MDG625: A daily high-resolution meteorological dataset derived by geopotential-guided attention network in Asia (1940-2023)

Zijiang Song[1,2], Zhixiang Cheng[1,2], Yuying Li[1,2], Shanshan Yu[1,2], Xiaowen Zhang[1,2], Lina Yuan[1,2], and Min Liu[1,2]

[1]Key Laboratory of Geographic Information Science, Ministry of Education, School of Geographic Sciences, East China Normal University, Shanghai 200241, China
[2]Key Laboratory of Spatial-temporal Big Data Analysis and Application of Natural Resources in Megacities, Ministry of Natural Resources, Shanghai, 200241, China

Correspondence: Lina Yuan (lnyuan@geoai.ecnu.edu.cn) and Min Liu (mliu@geo.ecnu.edu.cn)

Abstract.

The long-term and reliable meteorological reanalysis dataset with high spatial-temporal resolution is crucial for various hydrological and meteorological applications, especially in regions or periods with scarce in situ observations and with limited open-access data. Based on the ERA5 (produced by the European Centre for Medium-Range Weather Forecasts, 0.25°×0.25°, since 1940) and CLDAS (China Meteorological Administration Land Data Assimilation System, 0.0625°×0.0625°, since 2008), we proposed a novel downscaling method Geopotential-guide Attention Network (GeoAN) leveraging the high spatial resolution of CLDAS and the extended historical coverage of ERA5 and produced the daily multi-variable (2m temperature, surface pressure, and 10m wind speed) meteorological dataset MDG625 (Song et al., 2024). MDG625 (0.0625° Meteorological Dataset derived by GeoAN) covers most of Asia from 0.125° S to 64.875° N and 60.125° E to 160.125° E since 1940. Compared with other downscaling methods, GeoAN shows better performance with the R^2 (2m temperature, surface pressure, and 10m wind speed reached 0.990, 0.998, and 0.781, respectively). MDG625 demonstrates superior continuity and consistency from both spatial and temporal perspectives. We anticipate that this GeoAN method and this dataset MDG625 will aid in climate studies of Asia and will contribute to improving the accuracy of reanalysis products from the 1940s. The dataset (Song et al., 2024) is presented at https://doi.org/10.57760/sciencedb.17408 and the code can be found at https://github.com/songzijiang/GeoAN.

1 Introduction

As temperatures rise and extremes become more frequent, weather-related data analysis is becoming increasingly important (Berrang-Ford et al., 2011; Dietz et al., 2020; Taylor et al., 2013; Karl and Trenberth, 2003). Spatial resolution is crucial for geographic datasets. However, the distribution of in-situ stations is too sparse to produce a high-quality reanalysis dataset, especially for decades ago. For getting a higher resolution reanalysis dataset, downscaling is widely used in geoprocessing (Atkinson, 2013), especially in climate-related fields (Wang et al., 2021; Vogel et al., 2023; Tefera et al., 2024; Sun et al., 2024). The meteorological reanalysis dataset, which is obtained from in situ and remote sensing measurements, is important for agriculture, extreme weather forecasts, etc. Higher resolution of these data can better guide life and production. He et al. (2020) produced a meteorological dataset with a spatial resolution of $0.1°$ from 1979 in China. In this paper, the China Meteorological Forcing Dataset was proposed by fusing remote sensing products, reanalysis datasets, and in-situ station data. The most significant contribution of this work was using a larger number of stations to raise the quality of the dataset. A long-term gridded daily meteorological dataset for northwestern North America was proposed by Werner et al. (2019). The authors try to produce a dataset for training statistical downscaling schemes in Canada. The same in Italian, Bonanno et al. (2019) proposed the high-resolution meteorological dataset named MERIDA. MERIDA was produced by dynamical downscaling from the fifth-generation reanalysis dataset for the global climate and weather (ERA5) using WRF. Considering the optical limitation, high-resolution reanalysis data is expensive to produce, and historical high-resolution data is even harder to come by. High-quality and high-resolution data is necessary for various studies, to solve the contradiction, low-resolution (LR) data products being used to downscale into high-resolution (HR, also called as ground truth) are widely used (Hu et al.,

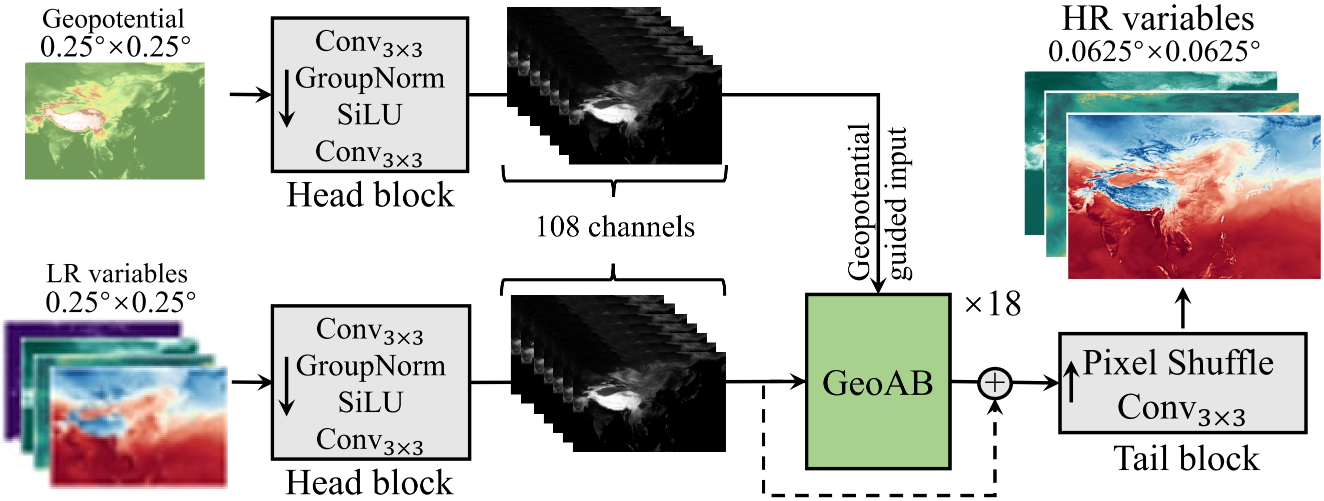

Figure 1. Sketch of the GeoAN. LR and HR denoted the low-resolution and high-resolution, respectively. The head block contains one group norm, one activation function, and two convolutions, which are abbreviated by $\mathrm{Conv}_{3\times3}$ meaning the kernel size is 3×3. SiLU is adopted as the activation function. The results of the two blocks of head in the diagram have the same channels of 108. GeoAB, which is repeated 18 times constricted by the hardware and the data amount, is the attention block for extracting deep information using geopotential. The pixel shuffle operation is performed after the convolution in the tail block to produce the high-resolution variables. Note that, the order of execution in each grey block (i.e., head and tail blocks) is along the arrows in the box.

2023; Zhong et al., 2023). The mainly used downscaling methods are categorized into statistical downscaling and dynamical downscaling. While the existing downscaling methods could produce high-resolution results, the results are unsatisfactory and unable to reconstruct detail and texture information(Murphy, 1999). Dynamical downscaling methods are usually based on Regional Climate Models (RCMs) with the initial fields produced by Global Climate Models (GCMs). Although the resolution of RCMs is higher than GCMs, the comprehension ability to understand the real world is not enough. It leads to a considerable bias (Teutschbein and Seibert, 2012), since the hard of establishing simulation equations, and cannot meet the needs of various related tasks. In another opinion, the computational cost of RCMs is huge, and it is an obstacle to producing a wider range of results for each calculation (Giorgi and Gutowski Jr, 2015; Di Luca et al., 2015). Compared with dynamical downscaling, statistical downscaling uses the mapping relationship between high-resolution and low-resolution from historical datasets to produce future datasets. The computational cost and bias of statistical downscaling are lower than dynamical downscaling methods.

Deep learning is a statistical method to build the bridge between input and output. Since Vaswani et al. (2017) proposed the transformer network, the ability of deep learning to harvest shadow information has gained a step. After that, the transformer block is widely used in diverse tasks including Super-Resolution (SR) (Liang et al., 2021; Zhang et al., 2022; Song and Zhong, 2022). Liang et al. (2021) proposed SwinIR and achieved impressive results in the SR task and be considered the benchmark for the SR task. The core algorithm of SwinIR is using no overlap windows to split the input feature to calculate the attention relationship inner each window and shift the windows by the step of the half-width of the windows. Song and Zhong (2022) proposed a novel network to harvest long-range information from global instead of inner the window. The experimental results on SR benchmarks (Bevilacqua et al., 2012; Martin et al., 2001; Huang et al., 2015; Matsui et al., 2017) show this strategy can achieve better results. Super-resolution tasks are similar to geographic downscaling tasks, using SR deep learning methods to downscale the geographic data can effectively avoid the problems encountered by former downscaling methods, such as high biases, regional sensitivity, high computational cost, etc. Deep learning methods use deep layers to bridge the relationship from low-resolution to high-resolution data and have robustness against the sensitivity. The computational cost is very low once the model has been trained, during the using step, and the deep learning method can nest a wide range easily. Shen et al. (2023) proposed a near-surface air temperature downscaling network SNCA-CLDASSD. In this model, Shen, et al. used two attention blocks to downscale the input data called Cross-Attention based on Light-CLDASSD. However, only near-surface air temperature is considered in this work and the network was built on CLDAS, which cannot cover long-term years. Liu et al. (2023) used the terrain to guide the deep learning network for the downscaling task called terrain-guided attention network (TGAN) in Southwest China. TGAN used the digital elevation model (DEM) to build high-resolution temperature (at 2 meters) results. The range of TGAN used begins in 2018 and TGAN cannot be used in the historical situation. Zhong et al. (2023) proposed a transformer-based learning method Uformer, which directly adds topography data, to achieve high-resolution meteorological variables in inner Mongolia province, China. Although topography data can help rebuild the high-resolution, directly adding into the input low-resolution will lose the characters of topography. All of the above, existing advanced deep learning methods of meteorological downscaling mostly used attention architecture (Transformer is one of the special attention architectures). However, now existing methods focus on one or two meteorological variables, while different variables have correlations and deep learning could handle multiple variables simultaneously. This way, not only saves the computational resources but also improves the performance of the model. Last and most important, there are no models that can cover a long-term and wide range of historical multiple variables.

In this paper, we propose a new attention-based network called the Geopotential-guided Attention Network (GeoAN), the structure of which is shown in Fig. 1 for downscaling meteorological variables, including temperature at 2m (T2m), pressure at the surface (PRS), and wind speed at 10m (WS10m) from $0.25°$ to $0.0625°$. The proposed GeoAN is guided by the geopotential, which makes the model learn information directly instead of spontaneously. The low-resolution input of the variables is organized from ERA5, provided by the European Centre for Medium-Range Weather Forecasts (ECMWF), with data ranging from 1940 to the present. The target used by the downscaling algorithm is derived from the China Meteorological Administration's Land Data Assimilation System (CLDAS) daily (Shi et al., 2014; Sun et al., 2020; Shi et al., 2011). The data quality and resolution of CLDAS are relatively high, but data is only available for China and the surrounding areas and for the years after 2008. After utilizing deep learning networks to construct the mapping relationship between ERA5 and CLDAS, a historical meteorological dataset since 1940 was produced. Our produced MDG625 makes up for the lack of the CLDAS before 2008 and increases the spatial resolution of ERA5. This dataset is valuable for a wide range of applications, including climate change studies and extreme weather events analysis.

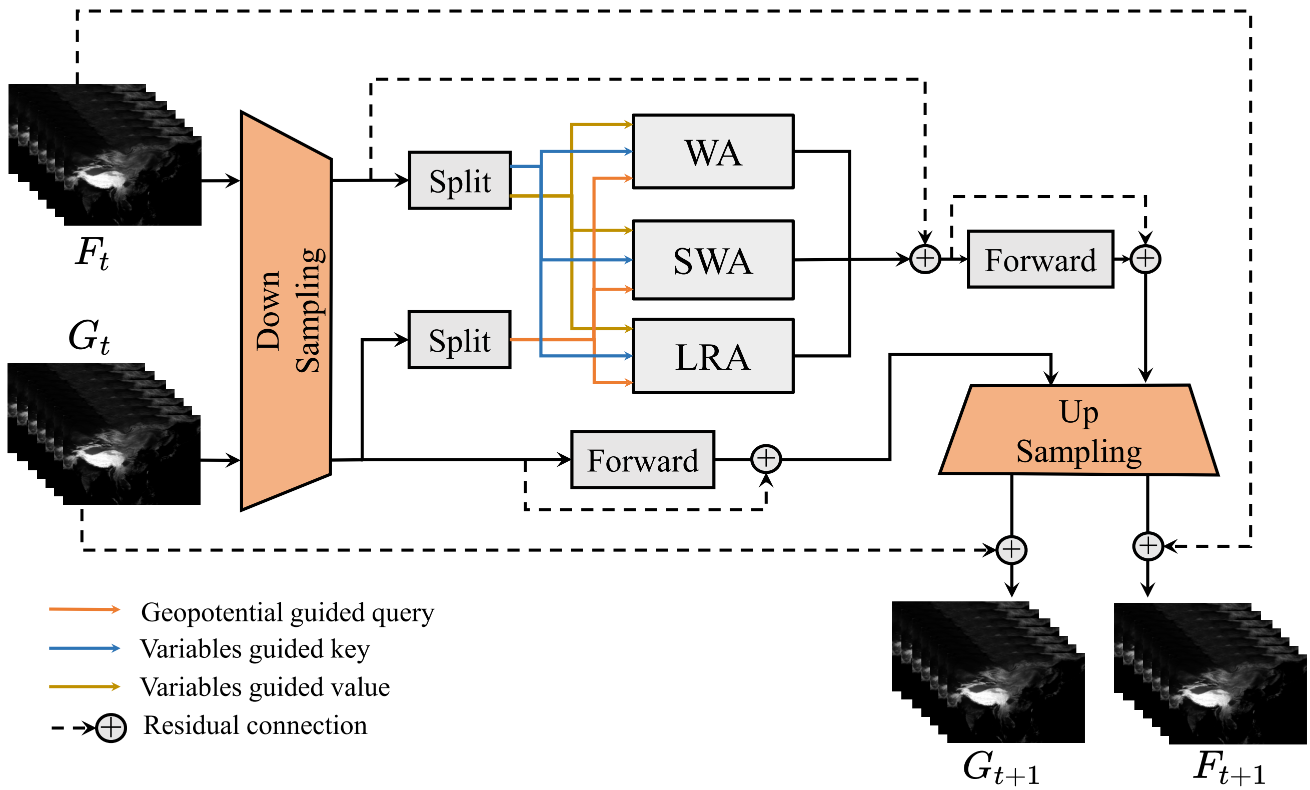

Figure 2. Sketch of the GeoAB, which repeated 18 times in GeoAN. GeoAB is the attention block to extract deep information. The query information of GeoAN is harvested from geopotential and the key and value are made from variable features. To make the loops, the outputs of the t^{th} GeoAB, i.e., F_{t+1} and G_{t+1}, are treated as the input of the $(t+1)^{th}$ GeoAB.

2 Data and methods

2.1 Data

The study area spans most of Asia (latitudes from 0.125° S to 64.875° N and longitudes from 60.125° E to 160.125° E), including China, Japan, India, etc. ERA5 is the fifth generation ECMWF reanalysis, provided by the ECMWF and used widely (Muñoz-Sabater et al., 2021; Hersbach et al., 2020; Jiang et al., 2021; Olauson, 2018; Cucchi et al., 2020), for the global climate and weather. ECMWF is a premier international organization, considered advanced in numerical weather prediction (NWP) models. The variables of PRS and T2m used in this work are listed in the ERA5 data list directly, and WS10m is calculated by U and V components of the wind at 10m. CLDAS, which uses multigrid variational analysis and multi-source precipitation fusion, is a reanalysis production provided by the China Meteorological Administration (CMA). The value in higher resolution in China is more reasonable than other datasets. In this paper, ERA5 is used as the low-resolution image (LR), and CLDAS is treated as the high-resolution image (HR, i.e., ground truth) to train the proposed model. The output of the downscaling network is called super-resolution images (SR).

There are four meteorological variables, temperature at 2m, pressure at the surface, wind speed at 10m, and daily total precipitation (TP) considered in GeoAN. Considering it is hard to process the downscale of TP, only three other variables are produced by GeoAN in MDG625. The period of dataset used to train the network is from 2020 to 2022. And, the period of the validation dataset is from January 1 to December 31, 2023. Note that, all times are in Coordinated Universal Time (UTC). CLDAS data is used as the high-resolution and the ERA5 data is used as the low-resolution input. The spatial resolution of ERA5 and CLDAS are $0.25°$ and $0.0625°$ respectively. The temporal resolution of these two datasets is calculated to one day, which is calculated by the mean of PRS (hPa), T2m (K), WS10m ($\mathrm{m \cdot s^{-1}}$) and the sum of TP (mm) over the whole day respectively using the original hourly data. The region is limited by CLDAS, i.e. latitudes from $0.125°$S to $64.875°$N, and the range of longitudes is from $60.125°$E to $160.125°$E. Because the grids of ERA5 and CLDAS do not overlap, thus the extent of ERA5 is a bit larger than CLDAS ($0.25°$S to $65°$N and $60°$E to $160.25°$E).

2.2 Geopotential-guided attention network

Geopotential ($\mathrm{m^2 \cdot s^{-2}}$) is the gravitational potential energy of a unit mass. Geopotential can reflect the elevation, latitude, pressure, etc. The value of geopotential used in this paper is obtained from the ERA5 dataset. Using geopotential to guide the attention calculation for downscaling can gain geographic semantic information, which is lacking in common deep learning networks.

As shown in Fig. 2, geopotential-guided attention is realized by the Geopotential-Guided Attention Block (GeoAB), which is the core unit of the GeoAN. The window attention (WA), shifted window attention (SWA), and long rang attention (LRA) are constructed from Song and Zhong (2022) and Song et al. (2022). The concepts of query, key, and value were used in transformer block Vaswani et al. (2017) to excavate the effects of attention. The query being produced from geopotential is different from the original transformer block, in which the query is produced from the input features. The key and value are, the same as the original block, harvested from the input features. For ease of understanding, normalization, residual operation, and other detailed parts are not listed in the formulas incidentally. The formulas are defined as follows, where F_t and G_t denoted the deep features of meteorological variables and geopotential at t^{th} loop respectively:

$$F_{t+1} = \mathcal{F}(\mathcal{A}(G_t, F_t)), \tag{1}$$

$$G_{t+1} = \mathcal{F}(G_t). \tag{2}$$

$\mathcal{F}(\cdot)$ and $\mathcal{A}(\cdot)$ denoted the forward and attention parts respectively, all forward parts did not share the parameters. After that, the WA and SWA were updated from Swin Transformer (Liu et al., 2021) to Swin Transformer V2 (Liu et al., 2022) comparing Song and Zhong (2022).

The network architecture is described in Fig. 1. The GeoAB is repeated 18 times to harvest more geographic information as shown in Eq. 1, Eq. 2, and Fig. 2, the definition of the network architecture is described as follows:

$$\mathrm{SR} = \mathcal{T}(\mathrm{GeoAB}^{18}[\mathcal{H}(LR), \mathcal{H}(G)]), \tag{3}$$

where LR, G, and SR denoted low-resolution variables, geopotential, and produced high-resolution variables, respectively. $\mathcal{H}(\cdot)$, $\mathcal{T}(\cdot)$, and $\mathrm{GeoAB}^{\mathrm{k}}[\cdot,\cdot]$ denoted head block, tail block, and GeoAB block, respectively. The GeoAB block is repeated for k times.

The batch size of the training step was 5, which is an unbalanced GPU distribution for 3 NVIDIA RTX 6000 Ada Generation (48G), considering the GPU memory limitation. There were 6 days of high-resolution data missing from 2020 to 2022. Thus there were only 1090 LR and SR pairs for training ($366+365+365-6 = 1090$). The learning rate was set to 10^{-4} and reduced by half at epochs 20, 40, 60, 80, 90 and 95. The network was trained for 100 epochs from the pre-trained models. Considering differences among the lines of latitude, the latitude-weighted loss was chosen to be the loss function, and the distortion of geographical coordinates with changes in latitude is fully taken into account (Bi et al., 2023; Rasp et al., 2020). The loss function is defined as follows:

$$loss = \frac{\sum_{i=1}^{\mathrm{H}}\sum_{j=1}^{\mathrm{W}}\sum_{c=1}^{\mathrm{C}} a_i \times |\mathrm{HR}_{i,j,c} - \mathrm{SR}_{i,j,c}|}{\mathrm{H}\times\mathrm{W}\times\mathrm{C}}, \tag{4}$$

where H, W, and C are 1040, 1600, and 4, respectively. $\mathrm{HR}_{i,j,c}$ and $\mathrm{SR}_{i,j,c}$ is the value at position of (i,j) of channel c in HR variables and SR variables. The a_i is latitude weight defined as:

$$a_i = \mathrm{H}\cdot\frac{\cos\theta_i}{\sum_{i=1}^{\mathrm{H}}\cos\theta_i}, \tag{5}$$

where θ_i is the latitude of the i^{th} line in the map of the variables in the form of $1040\times1600\times4$ (1040 and 1600 represent the pixel counts along latitude and longitude, and 4 represent WS10m, T2m, PRS, and TP, TP is only for assist and not in the results of the network.) For calculation purposes, the latitudes range is offset to 0 - $65°$N (i.e., $0 <= \theta_i < 65\frac{\pi}{180}$) replacing $0.125°$S - $64.875°$N.

3 Performance

3.1 Quantitative comparison

This section details experiments to evaluate the performance of GeoAN. For comparison, the classic algorithm bilinear interpolation, widely used in downscaling, is included. Additionally, two deep learning methods, U-Net (Ronneberger et al., 2015) and SwinIR (Liang et al., 2021), were employed for comparative analysis. The source code for both networks was obtained from their perspective GitHub repositories. To ensure a fair comparison, the U-Net architecture was modified for the downscaling task, resulting in a customized version referred to as U-Net Evolution (U-Net Evol.). The original U-Net implementation is available at https://github.com/milesial/Pytorch-UNet, while the SwinIR code can be accessed at https:

Table 1. A comparison of our proposed GeoAN with other downscaling methods. The bigger value stands for better performance, and the value in bold indicates the best performance in each metric. Considering the suitability of the downscaling task, PSNR, SSIM, and R^2 are chosen. All results are produced by the same environment and super parameters.

Methods	Variables	PSNR (dB) ↑	SSIM ↑	$\mathbf{R^2}$ ↑
Bilinear	**T2m**	27.920	0.900	0.939
	WS10m	21.271	0.747	0.582
	PRS	33.392	0.902	0.965
U-net (Evol.)	**T2m**	**35.471**	0.969	**0.991**
	WS10m	25.556	0.845	0.780
	PRS	40.008	0.969	0.990
SwinIR	**T2m**	34.042	0.956	0.988
	WS10m	24.452	0.825	0.745
	PRS	37.435	0.943	0.978
GeoAN (Ours)	**T2m**	35.054	**0.983**	0.990
	WS10m	**25.599**	**0.859**	**0.781**
	PRS	**47.251**	**0.996**	**0.998**

//github.com/JingyunLiang/SwinIR. To maintain consistency, all deep learning models were configured with equivalent parameters or computational complexity, and they were trained for 100 epochs using identical hyperparameters under the same environmental conditions.

As shown in Tab. 1, PSNR (Peak Signal-to-Noise Ratio), SSIM (Structural Similarity Index), and R^2 (Coefficient of Determination) are considered to evaluate the performance of the methods. PSNR and SSIM are the most commonly used metrics for measuring super-resolution algorithms. Compared to RMSE, R^2 or other numerical metrics, which only calculate the independent value of each pixel, a more holistic and detailed assessment is considered. A PSNR greater than 25dB is acceptable and greater than 30dB is considered a good result. In most metrics, GeoAN produces better results than others, however, in the T2m comparison, U-net (Evol.) got a higher result, further analyses about this part will be discussed in the appendix. Note that, the performance of SwinIR is worse than U-net (Evol.), this phenomenon may caused by that, the training step only contains 100 epochs considering the limitation of GPU, and it is not enough for an attention-based SwinIR or GeoAN. Even in this situation, GeoAN could outperform the other methods.

3.2 Visual comparison

Although GeoAN achieved better results in most cases in Tab. 1, a more subjective comparison also needs to be drawn directly. Visual results of PRS, T2m, WS10m are shown in Fig. 3, Fig. 4, and Fig. 5, respectively. We compare the 1st of each two months in 2023 (i.e., January, March, May, July, September, and November) and choose one day to display for each variable.

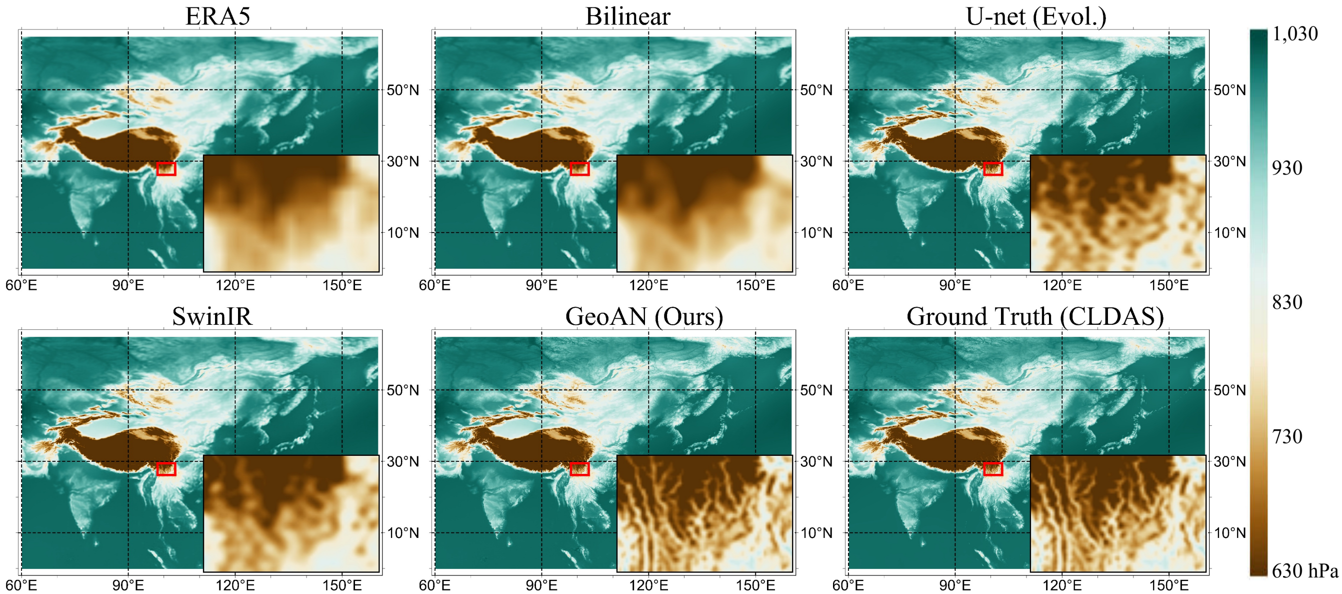

Figure 3. Pressure visual results of GeoAN and other downscaling algorithms on the 1st of November 2023. GroundTruth means the target high-resolution data (i.e., CLDAS), and ERA5 is the original low-resolution data. GeoAN is the deep learning method we proposed in this paper. The picture in the lower right corner of each subgraph is the detailed picture of the target area (i.e., red rectangle) respectively.

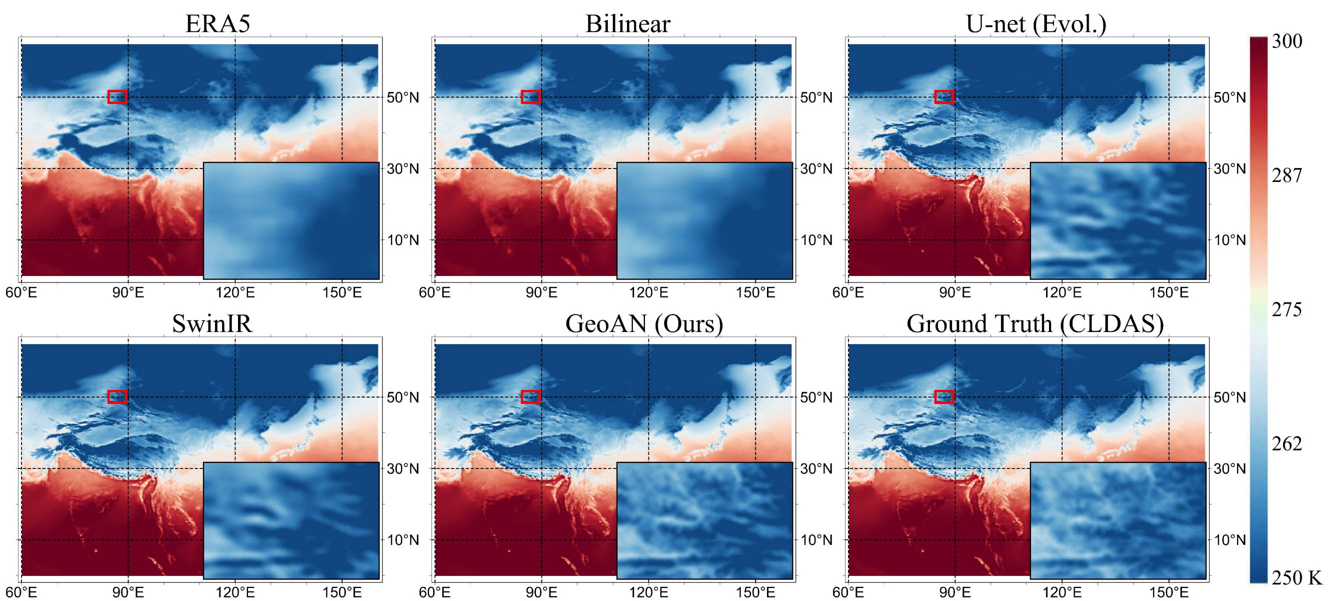

Figure 4. Temperature visual results of GeoAN and other downscaling algorithms on the 1st of January 2023.

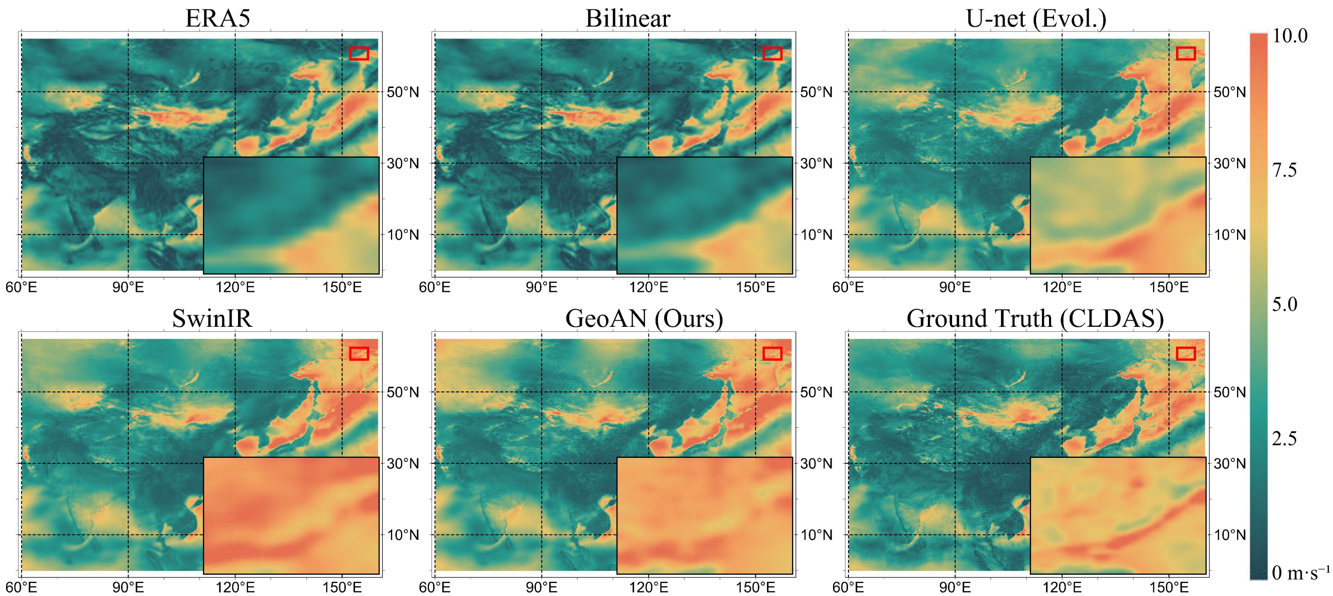

Figure 5. Wind speed visual results of GeoAN and other downscaling algorithms on the 1st of November 2023.

As shown in the figures, GeoAN can achieve the best results among all the compared algorithms. Especially, for extracting details, GeoAN has an excellent performance. Benefiting from the geopotential-guided attention and training from the historical data, enough geographic semantic information can be harvested by the neural network, and even the distorted detail parts, can be restored well by the GeoAN. Comparing other downscaling algorithms, the data produced by GeoAN can be treated as high-quality meteorological data.

4 Produced dataset

4.1 Historical meteorological data

The period of CLDAS is from 2008 to the present. The producing of CLDAS relies on the stations, which were rare before the 2000s in China (Tie et al., 2022). Due to the historical high-resolution meteorological data being hard to access, we used our proposed GeoAN, which can be guided by the geopotential well, to produce the historical meteorological data, called MDG625 (Meteorological Dataset with $0.0625°$ resolution produced by GeoAN), in the study area since 1940. MDG625 is valuable for historical meteorological studies in relevant areas. The comparison between similar datasets is in Tab. 2. The resolution of ERA5 and GLDAS is too low for various regional studies. The CLDAS dataset produced by CMA is not long enough in time series, and cannot do a long period study. The biggest difference is that MDG625 is driven by the deep learning method instead of numerical methods.

Table 2. A comparison of different datasets.

Datasets	Time period	Spatial resolution	Derived from	Sources
ERA5	1940 - present	$0.25° \times 0.25°$	Reanalysis method	ECMWF
CLDAS	2008 - present	$0.0625° \times 0.0625°$	Reanalysis method	CMA
GLDAS	1948 - present	$0.25° \times 0.25°$	Reanalysis method	NASA
MDG625	1940 - present	$0.0625° \times 0.0625°$	Deep learning method	Ours

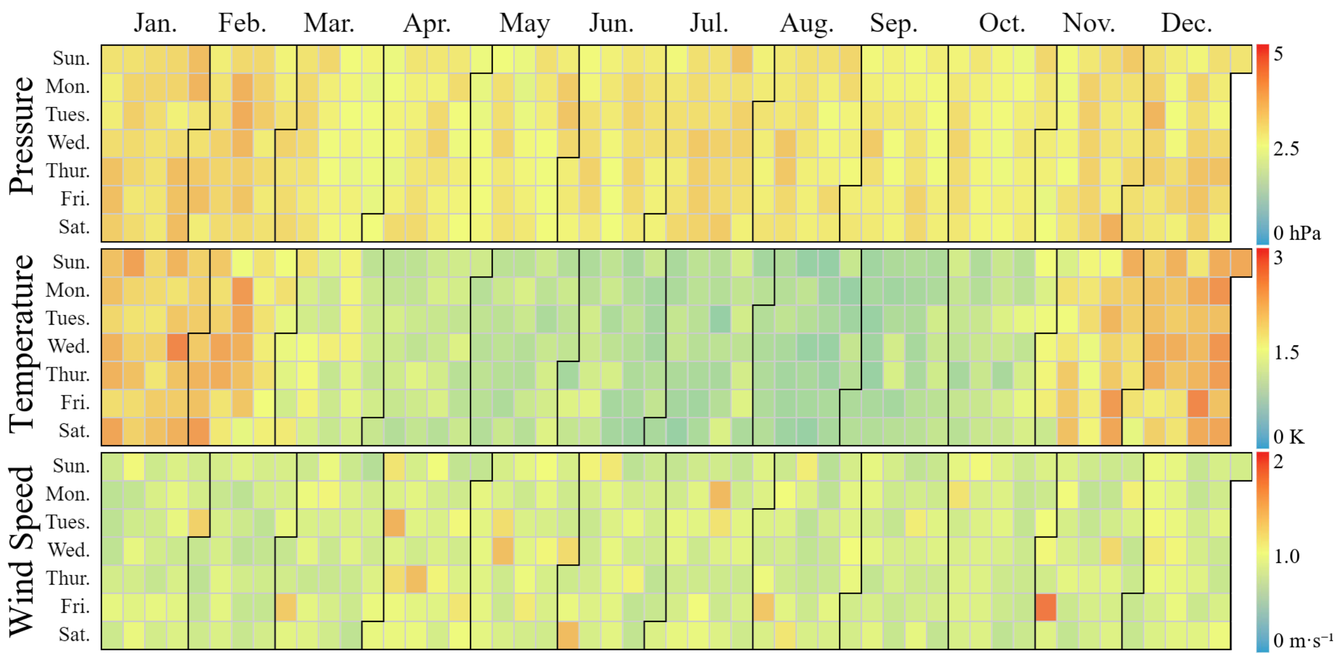

Figure 6. The daily average RMSE of MDG625 in 2023. From top to bottom are pressure (hPa), temperature (K) at 2m, and wind speed ($\mathrm{m \cdot s^{-1}}$) at 10m, respectively. The RMSE at a single day is calculated from the daily pixels.

Note that, there are two days variables abnormal in ERA5, i.e., '1965-11-29' and '2008-7-6'. The first day of MDG625 is '1940-1-1' and the index of this day is recorded as '0'. The index of each day means the number of days elapsed since the 1st of January 1940. A larger spatial extent dataset also can be produced by GeoAN, considering the pattern used in training steps, only the data in the study area is provided in MDG625.

4.2 Error distribution

Considering the period of CLDAS, and the data from 2020 to 2022 are used in the training step, the results of error distribution are calculated in 2023. The RMSE of the 2m temperature, surface pressure, and 10m wind speed are 1.40 K, 2.76 hPa, and 0.89 $\mathrm{m \cdot s^{-1}}$ respectively. To further evaluate the quality of MDG625 temporally and spatially, the error distributions of the

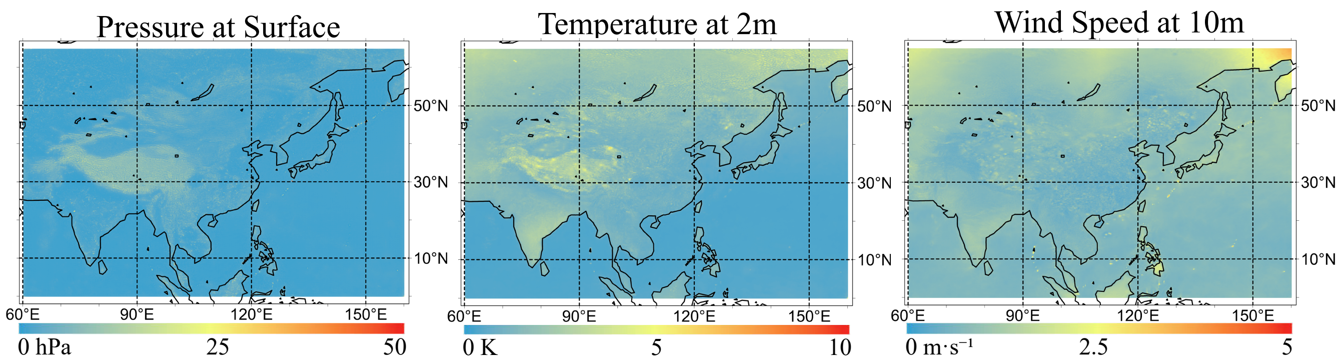

Figure 7. The RMSE map of three meteorological variables (PRS, T2m, and WS10m) between MDG625 and the ground truth. The RMSE is calculated from the whole year daily in 2023. Blue represents a smaller error and red represents the bad results.

variables (PRS, T2m, and WS10m) are analyzed in Fig. 6 and Fig. 7. As shown in Fig. 6, the variables of T2m in winter are not satisfied, and other variables performed well. However, although in the worst month of T2m (i.e., January), the difference to ground truth is around 3K, which is acceptable. The average RMSE of T2m is about 1K for the whole year. For PRS and WS10m, stable good performance throughout the whole year. This phenomenon may be caused by the seasonal change of the temperature which is hard for a statistical model to infer the right results without any extra season or date information. One thing to note, 9 days of data are missing ('2023-05-22', '2023-06-27', '2023-10-10', '2023-10-11', '2023-10-12', '2023-10-13', '2023-10-14', '2023-10-15', and '2023-10-16') in CLDAS by the time of the program run. The RMSE of the missing data is calculated by the mean of the nearest data that are not missing before and after.

Considering the spatial distribution as shown in Fig. 7, the mean RMSE across the whole year shows the condition of three variables. PRS and T2m show better results in marine than in the mainland. On the contrary, WS10m results on the mainland are better. For T2m, the results of the mainland performed worse than the marine, which may be caused by the specific heat capacity of water being higher than the land (e.g., soil, sand, etc.) Temperature variations over the oceans are lower in magnitude than over the continents, and the deep learning method is better at learning the patterns of small changes. For the other two variables, maybe the same reason caused the error distribution. No matter which variables, the results of the coastal plain (e.g., east of China) are relatively better. Lastly, even if the relative effectiveness of the results in some areas is unsatisfactory, the errors are still in a reasonable and acceptable range and the dataset can be used for various analyses.

4.3 Limitations

However, GeoAN achieved a great result and MDG625 is valuable for a wide range of applications, there are some limitations in this work. Limited by the computational expenses and the memory of GPU, we cannot use a larger model to train with more datasets. A larger model and more training datasets could help the performance a lot. The MDG625 produced the area in Asia, however, if the fixed position information was masked and more datasets to cover the other areas in the training step, the global results could be produced. A global high-resolution historical dataset can be very valuable. In addition, by our proposed

GeoAN, hourly variables also can be produced, but only daily data are provided in MDG625 since the hard disk capacity is not enough for such a large amount of data. Lastly, precipitation poses a challenge for statistical models, particularly in extreme cases (Zou et al., 2024; Sachindra et al., 2018; Xu et al., 2015). The dataset referenced in MDG625 includes three meteorological variables (PRS, T2M, and WS10m), but it lacks precipitation data. The low correlation of total precipitation (TP) between the ERA5 and CLDAS datasets is the primary reason why the GeoAN model struggles to accurately restore the spatial structure of precipitation. Inspired by the work of (Zou et al., 2024), more comprehensible algorithms could be implemented in GeoAN in the future, and we will try to compare a wider range of models to evaluate our performance.

5 Conclusions and discussion

Considering the rarity of long-term historical high-resolution meteorological data in Asia, MDG625 (Meteorological Dataset with $0.0625°$ resolution produced by a Geopotential-guide attention network) provided a solution using a deep geographic coupling attention network called the geopotential-guide attention network (GeoAN) within an acceptable error. GeoAN could learn the geopotential relationship directly, which is closely related to meteorological variables. This Strategy could help the network understand the geographic world more easily and produce a better result. MDG625 contains daily temperature at 2m, surface pressure, and 10m wind speed since 1940. Experimental results demonstrated the superior performance of the GeoAN and the satisfaction of MDG625. Our proposed MDG625 could make up for the lack of a long historical meteorological high-resolution dataset. For various meteorological researches is valuable. In future work, we will try to merge remote sensing and gauged precipitation to downscale the precipitation using other geographic principles and using more suitable variables instead of geopotential. In conclusion, more deep learning methods coupling geographic mechanisms may provide solutions to various geographic problems.

Code and data availability. The ERA5 data of ECMWF can be found at https://cds.climate.copernicus.eu. The high-resolution data, CLDAS, is provided by CMA at https://data.cma.cn. An education and research account is required to acquire the CLDAS data, this requirement is set by the CMA. The code and the generated dataset MDG625 (Song et al., 2024) can be found in the GitHub repository: https://github.com/songzijiang/GeoAN and ScienceDB repository: https://doi.org/10.57760/sciencedb.17408. Considering CLDAS is not public, and GeoAN was trained using CLDAS, the data of MDG625 for 2017-2023 are not offered in the repository.

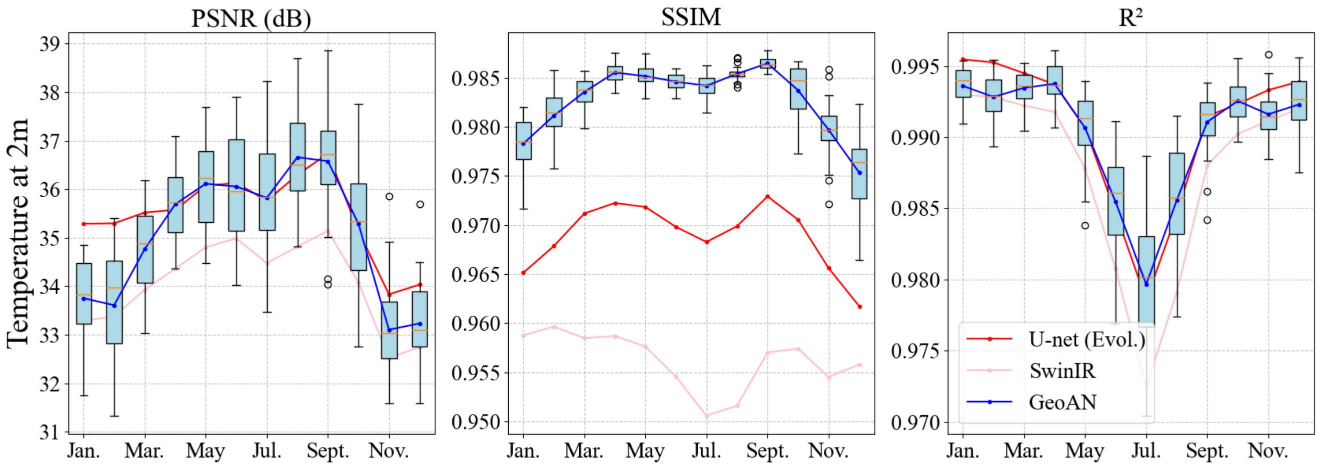

Figure A1. The monthly average statistic of GeoAN in 2023 compared with other methods on temperature (K) at 2m. The box plot is the distribution of GeoAN in each month.

Appendix A: T2m comparison against U-net (Evol.)

Shown in Tab. 1, U-net (Evol.) performed better than GeoAN in T2m on PSNR and R^2, to explain this situation, we analyzed the error distributions of T2m temporal and spatial in Fig. 6 and Fig. 7. In summer, the results of GeoAN have a similar performance with U-net (Evol.) on PSNR and R^2 as shown in Fig. A1. The higher the altitude, the more error in GeoAN observed refers to Fig. 7. Both in winter or high-altitude altitude, temperatures will be lower, we extrapolate that, the GeoAN performs badly in cold areas and periods. To verify whether PSNR and R^2 react to the real performance in cold environments, we conducted a full year of comparisons in 2023 at high altitudes area, the specific results are shown in Fig. A2. In comparison, the texture of GeoAN is clearer than U-net (Evol.), and the temperature values in each pixel of these two methods are close, the difference is almost negligible. However, the improvement in sharpness GeoAN brings is discernible to the naked eye. Refers to Fig. 6, the largest RMSE between ground truth and GeoAN in winter is around 3K, and the mean RMSE of T2m is 1.40K, consider the RMSE between CLDAS and in-situ stations is 1.8K, the bias in GeoAN totally could be accepted.

Appendix B: Discussion on how GeoAN restore the absent details

The PSNR of GeoAN on PRS is up to 47.251dB on the testing dataset. The leading performance compared with other methods of PRS is larger than other variables, and we find the PRS's outperformance of GeoAN in mountainous areas extremely obvious. GeoAN can produce the details that are not included in the input data as shown in Fig. 3. Considering the primal input data ERA5 lacks the details as shown in Fig. B1, it is reasonable to infer that deep learning methods can learn detailed information from the distribution of the ground truth in the training step. The less change in the variables, the better the results will be. The training step provides enough information on the detailed shape and the low-resolution input ERA5 of each test

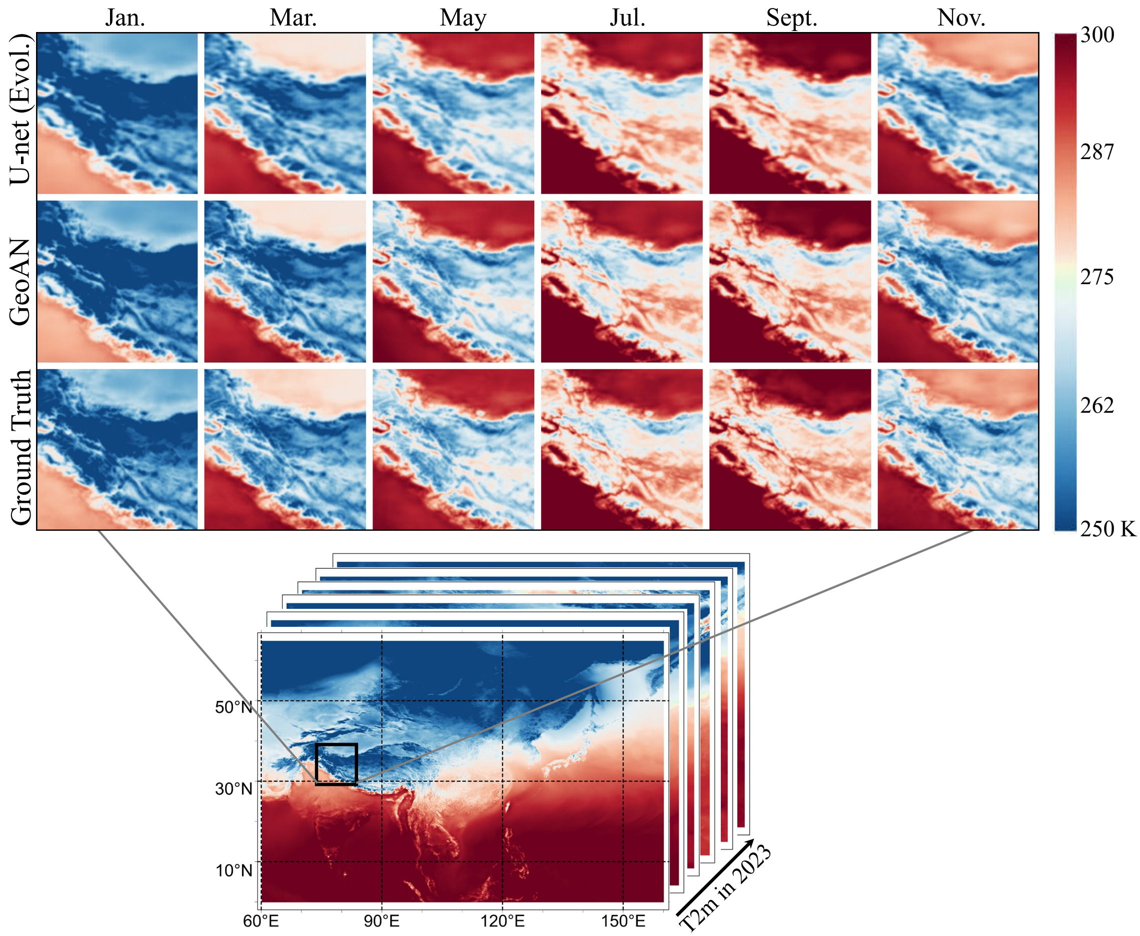

Figure A2. Temperature at 2m comparison between U-net (Evol.), GeoAN, and ground truth at high altitude areas (Himalayas areas) in 2023. Only the results of the first day of odd-numbered months are shown for convenient observation.

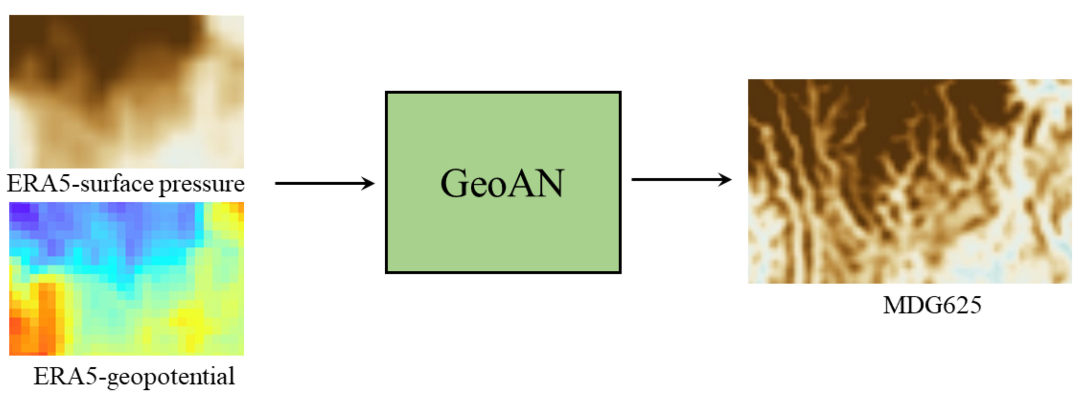

Figure B1. GeoAN can restore the details that are not included in the input data.

offers the value of each grid. This is the most important reason why GeoAN can produce fine details. We also found the change in pressure is much lower than other variables, and we infer that's why the metrics of pressure perform better than others. The change of temperature in one year is bigger than pressure and the PSNR of T2m is worse than PRS. Therefore, the performance of WS10m is worst can be inferred from the drastic irregular changes, because there is not enough information for the models to restore the details of WS10m in the training dataset.

Author contributions. SZJ and YLN designed the research. SZJ and CZX performed the experiments and code. SZJ wrote the manuscript. LYY, YSS, and ZXW process data analysis. LM and YLN provided the resource. LM and YLN supervised and reviewed the manuscript.

Competing interests. The contact author has declared that none of the authors has any competing interests.

Disclaimer. The MDG625 are produced by GeoAN, which is trained by ERA5 and CLDAS. To download and use the algorithms and datasets associated with this paper, please follow the relevant restrictions and requirements. The license of GeoAN is Apache-2.0 and of MDG625 is CC-BY 4.0.

Acknowledgements. This document is the results of the research project funded by the International Research Center of Big Data for Sustainable Development Goals (No. CBAS2022GSP07), and the National Natural Science Foundation of China (No. 42230505, 42206148). We thank China Meteorological Administration and European Centre for Medium-Range Weather Forecasts for the provided datasets. We thank all reviewers, editors, and others who helped with this paper.

References

Atkinson, P. M.: Downscaling in remote sensing, International Journal of Applied Earth Observation and Geoinformation, 22, 106–114, 2013.

Berrang-Ford, L., Ford, J. D., and Paterson, J.: Are we adapting to climate change?, Global Environmental Change, 21, 25–33, 2011.

Bevilacqua, M., Roumy, A., Guillemot, C., and Alberi-Morel, M. L.: Low-complexity single-image super-resolution based on nonnegative neighbor embedding, in: The British Machine Vision Conference, pp. 1–10, 2012.

Bi, K., Xie, L., Zhang, H., Chen, X., Gu, X., and Tian, Q.: Accurate medium-range global weather forecasting with 3D neural networks, Nature, 619, 533–538, 2023.

Bonanno, R., Lacavalla, M., and Sperati, S.: A new high-resolution meteorological reanalysis Italian dataset: MERIDA, Quarterly Journal of the Royal Meteorological Society, 145, 1756–1779, 2019.

Cucchi, M., Weedon, G. P., Amici, A., Bellouin, N., Lange, S., Müller Schmied, H., Hersbach, H., and Buontempo, C.: WFDE5: bias-adjusted ERA5 reanalysis data for impact studies, Earth System Science Data, 12, 2097–2120, 2020.

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

Dietz, T., Shwom, R. L., and Whitley, C. T.: Climate change and society, Annual Review of Sociology, 46, 135–158, 2020.

Giorgi, F. and Gutowski Jr, W. J.: Regional dynamical downscaling and the CORDEX initiative, Annual review of environment and resources, 40, 467–490, 2015.

He, J., Yang, K., Tang, W., Lu, H., Qin, J., Chen, Y., and Li, X.: The first high-resolution meteorological forcing dataset for land process studies over China, Scientific data, 7, 25, 2020.

Hersbach, H., Bell, B., Berrisford, P., Hirahara, S., Horányi, A., Muñoz-Sabater, J., Nicolas, J., Peubey, C., Radu, R., Schepers, D., et al.: The ERA5 global reanalysis, Quarterly Journal of the Royal Meteorological Society, 146, 1999–2049, 2020.

Hu, Y., Tang, R., Jiang, X., Li, Z.-L., Jiang, Y., Liu, M., Gao, C., and Zhou, X.: A physical method for downscaling land surface temperatures using surface energy balance theory, Remote Sensing of Environment, 286, 113 421, 2023.

Huang, J.-B., Singh, A., and Ahuja, N.: Single image super-resolution from transformed self-exemplars, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 5197–5206, 2015.

Jiang, Q., Li, W., Fan, Z., He, X., Sun, W., Chen, S., Wen, J., Gao, J., and Wang, J.: Evaluation of the ERA5 reanalysis precipitation dataset over Chinese Mainland, Journal of Hydrology, 595, 125 660, 2021.

Karl, T. R. and Trenberth, K. E.: Modern global climate change, Science, 302, 1719–1723, 2003.

Liang, J., Cao, J., Sun, G., Zhang, K., Van Gool, L., and Timofte, R.: Swinir: Image restoration using swin transformer, in: Proceedings of the IEEE/CVF International Conference on Computer Vision, pp. 1833–1844, 2021.

Liu, G., Zhang, R., Hang, R., Ge, L., Shi, C., and Liu, Q.: Statistical downscaling of temperature distributions in southwest China by using terrain-guided attention network, IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 16, 1678–1690, 2023.

Liu, Z., Lin, Y., Cao, Y., Hu, H., Wei, Y., Zhang, Z., Lin, S., and Guo, B.: Swin transformer: Hierarchical vision transformer using shifted windows, in: Proceedings of the IEEE/CVF International Conference on Computer Vision, pp. 10 012–10 022, 2021.

Liu, Z., Hu, H., Lin, Y., Yao, Z., Xie, Z., Wei, Y., Ning, J., Cao, Y., Zhang, Z., Dong, L., et al.: Swin transformer v2: Scaling up capacity and resolution, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 12 009–12 019, 2022.

Martin, D., Fowlkes, C., Tal, D., and Malik, J.: A database of human segmented natural images and its application to evaluating segmentation algorithms and measuring ecological statistics, in: Proceedings of the IEEE/CVF International Conference on Computer Vision, vol. 2, pp. 416–423, IEEE, 2001.

Matsui, Y., Ito, K., Aramaki, Y., Fujimoto, A., Ogawa, T., Yamasaki, T., and Aizawa, K.: Sketch-based manga retrieval using manga109 dataset, Multimedia Tools and Applications, 76, 21 811–21 838, 2017.

Muñoz-Sabater, J., Dutra, E., Agustí-Panareda, A., Albergel, C., Arduini, G., Balsamo, G., Boussetta, S., Choulga, M., Harrigan, S., Hersbach, H., et al.: ERA5-Land: A state-of-the-art global reanalysis dataset for land applications, Earth System Science Data, 13, 4349–4383, 2021.

Murphy, J.: An evaluation of statistical and dynamical techniques for downscaling local climate, Journal of Climate, 12, 2256–2284, 1999.

Olauson, J.: ERA5: The new champion of wind power modelling?, Renewable Energy, 126, 322–331, 2018.

Rasp, S., Dueben, P. D., Scher, S., Weyn, J. A., Mouatadid, S., and Thuerey, N.: WeatherBench: a benchmark data set for data-driven weather forecasting, Journal of Advances in Modeling Earth Systems, 12, e2020MS002 203, 2020.

Ronneberger, O., Fischer, P., and Brox, T.: U-net: Convolutional networks for biomedical image segmentation, in: Medical Image Computing and Computer-Assisted Intervention, pp. 234–241, Springer, 2015.

Sachindra, D., Ahmed, K., Rashid, M. M., Shahid, S., and Perera, B.: Statistical downscaling of precipitation using machine learning techniques, Atmospheric research, 212, 240–258, 2018.

Shen, Z., Shi, C., Shen, R., Tie, R., and Ge, L.: Spatial Downscaling of Near-Surface Air Temperature Based on Deep Learning Cross-Attention Mechanism, Remote Sensing, 15, 5084, 2023.

Shi, C., Xie, Z., Qian, H., Liang, M., and Yang, X.: China land soil moisture EnKF data assimilation based on satellite remote sensing data, Science China Earth Sciences, 54, 1430–1440, 2011.

Shi, C., Jiang, L., Zhang, T., Xu, B., and Han, S.: Status and plans of CMA land data assimilation system (CLDAS) project, in: EGU General Assembly Conference Abstracts, p. 5671, 2014.

Song, Z. and Zhong, B.: A Lightweight Local-Global Attention Network for Single Image Super-Resolution, in: Proceedings of the Asian Conference on Computer Vision, pp. 4395–4410, 2022.

Song, Z., Zhong, B., Ji, J., and Ma, K.-K.: A direction-decoupled non-local attention network for single image super-resolution, IEEE Signal Processing Letters, 29, 2218–2222, 2022.

Song, Z., Cheng, Z., Li, Y., Yu, S., Zhang, X., Yuan, L., and Liu, M.: MDG625: Meteorological Dataset with 0.0625° resolution produced by GeoAN, https://doi.org/10.57760/sciencedb.17408, 2024.

Sun, S., Shi, C., Pan, Y., Bai, L., Xu, B., Zhang, T., Han, S., and Jiang, L.: Applicability assessment of the 1998–2018 CLDAS multi-source precipitation fusion dataset over China, Journal of Meteorological Research, 34, 879–892, 2020.

Sun, Y., Deng, K., Ren, K., Liu, J., Deng, C., and Jin, Y.: Deep learning in statistical downscaling for deriving high spatial resolution gridded meteorological data: A systematic review, ISPRS Journal of Photogrammetry and Remote Sensing, 208, 14–38, 2024.

Taylor, R. G., Scanlon, B., Döll, P., Rodell, M., Van Beek, R., Wada, Y., Longuevergne, L., Leblanc, M., Famiglietti, J. S., Edmunds, M., et al.: Ground water and climate change, Nature Climate Change, 3, 322–329, 2013.

Tefera, G. W., Ray, R. L., and Wootten, A. M.: Evaluation of statistical downscaling techniques and projection of climate extremes in central Texas, USA, Weather and Climate Extremes, 43, 100 637, 2024.

Teutschbein, C. and Seibert, J.: Bias correction of regional climate model simulations for hydrological climate-change impact studies: Review and evaluation of different methods, Journal of hydrology, 456, 12–29, 2012.

Tie, R., Shi, C., Wan, G., Hu, X., Kang, L., and Ge, L.: CLDASSD: Reconstructing fine textures of the temperature field using super-resolution technology, Advances in Atmospheric Sciences, 39, 117–130, 2022.

Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, Ł., and Polosukhin, I.: Attention is all you need, Advances in Neural Information Processing Systems, 30, 2017.

Vogel, E., Johnson, F., Marshall, L., Bende-Michl, U., Wilson, L., Peter, J. R., Wasko, C., Srikanthan, S., Sharples, W., Dowdy, A., et al.: An evaluation framework for downscaling and bias correction in climate change impact studies, Journal of Hydrology, p. 129693, 2023.

Wang, F., Tian, D., Lowe, L., Kalin, L., and Lehrter, J.: Deep learning for daily precipitation and temperature downscaling, Water Resources Research, 57, e2020WR029 308, 2021.

Werner, A., Schnorbus, M., Shrestha, R., Cannon, A., Zwiers, F., Dayon, G., and Anslow, F.: A long-term, temporally consistent, gridded daily meteorological dataset for northwestern North America, Scientific Data, 6, 1–16, 2019.

Xu, S., Wu, C., Wang, L., Gonsamo, A., Shen, Y., and Niu, Z.: A new satellite-based monthly precipitation downscaling algorithm with non-stationary relationship between precipitation and land surface characteristics, Remote sensing of environment, 162, 119–140, 2015.

Zhang, X., Zeng, H., Guo, S., and Zhang, L.: Efficient Long-Range Attention Network for Image Super-resolution, in: European Conference on Computer Vision, 2022.

Zhong, X., Du, F., Chen, L., Wang, Z., and Li, H.: Investigating transformer-based models for spatial downscaling and correcting biases of near-surface temperature and wind-speed forecasts, Quarterly Journal of the Royal Meteorological Society, 2023.

Zou, W., Hu, G., Wiersma, P., Yin, S., Xiao, Y., Mariethoz, G., and Peleg, N.: Multiple-point geostatistics-based spatial downscaling of heavy rainfall fields, Journal of Hydrology, 632, 130 899, 2024.