Generating a long-term (2003 − 2020) hourly 0.25° global PM$_{2.5}$ dataset via spatiotemporal downscaling of CAMS with deep learning (DeepCAMS)

Yi Xiao [a,1], Yuan Wang [a,1], Qiangqiang Yuan [a,*], Jiang He [a], Liangpei Zhang [b]

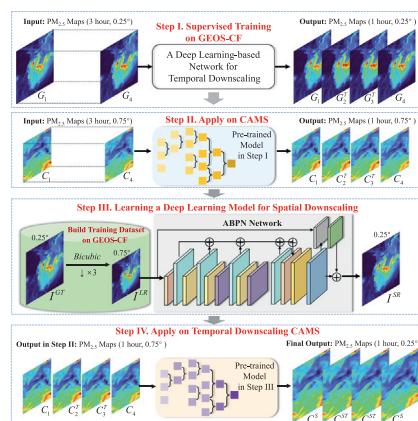[a] School of Geodesy and Geomatics, Wuhan University, Wuhan, Hubei 430079, China
[b] The Key Laboratory of Geospace Environment and Geodesy, Ministry of Education, Wuhan University, Wuhan, Hubei 430079, China

HIGHLIGHTS

- A deep learning framework (DeepCAMS) was developed to realize the spatiotemporal downscaling of CAMS.
- A long-term (2003–2020) hourly 0.25° global PM$_{2.5}$ dataset was proposed.
- DeepCAMS can maintain the spatial consistency and temporal continuity.

GRAPHICAL ABSTRACT

ARTICLE INFO

ABSTRACT

Generating a long-term high-spatiotemporal resolution global PM$_{2.5}$ dataset is of great significance for environmental management to mitigate the air pollution concerns worldwide. However, the current long-term (2003–2020) global reanalysis dataset Copernicus Atmosphere Monitoring Service (CAMS) reanalysis has drawbacks in fine-scale research due to its coarse spatiotemporal resolution (0.75°, 3-h). Hence, this paper developed a deep learning-based framework (DeepCAMS) to downscale CAMS PM$_{2.5}$ product on the spatiotemporal dimension for resolution enhancement. The nonlinear statistical downscaling from low-resolution (LR) to high-resolution (HR) data can be learned from the high quality (0.25°, hourly) but short-term (2018–2020) Goddard Earth Observing System composition forecast (GEOS-CF) system PM$_{2.5}$ product. Compared to the conventional spatiotemporal interpolation methods, simulation validations on GEOS-CF demonstrate that DeepCAMS is capable of producing accurate temporal variations with an improvement of Root-Mean-Squared Error (RMSE) of 0.84 (4.46 to 5.30) ug/m^3 and spatial details with an improvement of Mean Absolute Error (MAE) of 0.16 (0.34 to 0.50) ug/m^3. The real validations on CAMS reflect convincing spatial consistency and temporal continuity at both regional and global scales. Furthermore, the proposed dataset is validated with OpenAQ air quality data from 2017 to 2019, and the in-situ validations illustrate that the DeepCAMS maintains the consistent precision (R: 0.597) as the original CAMS (R: 0.593) while tripling the spatiotemporal resolution. The proposed dataset will be available at https://doi.org/10.5281/zenodo.6381600.

* Corresponding author.
E-mail addresses: xiao_yi@whu.edu.cn (Y. Xiao), yqiang86@gmail.com (Q. Yuan), jiang_he@whu.edu.cn (J. He), zlp62@whu.edu.cn (L. Zhang).
[1] Co-first author: Yi Xiao and Yuan Wang

1. Introduction

Fine particulate matter with aerodynamic diameter below 2.5 *um* ($PM_{2.5}$) (Yang et al., 2021; Thind et al., 2022), as a significant pollutant in the atmosphere, has become a global threat on human health (Lelieveld et al., 2015; Li et al., 2021), ecology (Zhai et al., 2014; Johnston et al., 2021), meteorology (Bae et al., 2021; Bai et al., 2022) and economy (Kocak and Celik, 2022; Zhu et al., 2022). Nowadays, many researchers and institutions are dedicated to the monitoring and prediction of $PM_{2.5}$ in hope of alleviating this global concern. To integrally study the spatial distribution and temporal variation of $PM_{2.5}$, the research community urgently needs a global scale long-term high spatiotemporal resolution dataset (Wang et al., 2021b, 2022c; Zhang et al., 2021a). Fortunately, the widely used reanalysis data is born to meet such requirements.

Existing reanalysis datasets mainly integrate the accuracy of various observations through sophisticated data assimilation and modeling techniques that can preserve spatial consistency and temporal continuity. The typical reanalysis products are the Copernicus Atmosphere Monitoring Service (CAMS) reanalysis (Inness et al., 2019) produced by the European Centre for Medium-Range Weather Forecasts (ECMWF) and the replay dataset in Goddard Earth Observing System composition forecast (GEOS-CF) system from NASA's Global Modeling and Assimilation Office (GMAO) (Keller et al., 2021). In many cases, they can serve as an effective data product in model evaluation, boundary conditions, and long-term tendency analysis (Zhang et al., 2022; Chen et al., 2019). However, these datasets still have bottlenecks in fine-resolution research or a long-time horizon. In practice, CAMS is long-term but has a lower spatial and temporal resolution (0.75° and 3-h) than GEOS-CF (0.25° and 1-h); GEOS-CF has higher resolution while covers short-term range (2018-now). On the one hand, the coarse spatial resolution hinders fine-scale research at the regional and global level, and the coarse temporal resolution is also not conducive to temporal evolution analysis. On the other hand, short-term products cannot reflect the long-term diachronic evolution of $PM_{2.5}$, which limits the application scope. Since numerical simulations of coupled physics and chemistry on a global scale require significant computational resources (Gu et al., 2022), conducting resolution enhancement on existing products can be a more efficient alternative. Hence, it is worthwhile to generate a long-term global high spatiotemporal resolution $PM_{2.5}$ dataset based on existing reanalysis products.

At present, numerous studies have applied statistical downscaling to atmospheric parameters for resolution enhancement. Early approaches were model-driven primarily. For example, Liu et al. (2018) adopted a simple downscaling method to obtain surface and near-surface air temperature data with high spatiotemporal resolution. Lv et al. (2017) designed a Bayesian-based statistical downscaling predicter to model the spatiotemporal relationships between AOD and $PM_{2.5}$. Varga-Balogh et al. (2020) produced regional-level $PM_{2.5}$ data in Budapest by downscaling CAMS by linear combination. Shen et al. (2017) downscaled the near-surface $PM_{2.5}$ maps through a Gaussian downscaling method. Subsequently, data-driven approaches appeared with remarkable performance. Nourani et al. (2018) proposed a data-driven ensemble model for downscaling of rainfall data. Yang et al. (2020) introduced random forests for high spatial resolution PM2.5 concentration mapping. Some works downscale directly on existing products. Pu and Yoo (2021) used the gradient boosting machine (GBM) to downscaledMERRA-2AOD into 1 km resolution and used the finer product as further auxiliary. In general, most statistical downscaling methods need to apply statistical models (predictors) to coarse-resolution data (predictands) for information transformation. Recently, deep learning methods based on Convolutional Neural Networks (CNN) have become a hot spot in statistical downscaling models. Thanks to the powerful nonlinear representation ability (Zhang et al., 2021b; He et al., 2022b; Wang et al., 2020, 2022a; Zheng et al., 2021, 2022) of CNN, we can explore complex statistical patterns of downscaling process. Naturally, by employing deep learning in spatiotemporal datasets, we are able to describe spatiotemporal features and achieve accurate spatiotemporal downscaling. For instance, Baño-Medina et al. (2020) achieved downscaling temperature and precipitation in Europe through deep learning neural models. Baño-Medina et al. (2021) used CNN networks for downscaling of climate change projections. Xue et al. (2019) developed a deep learning model to downscale meteorological variables. Wang et al. (2021a) introduced a super-resolution network named SRDRN to bridge the resolution gap between global scale and regional scale daily precipitation and temperature. Although some studies have used downscaling approach for meteorological parameter estimation, few efforts have been made to generate a resolution-enhanced product from a reanalysis dataset using spatiotemporal downscaling techniques.

Therefore, this paper proposes to simultaneously downscale the CAMS $PM_{2.5}$ product on spatial and temporal dimension based on deep learning approaches. The predictors used for statistical downscaling are trained on high resolution GEOS-CF products and applied to long-term but low-resolution CAMS products. After model training, we realize spatiotemporal downscaling of CAMS and produce a long-term global high-resolution $PM_{2.5}$ dataset. Specifically, DeepCAMS introduce a deep learning model named XVFI (Sim et al., 2021) used in video frame interpolation to realize $\times 3$ temporal downscaling and increase the temporal resolution of CAMS from 3-h to 1-h. To achieve spatial downscaling, an image super-resolution convolutional neural network named ABPN (Liu et al., 2019) is employed to improve spatial resolution from 0.75° to 0.25°. Experiments on GEOS-CF demonstrate that the DeepCAMS can predict the complex dynamic of $PM_{2.5}$ in the temporal dimension than the simple linear combination. Also, DeepCAMS effectively captures spatial nonlinearity and outputs fine spatial resolution results with rich details which are closest to the actual data. In the real experiment, we applied DeepCAMS to CAMS products to build a global hourly 0.25° $PM_{2.5}$ dataset from 2003 to 2020. The proposed dataset was validated with OpenAQ $PM_{2.5}$ air quality data from 2017 to 2019, and the results showed that the correlation coefficients were consistent with the original CAMS. To sum up, the innovation of this paper is as follows:

1) A deep learning framework (DeepCAMS) was developed to realize the spatiotemporal downscaling of reanalysis products. It bridges the resolution gap by introducing video frame interpolation and single-image super-resolution techniques.
2) We proposed a long-term (2003–2020) hourly 0.25° global $PM_{2.5}$ dataset through DeepCAMS. It can serve as a high quality product for fine-grained $PM_{2.5}$ research in long time series.
3) Extensive experiments demonstrate that DeepCAMS can explore the nonlinear spatiotemporal relationship and produce accurate downscaling results with spatial consistency and temporal continuity.

2. Data and method

2.1. Data source

This paper involves three datasets, including the $PM_{2.5}$ product from CAMS, $PM_{2.5}$ replay product from GEOS-CF, and OpenAQ in-situ $PM_{2.5}$ measurement. GEOS-CF is used for model training to learn the coarse-to-fine statistical downscaling relationship. Besides, DeepCAMS is tested on GEOS-CF in the simulation experiment. The pre-trained DeepCAMS is used to implement spatiotemporal downscaling on CAMS to obtain long-term global high-resolution $PM_{2.5}$ datasets. OpanAQ is used to verify the accuracy of the proposed dataset. These data sources will be described in detail below and summarized in Table 1.

Table 1
Summary of the PM2.5 data sources used in this paper.

Category	Source data	Years	Temporal resolution	Spatial resolution
Reanalysis data	CAMS	2003–2020	3-h	0.75° × 0.75°
	GEOS-CF	2018–2020	hourly	0.25° × 0.25°
In-situ data	OpenAQ	2017–2019	hourly	point

2.1.1. CAMS PM₂.₅ product

The Copernicus Atmosphere Monitoring Service (CAMS) reanalysis is proposed by the European Centre for Medium Range Weather Forecasts (ECMWF) (Kuenen et al., 2022; Jin et al., 2022). It is a global reanalysis atmospheric composition (AC) dataset which comprises three-dimensional time-consistent AC fields. This AC forecasts and analyses project can perform data assimilation and modeling for aerosols, chemical reactive gases and greenhouse gases based on the laws of physics and chemistry. Compared with earlier Monitoring Atmospheric Composition and Climate (MACC) reanalysis, its spatial resolution can reach $0.75°$ ($\sim 75\ km$). Temporally, CAMS can provide a 3-h analysis field, forecast field and hour-by-hour ground forecast field. Here, we downloaded and extracted 3-h $PM_{2.5}$ product with a spatial resolution of $0.75°$ from 2003 to 2020. Note that the raw data needs to be multiplied by 1×10^9 to convert the default unit kg/m^3 to ug/m^3. The dataset is available at https://ads.atmosphere.copernicus.eu/

2.1.2. GEOS-CF PM₂.₅ replay product

The Goddard Earth Observing System composition forecast (GEOS-CF) system is produced by NASA's Global Modeling and Assimilation Office (GMAO) (Colarco et al., 2010; Wang et al., 2022b). Compared with CAMS, GEOSCF has a higher spatial resolution ($0.25°$) and can give an hourly global constituent prediction. Thanks to the GEOSChem chemistry module, GEOS-CF offers replay (2018-now) and 5-days future forecasts of AC. Besides $PM_{2.5}$, many other critical meteorological products such as ozone (O_3), carbon monoxide (CO), nitrogen dioxide (NO_2) and sulfur dioxide (SO_2) are also available. Here, we use the 1-h $PM_{2.5}$ replay product with a spatial resolution of $0.25°$ from 2018 to 2020. In our paper, this high-resolution data is used for training set construction, model training and testing. The dataset can be downloaded freely at https://fluid.nccs.nasa.gov/cf/

2.1.3. OpenAQ PM₂.₅ air quality data

OpenAQ is an innovative open platform created by scientists and open data programmers. Its purpose is to effectively integrate air quality data publicly released by ground-based air quality monitors around the world. In addition, it is devoted to providing fast and uniform data access for researchers and the public worldwide (Hasenkopf et al., 2016; Hasenkopf, 2017). In OpenAQ, real-time air quality measurements such as $PM_{2.5}$,
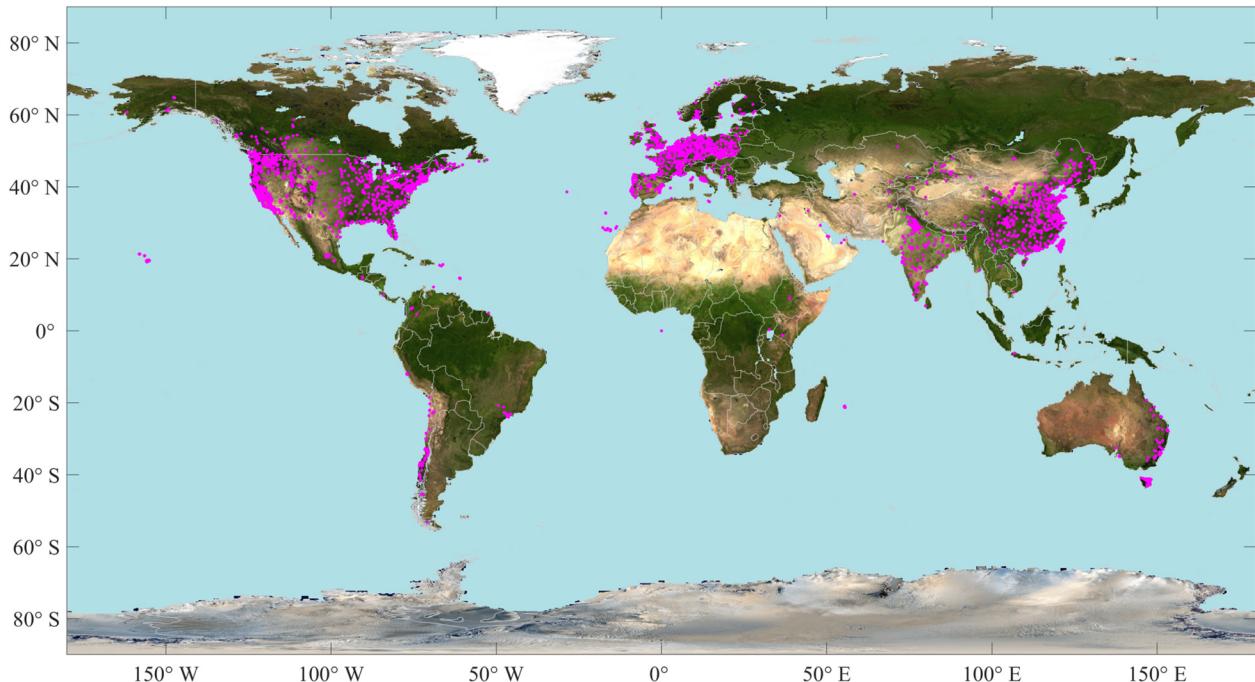
PM_{10}, SO_2, NO_2, and O_3) can be collected, stored, and published through an application programming interface (API). This paper uses $PM_{2.5}$ air quality data from 2017 to 2019. The data is available at https://openaq.org/. The global distribution of total ground stations from 2017 to 2019 is shown in Fig. 1. A total of 6656 stations were included from 2017 to 2019.

2.2. Method

DeepCAMS implements spatiotemporal downsampling by introducing two super-resolution (SR) networks. In the field of image processing, SR is known as a classic low-level vision task (Liu et al., 2021; Xiao et al., 2022a). The goal of SR is to generate high-resolution data from low-resolution data while retaining as much spatiotemporal details as possible. The overall framework of our DeepCAMS is exhibited in Fig. 2. To be specific, we utilize a video frame interpolation network named XVFI to realize temporal downscaling in steps I and II, and then we exploit a single image SR model termed ABPN to achieve spatial downscaling in steps III and IV. The network used for spatiotemporal downscaling is described in detail in the following subsections. Notably, during the design of our DeepCAMS, we did several exploratory experiments, which included the sequence of temporal downscaling and spatial downscaling. We found that performing spatial downscaling first yields a slightly lower performance than performing temporal downscaling in the first step. Therefore, we choose to perform temporal downscaling first, followed by spatial downscaling.

2.2.1. Temporal downscaling network

Video frame interpolation aims to synthesize the missing information between existing frames (Choi et al., 2020; Xiao et al., 2022b; Hu et al., 2022). This mission is similar to temporal downscaling. In Fig. 2, XVFI are trained to predict the non-exist $PM_{2.5}$ estimates $\{C_2, C_3\}$ between two CAMS $PM_{2.5}$ maps $\{C_1, C_4\}$ in adjacent time phase. To construct 3-h to 1-h training pairs for supervised learning, we artificially eliminated the existing PM2.5 maps and reduced the temporal resolution of GEOS-CF from 1-h $\{G_1, G_2, G_3, G_4\}$ to 3-h $\{G_1, G_4\}$. Since estimating $\{G_2, G_3\}$ from $\{G_1, G_4\}$ is equivalent to predicting $G_t(t = 0.33, 0.66)$ from $\{G_0, G_1\}$, we adopt this elaboration rule in the following description which is consistent with Sim et al. (2021).

Fig. 1. The diagram of the global distribution of OpenAQ stations in 2017, 2018, and 2019.
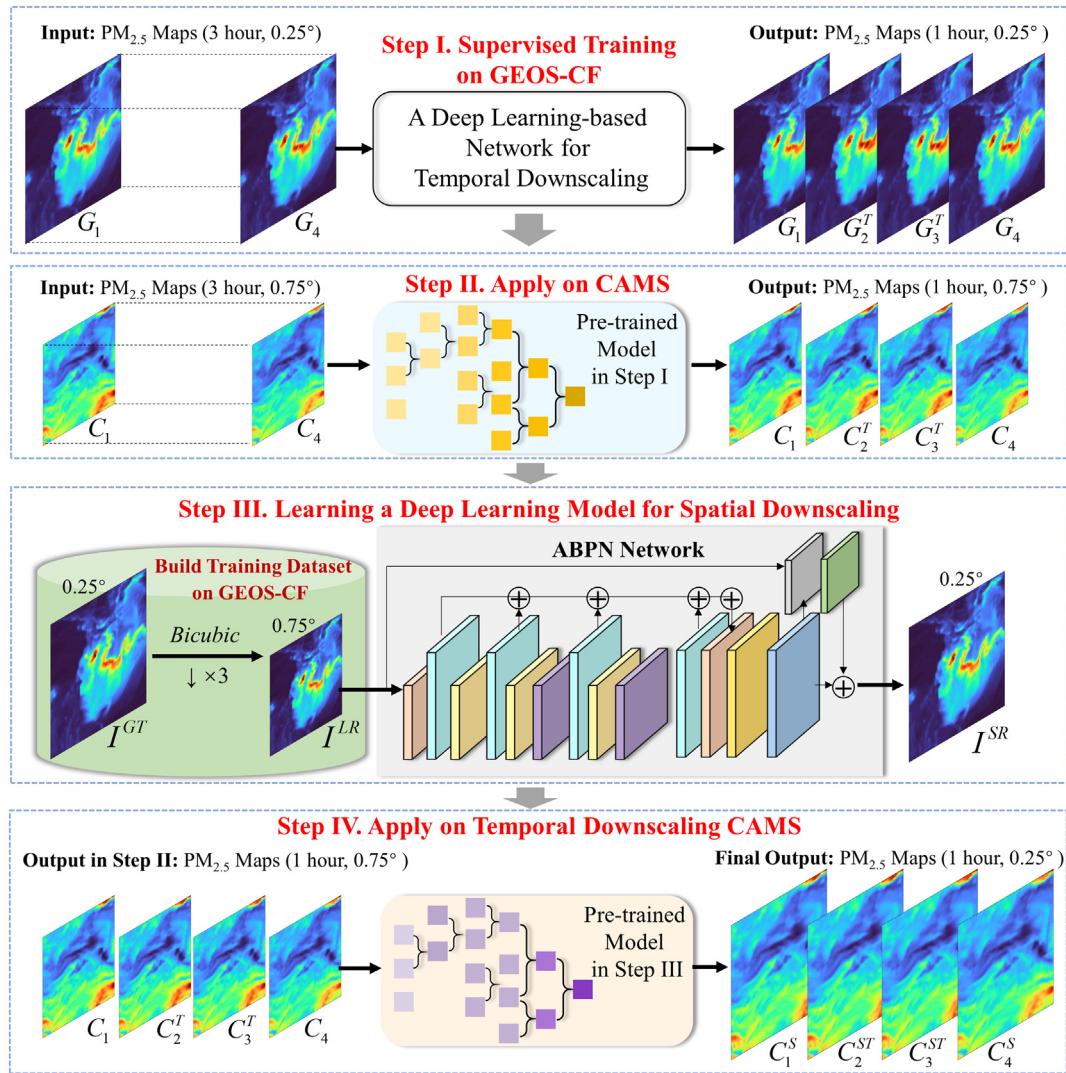
Fig. 2. The flowchart of our DeepCAMS. The framework takes 3-h 0.75° PM$_{2.5}$ product from CAMS as input and predict 1-h 0.25° high spatiotemporal resolution output.

XVFI is a sophisticated network, which is only briefly introduced here, and more details can be found in Sim et al. (2021). XVFI explore the relationship between missing information and existing estimates in a multiscale pyramid. Take the scale s for example, the input $\{G_0^s, G_1^s\}$ at scale s is the $1/2^s$ bicubic downsampled result of the original input $\{G_0, G_1\}$. After feature extraction, the input PM$_{2.5}$ maps $\{G_0^s, G_1^s\}$ (single channel) are converted to feature maps $\{I_0^s, I_1^s\}$ (32 channels) by 3×3 convolution. After feature extraction, XVFI design a BiFlowNet to predicts the coarse optical flow $\left\{ \widetilde{F}_{t0}^s, \widetilde{F}_{t1}^s \right\}$ between the latent PM$_{2.5}$ maps $\overline{G}_t^s (t = 0.33, 0.66)$ and $\{G_0^s, G_1^s\}$. Then, a TFlowNet is cascaded to refine the coarse optical flow and produce a fine-scale optical flow maps $\{F_{t0}^s, F_{t1}^s\}$ which are more accuracy to depict the motion relationships. After that, $\{F_{t0}^s, F_{t1}^s\}$ are used to warp $\{G_0^s, G_1^s\}$ and their corresponding feature $\{I_0^s, I_1^s\}$ by backward warping operation (Jaderberg et al., 2015), that is:

$$\overline{G}_{t0}^s = warp\left(\overline{G}_0^s, F_{t0}^s\right), \overline{I}_{t0}^s = warp\left(I_0^s, F_{t0}^s\right). \tag{1}$$

$$\overline{G}_{t1}^s = warp\left(\overline{G}_1^s, F_{t1}^s\right), \overline{I}_{t1}^s = warp\left(I_1^s, F_{t1}^s\right). \tag{2}$$
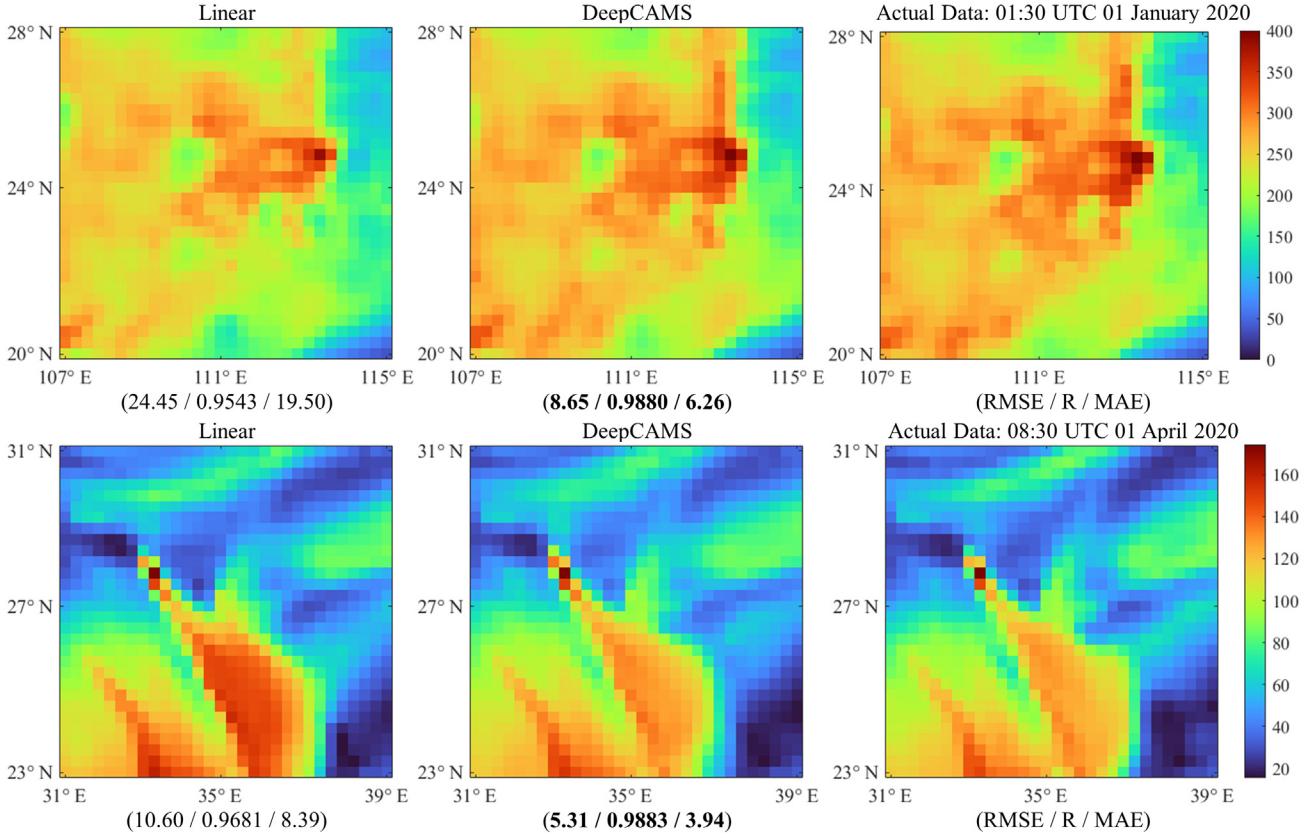
In this manner, the optical flow that encodes temporal PM$_{2.5}$ motion dynamics can be compensated to the existing PM$_{2.5}$ information to realize the information transformation. Afterwards,

$\left\{ G_0^s, G_1^s, \overline{G}_{t0}^s, \overline{G}_{t1}^s, F_{t0}^s, F_{t1}^s, I_0^s, I_1^s, \overline{I}_{t0}^s, \overline{I}_{t1}^s \right\}$ will be aggregated and send to a UNet-based RefinementBlock to generate occlusion mask m_t^s and

Table 2
The quantitative results of temporal downscaling. The R, RMSE and MAE ug/m^3 results in each row is the average result of all temporal interpolated samples (16 samples a day) in each month, and the final result is the average of total 12 months ($16 \times 365 = 5840$ samples). The best results are shown in bold.

	R ↑		RMSE ↓		MAE ↓	
Date	Linear	DeepCAMS	Linear	DeepCAMS	Linear	DeepCAMS
2020/01	0.9541	**0.9588**	9.30	**8.63**	0.66	0.35
2020/02	0.9773	**0.9834**	4.65	**3.89**	0.54	0.28
2020/03	0.9907	**0.9967**	2.37	**1.36**	0.50	0.26
2020/04	0.8929	**0.9019**	11.53	**10.91**	0.60	0.35
2020/05	0.9625	**0.9731**	4.34	**3.26**	0.50	0.27
2020/06	0.9804	**0.9917**	2.37	**1.53**	0.41	0.23
2020/07	0.9830	**0.9939**	2.89	**1.71**	0.48	0.25
2020/08	0.9816	**0.9922**	2.82	**1.83**	0.49	0.26
2020/09	0.8548	**0.8640**	13.60	**13.25**	0.68	0.43
2020/10	0.9646	**0.9735**	4.60	**3.86**	0.52	0.28
2020/11	0.9902	**0.9966**	2.15	**1.24**	0.43	0.22
2020/12	0.9828	**0.9877**	2.93	**2.06**	0.65	0.46
Average	0.9596	**0.9678**	5.30	**4.46**	0.54	0.30

Fig. 3. The qualitative results of temporal downscaling.

residual PM$_{2.5}$ maps $\tilde{G}_r^s$. Finally, the predicted missing PM$_{2.5}$ estimation can be formulated by:

$$\overline{G}_t^s = \frac{(1-t) \cdot m_t^s \cdot \tilde{G}_{t0}^s + t \cdot (1-m_t^s) \cdot \tilde{G}_{t1}^s}{(1-t) \cdot m_t^s + t \cdot (1-m_t^s)} + \tilde{G}_r^s. \tag{3}$$

The aforementioned process happens in scale s, and the final is the prediction in scale 0. We follow the settings in the.

Table 3
The quantitative results of spatial downscaling. The R, RMSE and MAE ug/m^3 results in each row is the average result of all samples (24 samples a day) in each month, and the final result is the average of total 12 months ($24 \times 365 = 8760$ samples). The best results are shown in bold.

Date	R ↑		RMSE ↓		MAE ↓	
	Bicubic	DeepCAMS	Bicubic	DeepCAMS	Bicubic	DeepCAMS
2020/01	0.9600	**0.9762**	9.30	**8.42**	0.64	**0.44**
2020/02	0.9782	**0.9836**	4.65	**3.89**	0.54	**0.36**
2020/03	0.9897	**0.9960**	2.51	**1.57**	0.45	**0.30**
2020/04	0.8501	**0.8981**	11.53	**10.91**	0.56	**0.40**
2020/05	0.9727	**0.9721**	3.89	**3.26**	0.43	**0.27**
2020/06	0.9878	**0.9955**	2.37	**1.12**	0.36	**0.24**
2020/07	0.9735	**0.9883**	3.54	**1.71**	0.49	**0.32**
2020/08	0.9700	**0.9888**	3.54	**1.83**	0.49	**0.31**
2020/09	0.8677	**0.8968**	13.60	**13.25**	0.71	**0.52**
2020/10	0.9582	**0.9687**	4.99	**3.86**	0.53	**0.34**
2020/11	0.9880	**0.9956**	2.39	**1.24**	0.41	**0.27**
2020/12	0.9875	**0.9948**	2.75	**1.77**	0.44	**0.29**
Average	0.9569	**0.9712**	5.42	**4.40**	0.50	0.34

XVFI and set $s = 5$. We adopt $\mathscr{L}_1$ loss function to constraint training process in each scale, which means:

$$\mathscr{L}_1 = \sum_{s=0}^{5} \left\| \overline{G}_t^s - G_t^s \right\|_1. \tag{4}$$

Here G_t^s is the $1/2^s$ bicubic downsampled result of the ground truth PM$_{2.5}$ map G_t. To better optimize the smoothness of the optical flow, we additionally introduce the first-order edge-aware smoothness loss at scale 0:

$$\mathscr{L}_s = \sum_{i=0}^{1} \exp\left(-e^2 |\nabla_x G_t|\right)^T \cdot \left| \nabla_x F_{ti}^0 \right|, \tag{5}$$

where $e = 150$ is an edge weighting factor and x represents a spatial coordinate. At last, the total loss function can be expressed as follows:

$$\mathscr{L}_{total} = \mathscr{L}_1 + \lambda_s \cdot \mathscr{L}_s. \tag{6}$$

The λ_s is set to 0.5 to balance the training process.

2.2.2. Spatial downscaling network

The spatial downscaling aims at predicting the high-resolution PM$_{2.5}$ concentration from its low-resolution counterpart (He et al., 2021, 2022a; Jiang et al., 2019; Yi et al., 2020). To build low-resolution (LR) to high-resolution (HR) PM$_{2.5}$ training pairs, we first spatially degenerate the GEOS-CF product from I^{GT} (0.25°) to I^{LR} (0.75°) through the *imresize* function in MATLAB. In the simulation experiment of GEOS-CF, the total of $24 \times 365 \times 2 = 17,520$ samples in 2018 and 2019 are used for training, and we tested the pre-trained model on GEOS-CF products in 2020. In the real experiment on CAMS, GEOS-CF data from 2018 to 2020 were all used for training. Although the PM$_{2.5}$ concentrations may be temporally correlated at each time phase, the time interval is usually at the hourly level and it is difficult to exploit the redundant information. To simplify
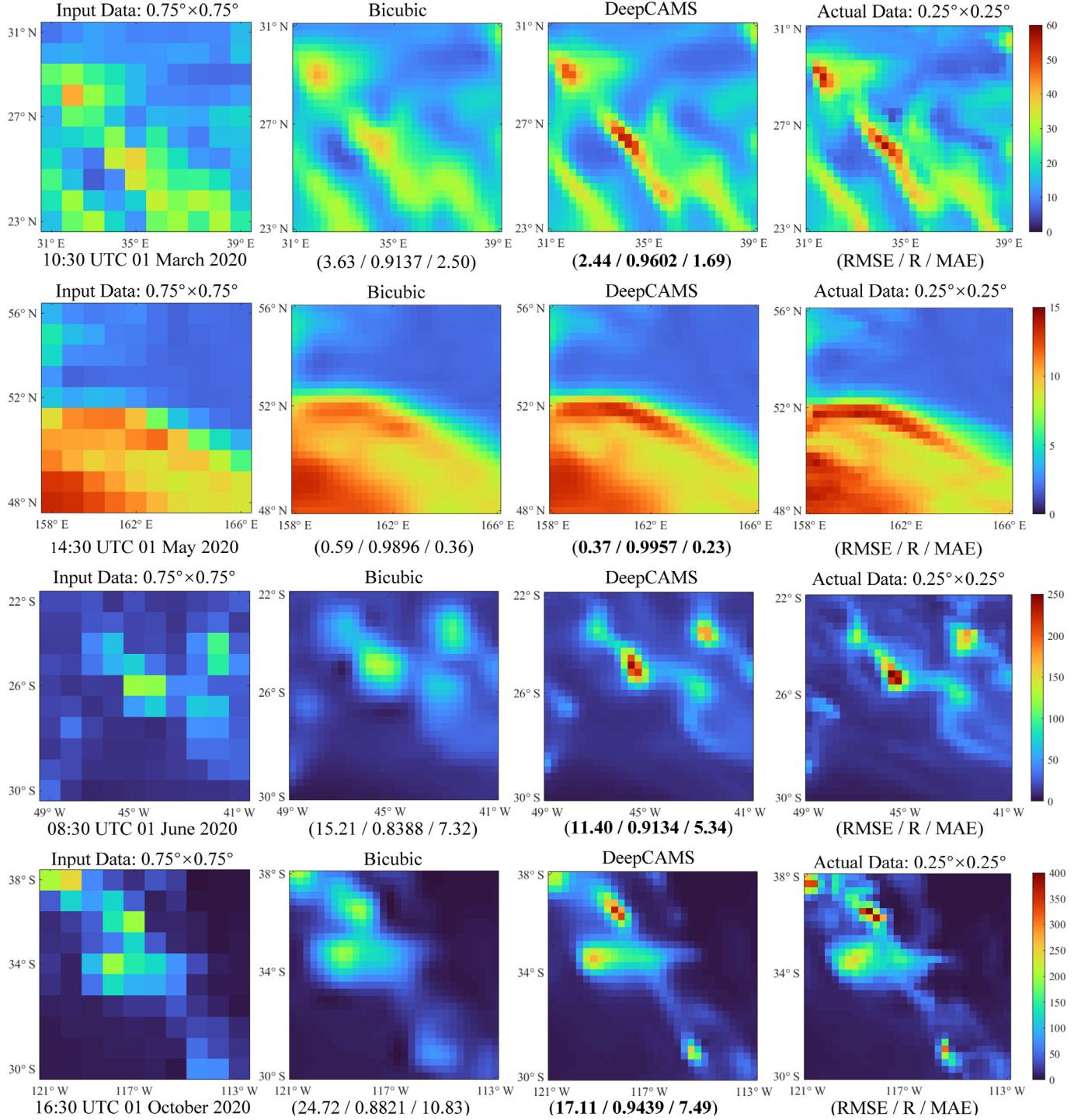
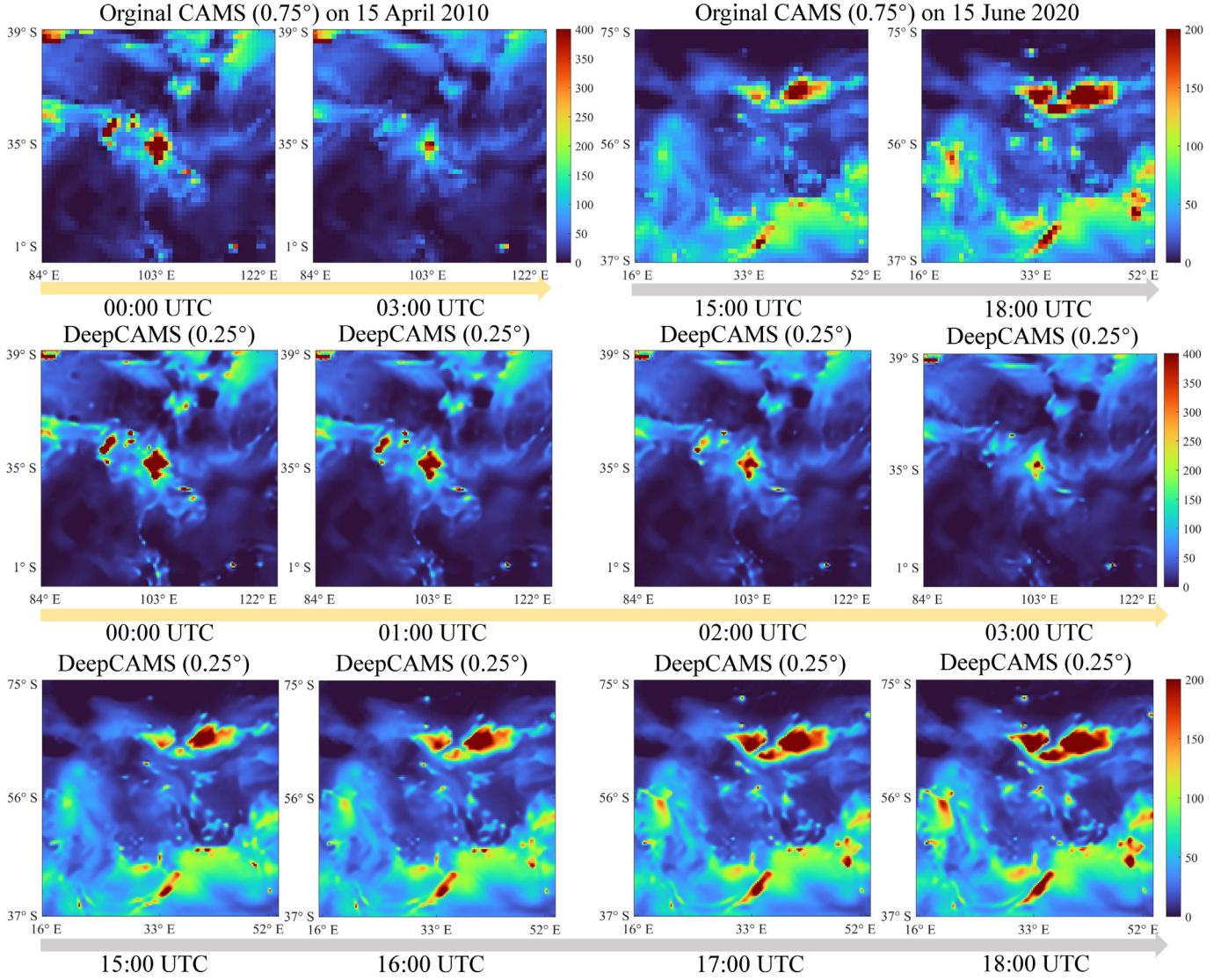**Fig. 4.** The qualitative results of spatial downscaling.

the framework, we did not adopt a multi-frame SR approach (Bhat et al., 2021; Haris et al., 2019) but used a single-frame SR to super-resolve PM$_{2.5}$ independently for each time phase.

ABPN is a state-of-the-art single image SR network whose core is learning recurrent up-downsampling backprojection to explore complex nonlinear mappings from LR PM$_{2.5}$ to HR PM$_{2.5}$. The network structure is shown in Fig. 2. After feature extraction, the input I^{LR} goes through several Enhanced Upsampling Back Projection blocks (EUBP) and Enhanced Downsampling Back Projection block (EDBP). Also, Spatial Attention Blocks (SAB) are injected to learn cross-correlation between features at

different levels. More details about ABPN can be found in Liu et al. (2019). Thanks to ABPN, our DeepCAMS can predict super-resolved PM$_{2.5}$ concentration I^{SR} from an LR input I^{LR}:

$$I^{SR} = ABPN\left(I^{LR}\right) \tag{7}$$

Follow the setting in Liu et al. (2019), the $\mathscr{L}_1$ distance between I^{SR} and ground truth I^{GT} was adopted as the objective function for network

Fig. 5. The real experiment on CAMS. The first row is the original CAMS product (0.75°, 3-h), and the second and third rows correspond to the spatiotemporal downscaling results (0.25°, hourly).

optimization. Also, we use back propagation (BP) to updating the network parameters:

$$\mathscr{L}_1 = \left\| I^{SR} - I^{GT} \right\|_1 \tag{8}$$

By applying the pre-trained ABPN in step III to the temporal downscaled CAMS outputs in step II, a high-spatiotemporal resolution PM$_{2.5}$ product can be obtained.

3. Experiment and analysis

3.1. Simulation experiment on GEOS-CF

Notably, in the simulation experiment on GEOS-CF, we only use the data in 2018 and 2019 to train our DeepCAMS and test on 2020. Therefore, the training set and test set are independent of each other and do not overlap.
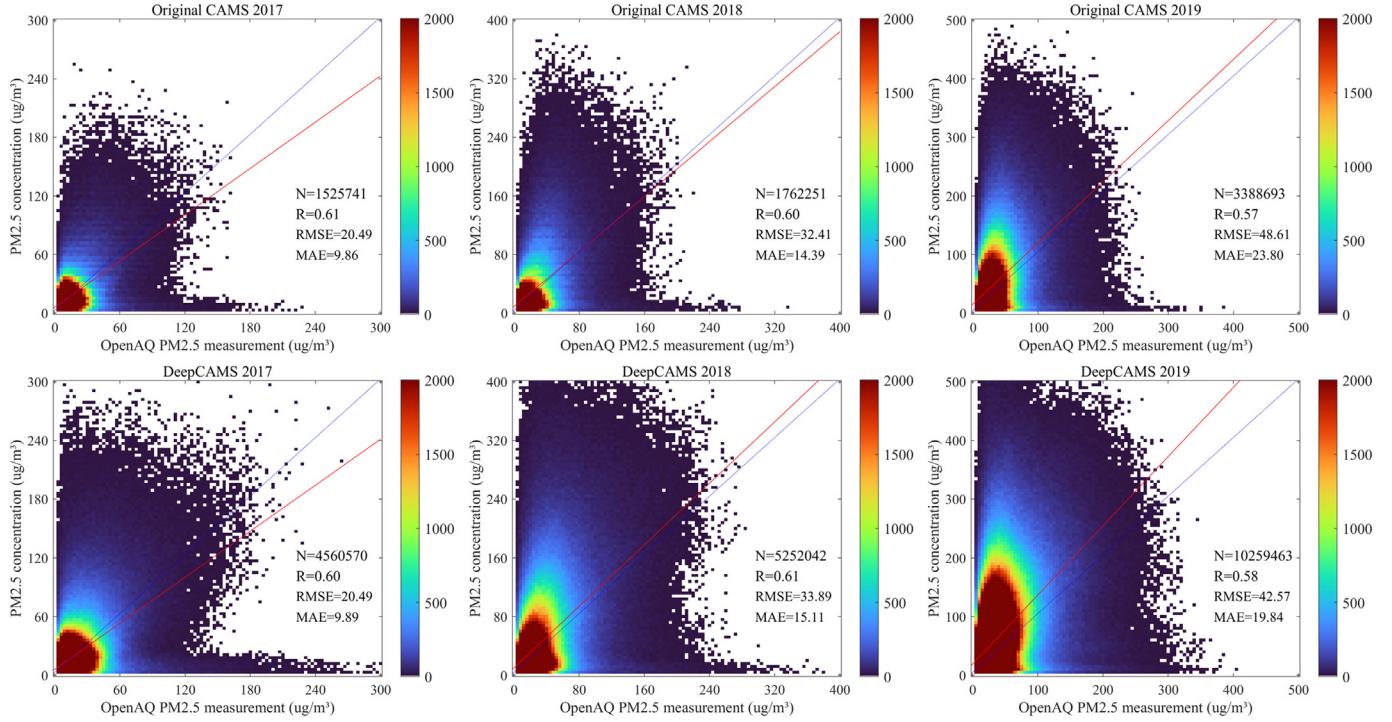
3.1.1. Temporal downscaling

We evaluate the performance of our DeepCAMS on both quantitative and qualitative aspects. Before that, a linear interpolation baseline is set for comparison, which means the intermediate PM$_{2.5}$ concentration $\overline{G}_t$ is

synthesized by linear combination of PM$_{2.5}$ maps $\{G_0, G_1\}$ at adjacent time phases according to the temporal distance:

$$\overline{G}_t = t \times G_0 + (1-t) \times G_1 \tag{9}$$

The quantitative results are shown in Table 2. Here we selected all the temporal interpolated results (16 samples a day) in each month to evaluate the predicted PM$_{2.5}$ concentrations on that month. Generally, our DeepCAMS outperforms linear interpolation on all metrics. Compared to linear, DeepCAMS is ahead -0.84 μg/m^3 in RMSE (Root Mean Square Error), decreases MAE (Mean Absolute Error) by nearly 44 %, and has a higher correlation (R) with the actual PM$_{2.5}$ distribution. That is, the result predicted by DeepCAMS is closest to the real PM$_{2.5}$ evolution in the temporal dimension. Since there are complex interactions between atmospheric components, it is tough for the linear method to describe such a nonlinear chemical reaction. Besides, simple temporal interpolation cannot accommodate various motion and physical phenomena. Hence, it naturally results in poor performance. Thanks to the nonlinearity representation ability of deep learning, DeepCAMS can implicitly learn the complicated tendency of PM$_{2.5}$ from adequate training samples.

The qualitative results can be found in Fig. 3, the performance gap between DeepCAMS and linear is pronounced. We zoomed in on a local
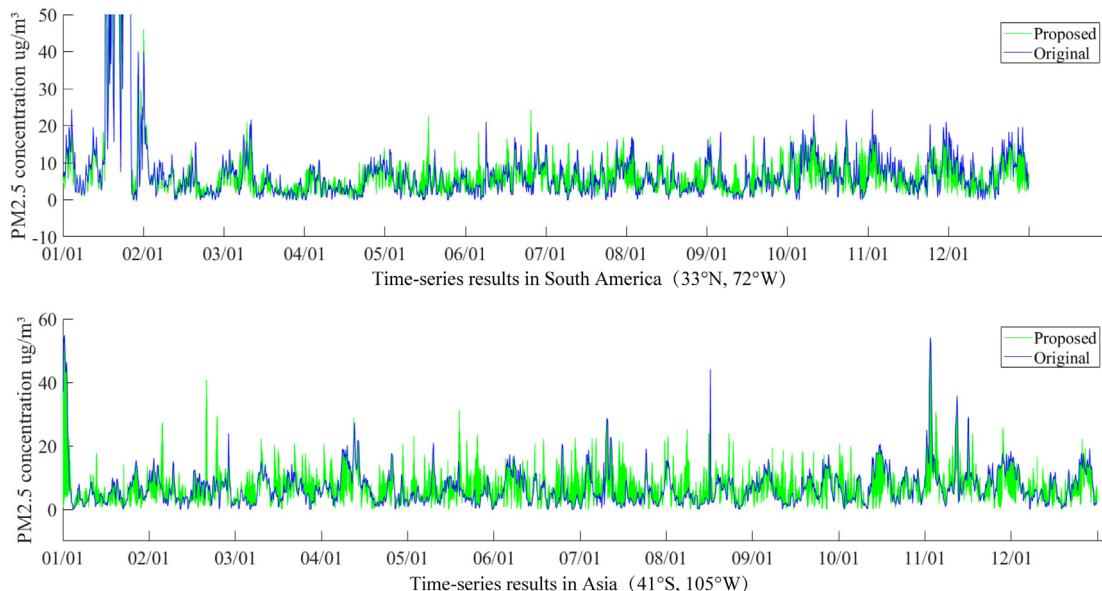
Fig. 6. The in-situ validation of the original CAMS and our DeepCAMS based on OpenAQ platform in 2017, 2018 and 2019.
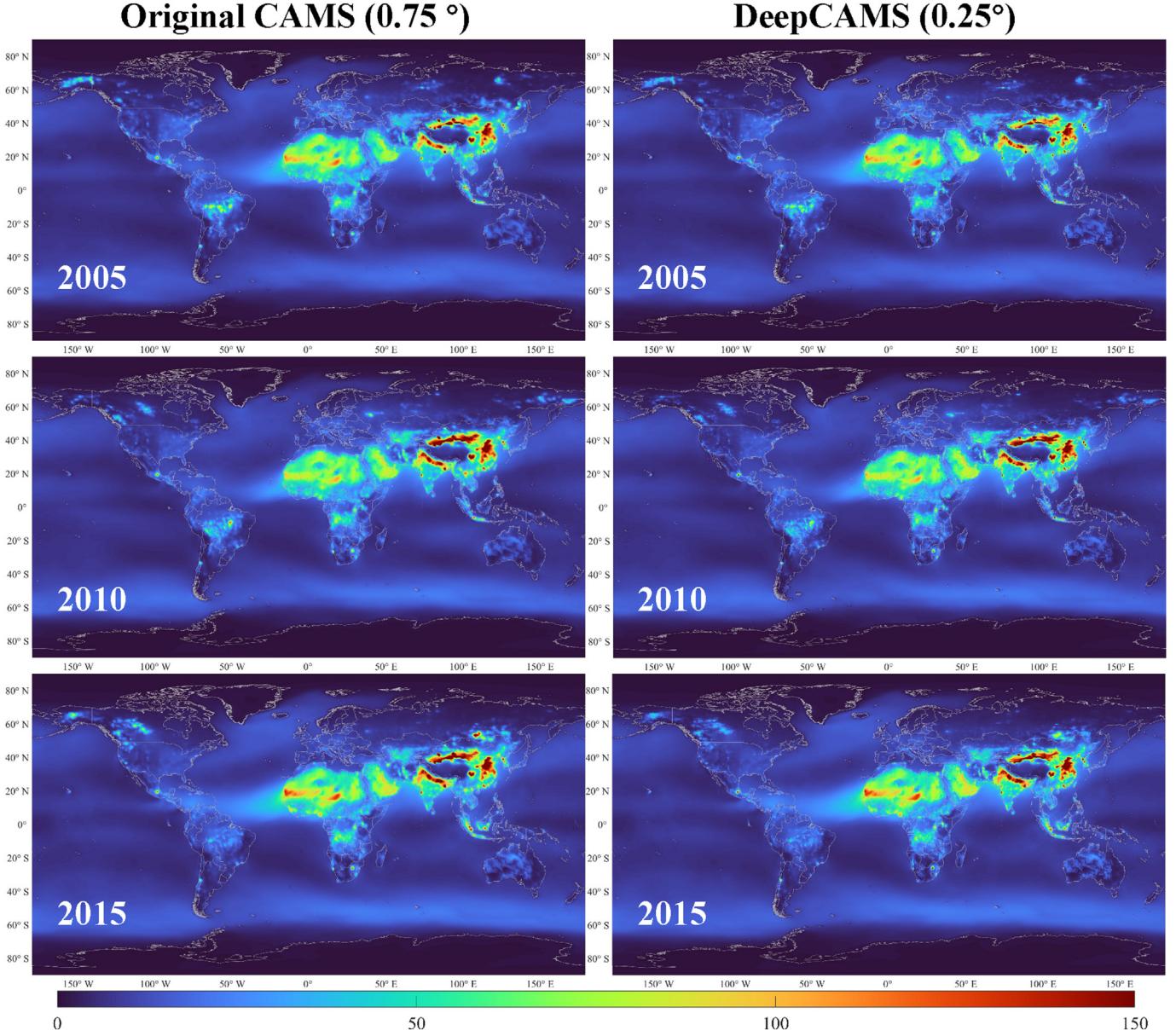
region at 01:30 on January 1, 2020. Our DeepCAMS receives lower RMSE (-15.8 μg/m^3) than linear method. It is easy to find that the linear produces underestimation to some extent. Similarly, linear has revealed an overestimation phenomenon for the forecast at 8:30 on April 1, 2020. This further demonstrates that it is not easy to reflect the diverse dynamics of PM$_{2.5}$ by linear modeling. Nevertheless, DeepCAMS yields accurate predictions, proving that a deep learning-based model can thoroughly learn the underlying nonlinear spatiotemporal relevance. All in all, the latent PM$_{2.5}$ estimates at the intermediate time phase can be precisely excavated from the existing PM$_{2.5}$ statistical distribution through our DeepCAMS.

3.1.2. Spatial downscaling

Spatial downscaling aims to generate more detail for fine-grained analysis. The R, RMSE, and MAE results are in Table 3, which illustrate the superiority of DeepCAMS. We set bicubic interpolation as the baseline and selected all samples for each day of each month for evaluation. Each day contains 24 hourly PM$_{2.5}$ samples, and the final average result is the mean of the total samples ($24 \times 365 = 8760$ samples) for 12 months. DeepCAMS still leads the baseline method on all metrics. That shows deep learning can explore the nonlinear spatial relationship of PM$_{2.5}$ and predict the high-resolution data from the low-resolution

Fig. 7. Time-series results of the original and proposed product in 2017 (top) and 2019 (bottom).

Original CAMS (0.75 °)   DeepCAMS (0.25°)

Fig. 8. The annual average (8760 samples) of our global hourly 0.25° PM$_{2.5}$ dataset (left) and the annual average (2920 samples) of the original 3-h 0.75° CAMS product (right). The proposed dataset has a good spatiotemporal consistency with the original CAMS.

statistical distribution. In addition to quantitative performance, we focus more on the visual effect of the spatial downscaling results. In Fig. 4, we locally enlarge some details for better observation. On the sample at 10:30 on March 1, 2020, DeepCAMS recovered a more realistic super-resolution result and obtained clearer texture and detailed information of PM$_{2.5}$ distribution. The same results can also be received in the other three samples. The Bicubic interpolation cannot find out the spatial regularity of PM$_{2.5}$. It weighted and combined the surrounding pixels in the sampling grid and obtained the result with severe oversmoothing and blurring. It is worth noting that the bicubic fails to recover the high-value part of PM$_{2.5}$. This may be because of the unbalanced spread and complex relationship between the high-value and the low-value PM$_{2.5}$. We discover that high concentrations of PM$_{2.5}$ tend to be concentrated in a local area in the spatial coordinate. However, bicubic has a "global-average" drawback, which will pull down the entire concentration. DeepCAMS learned this pattern from the training data and can restore the proper statistical property of PM$_{2.5}$.

3.2. Real experiment on CAMS

In the real experiment, we used all GEOS-CF data from 2018 to 2020 to train our DeepCAMS, then we apply the model trained on GEOS-CF to the CAMS dataset (0.75°, 3-h). In this manner, spatiotemporal relationships learned in high-resolution data can be migrated to low-resolution data. Since we have no ground truth hourly high-spatiotemporal resolution PM$_{2.5}$ product in real experiment, we demonstrate the effectiveness of DeepCAMS based on visual validation and in-situ PM$_{2.5}$ validation.

In Fig. 5, we display the results of the spatiotemporal downscaling of the CAMS products on April 15, 2010 (from 00:00 to 03:00 UTC) and on June 15, 2020 (from 15:00 to 18:00 UTC), with yellow and gray arrows representing the time axis, respectively. Some dynamic results can be found at https://github.com/XY-boy/DeepCAMS. In the spatial dimension, the reconstructed products have more apparent details compared with the original CAMS dataset, and the spatial distribution characteristics of PM$_{2.5}$ are correctly preserved. In the temporal dimension, DeepCAMS legitimately

predicts the temporal variation of $PM_{2.5}$ at the non-existing time phase. Note in the second row, the $PM_{2.5}$ concentration has a decreasing trend from 00:00 to time 03:00, and the high-value region tends to contract. DeepCAMS correctly reflects this spatiotemporal continuum and generates a convincing forecast. Similarly, in the third row, there is a clear tendency for $PM_{2.5}$ to become more concentrated from 15:00 to 18:00, and DeepCAMS still deduces a reasonable evolutionary procedure. To further verify the temporal consistency, two subregions (South America: 33°N, 72°W, Asia: 41°S, 105°W) were chosen to demonstrate the time series variation of $PM_{2.5}$ throughout the year. As shown in Fig. 7, the proposed dataset has periodically consistent variation with the original CAMS. In summary, DeepCAMS maintains CAMS's temporal continuity and spatial consistency and produces a product with more acceptable spatiotemporal resolution. In Fig. 8, we display the global $PM_{2.5}$ mapping results by averaging all samples of the corresponding year. In the previous local-scale results, we find that DeepCAMS has finer texture, and here on the global-scale aspects, DeepCAMS still produces a consistent distribution pattern compared with the original product. This fully illustrates that the proposed long-term global hourly high-spatiotemporal resolution $PM_{2.5}$ dataset has a satisfactory accuracy with the original product at the regional and global scales.

In Fig. 6, we used the OpenAQ $PM_{2.5}$ air quality data from 2017 to 2019 for validation. Correlation (R) was calculated to reveal the errors between the $PM_{2.5}$ concentrations measured by the ground stations and the DeepCAMS product. The scatter density plot for the validation explains that DeepCAMS maintains consistent accuracy with the original CAMS while tripling the temporal and spatial resolution. In the whole year of 2017, the original CAMS product has a precision of $R = 0.61$ and DeepCAMS reaches $R = 0.6$. In 2018 and 2019, the validation accuracy of the high-spatiotemporal resolution product generated by DeepCAMS was slightly improved compared with the original product, which indicates that the predicted missing $PM_{2.5}$ data at the non-existent temporal phase has a high agreement with the real situation. Overall, DeepCAMS has substantially improved the spatiotemporal resolution of CAMS while maintaining its accuracy.

4. Conclusion

This paper proposed a long-term global hourly 0.25° $PM_{2.5}$ dataset based on a deep learning framework (DeepCAMS). With the help of DeepCAMS, we conducted spatiotemporal downscaling in the global $PM_{2.5}$ data products provided by CAMS and tripled its resolution on spatial and temporal dimensions simultaneously.

For temporal downscaling, the temporal dynamic trends of $PM_{2.5}$ are deeply explored by a video frame interpolation model XVFI. The test results in GEOS-CF (2020) show that DeepCAMS can reasonably infer the various motions and complex physicochemical phenomena in the air. Compared to a temporal linear interpolation approach, DeepCAMS correctly predict the underlying evolution pattern of $PM_{2.5}$. Real experiments on CAMS data also prove that DeepCAMS yields temporally continuous and reasonable predictions. For the spatial downscaling, the low-resolution $PM_{2.5}$ map is mapped to high-resolution space by a single-image super-resolution model ABPN. Simulation experiments on GEOS-CF demonstrate that DeepCAMS can produce more spatial detail for fine-grained analysis. Furthermore, real experiments on CAMS illustrate that the proposed high-quality product has a spatially consistent distribution with the original CAMS product on regional and global scales. Finally, the validation with the in-situ OpenAQ $PM_{2.5}$ air quality data from 2017 to 2019 indicates that the proposed dataset maintains the accuracy of the original CAMS products. In short, the developed DeepCAMS is conductive to generate long-term global hourly 0.25° $PM_{2.5}$ product with the same credibility as the low-resolution CAMS dataset.

In future work, we plan to develop an end-to-end spatiotemporal downscaling framework to moderate the accumulation of errors due to multi-step processing. In addition, since the downscaling process does not introduce auxiliary variables that can improve the accuracy, we will try our best to further improve the accuracy by introducing more auxiliary variables (such as wind speed and temperature).

CRediT authorship contribution statement

Yi Xiao: Methodology, data collection, analysis of experiment, Writing-Original draft.
Yuan Wang: Methodology, data collection.
Qiangqiang Yuan: Supervision.
Jiang He: Data collection.
Liangpei Zhang: Supervision.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

Bae, M., Kim, B.U., Kim, H.C., Kim, J., Kim, S., 2021. Role of emissions and meteorology in the recent pm2.5 changes in China and South Korea from 2015 to 2018. Environ. Pollut. 270, 116233.

Bai, Y., Zhao, T., Hu, W., Zhou, Y., Xiong, J., Wang, Y., Liu, L., Shen, L., Kong, S., Meng, K., et al., 2022. Meteorological mechanism of regional pm2.5 transport building a receptor region for heavy air pollution over central China. Sci. Total Environ. 808, 151951.

Baño-Medina, J., Manzanas, R., Gutiérrez, J.M., 2020. Configuration and intercomparison of deep learning neural models for statistical downscaling. Geosci. Model Dev. 13 (4), 2109–2124.

Baño-Medina, J., Manzanas, R., Gutiérrez, J.M., 2021. On the suitability of deep convolutional neural networks for continental-wide downscaling of climate change projections. Clim. Dyn. 57 (11), 2941–2951.

Bhat, G., Danelljan, M., Yu, F., Van Gool, L., Timofte, R., 2021. Deep reparametrization of multi-frame super-resolution and denoising. Proceedings of the IEEE/CVF International Conference on Computer Vision, pp. 2460–2470.

Chen, S., Gan, T.Y., Tan, X., Shao, D., Zhu, J., 2019. Assessment of cfsr, era-interim, jra-55, merra-2, ncep-2 reanalysis data for drought analysis over China. Clim. Dyn. 53 (1), 737–757.

Choi, M., Kim, H., Han, B., Xu, N., Lee, K.M., 2020. Channel attention is all you need for video frame interpolation. Proceedings of the AAAI Conference on Artificial Intelligence, pp. 10663–10671.

Colarco, P., da Silva, A., Chin, M., Diehl, T., 2010. Online simulations of global aerosol distributions in the nasa geos-4 model and comparisons to satellite and ground-based aerosol optical depth. J.Geophys.Res.Atmos. 115.

Gu, J., Feng, J., Hao, X., Fang, T., Zhao, C., An, H., Chen, J., Xu, M., Li, J., Han, W., et al., 2022. Establishing a non-hydrostatic global atmospheric modeling system at 3-km horizontal resolution with aerosol feedbacks on the sunway supercomputer of China. Sci.Bull. 67 (11), 1170–1181.

Haris, M., Shakhnarovich, G., Ukita, N., 2019. Recurrent back-projection network for video super-resolution. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, pp. 3897–3906.

Hasenkopf, C.A., 2017. Sharing lessons-learned on effective open data, open-source practices from openaq, a global open air quality community. AGU Fall Meeting Abstracts pp. IN44A–08.

Hasenkopf, C.A., Flasher, J.C., Veerman, O., Scalamogna, A., Silva, D., Salmon, M., Buuralda, D., DeWitt, L.H., 2016. Stories from openaq, a global and grassroots open air quality community. AGU Fall Meeting Abstracts pp. PA43A–2162.

He, J., Li, J., Yuan, Q., Shen, H., Zhang, L., 2021. Spectral response function-guided deep optimization-driven network for spectral super-resolution. IEEE Transactions on Neural Networks and Learning Systems.

He, J., Yuan, Q., Li, J., Xiao, Y., Liu, X., Zou, Y., 2022a. Dster: a dense spectral transformer for remote sensing spectral super-resolution. Int. J. Appl. Earth Obs. Geoinf. 109, 102773.

He, J., Yuan, Q., Li, J., Zhang, L., 2022b. Ponet: a universal physical optimization-based spectral super-resolution network for arbitrary multispectral images. Inf.Fusion 80, 205–225.

Hu, M., Jiang, K., Liao, L., Xiao, J., Jiang, J., Wang, Z., 2022. Spatial-temporal space hand-in-hand: spatial-temporal video super-resolution via cycle-projected mutual learning. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, pp. 3574–3583.

Inness, A., Ades, M., Agusti-Panareda, A., Barré, J., Benedictow, A., Blechschmidt, A.M., Dominguez, J.J., Engelen, R., Eskes, H., Flemming, J., et al., 2019. The cams reanalysis of atmospheric composition. Atmos. Chem. Phys. 19 (6), 3515–3556.

Jaderberg, M., Simonyan, K., Zisserman, A., et al., 2015. Spatial transformer networks. Advances in Neural Information Processing Systems, p. 28.

Jiang, K., Wang, Z., Yi, P., Wang, G., Lu, T., Jiang, J., 2019. Edge-enhanced GAN for remote sensing image superresolution. IEEE Trans. Geosci. Remote Sens. 57 (8), 5799–5812.

Jin, C., Wang, Y., Li, T., Yuan, Q., 2022. Global validation and hybrid calibration of cams and merra-2 pm2.5 reanalysis products based on openaq platform. Atmos. Environ. 274, 118972.

Johnston, F.H., Borchers-Arriagada, N., Morgan, G.G., Jalaludin, B., Palmer, A.J., Williamson, G.J., Bowman, D.M., 2021. Unprecedented health costs of smoke-related pm2.5 from the 2019–20 Australian megafires. Nat.Sustain. 4 (1), 42–47.

Keller, C.A., Knowland, K.E., Duncan, B.N., Liu, J., Anderson, D.C., Das, S., Lucchesi, R.A., Lundgren, E.W., Nicely, J.M., Nielsen, E., et al., 2021. Description of the nasa geos composition forecast modeling system geos-cf v1. 0. J.Adv.Model.Earth Syst. 13, e2020MS002413.

Kocak, E., Celik, B., 2022. The nexus between access to energy, poverty reduction and pm2.5 in Sub-Saharan Africa: new evidence from the generalized method of moments estimators. Sci. Total Environ. 154377.

Kuenen, J., Dellaert, S., Visschedijk, A., Jalkanen, J.P., Super, I., Denier van der Gon, H., 2022. Cams-reg-v4: a state-of-the-art high-resolution European emission inventory for air quality modelling. Earth Syst.Sci.Data 14, 491–515.

Lelieveld, J., Evans, J.S., Fnais, M., Giannadaki, D., Pozzer, A., 2015. The contribution of outdoor air pollution sources to premature mortality on a global scale. Nature 525, 367–371.

Li, F., Yan, J., Wei, Y., Zeng, J., Wang, X., Chen, X., Zhang, C., Li, W., Chen, M., Lü, G., 2021. Pm2.5-bound heavy metals from the major cities in China: spatiotemporal distribution, fuzzy exposure assessment and health risk management. J. Clean. Prod. 286, 124967.

Liu, D., Grimmond, C., Tan, J., Ao, X., Peng, J., Cui, L., Ma, B., Hu, Y., Du, M., 2018. A new model to downscale urban and rural surface and air temperatures evaluated in Shanghai, China. J. Appl. Meteorol. Climatol. 57, 2267–2283.

Liu, Z.S., Wang, L.W., Li, C.T., Siu, W.C., Chan, Y.L., 2019. Image super-resolution via attention based back projection networks. 2019 IEEE/CVF International Conference on Computer Vision Workshop (ICCVW), IEEE, pp. 3517–3525.

Liu, A., Liu, Y., Gu, J., Qiao, Y., Dong, C., 2021. Blind Image Super-resolution: A Survey and Beyond arXiv preprint arXiv:2107.03055.

Lv, B., Hu, Y., Chang, H.H., Russell, A.G., Cai, J., Xu, B., Bai, Y., 2017. Daily estimation of ground-level pm2.5 concentrations at 4 km resolution over Beijing-Tianjin-Hebei by fusing modis aod and ground observations. Sci. Total Environ. 580, 235–244.

Nourani, V., Baghanam, A.H., Gokcekus, H., 2018. Data-driven ensemble model to statistically downscale rainfall using nonlinear predictor screening approach. J. Hydrol. 565, 538–551.

Pu, Q., Yoo, E.H., 2021. Ground pm2.5 prediction using imputed maiac aod with uncertainty quantification. Environ. Pollut. 274, 116574.

Shen, H., Tao, S., Chen, Y., Ciais, P., Güneralp, B., Ru, M., Zhong, Q., Yun, X., Zhu, X., Huang, T., et al., 2017. Urbanization-induced population migration has reduced ambient pm2.5 concentrations in China. ScienceAdvances 3, e1700300.

Sim, H., Oh, J., Kim, M., 2021. Xvfi: extreme video frame interpolation. Proceedings of the IEEE/CVF International Conference on Computer Vision, pp. 14489–14498.

Thind, M.P., Heath, G., Zhang, Y., Bhatt, A., 2022. Characterization factors and other air quality impact metrics: case study for pm2.5-emitting area sources from biofuel feedstock supply. Sci. Total Environ. 822, 153418.

Varga-Balogh, A., Leelőssy, Á., Lagzi, I., Mészáros, R., 2020. Time-dependent downscaling of pm2.5 predictions from cams air quality models to urban monitoring sites in Budapest. Atmosphere 11, 669.

Wang, J., Zhong, Y., Zheng, Z., Ma, A., Zhang, L., 2020. Rsnet: the search for remote sensing deep neural networks in recognition tasks. IEEE Trans. Geosci. Remote Sens. 59, 2520–2534.

Wang, F., Tian, D., Lowe, L., Kalin, L., Lehrter, J., 2021a. Deep learning for daily precipitation and temperature downscaling. Water Resour. Res. 57, e2020WR029308.

Wang, Y., Yuan, Q., Li, T., Zhu, L., Zhang, L., 2021b. Estimating daily full-coverage near surface o3, co, and no2 concentrations at a high spatial resolution over China based on s5p-tropomi and geos-fp. ISPRS J. Photogramm. Remote Sens. 175, 311–325.

Wang, J., Ma, A., Zhong, Y., Zheng, Z., Zhang, L., 2022a. Cross-sensor domain adaptation for high spatial resolution urban land-cover mapping: from airborne to spaceborne imagery. Remote Sens. Environ. 277, 113058.

Wang, Y., Yuan, Q., Li, T., Zhu, L., 2022b. Global spatiotemporal estimation of daily high-resolution surface carbon monoxide concentrations using deep forest. J. Clean. Prod. 131500.

Wang, Y., Yuan, Q., Zhu, L., Zhang, L., 2022c. Spatiotemporal estimation of hourly 2-km ground-level ozone over China based on himawari-8 using a self-adaptive geospatially local model. Front. 13, 101286.

Xiao, Y., Su, X., Yuan, Q., Liu, D., Shen, H., Zhang, L., 2022a. Satellite video super-resolution via multiscale deformable convolution alignment and temporal grouping projection. IEEE Trans. Geosci. Remote Sens. 60, 1–19. https://doi.org/10.1109/TGRS.2021.3107352.

Xiao, Y., Yuan, Q., He, J., Zhang, Q., Sun, J., Su, X., Wu, J., Zhang, L., 2022b. Space-time super-resolution for satellite video: a joint framework based on multi-scale spatial-temporal transformer. Int. J. Appl. Earth Obs. Geoinf. 108, 102731.

Xue, T., Zheng, Y., Tong, D., Zheng, B., Li, X., Zhu, T., Zhang, Q., 2019. Spatiotemporal continuous estimates of pm2.5 concentrations in China, 2000–2016: a machine learning method with inputs from satellites, chemical transport model, and ground observations. Environ. Int. 123, 345–357.

Yang, Q., Yuan, Q., Li, T., Yue, L., 2020. Mapping pm2.5 concentration at high resolution using a cascade random forest based downscaling model: evaluation and application. J. Clean. Prod. 277, 123887.

Yang, X., Zhang, T., Zhang, Y., Chen, H., Sang, S., 2021. Global burden of copd attributable to ambient pm2.5 in 204 countries and territories, 1990 to 2019: a systematic analysis for the global burden of disease study 2019. Sci. Total Environ. 796, 148819.

Yi, P., Wang, Z., Jiang, K., Jiang, J., Lu, T., Ma, J., 2020. A progressive fusion generative adversarial network for realistic and consistent video super-resolution. IEEE Trans. Pattern Anal. Mach. Intell. 44 (5), 2264–2280.

Zhai, Y., Liu, X., Chen, H., Xu, B., Zhu, L., Li, C., Zeng, G., 2014. Source identification and potential ecological risk assessment of heavy metals in pm2.5 from Changsha. Sci. Total Environ. 493, 109–115.

Zhang, Q., Yuan, Q., Li, J., Wang, Y., Sun, F., Zhang, L., 2021a. Generating seamless global daily amsr2 soil moisture (sgd-sm) long-term products for the years 2013–2019. Earth Syst.Sci.Data 13, 1385–1401.

Zhang, Q., Yuan, Q., Li, Z., Sun, F., Zhang, L., 2021b. Combined deep prior with low-rank tensor svd for thick cloud removal in multitemporal images. ISPRS J. Photogramm. Remote Sens. 177, 161–173.

Zhang, Y., Li, J., Li, J., Pan, X., Wang, W., Zhu, L., Wang, Z., Chen, X., Yang, W., Wang, Z., 2022. An intercomparison of ozone taken from the copernicus atmosphere monitoring service and the second modern-era retrospective analysis for research and applications over China during 2018 and 2019. J. Environ. Sci.

Zheng, Z., Zhong, Y., Tian, S., Ma, A., Zhang, L., 2022. ChangeMask: Deep multi-task encoder-transformer-decoder architecture for semantic change detection. ISPRS J. Photogramm. Remote Sens. 183, 228–239.

Zheng, Z., Zhong, Y., Wang, J., Ma, A., Zhang, L., 2021. Building damage assessment for rapid disaster response with a deep object-based semantic change detection framework: From natural disasters to man-made disasters. Remote Sens. Environ. 265, 112636.

Zhu, Y., Chen, G., Xu, L., Zhang, Y., Wang, Y., Liang, S., 2022. Inequality of household consumption and pm2.5 footprint across socioeconomic groups in China. Environ. Res. Lett.