

APPLICATION PAPER

Climate model-driven seasonal forecasting approach with deep learning

Alper Unal¹ , Busra Asan², Ismail Sezen¹, Bugra Yesilkaynak², Yusuf Aydin¹ , Mehmet Ilicak¹ and Gozde Unal³

¹Eurasia Institute of Earth Sciences, Istanbul Technical University, Istanbul, Türkiye

²Department of Computer Engineering, Istanbul Technical University, Istanbul, Türkiye

³Department of AI and Data Engineering, Istanbul Technical University, Istanbul, Türkiye

Corresponding author: Alper Unal; Email: alper.unal@itu.edu.tr

Received: 22 May 2023; **Accepted:** 30 May 2023

Keywords: climate change; deep neural networks; machine learning; seasonal forecast

Abstract

Understanding seasonal climatic conditions is critical for better management of resources such as water, energy, and agriculture. Recently, there has been a great interest in utilizing the power of Artificial Intelligence (AI) methods in climate studies. This paper presents cutting-edge deep-learning models (UNet++, ResNet, PSPNet, and DeepLabv3) trained by state-of-the-art global CMIP6 models to forecast global temperatures a month ahead using the ERA5 reanalysis dataset. ERA5 dataset was also used for fine-tuning as well performance analysis in the validation dataset. Ten different setups (with CMIP6 and CMIP6 + ERA5 fine-tuning) including six meteorological parameters (i.e., 2 m temperature, 10 m eastward component of wind, 10 m northward component of wind, geopotential height at 500 hPa, mean sea-level pressure, and precipitation flux) and elevation were used with both four different algorithms. For each model 14 different sequential and nonsequential temporal settings were used. The mean absolute error (MAE) analysis revealed that UNet++ with CMIP6 with 2 m temperature + elevation and ERA5 fine-tuning model with “Year 3 Month 2” temporal case provided the best outcome with an MAE of 0.7. Regression analysis over the validation dataset between the ERA5 data values and the corresponding AI model predictions revealed slope and R^2 values close to 1 suggesting a very good agreement. The AI model predicts significantly better than the mean CMIP6 ensemble between 2016 and 2021. Both models predict the summer months more accurately than the winter months.

Impact Statement

This paper discusses the use of novel developments in the machine-learning field in seasonal forecasting. Traditionally, climate models are used to simulate physical, chemical and biological processes in the atmosphere to generate climate projections. Machine-learning methods are gaining popularity in different fields. A team of computer scientists and earth scientists worked on this paper that investigates the use of machine-learning algorithms along with climate models for seasonal forecasts, which are critical for better resource management.

1. Introduction

Seasonal forecast is defined as a variety of potential climate changes that are likely to occur in the coming months and seasons (Pan et al., 2022). This is crucial for governments and decision makers to better manage natural resources such as water, energy, and agriculture, as well as protect human health (Yuan

et al., 2019; Talukder et al., 2021). For example, crop producers use seasonal forecasts to make decisions about the timing of planting and harvesting, field fertilization and water management (Klemm and McPherson, 2017). Weather plays an important role in energy supply and demand (Felice et al., 2015); hence, an accurate forecast of future weather conditions could increase the effectiveness and dependability of energy management at the local and national levels given the requirement to maintain the balance between electricity production and demand. Accurate forecasting of extreme events such as storms, heatwaves, droughts and floods is required to improve disaster preparedness (Liu et al., 2022).

Weather prediction at shorter timescales such as daily to monthly depends on the understanding of physical processes in the atmosphere as well as interactions among the atmosphere, oceans and land. An important distinction must be made between dynamical predictions, which use intricate physical numerical models and statistical predictions, that use regional historical relationships between physical variables like temperature and precipitation (Roads et al., 2003; Lorenzoni and Pidgeon, 2006; Troccoli, 2010; Klemm and McPherson, 2017; Franzke et al., 2022). Two approaches for constraining climate predictions based on past climate change include large ensembles of simulations from computationally efficient models and small ensembles of simulations from state-of-the-art coupled ocean–atmosphere General Circulation Models (GCMs; Stott and Forest, 2007). GCMs are frequently used in studies related to the impacts of large-scale climate change (Fujihara et al., 2008). High-resolution climate data from current global climate models are provided using Regional Climate Downscaling (RCD) techniques (Laprise, 2008; Scinocca et al., 2016). Several programs such as THORPEX, DEMETER and EUPORIAS have been launched in practice to work toward seasonal forecasting (Toth et al., 2007; Klemm and McPherson, 2017).

In 1995, Coupled Model Intercomparison Projects (CMIP) began as a comparison of a few pioneering global coupled climate models and their outputs are used by different organizations around the world, such as the IPCC to better understand past, present, and future climate change (Wang et al., 2021; Xu et al., 2021; Liu et al., 2022). CMIP6 is the most recent phase of the CMIP. The CMIP6 platform, which began in 2015, offers the most up-to-date multi-model datasets. Simulation outputs from more than 100 different climate models produced by more than 50 different modeling groups contributed to CMIP6. In addition to historical studies, seasonal forecasts for different emission scenarios are provided (Fan et al., 2020; Turnock et al., 2020; Zhang and Li, 2021; Liu et al., 2022).

In recent years, big data, effective supercomputers with graphics processing units (GPUs), and scientific interest in novel algorithms and optimization techniques proved to be significant turning points in machine-learning history. Machine learning has recently been a hot topic in climate studies. Tyagi et al. (2022) reviewed a number of studies that applied the different machine-learning/deep-learning algorithms in flash drought (FD) studies. Luo et al. (2022) used a Bayesian deep-learning approach to near-term climate prediction in the North Atlantic. Bochenek and Ustrnul (2022) investigated the top 500 scientific articles about machine learning in the field of climate and numerical weather prediction that have been published since 2018. Anochi et al. (2021) evaluated different machine-learning methods for precipitation prediction in South America. Zhang and Li (2021) and Feng et al. (2022) used deep-learning algorithms to downscale hydroclimatic data of CMIP6 simulations in China.

This study aimed to improve seasonal temperature forecasts using both climate models and machine-learning algorithms. Specifically, the objective was to utilize the power of CMIP6 physical models along with the European Centre for Medium-Range Weather Forecasts (ECMWF) Reanalysis 5th Generation (ERA5) dataset (created using data assimilation and model forecasts) while utilizing specifically deep neural network-learning methods for a better global seasonal forecast of 2 m temperature.

2. Materials and Methods

2.1. Training data

This study has utilized the monthly averaged meteorological parameters (i.e., 2 m temperature, 10 m eastward component of wind, 10 m northward component of wind, geopotential height at 500 hPa, mean sea-level pressure, and precipitation flux) from nine fully coupled Earth System Models (ESM) that

participated in the CMIP6 (Eyring et al., 2016). These models are ACCESS-CM2, CNRM-CM6-1-HR, GISS-E2-1-H, NorESM2-MM, CESM2-WACCM, EC-Earth3-Veg, MPI-ESM1-2-HR, MIROC-ES2L, and IPSL-CM6A-LR. For each ESM model, only one ensemble member was used and the time slice for the coupled models is chosen from the historical period (i.e., from year 1850 to 2014). For each ensemble, 1,700 data instances are separated for training and 100 instances for testing.

For validation and fine-tuning, monthly averaged 2 m temperature data from the ERA5 reanalysis dataset Copernicus Climate Change Service (C3S) (2017) have been used. The ERA5 dataset is partitioned into two different time slices: years 1973–2016 (400 samples) are used for fine-tuning, while years 2016–2021 (116 samples) are used for evaluation of the trained models. In order to estimate the performance of the deep-learning model (which is called as AI model from now on) between 2016 and 2021, we also used the multi-model mean of the CMIP6 models using the IPCC SSP5–8.5 scenario which is selected since there was no significant reduction in carbon emissions after 2014 when the historical simulations ended.

2.2. Model architecture

As this study focuses on spatiotemporally varying data, an encoder–decoder-based architecture UNet++ (Zhou et al., 2018), which is based on the original UNet architecture (Ronneberger et al., 2015) is adapted. In UNet++, the skip connections of the UNet are re-designed to minimize the semantic gap between the feature maps coming from the sub-networks of the encoder and the decoder, making the learning easier. Our model specifically employs convolutional neural network (CNN) layers due to the nature of the input data. We construct the UNet++ in order to perform a prediction task, which is explained below. The architecture includes a contracting path, that is, an encoder part, which summarizes the information by reducing the size of the input image and increasing the number of channels. This downsampling operation results in spatial information loss due to the compression of the input. The UNet++, as in the original UNet, introduces skip connections that reduce the information loss after the bottleneck layer and recovers fine-grained details. Skip connections aggregate information from different resolution levels in order to increase accuracy and speed up the convergence. In the expansive path, that is, the decoder part, skip connections concatenate the outputs of each downsampling layer to corresponding upsampling layers, aiming at image reconstruction that is at the same spatial resolution as that of the input. In addition to UNet++, by training several state-of-the-art CNN architectures a benchmark has been created. As the baseline model, the classical CNN architecture ResNet is used. Furthermore, similar to UNet++, results of the other segmentation models such as DeepLabV3 Chen et al. (2017) and PSPNet Zhao et al. (2017) are included in the benchmark.

Figure 1a depicts the block diagram of the neural network model that we construct for the seasonal forecast of temperature. In addition to 2 m temperature, an elevation map is stacked to the input as ancillary data to investigate whether it can improve the prediction performance. Along with elevation and temperature data, experiments are done by stacking more climate variables per month such as u_{10} (10 m eastward component of wind), v_{10} (10 m northward component of wind), z_{g500} (Geopotential height at 500 hPa), p_{sl} (mean sea-level pressure) and pr (precipitation flux) and giving into the network. We use two different experimental settings to evaluate the effectiveness of using fully historical input data versus periodic data. The first setting is designed to see the performance of the historical data stacked sequentially in a temporal manner such as from $t - 1$ to $t - 6$, $t - 12$ and up to $t - 36$. However in the second set, the target month and its neighbors are stacked in a yearly manner from historical data to assess the effect of periodicity. Months before the target prediction, the previous years' target month and months before and after them (e. g. neighboring months) are stacked as the input. The number of previous lag years as well as the number of preceding and succeeding months are selected as hyper-parameters. We have explored lag years 1 to 4 and preceding/succeeding 1 and 2 months. Two examples for these settings are given in 1b. The input to the Encoder network consists of the maps (temperature and elevation) corresponding to the month $t - 1$, which is one previous to the target month t , $t - 12 \pm \Delta t$, $t - 24 \pm \Delta t$, and $t - 36 \pm \Delta t$, where Δt could be either one of $\{1, 2\}$. Specifically, we selected $\Delta t = 2$ in order to account for possible seasonal monthly shifts. The overall concatenated input tensor goes through the UNet++ model, and a single prediction map for

Adam optimizer Kingma and Ba (2015). The model is optimized for 40 epochs and early stopping is used to avoid over-fitting. After each convolutional layer, a batch normalization layer is used. The batch size is chosen as 16. The training process is performed on an NVIDIA RTX A5000 GPU and the results are delivered on average after 3–4 hr of training.

During validation, the loss versus iterations is monitored and the model with the lowest validation error is selected. During inference/test time, the input is formed by the monthly temperature data coming from CMIP6 temperature maps that are stacked and given to the model in a feedforward evaluation. In each evaluation experiment, the target is defined as the month after the latest month in the input. After training with CMIP6 data, the selected model is further fine-tuned with ERA5 t2m data in order to increase the capability of the model in real-world forecasting scenarios. The fine-tuning process is carried out by following the same process as in the training. Monthly temperature data taken from ERA5 and the one-channel elevation map are stacked as a multichannel input and given to the network. During the fine-tuning stage, layers of the network are trained with the same learning rate and a weight decay of $1e - 5$ for 10 epochs with early stopping.

As an evaluation metric, the mean absolute error (MAE) (2) is chosen, where each x_{ij} and y_{ij} correspond to the predicted temperature and the ground truth temperature value of the corresponding grid, respectively. D refers to the number of longitudes, and M refers to the number of latitudes in the 2D spatial grid. All MAEs are summed and averaged across the temperature map to measure the overall error:

$$MAE = \frac{1}{D} \sum_{j=1}^D \frac{1}{M} \sum_{i=1}^M |x_{ij} - y_{ij}|. \tag{2}$$

3. Results and Discussion

Average MAE results over the validation dataset for temporal cases are provided in Table 1. As seen in this table, MAE ranges between 0.955 and 1.162. For comparison purposes, the MAE of the persistence forecast test (over the ERA5 validation dataset by copying the previous month’s temperature value as the target month’s prediction) is estimated as 2.62. This indicates that all models have improved the

Table 1. Mean Absolute Error (MAE) values as estimated for the entire domain (lat:192 × lon:288) for each simulation conducted: 6 models × 14 cases = 84 simulations

Temporal setting	CMIP6 train			CMIP6 train + ERA5 fine-tune		
	M1	M2	M3	M4	M5	M6
6 months	1.136	1.123	1.162	1.030	1.077	1.046
months	1.031	1.024	1.014	1.004	1.028	1.002
months	1.020	1.027	1.013	0.993	0.997	0.999
months	0.998	0.992	0.984	0.981	0.987	0.983
months	1.007	1.006	0.989	1.008	1.002	0.976
months	0.991	0.979	0.984	0.990	0.981	1.003
year 1 month	1.024	1.024	0.989	0.999	0.989	0.983
year 2 months	1.011	1.008	1.008	0.997	0.986	0.984
years 1 month	0.999	0.995	0.996	0.977	0.967	0.967
years 2 months	0.996	0.986	0.980	0.973	0.969	0.979
years 1 month	0.981	0.977	0.986	0.988	0.973	0.967
years 2 months	0.995	0.989	0.978	0.993	0.969	0.975
years 1 month	0.982	0.989	0.966	0.982	0.965	0.962
years 2 months	0.998	0.984	0.967	0.992	0.965	0.955

Note: M1 and M4 are UNet; M2 and M5 are UNet with elevation; M3 and M6 are UNet++ with elevation.

Cases	Africa	North America	Europe	Asia	Winter	Spring	Summer	Fall	Overall
yr0m6	14	14	13	14	14	14	14	12	14
yr0m12	11	13	6	12	12	10	11	11	12
yr0m18	12	11	11	13	10	11	12	13	13
yr0m24	9	7	14	8	7	10	10	9	8
yr0m30	13	10	10	11	13	12	13	10	11
yr0m36	6	7	3	5	3	7	5	7	6
yr1m1	10	12	7	11	12	8	8	14	10
yr1m2	7	8	1	6	6	4	7	3	5
yr2m1	8	9	13	7	9	13	9	4	9
yr2m2	5	5	8	9	8	6	6	9	7
yr3m1	4	4	9	2	5	5	4	6	4
yr3m2	1	3	2	1	4	1	1	1	1
yr4m1	3	2	4	4	1	3	3	5	3
yr4m2	2	1	5	3	2	2	2	2	2

Figure 2. MAE ranks of Model 6 for 14 temporal cases over four continents and four seasons.

MAE significantly with respect to the persistence forecast baseline. It should be noted that out of six models, for each temporal case, 50% (or seven cases) of the lowest MAEs occur for Model 6 (M6). Three of the minimum MAEs occur for M5, and another three cases occur for M4, whereas only one minimum MAE occurs for M2. It is clear that ERA5 fine-tune has improved the performance significantly and using UNet++ with elevation is the best available model. Therefore, Model 6 was selected for the rest of the analysis. In order to choose the best temporal case for Model 6, we estimate the MAEs and rank them for four main continents (i.e., Africa, North America, Europe, and Asia) as well as their distribution among different seasons (i.e., Winter, Spring, Summer, and Fall). These values along with overall (as estimated over continents and seasons) are given in Figure 2. As seen in this figure, sequential cases (such as month 6, month 12, and even month 36) have higher ranks (hence lower performances) as compared to nonsequential cases. This is possibly due to the fact that the latter is able to recognize the strong seasonality in the data while sequential cases lack this ability. Among the nonsequential cases “Year 3 Month 2” case has the best performance as it has the best MAE rank for Fall, Spring, and Summer (and fourth for Winter). This case is the best for Africa and Asia continents, second for Europe, and third for North America. The overall rank for this case is estimated to be number one as well. Therefore, “Year 3 Month 2” is selected as the temporal case for Model 6 (CMIP6 with ERA5 fine-tune with UNet++ with elevation) for the rest of the analysis.

For the selected temporal case we have conducted additional experiments where additional meteorological parameters were used. The results of the new experiments are given below in Table 2. This table indicates the MAE values as estimated for the entire domain (lat:192 × lon:288). According to the results presented in the table, “elevation + temperature” simulation has the minimum MAE among the examined variables. It should be noted that we have conducted simulations where we add meteorological parameters paired with temperature one by one and another simulation where all parameters were included. Although all parameter simulation with ERA5 fine-tune has the lowest MAE (0.991), it is still higher than the original model where elevation + temperature is used (0.975 with ERA5 fine-tune).

The selected model settings were used in other deep-learning architectures (ResNet, PSPNet, and DeepLabV3) to investigate their performance. The results are provided in Table 3. As seen in this table, UNet++ with an MAE of 0.975 has the best performance while other models’ MAE values range between 1.025 and 1.738. According to these results, the UNet++ was selected to be used in this study.

Table 2. Mean Absolute Error (MAE) values as estimated for the entire domain (lat:192 × lon: 288)

	CMIP6 train	CMIP6 train + ERA5 fine-tune
Variables	M1	M2
elevation + temperature	0.976	0.975
pr + temperature	1.017	1.008
psl + temperature	1.011	1.007
uas + temperature	1.006	1.011
vas + temperature	1.005	1.006
zg500 + temperature	1.153	1.136
elevation+pr + vas + uas + zg500 + psl + temperature	1.039	0.991

Note: Both M1 and M2 are UNet++. M1 is baseline, M2 is fine-tuned with ERA5. For this comparison, 7 ensembles instead of 9 used for training since vas and uas variables are not included in 2 ensembles of CMIP6.

Table 3. Mean absolute error (MAE) values as estimated for the entire domain (lat:192 × lon:288)

	CMIP6 train	CMIP6 train + ERA5 fine-tune
Models	M1	M2
UNet++	0.978	0.975
ResNet	1.198	1.738
PSPNet	1.042	1.139
DeepLabV3	1.081	1.025

Note: Both M1 and M2 are UNet++. M1 is baseline, M2 is fine-tuned with ERA5. All experiments are done with elevation and temperature.

We compared the performance of the selected artificial intelligence (AI) model to the ensemble mean of the CMIP6 models. Figure 3 shows the MAE for the AI model and the mean of the CMIP6 ensemble as a function of time for the selected continents during the study period from 2016 to 2022. In Africa, both the AI and the CMIP6 model means have a similar error level (i.e., 0.27 versus 0.34) (Figure 3a). This is probably due to the fact that the climate in Africa is fairly uniform as a function of latitude and both models capture the overall climatology well. The AI model performs better compared to the mean CMIP6 in Asia and Europe. Although the inter-annual variability in the error between two different models overlaps, AI has significantly lower bias values. We believe that since the AI model has been trained by the CMIP6 ensemble, the model might inherit similar inter-annual variability. AI model's MAE value for February 2020 for Europe, Asia, and North America is significantly lower as compared to the CMIP6 model (a difference of 3.2° for Europe, almost 1° for Asia, and 0.5° for North America), while they estimate closer values for Africa. The error in North America is the only place where the CMIP6 ensemble (0.7) is slightly better than the AI model (0.71) that needs further investigation. In both models, the error increases in the winter months, indicating that the models do not accurately represent the cold climate in the northern hemisphere.

The spatial distributions of the MAE fields for summer and winter for the AI and mean CMIP6 models are shown in Figure 4. Summer-time MAE in the AI model is fairly uniform and approximately 1.5°C over the continents (Figure 4a1). In contrast, the mean CMIP6 shows a relatively larger error (up to 5°C) in high-topography regions such as the Himalayas in Asia, the Andes in South America, the Rockies in North America, and the Alps in Europe (Figure 4a2). The MAE pattern in winter of the AI model is similar to the mean CMIP6 in high latitudes in the northern hemisphere (Figure 4b1,b2). This indicates that large-scale Jetstream bias from the CMIP6 models is responsible for the AI's poor performance over Siberia and northern America. Once again, the error in winter is larger than in summer in both models, as we have shown in Figure 3. The performance of the AI model in terms of MAE is significantly better than that of the mean CMIP6 in both summer and winter.

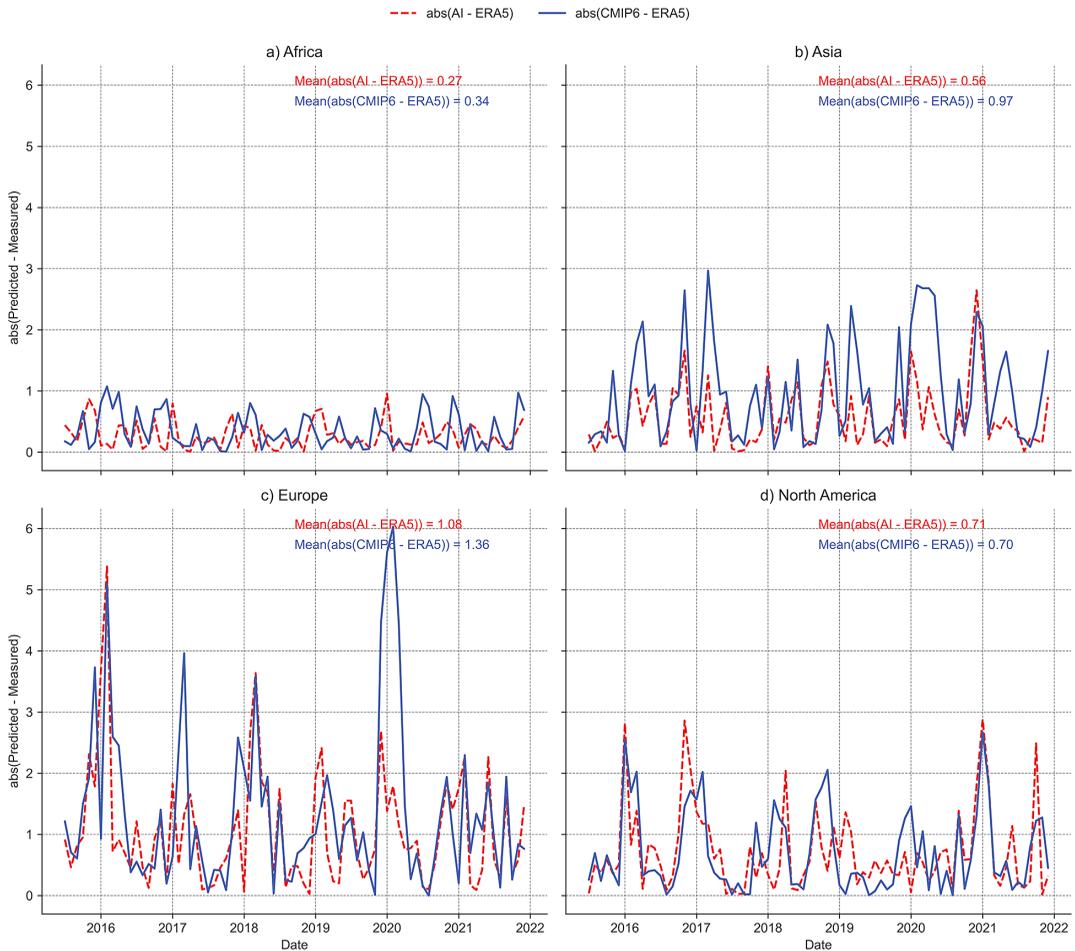


Figure 3. MAE results of AI and CMIP6 models for four different continents (a) Africa (b) Asia (c) Europe and (d) North America as estimated over the validation dataset.

Next, at every grid point of the global domain (lat: 192, lon: 288), we calculated the temperature anomalies for each month to remove the mean of the month of that grid point. Then, we computed the scatter plot of absolute errors (AEs) of the mean CMIP6 and AI models for all grids as a function of these temperature anomalies (Figure 5a). The AI model performs better when the temperature anomalies are between -5°C and 5°C indicating that if a particular month is around the monthly mean, then the AI model predicts significantly better than the CMIP6 mean. However, if the month is part of an extreme event such as very cold ($\Delta T \approx -10^{\circ}\text{C}$) or very hot ($\Delta T \approx 10^{\circ}\text{C}$), AI's performance is getting closer to the CMIP6 mean. To better understand the performance of the selected model, the box plots of the calculated AE based on temperature anomalies are given in Figure 5b. In all AI versus CMIP6 error bars for each temperature bin, AI model has significantly lower error values (for the median values and 25th and 75th percentiles). This outcome is even more pronounced, especially for the bins between -2 and $+2$ (as shown on the x-axis).

In addition to scatter plot and box plots of AE values, other statistical values such as R^2 are used for understanding the relationship between observed and predicted temperature for all continents based on the results of the selected scenario. The results indicate that predicted values are well-fitting with the observed values in all continents. The R^2 values of all continents are close to 1, which imply the power of the AI model in seasonally predicting temperature around the world (see Supplementary Material S1).

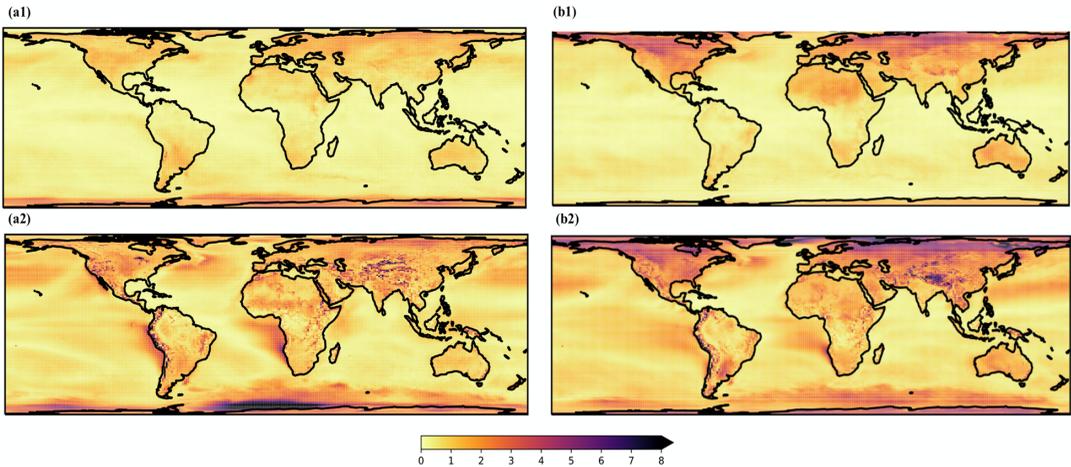


Figure 4. (a) MAE fields of AI model in Summer (a1); CMIP6 model in Summer (a2); AI model in Winter (b1); and CMIP6 model in Winter (b2) for the validation dataset.

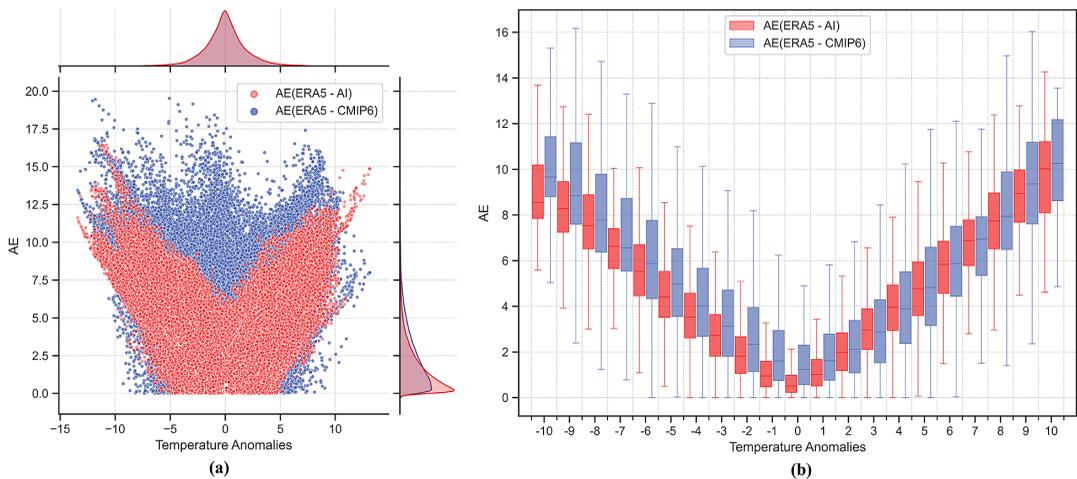


Figure 5. Absolute error plots of CMIP6 and AI model results for the validation dataset: (a) Scatter (b) Box plots.

4. Conclusion

We employed an advanced encoder–decoder model (UNet++) trained by state-of-the-art global CMIP6 ESM to forecast global temperatures a month ahead using the ERA5 reanalysis dataset. This study is a proof of concept for the use of this model in a complex climate system. We found that the deep-learning model predicts significantly better than the mean CMIP6 ensemble between 2016 and 2021. The AI model predicts the summer months more accurately than the winter months, similar to the mean CMIP6. Comparison of the results to other algorithms such as ResNet, PSPNet, and DeepLabv3 revealed that UNet++ provided the lowest MAE. Developed model is also used with additional meteorological parameters (u10, v10, zg500, psl, and pr) alongside 2 m temperature and elevation. The findings are striking since the simplest model (2 m temperature + elevation) provided the best result. In the future, we plan to improve our forecast time to seasonal predictions, that is, 3 months ahead.

Author contribution. Conceptualization: A.U., G.U., M.I.; Data curation: I.S., B.Y.; Data visualization: B.A., I.S., Y.A., A.U.; Methodology: A.U., G.U., M.I.; Software: B.A., I.S., B.Y., A.U.; Writing—original draft: B.A., Y.A.; Writing—review and editing: A.U., M.I., G.U. All authors approved the final submitted draft.

Competing interest. The authors declare none.

Data availability statement. Replication data and code can be found in Github: <https://www.github.com/ituvisionlab/climate-ai>.

Funding statement. This work received no specific grant from any funding agency, commercial or not-for-profit sectors.

Provenance statement. This article is part of the Climate Informatics 2023 proceedings and was accepted in *Environmental Data Science* on the basis of the Climate Informatics peer review process.

Supplementary material. The supplementary material for this article can be found at <https://doi.org/10.1017/eds.2023.24>.

References

- Anochi JA, de Almeida VA and de Campos Velho HF** (2021) Machine learning for climate precipitation prediction modeling over South America. *Remote Sensing* 13, 2468. <https://doi.org/10.3390/rs13132468>
- Bochenek B and Ustrnul Z** (2022) Machine learning in weather prediction and climate analyses—Applications and perspectives. *Atmosphere* 13, 180. <https://doi.org/10.3390/atmos13020180>
- Chen LC, George P, Florian S and Adam H** (2017) Rethinking atrous convolution for semantic image segmentation. Volume, abs/1706.05587, Preprint, arXiv:1706.05587.
- Copernicus Climate Change Service (C3S)** (2017) *ERA5: Fifth generation of ECMWF atmospheric reanalyses of the global climate. Copernicus Climate Change Service Climate Data Store (CDS) 2022*. Available at <https://cds.climate.copernicus.eu/cdsapp#!/home> (accessed September 2021).
- Eyring V, Bony S, Meehl GA, Senior CA, Stevens B, Stouffer RJ and Taylor KE** (2016) Overview of the coupled model Intercomparison project phase 6 (CMIP6) experimental design and organization. *Geoscientific Model Development* 9(2016), 1937–1958. <https://doi.org/10.5194/gmd-9-1937-2016>
- Fan X, Duan Q, Shen C, Wu Y and Xing C** (2020) Global surface air temperatures in CMIP6: Historical performance and future changes. *Environmental Research Letters* 15(10), 104056. <https://doi.org/10.1088/1748-9326/abb051>
- Felice MD, Alessandri A and Catalano F** (2015) Seasonal climate forecasts for medium-term electricity demand forecasting. *Applied Energy* 137, 435–444. <https://doi.org/10.1016/j.apenergy.2014.10.030>
- Feng D, Wang G, Wei X, Amankwah SOY, Hu Y, Luo Z, Hagan DFT and Ullah W** (2022) Merging and downscaling soil moisture data from CMIP6 projections using deep learning method. *Frontiers in Environmental Science* 10, 847475. <https://doi.org/10.3389/fenvs.2022.847475>
- Franzke CLE, Blender R, O’Kane TJ and Lembo V** (2022) Stochastic methods and complexity science in climate research and modeling. *Frontiers in Physics* 10, 931596. <https://doi.org/10.3389/fphy.2022.931596>
- Fujihara Y, Tanaka K, Watanabe T, Nagano T and Kojiri T** (2008) Assessing the impacts of climate change on the water resources of the Seyhan River basin in Turkey: Use of dynamically downscaled data for hydrologic simulations *Journal of Hydrology* 353(1–2), 33–48. <https://doi.org/10.1016/j.jhydrol.2008.01.024>
- Kingma DP and Ba J** (2015) Adam: A method for stochastic optimization. In *3rd International Conference for Learning Representations (ICLR)*, May 7–9, 2015, San Diego, Ithaca, NY.
- Klemm T and McPherson RA** (2017) The development of seasonal climate forecasting for agricultural producers. *Agricultural and Forest Meteorology* 232, 384–399. <https://doi.org/10.1016/j.agrformet.2016.09.005>
- Laprise R** (2008) Regional climate modelling. *Journal of Computational Physics* 227(7), 3641–3666. <https://doi.org/10.1016/j.jcp.2006.10.024>
- Liu Z, Huang J, Xiao X and Tong X** (2022) The capability of CMIP6 models on seasonal precipitation extremes over Central Asia. *Atmospheric Research* 278, 106364. <https://doi.org/10.1016/j.atmosres.2022.106364>
- Lorenzoni I and Pidgeon NF** (2006) Public views on climate change: European and USA perspectives. *Climatic Change* 77, 73–95. <https://doi.org/10.1007/s10584-006-9072-z>
- Luo X, Nadiga BT, Park JH, Ren Y, Xu W and Yoo S** (2022) A Bayesian deep learning approach to near-term climate prediction. *Journal of Advances in Modeling Earth Systems* 14, e2022MS003058. <https://doi.org/10.1029/2022MS003058>
- Pan B, Anderson GJ, Goncalves A, Lucas DD, Bonfils CJW and Lee J** (2022) Improving seasonal forecast using probabilistic deep learning. *Journal of Advances in Modeling Earth Systems* 14, e2021MS002766. <https://doi.org/10.1029/2021MS002766>
- Roads J, Chen S-C and Kanamitsu M** (2003) U.S. regional climate simulations and seasonal forecasts. *Journal of Geophysical Research Atmosphere* 108(D16), 8606. <https://doi.org/10.1029/2002JD002232>
- Ronneberger O, Fischer P and Brox T** (2015) U-Net: Convolutional networks for biomedical image segmentation. In *Medical Image Computing and Computer-Assisted Intervention (MICCAI)*, Vol. 9351. Cham: Springer.

- Scinocca JF, Kharin VV, Jiao Y, Qian MW, Lazare M, Solheim L and Flato GM** (2016) Coordinated global and regional climate modeling *Journal of Climate* 29(1), 17–35. <https://doi.org/10.1175/JCLI-D-15-0161.1>
- Stott PA and Forest C** (2007) Ensemble climate predictions using climate models and observational constraints. *Philosophical Transactions of the Royal Society A* 365, 2029–2052 <http://doi.org/10.1098/rsta.2007.2075>
- Talukder B, Ganguli N, Matthew R, vanLoon GW, Hipel KW and Orbinski J** (2021) Climate change-triggered land degradation and planetary health: A review. *Land Degradation & Development* 32, 4509–4522. <https://doi.org/10.1002/ldr.4056>
- Toth Z, Pena M and Vintzileso A** (2007) Bridging the gap between weather and climate forecasting, research priorities for intraseasonal prediction *Bulletin of the American Meteorological Society* 88, 1427–1440. <https://doi.org/10.1175/BAMS-88-9-1427>
- Troccoli A** (2010) Seasonal climate forecasting. *Meteorological Applications* 17, 251–268. <https://doi.org/10.1002/met.184>
- Turnock ST, Allen RJ, Andrews M, Bauer SE, Deushi M, Emmons L, Good P, Horowitz L, John JG, Michou M, Nabat P, Naik V, Neubauer D, O'Connor FM, Olivie D, Oshima N, Schulz M, Sellar A, Shim S, Takemura T, Tilmes S, Tsigaridis K, Wu T and Zhang J** (2020) Historical and future changes in air pollutants from CMIP6 models. *Atmospheric Chemistry and Physics* 20, 14547–14579. <https://doi.org/10.5194/acp-20-14547-2020>
- Tyagi S, Zhang X, Saraswat D, Sahany S, Mishra SK and Niyogi D** (2022) Flash drought: Review of concept, prediction and the potential for machine learning, deep learning methods. *Earth's Future* 10, e2022EF002723. <https://doi.org/10.1029/2022EF002723>
- Wang D, Liu J, Shao W, Mei C, Su X and Wang H** (2021) Comparison of CMIP5 and CMIP6 multi-model ensemble for precipitation downscaling results and observational data: The case of Hanjiang River basin. *Atmosphere* 12, 867. <https://doi.org/10.3390/atmos12070867>
- Xu Z, Han Y, Tam CY, Yang ZL and Fu C** (2021) Bias-corrected CMIP6 global dataset for dynamical downscaling of the historical and future climate (1979–2100) *Scientific Data* 8, 293. <https://doi.org/10.1038/s41597-021-01079-3>
- Yuan N, Huang Y and Duan J** (2019) On climate prediction: How much can we expect from climate memory? *Climate Dynamics* 52, 855–864. <https://doi.org/10.1007/s00382-018-4168-5>
- Zhang S and Li X** (2021) Future projections of offshore wind energy resources in China using CMIP6 simulations and a deep learning-based downscaling method *Energy* 217, 119321. <https://doi.org/10.1016/j.energy.2020.119321>
- Zhao H, Shi J, Qi X, Wang X and Jia J** (2017) Pyramid scene parsing network. In *Computer Vision and Pattern Recognition (CVPR)*.
- Zhou Z, Md Siddiquee MR, Tajbakhsh N and Liang J** (2018) UNet++: A nested U-Net architecture for medical image segmentation. In *Deep Learning in Medical Image Analysis (DLMIA)*. Cham: Springer.