CLIMLORA: FREQUENCY-AWARE DUAL-LORA FOR FINE-SCALE AND PHYSICS-DECOUPLED CLIMATE DOWNSCALING

Anonymous authors
Paper under double-blind review

ABSTRACT

High-resolution meteorological fields are essential for extreme-weather analysis and regional climate studies. In recent years, machine-learning-based super-resolution methods have provided an efficient approach to climate downscaling. However, existing methods typically optimize data reconstruction and physical constraints within a shared parameter space, coupling the two objectives and making it more difficult to achieve both reconstruction accuracy and physical structural consistency. To address this challenge, we propose ClimLoRA, a frequency-aware Dual-LoRA framework for climate downscaling that separates parameter updates for fine-scale reconstruction and physical structure correction and coordinates the two objectives through sequential training. We first formulate climate downscaling as learning the residual between low-resolution inputs and high-resolution targets, and then decouple the learning objectives into two LoRA parameter spaces: one learns fine-scale reconstruction through high-frequency modeling and spherical harmonic spectral constraints, while the other promotes physical consistency through physical structural constraints on spherical grids. We conduct global multivariable joint downscaling experiments under both ERA5-to-ERA5 in-domain and MPI-ESM-to-ERA5 cross-source settings, covering near-surface and upper-air meteorological variables. Experimental results show that ClimLoRA achieves leading overall downscaling performance while maintaining strong reconstruction capability in global downscaling.

1 INTRODUCTION

High-resolution meteorological information provides an essential basis for weather forecasting, extreme-weather analysis, and regional climate studies (Bi et al., 2023), while also supporting applications in agriculture, environmental protection, and hydrological management (Buster et al., 2024). However, the limited spatial resolution of meteorological data restricts the characterization of local atmospheric variations. Meteorological downscaling provides an effective way to bridge this gap by recovering fine-scale information from coarse-resolution fields (Vandal et al., 2017; Srivastava et al., 2024; Guan et al., 2026). Traditional downscaling methods typically rely on dynamical models or spatial interpolation to recover fine-scale information from coarse-resolution inputs (Sun et al., 2024), but still face challenges in complex scenarios and high-resolution reconstruction.

In recent years, artificial intelligence techniques have shown strong potential for meteorological downscaling (Vandal et al., 2017; Nguyen et al., 2023; Srivastava et al., 2024). SwinRDM (Chen et al., 2023) employs diffusion models to reconstruct meteorological fields at higher spatial resolutions. GeoFAR (Xu et al., 2026) recovers fine-scale climate information through frequency-aware modeling and geographic representations. SGD (Tu et al., 2025) further combines GridSat satellite observations with low-resolution ERA5 fields to guide diffusion models in generating high-resolution meteorological states. These studies demonstrate that deep generative models can effectively recover spatial details lost during downscaling.

However, existing methods mainly focus on the reconstruction accuracy of high-resolution meteorological fields, while physical consistency remains insufficiently modeled (Harder et al., 2023; Luo et al., 2025; Saccardi et al., 2025). A direct approach is to incorporate physical constraints

Reviewers: please read the Reviewer Guidelines (iclr.cc/Conferences/2027/ReviewerGuidelines) and the AI Policy for Reviewers (iclr.cc/Conferences/2027/AIPolicyForReviewers). If you used AI to expand, edit, or polish your review, please provide the input text to the LLM. Better still, consider skipping the LLM and submitting your original text: we, and the authors, are much more interested in your unedited thoughts than in what an LLM has to say. AI-assisted or not, you are putting your name and reputation behind your review: LLM-generated falsehoods, hallucinations or misrepresentations are subject to disciplinary action, which may include desk-rejecting all papers you have authored.

into the reconstruction objective and jointly optimize them within the same parameter space (Bastek et al., 2025). However, the two objectives may exhibit different gradient directions and magnitudes, leading to optimization interference (Yu et al., 2020; Ding et al., 2023). Even when the data objective and physical constraints are not inherently conflicting, shared parameters may still struggle to satisfy both optimization demands simultaneously. Minimizing reconstruction error alone cannot guarantee the physical consistency of the generated fields, while corrections introduced by physical constraints may affect the learned reconstruction performance. Therefore, reducing optimization interference and coordinating fine-scale reconstruction with physical-structure optimization remain key challenges in meteorological downscaling.

To balance prediction accuracy and physical consistency in meteorological downscaling, we propose a frequency-aware framework, ClimLoRA, as illustrated in Figure 1. Built upon a Base model trained for low-frequency residual reconstruction, ClimLoRA separates the parameters associated with data fitting and physical constraints into two independent LoRA parameter spaces to reduce optimization interference. We first train a low-frequency reconstruction Base to learn the large-scale residual structures between low-resolution and high-resolution meteorological fields. We then freeze the Base and train Fine-scale LoRA, which recovers local details through high-frequency constraints and preserves multi-scale spectral structures consistent with global spherical geometry through a spherical harmonic spectral constraint. Based on this stage, we further freeze the Base and Fine-scale LoRA and train only Physics LoRA, which learns physical structure corrections from wind-field vorticity and divergence, as well as the horizontal gradients of temperature and pressure in spherical coordinates. During inference, the scaling factors of the two branches can be adjusted independently, enabling control over the relative contributions of fine-scale reconstruction and physical correction without retraining. In summary, our main contributions are:

- We propose a climate downscaling framework that separates fine-scale reconstruction and physical-consistency optimization into independent LoRA parameter spaces to reduce optimization interference.
- We design Fine-scale LoRA and Physics LoRA, which recover fine-scale structures through high-frequency and spherical harmonic spectral constraints, and learn physical-consistency corrections through spherical physical structure modeling, respectively.
- We conduct global multivariable joint downscaling experiments in the in-domain ERA5-to-ERA5 and cross-source MPI-ESM-to-ERA5 settings, validating the effectiveness of ClimLoRA across different data sources and meteorological variables.

2 Related work

2.1 Climate Downscaling

Traditional meteorological downscaling methods mainly include dynamical and statistical approaches. Dynamical downscaling uses global climate models to drive regional climate models, providing a clear physical basis but at high computational cost; statistical downscaling learns coarse-to-fine relationships from historical observations, offering higher efficiency but limited capability in modeling complex nonlinear relationships. With the development of deep learning, CNNs, Transformers, and generative models have been widely applied to meteorological downscaling (Vandal et al., 2017; Liu et al., 2020; Yang et al., 2024). DeepSD (Vandal et al., 2017) employs cascaded convolutional networks, while DSFNO Yang et al. (2024) models cross-scale dependencies through frequency-domain representations. Generative approaches characterize fine-scale variability and uncertainty by learning conditional distributions, including GAN-based methods (Stengel et al., 2020; Lopez-Gomez et al., 2025) and diffusion models Watt & Mansfield (2024). SwinRDM Chen et al. (2023) uses conditional diffusion to generate high-resolution meteorological fields, while CorrDiff Mardani et al. (2025) decomposes downscaling into conditional-mean prediction and stochastic residual generation. SGD (Tu et al., 2025) incorporates satellite observations to guide diffusion generation, whereas GeoFAR (Xu et al., 2026) enhances fine-scale reconstruction through frequency-aware representations and geographic information. Despite these advances, jointly recovering local fine-scale structures and maintaining physical consistency remains an open challenge.

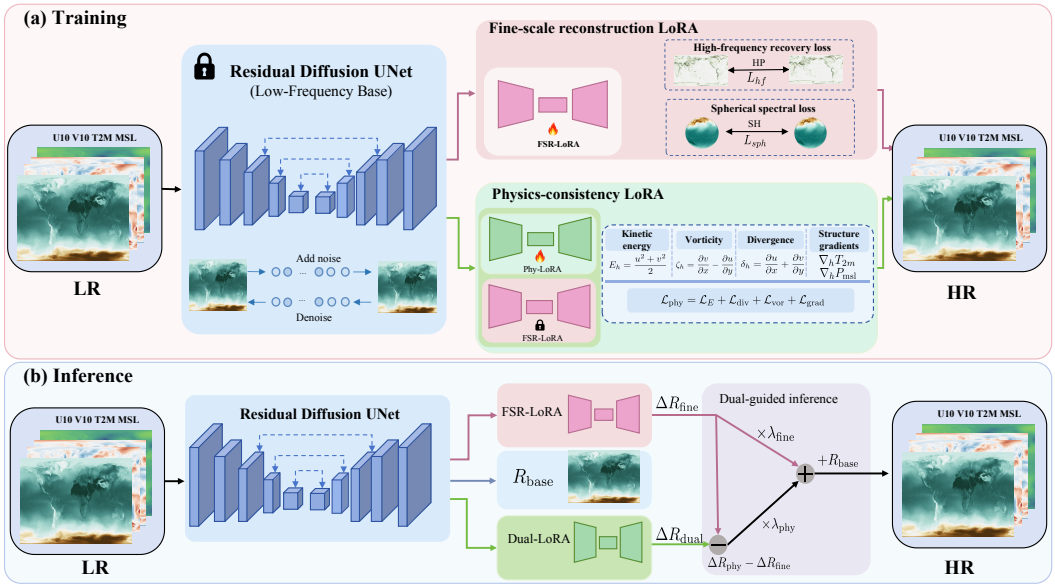


Figure 1: Overview of the ClimLoRA framework. (a) Training: The low-frequency Base reconstructs the large-scale structures of meteorological fields. With the Base frozen, FSR-LoRA recovers fine-scale spatial and spectral information through high-frequency reconstruction and spherical harmonic spectral constraints. We then freeze both the Base and FSR-LoRA and train only Phy-LoRA to learn physical structure corrections using horizontal kinetic energy, divergence, vorticity, and related quantities. (b) Inference: The trained branches are combined through dual-guided inference without retraining. The coefficients λ_{fine} and λ_{phy} control the contributions of fine-scale recovery and physical correction, respectively, enabling adjustable meteorological downscaling.

2.2 Physics-informed neural networks.

Integrating physical knowledge with machine learning has attracted growing interest in Earth system modeling, primarily through physical constraints and hybrid modeling Verma et al. (2024). The former incorporates known physical relationships as soft or hard constraints Bastek et al. (2025); Blanke et al. (2026). Harder et al. (2023) enforce consistency between high-resolution outputs and low-resolution inputs through hard constraints such as SmCL. PhyDL-NWP (Luo et al., 2025) computes physics-related terms using automatic differentiation and uses physics-based losses to guide weather forecasting and continuous-space downscaling. Hybrid modeling embeds physical computations into neural networks. NeuralGCM (Kochkov et al., 2024) combines a differentiable atmospheric dynamical core with learned physical parameterizations, while WeatherGFT (Xu et al., 2024) combines PDE-based evolution with neural correction modules to enable weather forecasting at finer temporal scales.These methods demonstrate the complementarity between physical knowledge and data-driven modeling, but reconstruction and physical objectives optimized over shared parameters may interfere due to differences in gradient directions and magnitudes. To address this issue, we use low-rank adaptation to separate the parameter updates of the two objectives, preventing the physical objective from directly modifying the learned reconstruction parameters.

2.3 Diffusion-Based Super-Resolution

Diffusion models recover fine-scale information in super-resolution tasks through generative modeling Wang et al. (2024). SR3 (Saharia et al., 2022) formulates super-resolution as a conditional diffusion process and reconstructs high-resolution images through iterative denoising. ResShift (Yue et al., 2023) connects low-quality and high-quality images through residual shifting, while OSEDiff (Wu et al., 2024) leverages the prior of pretrained Stable Diffusion to achieve one-step super-resolution directly from low-quality inputs. These studies provide a foundation for efficient residual reconstruction, but mainly focus on reconstruction and perceptual quality in natural images. Meteorological downscaling requires not only fine-scale recovery but also the preservation of spherical

geometry and physical structures Bonev et al. (2023); Saccardi et al. (2025). Therefore, we separate low-frequency and high-frequency reconstruction and introduce a spherical harmonic spectral constraint to preserve the spherical spectral structure of global meteorological fields.

3 METHOD

We propose ClimLoRA, a frequency-aware dual-LoRA framework that formulates meteorological downscaling as residual learning and separates fine-scale spatial–spectral reconstruction from physical structure correction through independent low-rank parameter updates. We denote the low-resolution input and high-resolution reference field by X^L and X^H, respectively.

3.1 RESIDUAL-BASED CLIMATE DOWNSCALING

Conventional diffusion models typically recover high-resolution fields through multi-step reverse sampling (Ho et al., 2020; Saharia et al., 2022), with the forward diffusion and reverse denoising processes detailed in Appendix A.4.1. Inspired by the one-step diffusion super-resolution method OSEDiff (Wu et al., 2024), we adopt a fixed timestep t to directly predict from the low-resolution meteorological field X^L in a single step as $\hat{X}^H = [X^L - \sqrt{1-\bar{\alpha}_t}\,\epsilon_\theta(X^L, t)]/\sqrt{\bar{\alpha}_t}$, and reparameterize this mapping into a residual form:

$$\hat{X}^H = X^L - R_\theta(X^L). \tag{1}$$

Here, $R_\theta(X^L)$ denotes the residual predicted by the model. To recover large-scale structures, we train the Base model primarily for low-frequency residual reconstruction. The target residual is $R^* = X^L - X^H$, whose low-frequency component is

$$R_{\text{low}} = \text{LP}(R^*), \tag{2}$$

where $\text{LP}(\cdot)$ denotes the low-pass operator.

3.2 DUAL-LORA TRAINING

Existing diffusion-based meteorological downscaling methods primarily rely on data-driven objectives Mardani et al. (2025), lacking explicit constraints on the physical structures of meteorological fields. Directly introducing a physics loss may cause optimization interference between reconstruction and physical objectives. Inspired by the application of LoRA to task adaptation (Hu et al., 2021; Jones et al., 2024; Shah et al., 2024), we propose a decoupled dual-LoRA training strategy that introduces Fine-Scale Reconstruction LoRA (FSR-LoRA) and Physics LoRA (Phy-LoRA) into the pretrained Base model to learn fine-scale recovery and physical consistency corrections, respectively. Specifically, we freeze the Base parameters θ_b^* and train FSR-LoRA using high-frequency reconstruction and spherical harmonic spectral losses. Its parameter set is

$$\theta_{\text{FSR}} = \{\theta_b^*, \Delta\theta_{\text{FSR}}\}. \tag{3}$$

Here, $\Delta\theta_{\text{FSR}}$ denotes the trainable parameters. FSR-LoRA recovers local fine-scale information on top of the base prediction and improves the spherical spectral distribution of global meteorological fields. We then freeze the FSR-LoRA parameters $\Delta\theta_{\text{FSR}}^*$ and train Phy-LoRA using the physical structural loss. The complete parameter set is

$$\theta_{\text{Dual}} = \{\theta_b^*, \Delta\theta_{\text{FSR}}^*, \Delta\theta_{\text{Phy}}\}. \tag{4}$$

During this stage, only $\Delta\theta_{\text{Phy}}$ is updated to learn physical structural corrections from horizontal kinetic energy, divergence, relative vorticity, and related diagnostics, preventing the physical objective from directly modifying the learned reconstruction parameters.

3.3 FINE-SCALE RECONSTRUCTION LORA

The Base model recovers large-scale structures of meteorological fields through low-frequency constraints. We then freeze its parameters and train Fine-Scale Reconstruction LoRA (FSR-LoRA) to recover the remaining fine-scale information in both the spatial and spectral domains. Numerical errors in grid space alone may not fully capture fine-scale variations Yan et al. (2024). Moreover,

global meteorological fields are defined on a sphere, requiring a frequency representation consistent with spherical geometry to describe multiscale structures. Spherical harmonics characterize variations at different spatial scales through different harmonic degrees Bonev et al. (2023); Rußwurm et al. (2024). We therefore constrain fine-scale corrections in the spatial domain and match spectral amplitudes in the spherical harmonic domain, forming a joint spatial–spectral constraint.

Spatial Reconstruction. Let $\hat{X}_b^H = X^L - R_b$ denote the Base prediction, and let $T_{\mathrm{hf}} = \mathrm{HP}(X^H - \hat{X}_b^H)$ represent the high-frequency information not recovered by the Base. Given the raw residual prediction $\tilde{R}_f$ after introducing FSR-LoRA, we extract $F_{\mathrm{hf}} = -\mathrm{HP}(\tilde{R}_f - R_b)$ as the fine-scale correction, yielding the predicted field $\hat{X}_f^H = \hat{X}_b^H + F_{\mathrm{hf}}$. The high-frequency reconstruction loss is defined as

$$\mathcal{L}_{\mathrm{hf}} = \frac{1}{N} \left\| F_{\mathrm{hf}} - T_{\mathrm{hf}} \right\|_2^2, \tag{5}$$

where $\mathrm{HP}(\cdot)$ denotes the high-pass operator and N is the number of field elements included in the loss. This constraint guides FSR-LoRA to recover high-frequency information missing from the Base while preserving the learned large-scale structures.

Spherical Spectral Representation. The spatial constraint focuses on local variations, whereas the spectral constraint further characterizes the multiscale distribution of global meteorological fields. We use the spherical harmonic transform (SHT) to map meteorological fields from the latitude–longitude grid to the spherical frequency domain. For a meteorological variable $X(\lambda, \phi)$ defined on the sphere, its spherical harmonic coefficients are

$$C_{l,m} = \mathrm{SHT}(X)_{l,m} = \int_{\mathbb{S}^2} X(\lambda, \phi) Y_{l,m}(\lambda, \phi) \, \mathrm{d}\Omega, \tag{6}$$

where λ and ϕ denote longitude and latitude, respectively, $Y_{l,m}$ is an orthonormal real spherical harmonic basis function, and l and m denote the harmonic degree and order. The spherical area element is $\mathrm{d}\Omega = \cos\phi \, \mathrm{d}\phi \, \mathrm{d}\lambda$. Lower-degree components primarily describe large-scale structures, while higher-degree components characterize variations at smaller spatial scales. Based on this multiscale representation, we apply the SHT to the predicted field $\hat{X}_f^H$ and the reference field X^H, obtaining coefficients $\hat{C}_{l,m}$ and $C_{l,m}$, respectively. The spherical harmonic spectral loss is defined as

$$\mathcal{L}_{\mathrm{sph}} = \frac{1}{|\Omega|} \sum_{l,m} w_l \left| \log(1 + |\hat{C}_{l,m}|) - \log(1 + |C_{l,m}|) \right|, \tag{7}$$

where $|\Omega|$ denotes the number of spherical harmonic coefficients included in the comparison, and w_l balances the contributions of different scales. The training objective of FSR-LoRA is

$$\mathcal{L}_{\mathrm{FSR}} = \lambda_{\mathrm{hf}} \mathcal{L}_{\mathrm{hf}} + \lambda_{\mathrm{sph}} \mathcal{L}_{\mathrm{sph}}, \tag{8}$$

where λ_{hf} and λ_{sph} control the weights of the two losses, allowing fine-scale reconstruction to account for both spatial errors and spectral discrepancies.

3.4 Physics-consistency LoRA

Although Fine-scale LoRA recovers fine-scale structures in meteorological fields, high-frequency details do not necessarily correspond to realistic small-scale dynamical structures. Generative models may produce details at high wavenumbers that are present in the spectrum but dynamically inconsistent (Saccardi et al., 2025). Meanwhile, introducing physical losses into a shared parameter space may cause optimization interference between reconstruction and physical objectives. Therefore, we introduce Physics-consistency LoRA (Phy-LoRA), which learns physical structural corrections on top of the frozen Base and Fine-scale LoRA, preventing the physical objective from directly modifying the learned reconstruction parameters.

Meteorological Structural Representation. This work focuses on single-time-step spatial downscaling without explicitly modeling temporal evolution. We therefore construct the physical structural operator using horizontal kinetic energy E_h, divergence δ_h, relative vorticity ζ_h, and the gradients of temperature T_{2m} and pressure P_{msl}:

$$\Phi_{\mathrm{phy}}(X) = \left[E_h, \, \delta_h, \, \zeta_h, \, \nabla_h T_{2m}, \, \nabla_h P_{\mathrm{msl}} \right]. \tag{9}$$

Table 1: ERA5-to-ERA5 global downscaling performance from 1° to 0.25° across four near-surface meteorological variables.

Method	U10			V10			T2M			MSL		
	RMSE	MB	LFD	RMSE	MB	LFD	RMSE	MB	LFD	RMSE	MB	LFD
Bilinear	0.6091	**-0.0003**	12.8101	0.6341	**-0.0003**	12.8806	0.7734	-0.0015	13.3081	40.0044	0.1196	21.1115
DeepSD	0.4017	0.0154	11.9901	0.3822	0.0052	11.8946	0.6023	-0.0612	12.8027	25.7888	0.4941	20.2735
ViT	0.3313	0.0078	11.6187	0.3202	-0.0023	11.5519	0.4961	-0.0073	12.4269	20.7784	0.4534	19.8699
CorrDiff	0.2922	0.0019	11.3672	0.2824	-0.0047	11.2998	0.4215	-0.0032	12.0940	19.1124	0.4293	19.6499
SRGAN	0.4266	-0.0191	12.0991	0.4201	0.0169	12.0750	0.6763	-0.0475	13.0215	31.9163	2.6724	20.6916
DSFNO	0.4318	-0.0007	12.1385	0.4125	0.0289	12.0475	0.6251	-0.0241	12.8807	27.4225	0.4683	20.3959
GeoFAR	0.2816	0.0029	11.2872	0.2720	-0.0035	11.2229	0.4038	-0.0055	12.0088	17.8289	-0.5889	19.5066
ClimLoRA	**0.2677**	0.0018	**11.1955**	**0.2599**	0.0011	**11.1390**	**0.3799**	**0.0012**	**11.8915**	**15.9568**	**0.1113**	**19.3229**

Since global meteorological fields are defined on a spherical latitude–longitude grid, we compute spatial derivatives in spherical coordinates. For a meteorological variable q, its horizontal gradient is defined as

$$\nabla_h q = \left(\frac{1}{a \cos\phi} \frac{\partial q}{\partial \lambda}, \frac{1}{a} \frac{\partial q}{\partial \phi} \right), \tag{10}$$

where a denotes the Earth's radius, and λ and ϕ denote longitude and latitude in radians, respectively. This operator characterizes the local eastward and northward spatial variations of temperature and pressure. For the horizontal wind field $\mathbf{v}_h = (u, v)$, we use horizontal kinetic energy, divergence, and relative vorticity to characterize its dynamical structure:

$$E_h = \frac{u^2 + v^2}{2}, \qquad \delta_h = \nabla_h \cdot \mathbf{v}_h = \frac{\partial u}{\partial x} + \frac{\partial v}{\partial y}, \qquad \zeta_h = \nabla_h \times \mathbf{v}_h = \frac{\partial v}{\partial x} - \frac{\partial u}{\partial y}. \tag{11}$$

Here, E_h describes the wind-speed magnitude and the coupling between the two wind components, while δ_h and ζ_h characterize convergence–divergence and rotational structures, respectively. For temperature and pressure, we use the spherical horizontal gradients $\nabla_h T_{2m}$ and $\nabla_h P_{\mathrm{msl}}$ to characterize local variations and gradient structures. Spherical formulations and experiment-specific constraints are given in Appendices A.4.2 and A.4.3.

Physical Structural Optimization. Let $\hat{X}^H_{\mathrm{Dual}}$ and X^H denote the Phy-LoRA prediction and high-resolution reference, respectively. We define the physical loss as

$$\mathcal{L}_{\mathrm{phy}} = \sum_k \omega_k \left\| \Phi_k(\hat{X}^H_{\mathrm{Dual}}) - \Phi_k(X^H) \right\|_1, \tag{12}$$

where ω_k controls the contribution of each physical quantity. This objective updates only Phy-LoRA, encouraging the predicted physical structures to approach those of the reference field.

3.5 Adjustable Dual-LoRA Inference

Different meteorological applications emphasize different aspects of the downscaled fields. For example, global and regional weather prediction focuses more on overall reconstruction accuracy, while extreme-weather analysis relies more on structures such as pressure gradients and wind convergence. To accommodate these different requirements, we introduce two guidance factors at inference time to control the contributions of the two LoRA branches, as illustrated in Figure 1(b):

$$R_\theta^{\mathrm{adapt}}(X^L) = R_b(X^L) + \lambda_{\mathrm{fine}} \left[R_f(X^L) - R_b(X^L) \right] + \lambda_{\mathrm{phy}} \left[R_{\mathrm{Dual}}(X^L) - R_f(X^L) \right]. \tag{13}$$

Here, R_b, R_f, and R_{Dual} denote the residual predictions of the Base model, Base+FSR-LoRA, and the full ClimLoRA model, respectively. The two difference terms represent fine-scale and physical corrections, with their contributions controlled by λ_{fine} and λ_{phy}, respectively.

4 Experiments

4.1 Experimental Setup

Datasets and experimental settings. We conduct experiments on the ERA5 reanalysis dataset Hersbach et al. (2020) and the MPI-ESM climate model dataset Eyring et al. (2016), constructing in-domain and cross-source downscaling tasks to evaluate performance across different data sources, meteorological variables, and resolutions.

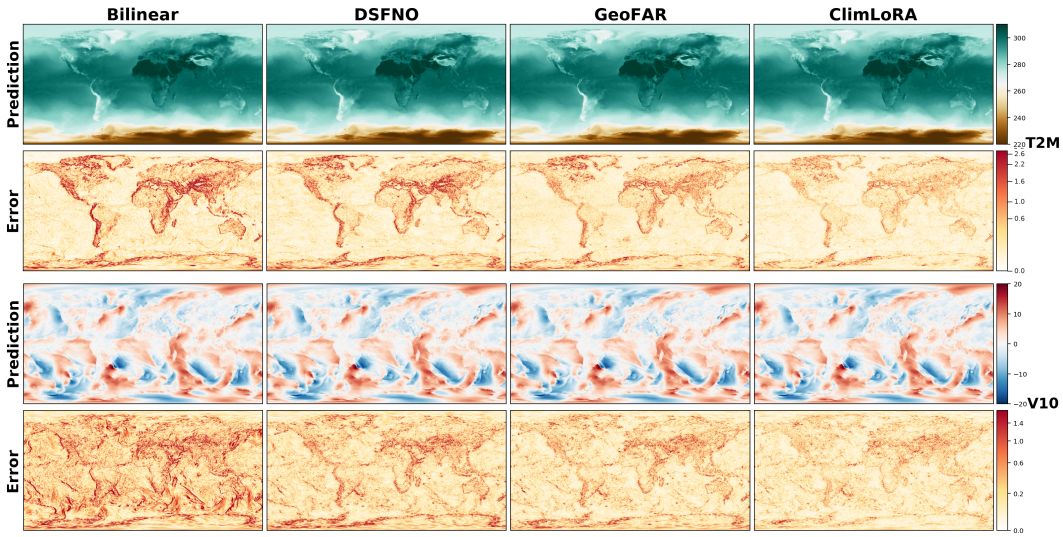


Figure 2: Qualitative comparison of ERA5-to-ERA5 downscaling from 1° to 0.25° for V10 and T2M at 12:00 UTC on July 1, 2020. Predicted meteorological fields and corresponding error maps are shown for Bilinear, DSFNO, GeoFAR, and ClimLoRA.

• In-domain downscaling (ERA5-to-ERA5). We use 0.25° ERA5 as the high-resolution target and regrid it to 1° as the low-resolution input, constructing a $4\times$ downscaling task. We jointly predict four near-surface variables: 10-m zonal wind (U10), 10-m meridional wind (V10), 2-m temperature (T2M), and mean sea-level pressure (MSL). This setting evaluates the model's downscaling capability within the same data source.

• Cross-source downscaling (MPI-ESM-to-ERA5). Following the cross-source downscaling setting of ClimaX (Nguyen et al., 2023), we use MPI-ESM as the input and ERA5 as the high-resolution reference, and further extend the target grid to 0.25°. We regrid the MPI-ESM inputs to 1° and 2°, constructing $4\times$ and $8\times$ grid downscaling tasks, respectively, and jointly predict Z500, T850, U10, V10, and T2M. This setting evaluates cross-source downscaling performance between climate-model outputs and reanalysis data, with data processing and split details provided in the appendix A.1.

Comparison methods. We compare the proposed method with general super-resolution methods and climate downscaling methods. The general methods include Bilinear, ViT Dosovitskiy et al. (2021), and SRGAN Ledig et al. (2017), while the climate downscaling methods include DeepSD (Vandal et al., 2017), CorrDiff(Mardani et al., 2025), DSFNO (Yang et al., 2024) and GeoFAR (Xu et al., 2026).

Evaluation metrics. We use root mean square error (RMSE) and mean bias (MB) to measure numerical error and systematic bias, respectively, and Log Frequency Distance (LFD) for frequency-aware evaluation. Lower RMSE and LFD, and an MB closer to zero, indicate better downscaling performance.

4.2 Comparative Performance

Global downscaling. Table 1 presents the global downscaling results for four near-surface meteorological variables under the ERA5-to-ERA5 setting. ClimLoRA achieves the lowest RMSE and LFD across all variables, outperforming general super-resolution methods such as ViT and SRGAN, as well as climate downscaling methods including CorrDiff, DSFNO, and GeoFAR. Compared with the residual diffusion method CorrDiff, ClimLoRA reduces RMSE by approximately 9.9% and 16.5% for T2M and MSL, respectively. Compared with the strongest baseline, GeoFAR, ClimLoRA achieves lower RMSE and LFD across all variables, with RMSE reductions of approximately 5.9% and 10.5% and LFD reductions of approximately 1.0% and 0.9% for T2M and MSL, respectively. These results demonstrate that ClimLoRA can simultaneously improve numerical accuracy and frequency-domain consistency.

Table 2: Performance of $4\times$ global cross-source downscaling from 1° MPI-ESM inputs to 0.25° ERA5 targets across different meteorological variables.

Method	Z500			T850			U10			V10			T2M		
	RMSE	MB	LFD	RMSE	MB	LFD	RMSE	MB	LFD	RMSE	MB	LFD	RMSE	MB	LFD
Bilinear	107.2651	-0.0939	22.9014	1.3837	0.0062	14.3200	2.1497	-0.0368	15.2412	1.8901	0.0353	15.0318	2.3323	0.0021	15.3554
DeepSD	108.7011	-1.1993	22.9328	1.4057	-0.0207	14.3540	2.0569	0.1260	15.1570	1.7463	0.0527	14.8799	2.2121	-0.0129	15.3023
ViT	88.0472	1.5448	22.5464	1.3365	-0.0098	14.2940	1.7613	**0.0057**	14.8715	1.5570	0.0362	14.6644	1.8076	-0.0374	14.9362
CorrDiff	86.3565	-2.0951	22.4935	1.3305	-0.0029	14.2862	1.4955	0.0224	14.5395	1.3658	-0.0072	14.4039	1.5100	-0.0019	14.5644
SRGAN	102.2618	7.9371	22.7751	1.4683	-0.0351	14.4007	2.0984	0.0603	15.1919	1.7752	-0.0143	14.9119	2.2292	0.0083	15.2866
DSFNO	92.7999	**0.0847**	22.6281	1.3586	0.0032	14.3299	1.6080	0.0301	14.6920	1.4736	-0.0385	14.5537	1.5948	-0.0074	14.6770
GeoFAR	85.8981	-0.6632	22.4926	1.3211	0.0036	14.2854	1.4037	-0.0110	14.4200	1.2967	-0.0060	14.3016	1.3843	-0.0155	14.3861
ClimLoRA	**82.8565**	0.4647	**22.4158**	**1.2033**	**-0.0026**	**14.1535**	**1.2797**	0.0171	**14.2851**	**1.1796**	**0.0047**	**14.0740**	**1.2746**	**-0.0012**	**14.3696**

Cross-source downscaling. We further evaluate cross-source downscaling from MPI-ESM to ERA5, as shown in Table 2. ClimLoRA achieves the lowest RMSE and LFD on Z500, T850, U10, V10, and T2M. Compared with GeoFAR, the RMSE of Z500 decreases from 85.8981 to 82.8565, while that of T850 decreases from 1.3211 to 1.2033, with consistent improvements on the remaining variables. These results show that ClimLoRA is effective not only for near-surface variables but also for upper-air variables strongly influenced by atmospheric dynamics. Despite the distribution shift between MPI-ESM and ERA5, ClimLoRA maintains lower reconstruction and frequency-domain errors, demonstrating stronger cross-source generalization. Results for the more challenging $8\times$ cross-source setting are provided in Appendix A.5.1.

Qualitative comparison. Figure 2 presents the predicted fields and error distributions of V10 and T2M under the ERA5-to-ERA5 setting. Compared with Bilinear, DSFNO, and GeoFAR, ClimLoRA exhibits lower errors across multiple regions. For V10, DSFNO shows pronounced banded errors at Southern Hemisphere mid-to-high latitudes, whereas ClimLoRA reduces these errors. Qualitative results for Z500 and T850 under the cross-source setting are provided in Appendix Figure 7. Unlike in-domain downscaling, the cross-source task additionally involves distribution differences between climate-model inputs and reanalysis targets. ClimLoRA also exhibits lower errors in some mid-to-high-latitude regions under this setting. These results support the downscaling capability of ClimLoRA in both in-domain and cross-source settings.

4.3 Ablation Study

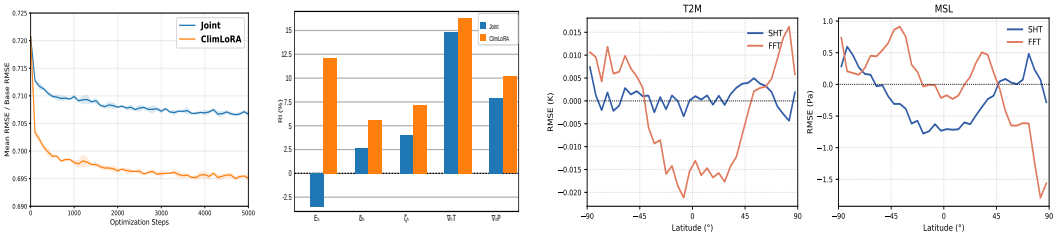
Effectiveness of individual components. Table 3 compares the contributions of different components. Adding FSR-LoRA reduces the RMSE for all variables, with U10 and T2M decreasing from 0.3338 and 0.4516 to 0.2709 and 0.3882, respectively. Removing the high-frequency constraint (NO HF) increases the RMSE of T2M and MSL to 0.4308 and 19.8896, respectively. Removing the spherical harmonic spectral constraint (NO Spherical) increases these values to 0.3977 and 18.3969, indicating that both constraints contribute to fine-scale recovery. Adding Phy-LoRA reduces the RMSE of U10, V10, T2M, and MSL by approximately 1.2%, 1.8%, 2.1%, and 7.7%, respectively, compared with FSR-LoRA. LFD also decreases across all four variables, indicating that physical structure correction further improves reconstruction accuracy.

Frequency representation. We replace the spherical harmonic spectral constraint with a two-dimensional FFT constraint while keeping the remaining settings unchanged. The spherical harmonic representation achieves lower RMSE and LFD for all variables. In particular, the RMSE/LFD of T2M decreases from $0.4072/12.0323$ to $0.3882/11.9359$, while the RMSE of MSL decreases from 17.8631 to 17.2929. Figure 3(c,d) shows the corresponding latitude-wise error distributions. These results indicate that, compared with FFT, a spherical harmonic spectral constraint consistent with spherical geometry better supports global meteorological field reconstruction.

Spherical physical operator. We replace the spherical physical operator with planar-grid differences. Compared with the planar operator, the spherical operator accounts for latitude-dependent spatial scales and spherical geometric, achieving lower RMSE and LFD for all variables. Specifically, the RMSE of V10 decreases from 0.2623 to 0.2599, while the RMSE/LFD of MSL decreases from $16.2733/19.3659$ to $15.9568/19.3229$. These results indicate that accounting for spherical geometry in the physical operator improves reconstruction accuracy and spectral consistency.

Table 3: Ablation study of ClimLoRA on ERA5-to-ERA5 downscaling.

Method	U10		V10		T2M		MSL	
	RMSE	LFD	RMSE	LFD	RMSE	LFD	RMSE	LFD
Base	0.3338	11.6118	0.3143	11.4946	0.4516	12.2367	19.9401	19.7539
+ FSR-LoRA (NO Spherical)	0.2717	11.2216	0.2663	11.1862	0.3977	12.0332	18.3969	19.5987
+ FSR-LoRA (NO HF)	0.2937	11.3798	0.2889	11.3477	0.4308	12.1436	19.8896	19.6984
+ FSR-LoRA (FFT)	0.2811	11.2948	0.2708	11.2207	0.4072	12.0323	17.8631	19.5598
+ FSR-LoRA	0.2709	11.2165	0.2647	11.1747	0.3882	11.9359	17.2929	19.4660
+ FSR-LoRA + Phy-LoRA (Planar)	0.2689	11.1962	0.2623	11.1535	0.3828	11.9076	16.2733	19.3659
+ FSR-LoRA + Phy-LoRA	**0.2677**	**11.1955**	**0.2599**	**11.1390**	**0.3799**	**11.8915**	**15.9568**	**19.3229**



(a) Optimization strategy (b) Physical consistency (c) T2M (d) MSL

Figure 3: **Physical update strategies and spectral representations.** (a) Reconstruction RMSE normalized by the Base RMSE under Joint Optimization and ClimLoRA. (b) Relative improvement in physical diagnostic errors over the Base. (c, d) Latitude-wise error profiles for T2M and MSL using SHT- and FFT-based spectral constraints.

4.4 Physical Update Strategy Analysis

We compare two physical update strategies. Joint Optimization optimizes high-frequency reconstruction, spherical harmonic spectral, and physical objectives within a shared LoRA parameter space; ClimLoRA adopts a decoupled sequential update strategy, where FSR-LoRA is trained first, followed by freezing the Base and FSR-LoRA and optimizing physical structures only through Phy-LoRA. For a fair comparison, Joint Optimization and ClimLoRA are matched in trainable parameter count and total training budget. Figure 3(a) shows the reconstruction error trajectories of the two strategies. As training progresses, the normalized reconstruction error of ClimLoRA decreases to approximately 0.695, lower than the approximately 0.707 achieved by Joint Optimization. Figure 3(b) shows the improvements in physical structural errors relative to the same Base. ClimLoRA outperforms Joint Optimization across all five metrics: the improvement in horizontal kinetic energy increases from -3.6% to 12.1%, while those in divergence and vorticity increase from 2.7% and 4.0% to 5.6% and 7.2%, respectively. These results show that, compared with joint optimization within a single LoRA module, ClimLoRA's sequential updates reduce reconstruction and physical diagnostic errors, supporting the use of separate parameter updates to coordinate the two objectives.

5 Conclusion

In this work, we propose ClimLoRA, a climate downscaling framework that separates fine-scale reconstruction from physical structure correction through independent LoRA parameter updates, thereby reducing optimization interference between the objectives. Built upon a low-frequency Base, Fine-scale LoRA recovers details through high-frequency reconstruction and spherical harmonic spectral constraints, while Physics LoRA learns spherical physical structure corrections with the reconstruction parameters frozen. Global downscaling experiments under the ERA5-to-ERA5 in-domain and MPI-ESM-to-ERA5 cross-source settings validate the effectiveness of the framework, demonstrating improved reconstruction accuracy and physical structural consistency despite distribution differences between climate-model and reanalysis data. The method still incurs additional computational overhead, and its physical constraints are limited to single-time-step spatial structures. Future work will explore more efficient inference and spatiotemporal physical consistency modeling.

AI USE STATEMENT

We used generative AI to draft and polish the manuscript. All AI-assisted content was reviewed by the authors, who take full responsibility for the final paper.

REPRODUCIBILITY STATEMENT

Methodological and experimental details are provided in the paper to support reproducibility. **Datasets.** Our work utilizes two publicly available climate datasets. ERA5 reanalysis is provided by the European Centre for Medium-Range Weather Forecasts (ECMWF) (Hersbach et al., 2020), while MPI-ESM data are obtained from the CMIP6 archive (Eyring et al., 2016). **Experimental Details.** Section 4.1 describes the overall experimental setup, including the in-domain ERA5-to-ERA5 and cross-source MPI-ESM-to-ERA5 downscaling settings. Further implementation details are provided in Appendix A.2, including the model hyperparameters.

REFERENCES

Jan-Hendrik Bastek, WaiChing Sun, and Dennis Kochmann. Physics-informed diffusion models. In *International Conference on Learning Representations*, volume 2025, pp. 3360–3385, 2025.

Kaifeng Bi, Lingxi Xie, Hengheng Zhang, Xin Chen, Xiaotao Gu, and Qi Tian. Accurate medium-range global weather forecasting with 3d neural networks. *Nature*, 619(7970):533–538, 2023.

Matthieu Blanke, Yongquan Qu, Sara Shamekh, and Pierre Gentine. Strictly constrained generative modeling via split augmented langevin sampling. In *International Conference on Learning Representations*, volume 2026, pp. 145927–145965, 2026.

Boris Bonev, Thorsten Kurth, Christian Hundt, Jaideep Pathak, Maximilian Baust, Karthik Kashinath, and Anima Anandkumar. Spherical fourier neural operators: Learning stable dynamics on the sphere. In *International conference on machine learning*, pp. 2806–2823. PMLR, 2023.

Grant Buster, Brandon N Benton, Andrew Glaws, and Ryan N King. High-resolution meteorology with climate change impacts from global climate model data using generative machine learning. *Nature Energy*, 9(7):894–906, 2024.

Lei Chen, Fei Du, Yuan Hu, Zhibin Wang, and Fan Wang. Swinrdm: integrate swinrnn with diffusion model towards high-resolution and high-quality weather forecasting. In *Proceedings of the AAAI conference on artificial intelligence*, volume 37, pp. 322–330, 2023.

Chuntao Ding, Zhichao Lu, Shangguang Wang, Ran Cheng, and Vishnu Naresh Boddeti. Mitigating task interference in multi-task learning via explicit task routing with non-learnable primitives. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 7756–7765, 2023.

Alexey Dosovitskiy, Lucas Beyer, Alexander Kolesnikov, Dirk Weissenborn, Xiaohua Zhai, Thomas Unterthiner, Mostafa Dehghani, Matthias Minderer, Georg Heigold, Sylvain Gelly, Jakob Uszkoreit, and Neil Houlsby. An image is worth 16x16 words: Transformers for image recognition at scale. In *International Conference on Learning Representations*, 2021.

Veronika Eyring, Sandrine Bony, Gerald A Meehl, Catherine A Senior, Bjorn Stevens, Ronald J Stouffer, and Karl E Taylor. Overview of the coupled model intercomparison project phase 6 (cmip6) experimental design and organization. *Geoscientific Model Development*, 9(5):1937–1958, 2016.

Haiwen Guan, Dibyajyoti Chakraborty, Moein Darman, Troy Arcomano, Ashesh Chattopadhyay, and Romit Maulik. High-resolution climate projections using diffusion-based downscaling of a lightweight climate emulator. *arXiv preprint arXiv:2602.13416*, 2026.

Paula Harder, Alex Hernandez-Garcia, Venkatesh Ramesh, Qidong Yang, Prasanna Sattegeri, Daniela Szwarcman, Campbell Watson, and David Rolnick. Hard-constrained deep learning for climate downscaling. *Journal of Machine Learning Research*, 24(365):1–40, 2023.

Hans Hersbach, Bill Bell, Paul Berrisford, Shoji Hirahara, András Horányi, Joaquín Muñoz-Sabater, Julien Nicolas, Carole Peubey, Raluca Radu, Dinand Schepers, et al. The era5 global reanalysis. *Quarterly journal of the royal meteorological society*, 146(730):1999–2049, 2020.

Jonathan Ho, Ajay Jain, and Pieter Abbeel. Denoising diffusion probabilistic models. *Advances in neural information processing systems*, 33:6840–6851, 2020.

Edward J Hu, Yelong Shen, Phillip Wallis, Zeyuan Allen-Zhu, Yuanzhi Li, Shean Wang, Lu Wang, and Weizhu Chen. Lora: Low-rank adaptation of large language models. *arXiv preprint arXiv:2106.09685*, 2021.

Maxwell Jones, Sheng-Yu Wang, Nupur Kumari, David Bau, and Jun-Yan Zhu. Customizing text-to-image models with a single image pair. In *SIGGRAPH Asia 2024 conference papers*, pp. 1–13, 2024.

Dmitrii Kochkov, Janni Yuval, Ian Langmore, Peter Norgaard, Jamie Smith, Griffin Mooers, Milan Klöwer, James Lottes, Stephan Rasp, Peter Düben, et al. Neural general circulation models for weather and climate. *Nature*, 632(8027):1060–1066, 2024.

Christian Ledig, Lucas Theis, Ferenc Huszár, Jose Caballero, Andrew Cunningham, Alejandro Acosta, Andrew Aitken, Alykhan Tejani, Johannes Totz, Zehan Wang, et al. Photo-realistic single image super-resolution using a generative adversarial network. In *2017 IEEE conference on computer vision and pattern recognition (CVPR)*, pp. 105–114. IEEE, 2017.

Yumin Liu, Auroop R Ganguly, and Jennifer Dy. Climate downscaling using ynet: A deep convolutional network with skip connections and fusion. In *Proceedings of the 26th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*, pp. 3145–3153, 2020.

Ignacio Lopez-Gomez, Zhong Yi Wan, Leonardo Zepeda-Núñez, Tapio Schneider, John Anderson, and Fei Sha. Dynamical-generative downscaling of climate model ensembles. *Proceedings of the National Academy of Sciences*, 122(17):e2420288122, 2025.

Yingtao Luo, Shikai Fang, Binqing Wu, Qingsong Wen, and Liang Sun. Physics-guided learning of meteorological dynamics for weather downscaling and forecasting. In *Proceedings of the 31st ACM SIGKDD Conference on Knowledge Discovery and Data Mining V. 2*, pp. 2010–2020, 2025.

Morteza Mardani, Noah Brenowitz, Yair Cohen, Jaideep Pathak, Chieh-Yu Chen, Cheng-Chin Liu, Arash Vahdat, Mohammad Amin Nabian, Tao Ge, Akshay Subramaniam, et al. Residual corrective diffusion modeling for km-scale atmospheric downscaling. *Communications Earth & Environment*, 6(1):124, 2025.

Tung Nguyen, Johannes Brandstetter, Ashish Kapoor, Jayesh K Gupta, and Aditya Grover. Climax: A foundation model for weather and climate. In *1st Workshop on the Synergy of Scientific and Machine Learning Modeling @ ICML2023*, 2023. URL https://openreview.net/forum?id=5CZyPMgjUu.

Marc Rußwurm, Konstantin Klemmer, Esther Rolf, Robin Zbinden, and Devis Tuia. Geographic location encoding with spherical harmonics and sinusoidal representation networks. In *International Conference on Learning Representations*, volume 2024, pp. 1746–1759, 2024.

Carlo Saccardi, Maximilian Pierzyna, Haitz Sáez de Ocáriz Borde, Simone Monaco, Cristian Meo, Pietro Liò, Rudolf Saathof, Geethu Joseph, and Justin Dauwels. Assessing the geographic generalization and physical consistency of generative models for climate downscaling. *arXiv preprint arXiv:2510.13722*, 2025.

Chitwan Saharia, Jonathan Ho, William Chan, Tim Salimans, David J Fleet, and Mohammad Norouzi. Image super-resolution via iterative refinement. *IEEE transactions on pattern analysis and machine intelligence*, 45(4):4713–4726, 2022.

Viraj Shah, Nataniel Ruiz, Forrester Cole, Erika Lu, Svetlana Lazebnik, Yuanzhen Li, and Varun Jampani. Ziplora: Any subject in any style by effectively merging loras. In *European Conference on Computer Vision*, pp. 422–438. Springer, 2024.

Prakhar Srivastava, Ruihan Yang, Gavin Kerrigan, Gideon Dresdner, Jeremy McGibbon, Christopher Bretherton, and Stephan Mandt. Precipitation downscaling with spatiotemporal video diffusion. *Advances in Neural Information Processing Systems*, 37:56374–56400, 2024.

Karen Stengel, Andrew Glaws, Dylan Hettinger, and Ryan N King. Adversarial super-resolution of climatological wind and solar data. *Proceedings of the National Academy of Sciences*, 117(29): 16805–16815, 2020.

Yongjian Sun, Kefeng Deng, Kaijun Ren, Jia Liu, Chongjiu Deng, and Yongjun Jin. 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.

Siwei Tu, Ben Fei, Weidong Yang, Fenghua Ling, Hao Chen, Zili Liu, Kun Chen, Hang Fan, Wanli Ouyang, and Lei Bai. Satellite observations guided diffusion model for accurate meteorological states at arbitrary resolution. In *2025 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 28071–28079. IEEE, 2025.

Thomas Vandal, Evan Kodra, Sangram Ganguly, Andrew Michaelis, Ramakrishna Nemani, and Auroop R Ganguly. Deepsd: Generating high resolution climate change projections through single image super-resolution. In *Proceedings of the 23rd acm sigkdd international conference on knowledge discovery and data mining*, pp. 1663–1672, 2017.

Yogesh Verma, Markus Heinonen, and Vikas Garg. Climode: Climate and weather forecasting with physics-informed neural odes. In *International Conference on Learning Representations*, volume 2024, pp. 8408–8430, 2024.

Yufei Wang, Wenhan Yang, Xinyuan Chen, Yaohui Wang, Lanqing Guo, Lap-Pui Chau, Ziwei Liu, Yu Qiao, Alex C Kot, and Bihan Wen. Sinsr: diffusion-based image super-resolution in a single step. In *2024 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 25796–25805. IEEE, 2024.

Robbie A Watt and Laura A Mansfield. Generative diffusion-based downscaling for climate. *arXiv preprint arXiv:2404.17752*, 2024.

Rongyuan Wu, Lingchen Sun, Zhiyuan Ma, and Lei Zhang. One-step effective diffusion network for real-world image super-resolution. *Advances in Neural Information Processing Systems*, 37: 92529–92553, 2024.

Chang Xu, Gencer Sumbul, Li Mi, Robin Zbinden, and Devis Tuia. Geofar: Geography-informed frequency-aware super-resolution for climate data. In *The Fourteenth International Conference on Learning Representations*, 2026.

Wanghan Xu, Fenghua Ling, Wenlong Zhang, Tao Han, Hao Chen, Wanli Ouyang, and Lei Bai. Generalizing weather forecast to fine-grained temporal scales via physics-ai hybrid modeling. *Advances in Neural Information Processing Systems*, 37:23325–23351, 2024.

Chiu-Wai Yan, Shi Q Foo, Van H Trinh, Dit-Yan Yeung, Ka-Hing Wong, and Wai-Kin Wong. Fourier amplitude and correlation loss: Beyond using l2 loss for skillful precipitation nowcasting. *Advances in Neural Information Processing Systems*, 37:100007–100041, 2024.

Qidong Yang, Alex Hernandez-Garcia, Paula Harder, Venkatesh Ramesh, Prasanna Sattigeri, Daniela Szwarcman, Campbell D Watson, and David Rolnick. Fourier neural operators for arbitrary resolution climate data downscaling. *Journal of Machine Learning Research*, 25(420):1–30, 2024.

Tianhe Yu, Saurabh Kumar, Abhishek Gupta, Sergey Levine, Karol Hausman, and Chelsea Finn. Gradient surgery for multi-task learning. *Advances in neural information processing systems*, 33: 5824–5836, 2020.

Zongsheng Yue, Jianyi Wang, and Chen Change Loy. Resshift: Efficient diffusion model for image super-resolution by residual shifting. *Advances in neural information processing systems*, 36: 13294–13307, 2023.

A APPENDIX

A.1 DATASETS

ERA5 Reanalysis. ERA5 reanalysis (Hersbach et al., 2020) is maintained by the European Centre for Medium-Range Weather Forecasts (ECMWF). ERA5 combines diverse observations with forecasts from the Integrated Forecasting System (IFS) to provide estimates of atmospheric and surface states. In its native format, ERA5 provides hourly data from 1979 to the present on a global 0.25° grid, covering a wide range of surface variables and atmospheric variables at 37 pressure levels. We select 10-m zonal wind (U10), 10-m meridional wind (V10), 2-m temperature (T2M), and mean sea-level pressure (MSL) for ERA5-to-ERA5 global downscaling experiments. The 0.25° ERA5 fields are used as high-resolution targets and bilinearly regridded to 1° as low-resolution inputs, corresponding to a $4\times$ downscaling factor. All variables are sampled at 6-hour intervals. We use data from 2010–2018 for training, 2019 for validation, and 2020 for testing.

MPI-ESM. CMIP6 (Eyring et al., 2016), organized by the World Climate Research Programme (WCRP), provides long-term climate simulations for the development and evaluation of climate models. MPI-ESM is a climate model developed by the Max Planck Institute for Meteorology, covering multiple surface and upper-air atmospheric variables with a 6-hour temporal resolution. Following the MPI-ESM-to-ERA5 cross-source downscaling setting of ClimaX (Nguyen et al., 2023), we select Z500, T850, U10, V10, and T2M and align them with ERA5 in time and variable definitions. ClimaX uses 5.625° MPI-ESM inputs and 1.40625° ERA5 targets; in this work, we further use 0.25° ERA5 as the high-resolution reference. The original MPI-ESM fields have a spatial resolution of 5.625° and are bilinearly regridded to 1° and 2° as inputs, forming $4\times$ and $8\times$ grid downscaling tasks, respectively. MPI-ESM and ERA5 are both sampled at 6-hour intervals, and corresponding samples are constructed using the same timestamps. We use data from 1979–2016 for training, 2017 for validation, and 2018 for testing.

A.2 MODEL AND TRAINING CONFIGURATION

Network Architecture. ClimLoRA adopts a conditional residual U-Net. The low-resolution input is first bilinearly regridded to the target grid and then fed into the network for residual prediction. The U-Net contains four resolution levels with a base channel width of 32 and channel multipliers of $[1, 2, 2, 4]$. Each encoder level contains two residual blocks, while the bottleneck employs a self-attention module with four attention heads. FSR-LoRA and Phy-LoRA are applied only to all decoder blocks and the final output convolution of the U-Net, with independent low-rank parameters maintained for the two branches. Both branches use a LoRA rank of 64 and an alpha value of 128, corresponding to a scaling factor of 2.

Training Configuration. ClimLoRA is trained in a stage-wise manner. In the FSR stage, the remaining Base parameters are frozen, and the learning rate of FSR-LoRA is set to 10^{-4}.This stage uses the high-frequency reconstruction loss and spherical harmonic spectral loss with weights of 0.6 and 0.4, respectively, and the spherical harmonic transform is truncated at order 64. In the subsequent Phy stage, all updated Base parameters and FSR-LoRA are frozen, and only Phy-LoRA is optimized with a learning rate of 10^{-5}. The physical loss consists of a wind-field structural spectral loss and a scalar-gradient loss, whose weights after fixed-scale normalization are set to 1.0 and 0.4, respectively. The wind-field structural loss equally averages the errors of horizontal kinetic energy, divergence, and relative vorticity, while the scalar-gradient loss equally averages the horizontal gradient errors of all non-wind variables except U10 and V10. Both LoRA branches employ deterministic residual prediction at a fixed timestep $t = 1$ and are optimized using AdamW with $\beta_1 = 0.9$, $\beta_2 = 0.999$. Training is conducted on an NVIDIA GeForce RTX 4090 GPU with a batch size of 1 for a total of 50 epochs.

Comparison Methods. We reimplement and evaluate a range of generic super-resolution methods and climate-oriented downscaling methods, including Nearest, Bilinear, ViT, SRGAN, DeepSD, DSFNO, and GeoFAR [ViT]. All learning-based methods are trained and evaluated using the same data splits, input and output variables, and downscaling settings.

ViT. Vision Transformer (ViT) (Dosovitskiy et al., 2021) divides the input climate field into spatial patches and models long-range dependencies among different regions through Transformer en-

coders. Compared with local convolutional networks, ViT enables spatial interactions over a larger range and is therefore suitable for modeling large-scale structures in climate fields.

SRGAN. SRGAN (Ledig et al., 2017) employs adversarial training between a generator and a discriminator to recover high-resolution details, and combines content and adversarial losses to improve local structures. We construct the generator with 16 residual blocks and the discriminator with 8 convolutional blocks.

CorrDiff. CorrDiff (Mardani et al., 2025) also adopts residual modeling, decomposing downscaling into conditional-mean prediction and stochastic residual generation. A deterministic regression model first predicts the conditional mean of the high-resolution field, and a conditional diffusion model then generates the residual. The final prediction is obtained by adding the two components. In our implementation, CorrDiff and ClimLoRA use the same prediction network architecture.

DeepSD. DeepSD (Vandal et al., 2017) is a climate downscaling method based on stacked convolutional super-resolution networks. It decomposes large-scale resolution enhancement into multiple consecutive stages, where each stage employs an SRCNN to progressively increase the spatial resolution, thereby reconstructing high-resolution climate fields from low-resolution inputs. We train DeepSD following its original network design.

DSFNO. DSFNO (Yang et al., 2024) formulates climate downscaling as a resolution-invariant operator learning problem in Fourier space. It models relationships among different spatial frequency components through Fourier modes and reconstructs high-resolution climate fields. Its frequency-domain operator captures global spatial dependencies and supports inference at different target resolutions.

GeoFAR [ViT]. GeoFAR (Xu et al., 2026)enhances climate super-resolution from both frequency and geographic perspectives. Its frequency-aware convolution kernel (FCK) models different frequency components, while the geographic-aware implicit neural representation (Geo-INR) incorporates spatial locations and geographic information for continuous high-resolution reconstruction. We adopt the GeoFAR [ViT] configuration, using ViT as the backbone while retaining the FCK and Geo-INR modules, and retrain it under the same training settings.

A.3 Evaluation Metrics

We evaluate the downscaling results using RMSE and MB for spatial reconstruction, LFD for frequency-domain discrepancy. **Root Mean Square Error (RMSE).** RMSE measures the pointwise numerical difference between the predicted and ground-truth meteorological fields and is defined as

$$\mathrm{RMSE} = \sqrt{\frac{1}{NHW}\sum_{k=1}^{N}\sum_{i=1}^{H}\sum_{j=1}^{W}\left(\hat{X}_{k,i,j} - X_{k,i,j}\right)^2}, \tag{14}$$

where N denotes the number of samples, H and W denote the spatial height and width of the meteorological field, $X_{k,i,j}$ is the ground-truth value at grid point (i, j) of sample k, and $\hat{X}_{k,i,j}$ is the corresponding prediction. We use the standard RMSE without latitude weighting. A lower RMSE indicates higher numerical accuracy.

Mean Bias (MB). MB measures the overall systematic bias between the predicted and ground-truth fields and is defined as

$$\mathrm{MB} = \frac{1}{NHW}\sum_{k=1}^{N}\sum_{i=1}^{H}\sum_{j=1}^{W}\left(\hat{X}_{k,i,j} - X_{k,i,j}\right). \tag{15}$$

The notation follows that used for RMSE. A positive MB indicates overall overestimation, while a negative MB indicates underestimation. An MB closer to zero indicates a smaller systematic bias.

Log Frequency Distance (LFD). We use LFD to measure the discrepancy between the predicted and ground-truth fields in the Fourier domain:

$$\mathrm{LFD} = \frac{1}{N}\sum_{k=1}^{N}\log\left[1 + \frac{1}{|\mathcal{K}|}\sum_{(u,v)\in\mathcal{K}}\left|\mathcal{F}(\hat{X}_k)_{u,v} - \mathcal{F}(X_k)_{u,v}\right|^2\right], \tag{16}$$

where N is the number of samples, $\mathcal{F}(\cdot)$ denotes the unnormalized forward two-dimensional discrete Fourier transform, (u, v) are frequency indices, and $\mathcal{K}$ is the set of indices included in the evaluation. For a meteorological field of size $H \times W$, we remove the first and last indices along both frequency axes of the unshifted FFT output, giving $\mathcal{K} = \{1, \ldots, H-2\} \times \{1, \ldots, W-2\}$. The logarithm compresses the dynamic range of the Fourier-domain error. A lower LFD indicates smaller prediction errors at the evaluated frequency indices. RMSE, MB, and LFD are computed independently for each meteorological variable.

A.4 Additional Method Details

A.4.1 One-Step Residual Modeling

This section describes the diffusion process and the residual formulation used in the main text.

Forward diffusion. Given a high-resolution meteorological field X_0, the forward process progressively adds Gaussian noise over T timesteps. The noisy field at timestep t can be sampled directly as

$$X_t = \sqrt{\bar{\alpha}_t} X_0 + \sqrt{1 - \bar{\alpha}_t}\epsilon, \qquad \epsilon \sim \mathcal{N}(0, I), \tag{17}$$

where $\bar{\alpha}_t = \prod_{s=1}^{t}(1 - \beta_s)$ and β_s is the noise variance at step s. When $\bar{\alpha}_T$ approaches zero, X_T approximately follows a standard Gaussian distribution.

Conditional reverse denoising. For meteorological downscaling, the reverse process is conditioned on the low-resolution field X^L. The network $\epsilon_\theta(X_t, X^L, t)$ predicts the noise in X_t, from which the clean field is estimated as

$$\hat{X}_0 = \frac{X_t - \sqrt{1 - \bar{\alpha}_t}\,\epsilon_\theta(X_t, X^L, t)}{\sqrt{\bar{\alpha}_t}}. \tag{18}$$

The reverse sampler then uses $\hat{X}_0$ and X_t to obtain X_{t-1}. Repeating this process produces the high-resolution prediction, with computational cost increasing with the number of denoising steps.

One-step residual formulation. Inspired by OSEDiff (Wu et al., 2024), we use X^L directly as the input at a fixed timestep t:

$$\hat{X}^H = \frac{X^L - \sqrt{1 - \bar{\alpha}_t}\,\epsilon_\theta(X^L, t)}{\sqrt{\bar{\alpha}_t}}. \tag{19}$$

This formulation does not assume that X^L is sampled from the forward diffusion process. With t fixed, we define

$$R_\theta(X^L) = \left(1 - \frac{1}{\sqrt{\bar{\alpha}_t}}\right) X^L + \sqrt{\frac{1 - \bar{\alpha}_t}{\bar{\alpha}_t}}\,\epsilon_\theta(X^L, t), \tag{20}$$

which yields the equivalent prediction $\hat{X}^H = X^L - R_\theta(X^L)$. The corresponding target residual is $R^* = X^L - X^H$, where X^L has been regridded to the target grid. As described in the main text, the Base model primarily learns its low-frequency component $R_{\text{low}} = \text{LP}(R^*)$ to provide an initial prediction for subsequent fine-scale reconstruction.

A.4.2 Spherical Physical Constraints

This section derives the spherical expressions of the horizontal gradient of scalar fields, horizontal divergence, and relative vorticity used in Phy-LoRA. These quantities characterize the spatial physical structures constrained during training. Figure 4 illustrates the geometric differences between spherical and planar computation. We consider a sphere with radius a, parameterized by longitude λ and latitude ϕ, with all angular quantities measured in radians.

Spherical distance and scalar gradient. At a fixed latitude ϕ, the radius of the corresponding latitude circle is $a\cos\phi$. Therefore, an infinitesimal longitudinal displacement $d\lambda$ corresponds to the eastward physical distance

$$dx = a\cos\phi\, d\lambda. \tag{21}$$

Similarly, an infinitesimal latitudinal displacement $d\phi$ corresponds to the northward physical distance

$$dy = a\, d\phi. \tag{22}$$

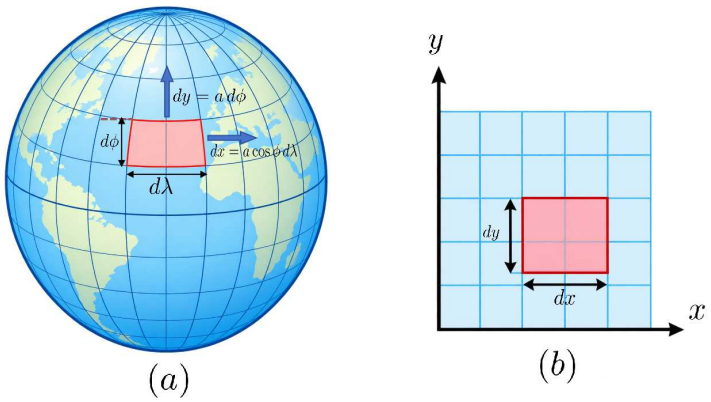


Figure 4: Spherical and planar physical operators. (a): spherical computation accounts for latitude-dependent spatial scales and geometric corrections. (b): planar computation assumes fixed grid spacing and omits these corrections.

Here, x and y denote the local eastward and northward distance coordinates, rather than global Cartesian coordinates over the sphere. For a scalar field $q(\lambda, \phi)$ defined on the sphere, applying the chain rule gives the eastward and northward spatial derivatives as

$$\frac{\partial q}{\partial x} = \frac{\partial q}{\partial \lambda}\frac{\partial \lambda}{\partial x} = \frac{1}{a\cos\phi}\frac{\partial q}{\partial \lambda}, \tag{23}$$

and

$$\frac{\partial q}{\partial y} = \frac{\partial q}{\partial \phi}\frac{\partial \phi}{\partial y} = \frac{1}{a}\frac{\partial q}{\partial \phi}. \tag{24}$$

Therefore, the components of the spherical horizontal gradient in the local eastward and northward directions are

$$\nabla_h q = \left(\frac{1}{a\cos\phi}\frac{\partial q}{\partial \lambda}, \frac{1}{a}\frac{\partial q}{\partial \phi}\right). \tag{25}$$

This expression is used to compute the horizontal gradients of near-surface temperature T_{2m} and mean sea-level pressure P_{msl}.

Spherical horizontal divergence. For the horizontal wind field $\mathbf{v}_h = (u, v)$, converting angular derivatives to physical-distance derivatives alone is insufficient to obtain the complete spherical divergence, since the variation of latitude-circle length with latitude must also be considered.

Consider an infinitesimal spherical grid cell of size $d\lambda \times d\phi$, whose area is

$$dA = a^2 \cos\phi\, d\lambda\, d\phi. \tag{26}$$

The net flux through the east and west boundaries is

$$a\frac{\partial u}{\partial \lambda}\, d\lambda\, d\phi, \tag{27}$$

while the net flux through the north and south boundaries is

$$a\frac{\partial (v\cos\phi)}{\partial \phi}\, d\lambda\, d\phi. \tag{28}$$

Dividing the total net outward flux by the grid-cell area gives the spherical horizontal divergence:

$$\delta_h = \frac{1}{a\cos\phi}\left[\frac{\partial u}{\partial \lambda} + \frac{\partial (v\cos\phi)}{\partial \phi}\right]. \tag{29}$$

Expanding the second term yields

$$\delta_h = \frac{1}{a\cos\phi}\frac{\partial u}{\partial \lambda} + \frac{1}{a}\frac{\partial v}{\partial \phi} - \frac{v\tan\phi}{a}. \tag{30}$$

Here, $-v\tan\phi/a$ is the spherical geometric correction term, which reflects the effect of latitude-dependent latitude-circle length on horizontal convergence and divergence.

Spherical relative vorticity. The radial component of relative vorticity can be defined by the circulation per unit area:

$$\zeta_h = \lim_{\Delta A \to 0} \frac{1}{\Delta A} \oint_{\partial A} \mathbf{v}_h \cdot d\boldsymbol{\ell}, \tag{31}$$

where the integration direction follows the right-hand rule with respect to the outward normal of the sphere. For the spherical grid cell defined above, the circulation contribution from the east and west boundaries is

$$a\frac{\partial v}{\partial \lambda}\, d\lambda\, d\phi, \tag{32}$$

while that from the north and south boundaries is

$$-a\frac{\partial (u\cos\phi)}{\partial \phi}\, d\lambda\, d\phi. \tag{33}$$

Dividing the total circulation by the grid-cell area gives the spherical relative vorticity:

$$\zeta_h = \frac{1}{a\cos\phi}\left[\frac{\partial v}{\partial \lambda} - \frac{\partial (u\cos\phi)}{\partial \phi}\right]. \tag{34}$$

Expanding the second term gives

$$\zeta_h = \frac{1}{a\cos\phi}\frac{\partial v}{\partial \lambda} - \frac{1}{a}\frac{\partial u}{\partial \phi} + \frac{u\tan\phi}{a}. \tag{35}$$

Here, $u\tan\phi/a$ is the spherical geometric correction term in the relative vorticity, accounting for the geometric effect induced by latitude variation.

In summary, the horizontal gradient of a scalar field converts angular variations along longitude and latitude into variations with respect to physical distance through the spherical metric. Horizontal divergence and relative vorticity further incorporate geometric correction terms arising from the latitude-dependent length of latitude circles. In Phy-LoRA, we use these spherical expressions to compute the horizontal gradients of T_{2m} and P_{msl}, as well as the divergence and relative vorticity of the horizontal wind field, thereby characterizing scalar spatial variations, convergence and divergence structures, and rotational structures in global meteorological fields.

A.4.3 Physical Constraints for In-Domain and Cross-Source Experiments

Both settings use horizontal kinetic energy, horizontal divergence, and relative vorticity to constrain the wind field, while spherical horizontal gradients are used to characterize spatial variations of scalar fields.

In-domain downscaling (ERA5-to-ERA5). This setting jointly predicts U10, V10, T2M, and MSL. For the horizontal wind field composed of U10 and V10, we use horizontal kinetic energy E_h, horizontal divergence δ_h, and relative vorticity ζ_h to characterize wind-speed magnitude, convergence–divergence, and rotational structures, respectively. For T2M and MSL, spherical horizontal gradients are used to constrain the spatial variations of temperature and pressure. The corresponding physical structural operator is

$$\Phi_{\mathrm{in}}(X) = \left[E_h,\, \delta_h,\, \zeta_h,\, \nabla_h T_{2m},\, \nabla_h P_{\mathrm{msl}}\right]. \tag{36}$$

Cross-source downscaling (MPI-ESM-to-ERA5). This setting jointly predicts Z500, T850, U10, V10, and T2M. The wind field uses the same three physical diagnostics described above. For Z500, T850, and T2M, spherical horizontal gradient constraints are applied, yielding

$$\Phi_{\mathrm{cross}}(X) = \left[E_h,\, \delta_h,\, \zeta_h,\, \nabla_h Z_{500},\, \nabla_h T_{850},\, \nabla_h T_{2m}\right]. \tag{37}$$

Here, $\nabla_h Z_{500}$ characterizes the spatial variations of the geopotential-height field, while $\nabla_h T_{850}$ and $\nabla_h T_{2m}$ characterize the spatial variations of upper-air and near-surface temperature fields, respectively.

Table 4: Performance of $8\times$ global cross-source downscaling from 2° MPI-ESM inputs to 0.25° ERA5 targets across different meteorological variables. Lower RMSE and LFD, and MB closer to zero, indicate better performance. Best results are shown in bold.

Method	Z500			T850			U10			V10			T2M		
	RMSE	MB	LFD	RMSE	MB	LFD	RMSE	MB	LFD	RMSE	MB	LFD	RMSE	MB	LFD
Bilinear	119.4930	**-0.2225**	23.1200	1.4226	**0.0052**	14.3736	2.1809	-0.0360	15.2711	1.9528	0.0358	15.0927	2.3739	0.0040	15.3915
ViT	89.6243	-2.6077	22.5802	1.3671	-0.0174	14.3278	1.7995	-0.0952	14.9091	1.5794	-0.0237	14.6906	1.8471	-0.0333	14.9779
DeepSD	107.7923	1.8577	22.9134	1.4128	-0.0884	14.3573	2.0607	0.0966	15.1655	1.7577	-0.0577	14.8918	2.2200	-0.0841	15.3067
CorrDiff	91.8154	1.4100	22.6034	1.3719	0.0174	14.3337	1.5587	0.0030	14.6219	1.4293	0.0099	14.4909	1.6029	-0.0172	14.6811
SRGAN	106.3857	8.2178	22.8596	1.6617	-0.0244	14.6334	2.1114	-0.0483	15.1852	1.7823	-0.0667	14.9151	2.3175	-0.0215	15.3703
DSFNO	96.2482	8.1645	22.6852	1.3785	0.0447	14.3458	1.6305	-0.0277	14.7198	1.4975	0.0365	14.5854	1.6439	0.0066	14.7375
GeoFAR	84.2520	0.6136	22.4202	1.3466	-0.0323	14.3162	1.4712	0.0098	14.5045	1.3450	**0.0031**	14.3700	1.4475	-0.0337	14.4728
ClimLoRA	**81.8283**	2.0764	**21.5546**	**1.2166**	-0.0136	**14.1680**	**1.3090**	**-0.0029**	**14.2950**	**1.2066**	0.0182	**14.1345**	**1.3183**	**-0.0019**	**14.3838**

A.5 Additional Experiments and Analysis

A.5.1 Cross-Source Downscaling at $8\times$

We further evaluate ClimLoRA on $8\times$ global cross-source downscaling using MPI-ESM inputs regridded to 2° and ERA5 targets at 0.25°. The downscaling factor refers to the ratio of input and target grid spacings. We jointly predict Z500, T850, U10, V10, and T2M. Table 4 reports the results. ClimLoRA achieves the lowest RMSE across all five variables. Compared with GeoFAR, RMSE decreases from 84.2520 to 81.8283 for Z500 and from 1.3466 to 1.2166 for T850. The reductions for U10, V10, and T2M are approximately 11.0%, 10.3%, and 8.9%, respectively. These results demonstrate improved reconstruction accuracy across both surface and upper-air variables at the $8\times$ grid downscaling factor. ClimLoRA also achieves the lowest LFD across all five variables. Although its absolute MB is not consistently the smallest across all variables, ClimLoRA achieves the lowest absolute MB for U10 and T2M. Overall, ClimLoRA improves reconstruction accuracy and reduces frequency-domain errors across all evaluated variables in this cross-source setting.

A.5.2 Adjustable Inference and Coefficient Sensitivity

We analyze the effects of the two inference guidance coefficients, λ_{fine} and λ_{phy}, on ERA5-to-ERA5 downscaling. Table 5 varies one coefficient while fixing the other at 1.0. Increasing λ_{fine} from 0 to 1.0 consistently improves both RMSE and LFD across all variables. For example, the RMSE of U10 decreases from 0.3776 to 0.2677, while that of MSL decreases from 21.8605 to 15.9568. Further increasing λ_{fine} beyond 1.0 degrades performance, indicating that excessive fine-scale correction can introduce additional errors.

A similar trend is observed for λ_{phy}. Increasing it from 0 to 0.8 reduces the RMSE of U10, V10, and T2M from 0.2709, 0.2647, and 0.3882 to 0.2669, 0.2593, and 0.3793, respectively. For MSL, the best RMSE and LFD are obtained at $\lambda_{\text{phy}} = 1.0$. Overall, the best performance is concentrated around 0.8–1.0, while larger coefficients gradually degrade reconstruction. We therefore use $\lambda_{\text{fine}} = \lambda_{\text{phy}} = 1.0$ as the default setting, which provides consistently strong performance across variables.

A.5.3 Additional Visualizations and Frequency Analysis

Latitude-Wise Analysis of Frequency Representations We further compare the latitude-wise errors of spherical harmonic transform (SHT) and two-dimensional Fourier transform (FFT) on the ERA5-to-ERA5 downscaling setting, using the average results over the full 2020 test year. As shown in Fig. 5, SHT yields more stable latitude-wise error profiles for both T2M and MSL, whereas FFT exhibits stronger latitude-dependent variations, particularly at high latitudes. These results indicate that spherical harmonic representations better accommodate the spherical geometry of global meteorological fields and reduce latitude-dependent errors associated with planar Fourier modeling.

Additional qualitative results. We further provide qualitative visualizations for both in-domain and cross-source downscaling across additional meteorological variables and timestamps. As shown in Figs. 6–8, these examples include U10 and MSL for in-domain downscaling, as well as Z500, T850, U10, V10, and T2M for cross-source downscaling, showing both predicted fields and the corresponding absolute errors.

Table 5: Effect of guidance coefficients on ERA5-to-ERA5 downscaling. Each block varies one coefficient while fixing the other at 1.0. The best result within each block is shown in bold. Lower RMSE and LFD indicate better performance.

λ_{fine}	λ_{phy}	U10		V10		T2M		MSL	
		RMSE ↓	LFD ↓	RMSE ↓	LFD ↓	RMSE ↓	LFD ↓	RMSE ↓	LFD ↓
0.0	1.0	0.3776	11.8560	0.3530	11.7251	0.4957	12.4250	21.8605	19.9212
0.2	1.0	0.3411	11.6586	0.3214	11.5434	0.4572	12.2628	19.8277	19.7313
0.5	1.0	0.2971	11.3930	0.2839	11.3055	0.4115	12.0517	17.4244	19.4831
0.8	1.0	0.2716	11.2226	0.2627	11.1591	0.3849	11.9177	16.0991	19.3358
1.0	1.0	**0.2677**	**11.1955**	**0.2599**	**11.1390**	**0.3799**	**11.8915**	**15.9568**	**19.3229**
1.2	1.0	0.2749	11.2465	0.2665	11.1869	0.3856	11.9215	16.4347	19.3828
1.5	1.0	0.3045	11.4426	0.2924	11.3641	0.4131	12.0596	18.1906	19.5804
1.0	0.0	0.2709	11.2165	0.2647	11.1747	0.3882	11.9359	17.2929	19.4660
1.0	0.2	0.2688	11.2028	0.2623	11.1568	0.3843	11.9162	16.7872	19.4152
1.0	0.5	0.2670	11.1909	0.2600	11.1398	0.3806	11.8965	16.2414	19.3581
1.0	0.8	**0.2669**	**11.1900**	**0.2593**	**11.1351**	**0.3793**	**11.8892**	15.9735	19.3274
1.0	1.0	0.2677	11.1955	0.2599	11.1390	0.3799	11.8915	**15.9568**	**19.3229**
1.0	1.2	0.2693	11.2057	0.2612	11.1484	0.3816	11.8993	16.0705	19.3315
1.0	1.5	0.2731	11.2297	0.2645	11.1724	0.3861	11.9213	16.4788	19.3680

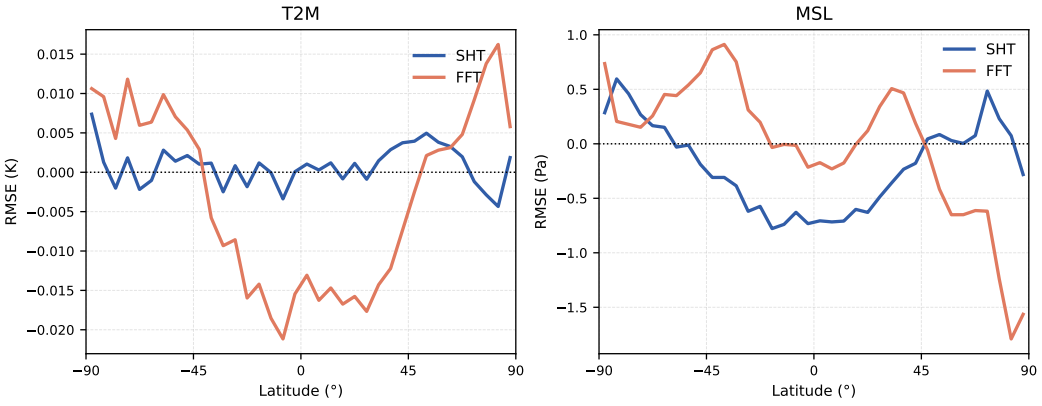


Figure 5: Latitude-wise RMSE differences of SHT and FFT for T2M and MSL, averaged over the 2020 test year. SHT exhibits more stable error variations across latitudes, while FFT shows stronger latitude-dependent fluctuations, particularly at high latitudes.

A.6 Relative Improvement in Physical Structural Errors

To compare the effects of different update strategies on physical structures, we compute the relative improvement (RI) in the error of each physical structural quantity using the same Base model as the reference:

$$\mathrm{RI}_k = \frac{e_k^{\mathrm{Base}} - e_k^{\mathrm{Method}}}{e_k^{\mathrm{Base}}} \times 100\%. \tag{38}$$

Here, e_k^{Base} and e_k^{Method} denote the errors of the Base and the evaluated method for the k-th physical structural quantity, respectively. These quantities include horizontal kinetic energy, horizontal divergence, relative vorticity, and the horizontal gradients of near-surface temperature and mean sea-level pressure. All structural quantities are computed after restoring physical units and compared with the high-resolution reference fields using the same test samples, spherical operators, and error metric. Positive RI indicates a lower error than the Base, while negative RI indicates increased error. Higher RI indicates greater relative improvement in physical structural errors.

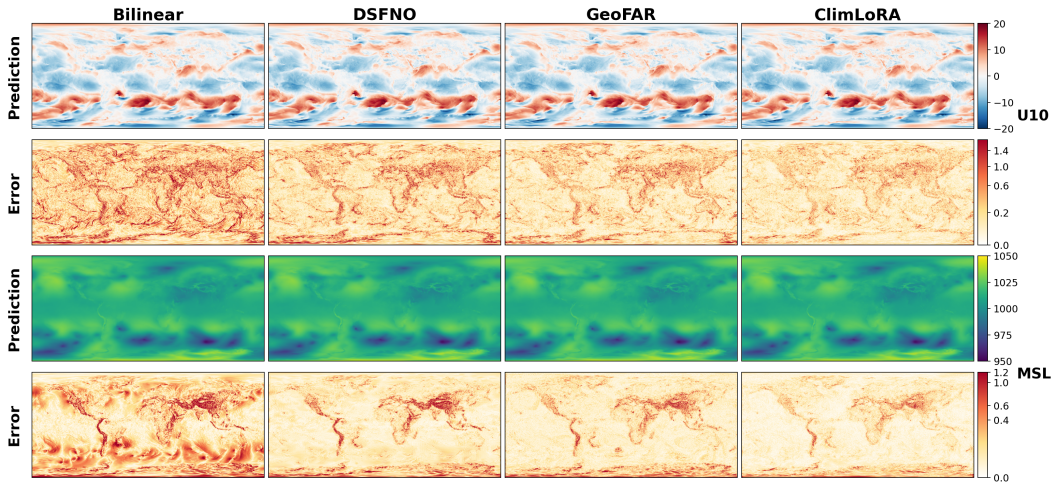


Figure 6: Qualitative comparison of same-source U10 and MSL downscaling from ERA5 (1°) to ERA5 (0.25°) at 12:00 UTC on July 1, 2020. The top row shows the predicted fields from different methods, while the bottom row shows the corresponding absolute errors with respect to the ground truth.

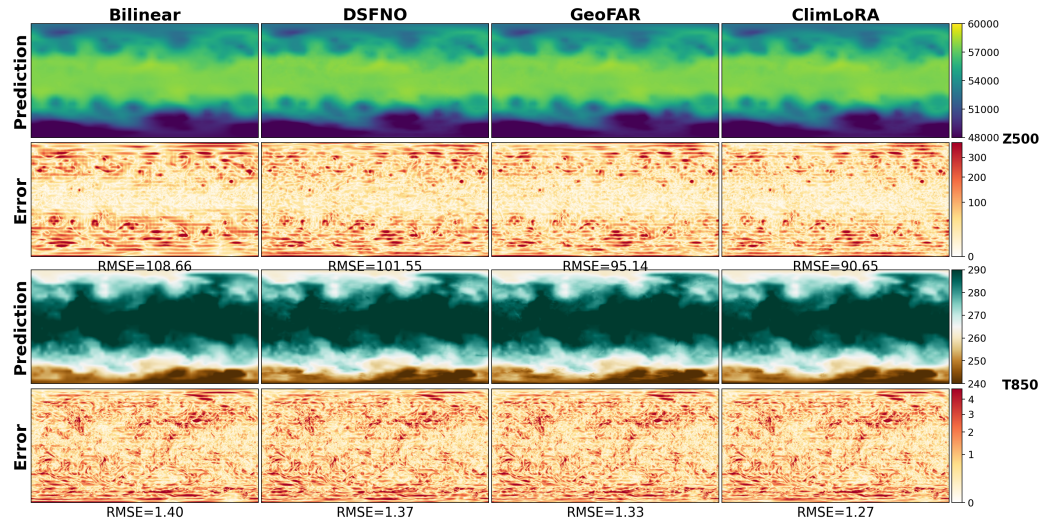


Figure 7: Cross-source downscaling of Z500 and T850 from MPI-ESM (1°) to ERA5 (0.25°) at 18:00 UTC on June 10, 2018. Predictions and absolute errors are shown.

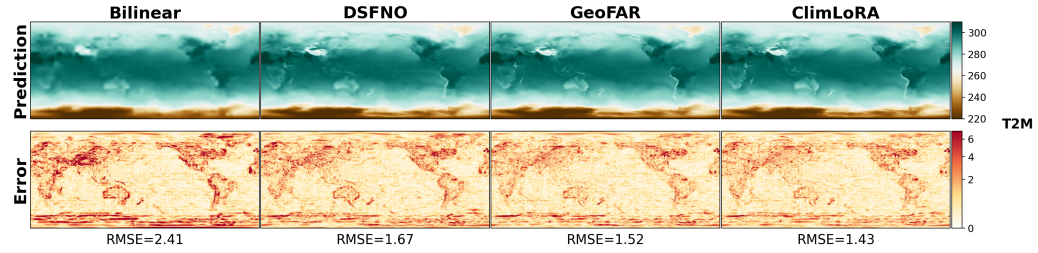


Figure 8: Cross-source downscaling of T2M from MPI-ESM (1°) to ERA5 (0.25°) at 18:00 UTC on May 28, 2018. Predictions and absolute errors are shown.

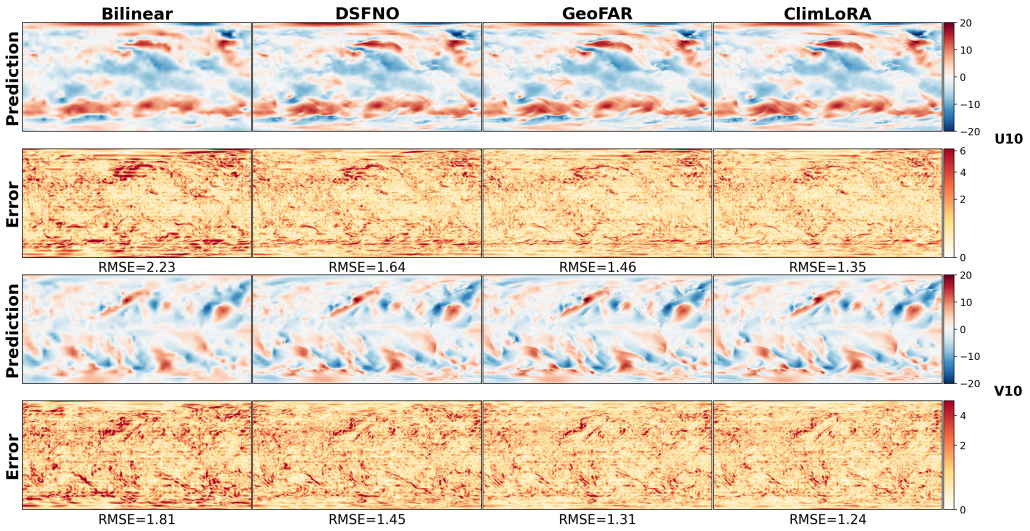


Figure 9: Cross-source downscaling of U10 and V10 from MPI-ESM ($1°$) to ERA5 ($0.25°$) at 00:00 UTC on March 5, 2018. Predictions and absolute errors are shown.