Li, L.: warming since the pre-industrial period, , doi:10.1038/s41558-020-00965-9, 2019.

Goodfellow, I., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., Courville, A. and Bengio, Y.: Generative adversarial networks, Commun. ACM, 63(11), 139–144, doi:10.1145/3422622, 2014.

Gudmundsson, L.,J.B. Bremnes,J.E.Haugen, and T. Engen-Skaugen.: Technical note: Downscaling RCM precipitation to the station scale using statistical transformations—Acomparison ofmethods, Hydrol. Earth Syst. Sci., 16(9), 3383–3390, doi:10.5194/hess-16-3383-2012, 2012.

Guti érrez, J. M., D. San-Mart ń, S. Brands, R. Manzanas, and S. Herrera.: Reassessing statistical downscaling techniques for their robust application under climate change conditions, J. Clim., 26(1), 171–188, doi:10.1175/JCLI-D-11-00687.1, 2013.

Ham, Y. G., Kim, J. H. and Luo, J. J.: Deep learning for multi-year ENSO forecasts, Nature, 573(7775), 568–572, doi:10.1038/s41586-019-1559-7, 2019.

Harris, I., Osborn, T. J., Jones, P. and Lister, D.: Version 4 of the CRU TS monthly high-resolution gridded multivariate climate dataset, Sci. Data, 7(1), 1–19, doi:10.1038/s41597-020-0453-3, 2020.

Hasegawa, T., Sakurai, G., Fujimori, S., Takahashi, K., Hijioka, Y. and Masui, T.: Extreme climate events increase risk of global food insecurity and adaptation needs, Nat. Food, 2(8), 587–595, doi:10.1038/s43016-021-00335-4, 2021.

https://doi.org/10.5194/essd-2021-418
Preprint. Discussion started: 10 December 2021

He, K., Zhang, X., Ren, S. and Sun, J.: Deep residual learning for image recognition, Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit., 2016-Decem, 770–778, doi:10.1109/CVPR.2016.90, 2016.

Ho, C. K., Stephenson, D. B., Collins, M., Ferro, C. A. T. and Brown, S. J.: Calibration strategies a source of additional uncertainty in climate change projections, Bull. Am. Meteorol. Soc., 93(1), 21–26, doi:10.1175/2011BAMS3110.1, 2012.

Huang, B., Thorne, P. W., Banzon, V. F., Boyer, T., Chepurin, G., Lawrimore, J. H., Menne, M. J., Smith, T. M., Vose, R. S. and Zhang, H. M.: Extended reconstructed Sea surface temperature, Version 5 (ERSSTv5): Upgrades, validations, and intercomparisons, J. Clim., 30(20), 8179–8205, doi:10.1175/JCLI-D-16-0836.1, 2017a.

Huang, G., Liu, Z., Van Der Maaten, L. and Weinberger, K. Q.: Densely connected convolutional networks, Proc. - 30th IEEE Conf. Comput. Vis. Pattern Recognition, CVPR 2017, 2017-Janua, 2261–2269, doi:10.1109/CVPR.2017.243, 2017b.

Huang, X.: Deep-learning based climate downscaling using the super-resolution method: a 2, Geosci. Model Dev., (September), 1–18 [online] Available from: https://doi.org/10.5194/gmd-2020-214, 2020.

Jakob Themeßl, M., Gobiet, A. and Leuprecht, A.: Empirical-statistical downscaling and error correction of daily precipitation from regional climate models, Int. J. Climatol., 31(10), 1530–1544, doi:10.1002/joc.2168, 2011.

Jevrejeva, S., Jackson, L. P., Riva, R. E. M., Grinsted, A. and Moore, J. C.: Coastal sea level rise with warming above 2 °C, Proc. Natl. Acad. Sci. U. S. A., 113(47), 13342–13347, doi:10.1073/pnas.1605312113, 2016.

Kadow, C., Hall, D. M. and Ulbrich, U.: Artificial intelligence reconstructs missing climate information, Nat. Geosci., 13(6), 408–413, doi:10.1038/s41561-020-0582-5, 2020.

Kannan, S., Ghosh, S.: A nonparametric kernel regression model for downscaling multisite daily precipitation in the Mahanadi basin. Water Resour. Res. 49 (3), 1360–1385, 2013.

Koch, J., Demirel, M. C. and Stisen, S.: The SPAtial EFficiency metric (SPAEF): Multiple-component evaluation of spatial patterns for optimization of hydrological models, Geosci. Model Dev., 11(5), 1873–1886, doi:10.5194/gmd-11-1873-2018, 2018.

Kriegler, E., Bauer, N., Popp, A., Humpenöder, F., Leimbach, M., Strefler, J., Baumstark, L., Bodirsky, B. L., Hilaire, J., Klein, D., Mouratiadou, I., Weindl, I., Bertram, C., Dietrich, J. P., Luderer, G., Pehl, M., Pietzcker, R., Piontek, F., Lotze-Campen, H., Biewald, A., Bonsch, M., Giannousakis, A., Kreidenweis, U., Müller, C., Rolinski, S., Schultes, A., Schwanitz, J., Stevanovic, M., Calvin, K., Emmerling, J., Fujimori, S. and Edenhofer, O.: Fossil-fueled development (SSP5): An energy and resource intensive scenario for the 21st century, Glob. Environ. Chang., 42(September 2018), 297–315, doi:10.1016/j.gloenvcha.2016.05.015, 2017.

Lafferty, D. C., Sriver, R. L., Haqiqi, I., Hertel, T. W., Keller, K. and Nicholas, R. E.: Statistically bias-corrected and downscaled climate models underestimate the adverse effects of extreme heat on U.S. maize yields, Commun. Earth Environ., 2(1), doi:10.1038/s43247-021-00266-9, 2021.

https://doi.org/10.5194/essd-2021-418
Preprint. Discussion started: 10 December 2021

Ledig, C., Theis, L., Huszár, F., Caballero, J., Cunningham, A., Acosta, A., Aitken, A., Tejani, A., Totz, J., Wang, Z. and Shi, W.: Photo-realistic single image super-resolution using a generative adversarial network, Proc. - 30th IEEE Conf. Comput. Vis. Pattern Recognition, CVPR 2017, 2017-Janua, 105–114, doi:10.1109/CVPR.2017.19, 2017.

Lovino, M. A. , Pierrestegui, M. J. , OV Müller, Berbery, E. H. , GV Müller, & Pasten, M.: Evaluation of historical cmip6 model simulations and future projections of temperature and precipitation in paraguay. Climatic Change, 164, 2021.

Maraun, D., Widmann, M.: Structure of statistical downscaling methods. In: Statistical Downscaling and Bias Correction for Climate Research. Cambridge University Press, Cambridge, pp. 135–140, 2018.

Miao, C., Ashouri, H., Hsu, K. L., Sorooshian, S. and Duan, Q.: Evaluation of the PERSIANN-CDR daily rainfall estimates in capturing the behavior of extreme precipitation events over China, J. Hydrometeorol., 16(3), 1387–1396, doi:10.1175/JHM-D-14-0174.1, 2015.

Miao, C., Su, L., Sun, Q. and Duan, Q.:, A nonstationary bias-correction technique to remove bias in GCM simulations,J. Geophys. Res. Atmos. Res., 175(4449), 238, doi:10.1038/175238c0, 2016.

Moss, R. H., Edmonds, J. A., Hibbard, K. A., Manning, M. R., Rose, S. K., Van Vuuren, D. P., Carter, T. R., Emori, S., Kainuma, M., Kram, T., Meehl, G. A., Mitchell, J. F. B., Nakicenovic, N., Riahi, K., Smith, S. J., Stouffer, R. J., Thomson, A. M., Weyant, J. P. and Wilbanks, T. J.: The next generation of scenarios for climate change research and assessment, Nature, 463(7282), 747–756, doi:10.1038/nature08823, 2010.

Naumann, G., Cammalleri, C., Mentaschi, L. and Feyen, L.: Increased economic drought impacts in Europe with anthropogenic warming, Nat. Clim. Chang., 11(6), 485–491, doi:10.1038/s41558-021-01044-3, 2021.

Norris, J., Hall, A., Chen, D., Thackeray, C. W., & Madakumbura, G. D., Assessing the representation of synoptic variability associated with California extreme precipitation in CMIP6 models. Journal of Geophysical Research: Atmospheres, 126, e2020JD033938, 2021.

O'Neill, B. C., Tebaldi, C., Van Vuuren, D. P., Eyring, V., Friedlingstein, P., Hurtt, G., Knutti, R., Kriegler, E., Lamarque, J. F., Lowe, J., Meehl, G. A., Moss, R., Riahi, K. and Sanderson, B. M.: The Scenario Model Intercomparison Project (ScenarioMIP) for CMIP6, Geosci. Model Dev., 9(9), 3461–3482, doi:10.5194/gmd-9-3461-2016, 2016.

Papalexiou, S. M., Rajulapati, C. R., Clark, M. P. and Lehner, F.: Robustness of CMIP6 Historical Global Mean Temperature Simulations: Trends, Long-Term Persistence, Autocorrelation, and Distributional Shape, Earth's Future., 8(10), doi:10.1029/2020EF001667, 2020.

Parsons, L. A., Brennan, M. K., Wills, R. C. J. and Proistosescu, C.: Magnitudes and Spatial Patterns of Interdecadal Temperature Variability in CMIP6, Geophys. Res. Lett., 47(7), doi:10.1029/2019GL086588, 2020.

Phillips, T. J. and Gleckler, P. J.: Evaluation of continental precipitation in 20th century climate simulations: The utility of multimodel statistics, Water Resour. Res., 42(3), doi:10.1029/2005WR004313, 2006.

Qi, W. H., Xiong, Y., Li, Y., Zhang, G., Hu, H. and Wei, Y.: Deformable Conv, Iccv 2017, 6003 [online] Available from: https://github.com/msracver/Deformable-ConvNets, 2017.

Quante, L., Willner, S. N., Middelanis, R. and Levermann, A.: Regions of intensification of extreme snowfall under future warming, Sci. Rep., 11(1), 1–9, doi:10.1038/s41598-021-95979-4, 2021.

Reichstein, M., Camps-Valls, G., Stevens, B., Jung, M., Denzler, J., Carvalhais, N. and Prabhat: Deep learning and process understanding for data-driven Earth system science, Nature, 566(7743), 195–204, doi:10.1038/s41586-019-0912-1, 2019.

Renard, D. and Tilman, D.: National food production stabilized by crop diversity, Nature, 571(7764), 257–260, doi:10.1038/s41586-019-1316-y, 2019.

Riahi, K., van Vuuren, D. P., Kriegler, E., Edmonds, J., Bauer, N., Calvin, K., Dellink, R., Fricko, O., Lutz, W., Popp, A., Crespo Cuaresma, J., Leimbach, M., Jiang, L., Kram, T., Rao, S., Emmerling, J., Ebi, K., Hasegawa, T., Havlik, P., Humpenöder, F., Da Silva, A., Smith, S., Stehfest, E., Bosetti, V., Eom, J., Gernaat, D., Masui, T., Rogelj, J., Strefler, J., Drouet, L., Krey, V., Luderer, G., Harmsen, M., Takahashi, K., Baumstark, L., Doelman, J., Kainuma, M., Klimont, Z., Marangoni, G., Lotze-Campen, H. and Obersteiner, M.: The Shared Socioeconomic Pathways and their Energy, Land Use, and Greenhouse Gas Emissions Implications: An Overview, 1–38, 2016.

Seif, G. and Androutsos, D.: Large receptive field networks for high-scale image super-resolution, IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit. Work., 2018-June, 876–885, doi:10.1109/CVPRW.2018.00120, 2018.

Shi, W., Caballero, J., Huszar, F., Totz, J., Aitken, A. P., Bishop, R., Rueckert, D. and Wang, Z.: Real-Time Single Image and Video Super-Resolution Using an Efficient Sub-Pixel Convolutional Neural Network, Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit., 2016-Decem, 1874–1883, doi:10.1109/CVPR.2016.207, 2016.

Shrestha, R. R., Schnorbus, M. A., Werner, A. T. and Zwiers, F. W.: Evaluating hydroclimatic change signals from statistically and dynamically downscaled GCMs and hydrologic models, J. Hydrometeorol., 15(2), 844–860, doi:10.1175/JHM-D-13-030.1, 2014.

Stockhause, M. and Lautenschlager, M.: CMIP6 data citation of evolving data, Data Sci. J., 16, 1–13, doi:10.5334/dsj-2017-030, 2017.

Tang, G. , Clark, M. P. , & Papalexiou, S. M.: Sc-earth: a station-based serially complete earth dataset from 1950 to 2019. *Journal of Climate*, 1-47,2021.

Tang, G., Clark, M. P., Papalexiou, S. M., Newman, A. J., Wood, A. W., Brunet, D., and Whitfield, P. H.: EMDNA: an Ensemble Meteorological Dataset for North America, Earth Syst. Sci. Data, 13, 3337–3362, https :// doi.org/ 10.5194/ essd-13-3337-2021, 2021.

Taylor, Karl E . Summarizing multiple aspects of model performance in a single diagram[J]. Journal of Geophysical Research Atmospheres, 106(D7):7183-7192, 2001.

Tian, L., Sun, J., Chang, J., Xia, J., Zhang, Z., Kolomenskii, A. A., Schuessler, H. A. and Zhang, S.: Retrieval of gas concentrations in optical spectroscopy with deep learning, Meas. J. Int. Meas. Confed., 182(July), 109739, doi:10.1016/j.measurement.2021.109739, 2021.

https://doi.org/10.5194/essd-2021-418
Preprint. Discussion started: 10 December 2021

Toms, B. A., Barnes, E. A. and Ebert-Uphoff, I.: Physically Interpretable Neural Networks for the Geosciences: Applications to Earth System Variability, J. Adv. Model. Earth Syst., 12(9), 1–20, doi:10.1029/2019MS002002, 2020.

Tong, X. Y., Xia, G. S., Lu, Q., Shen, H., Li, S., You, S. and Zhang, L.: Land-cover classification with high-resolution remote sensing images using transferable deep models, Remote Sens. Environ., 237, 111322, doi:10.1016/j.rse.2019.111322, 2020.

Touma, D., Stevenson, S., Lehner, F. and Coats, S.: Human-driven greenhouse gas and aerosol emissions cause distinct regional impacts on extreme fire weather, Nat. Commun., 12(1), 1–8, doi:10.1038/s41467-020-20570-w, 2021.

Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, Ł. and Polosukhin, I.: Attention is all you need, Adv. Neural Inf. Process. Syst., 2017-Decem(Nips), 5999–6009, 2017.

van Vuuren, D. P., Kriegler, E., O'Neill, B. C., Ebi, K. L., Riahi, K., Carter, T. R., Edmonds, J., Hallegatte, S., Kram, T., Mathur, R. and Winkler, H.: A new scenario framework for Climate Change Research: Scenario matrix architecture, Clim. Change, 122(3), 373–386, doi:10.1007/s10584-013-0906-1, 2014.

Wang, H., Liao, K., Yan, B. and Ye, R.: Deep residual network for single image super-resolution, ACM Int. Conf. Proceeding Ser., 66–70, doi:10.1145/3341016.3341030, 2019a.

Wang, J., Yang, B., Ljungqvist, F. C. and Zhao, Y.: The relationship between the Atlantic Multidecadal Oscillation and temperature variability in China during the last millennium, J. Quat. Sci., 28(7), 653–658, doi:10.1002/jqs.2658, 2013.

Wang, X., Chan, K. C. K., Yu, K., Dong, C. and Loy, C. C.: EDVR: Video restoration with enhanced deformable convolutional networks, IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit. Work., 2019-June, 1954–1963, doi:10.1109/CVPRW.2019.00247, 2019b.

Wang, X., Yu, K., Wu, S., Gu, J., Liu, Y., Dong, C., Qiao, Y. and Loy, C. C.: ESRGAN: Enhanced super-resolution generative adversarial networks, Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics), 11133 LNCS, 63–79, doi:10.1007/978-3-030-11021-5_5, 2019c.

Wei, X., Wang, G., Feng, D., Duan, Z., Hagan, F.T,D., Tao, L., Miao, L., Su, B., Tong, J.: Deep-Learning-Based Harmonization and Super-Resolution of Near-Surface Air Temperature from CMIP6 Models (1850-2100),Zenodo, https://doi.org/10.5281/zenodo.5746632,2021

White, R. H., and R. Toumi.: The limitations of bias correcting regional climate model inputs, Geophys. Res. Lett., 40, 2907–2912, doi:10.1002/grl.50612, 2013.

Wunderling, N., Willeit, M., Donges, J. F. and Winkelmann, R.: Global warming due to loss of large ice masses and Arctic summer sea ice, Nat. Commun., 11(1), 1–8, doi:10.1038/s41467-020-18934-3, 2020.

Xu, J., Chae, Y., Stenger, B. and Datta, A.: Dense bynet: Residual dense network for image super resolution, Proc. - Int. Conf. Image Process. ICIP, 71–75, doi:10.1109/ICIP.2018.8451696, 2018.

Yamaguchi, M., Chan, J. C. L., Moon, I. J., Yoshida, K. and Mizuta, R.: Global warming changes tropical cyclone translation speed, Nat. Commun., 11(1), 1–7, doi:10.1038/s41467-019-13902-y, 2020.

Yoon, Y., Jeon, H. G., Yoo, D., Lee, J. Y. and Kweon, I. S.: Learning a Deep Convolutional Network for Light-Field Image Super-Resolution, Proc. IEEE Int. Conf. Comput. Vis., 2015-Febru, 57–65, doi:10.1109/ICCVW.2015.17, 2015.

You, Q., Cai, Z., Wu, F., Jiang, Z., Pepin, N. and Shen, S. S. P.: Temperature dataset of CMIP6 models over China: evaluation, trend and uncertainty, Clim. Dyn., 57(1–2), 17–35, doi:10.1007/s00382-021-05691-2, 2021.

Yuan, Q., Shen, H., Li, T., Li, Z., Li, S., Jiang, Y., Xu, H., Tan, W., Yang, Q., Wang, J., Gao, J. and Zhang, L.: Deep learning in environmental remote sensing: Achievements and challenges, Remote Sens. Environ., 241, doi:10.1016/j.rse.2020.111716, 2020.

Yuan, X., Hamdi, R., Ochege, F. U., Kurban, A. and De Maeyer, P.: The sensitivity of global surface air temperature to vegetation greenness, Int. J. Climatol., 41(1), 483–496, doi:10.1002/joc.6633, 2021.

Zang, Z. D. L. G. 1: Superior PM 2 . 5 Estimation by Integrating Aerosol Fine Mode Data from the Himawari-8 Satellite in Deep and Classical Machine Learning Models, Remote Sens., 2021.

Zhang, Y., Li, K., Li, K., Wang, L., Zhong, B. and Fu, Y.: Image super-resolution using very deep residual channel attention networks, Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics), 11211 LNCS, 294–310, doi:10.1007/978-3-030-01234-2_18, 2018.

Zheng, G., Li, X., Zhang, R. H. and Liu, B.: Purely satellite data-driven deep learning forecast of complicated tropical instability waves, Sci. Adv., 6(29), doi:10.1126/sciadv.aba1482, 2020.

Zhukov, A. and Gushcha, S.: Stop Global Warming, J. Educ. Heal. Sport, 10(3), 18, doi:10.12775/jehs.2020.10.03.002, 2020.