NEURAL INFORMATION PROCESSING SYSTEMS

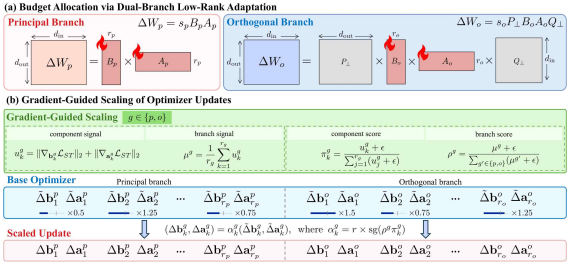
NeurIPS 2026

MLV Machine Learning & Vision Lab | KAIST

PROLA: Principal-Orthogonal Low-rank Adaptation for Predictive Spatiotemporal Weather Downscaling

Minseo Yoon · Minseong Bae · Sojin Lee · Hyunwoo J. Kim

Korea Advanced Institute of Science and Technology (KAIST)





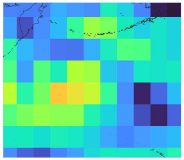
1. Background & motivation

Predictive spatiotemporal weather downscaling

North Pacific wind: resolve the structure and its evolution

Coarse history

5.625° grid · states 6h apart

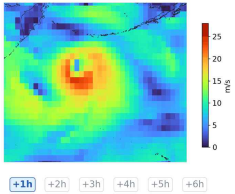


t – 6h t

PSWD

Fine-resolution future

1.5° grid · hourly targets



Spatial refinement + temporal refinement + future prediction

1. Background & motivation

Two adaptation needs under a fixed rank budget

Pretrained-aligned adaptation



Start from leading principal directions

Task-specific corrections



Reserve capacity outside the selected principal directions

Allocate a fixed rank budget to pretrained-aligned adaptation and orthogonal corrections.

1. Background & motivation

Three research questions

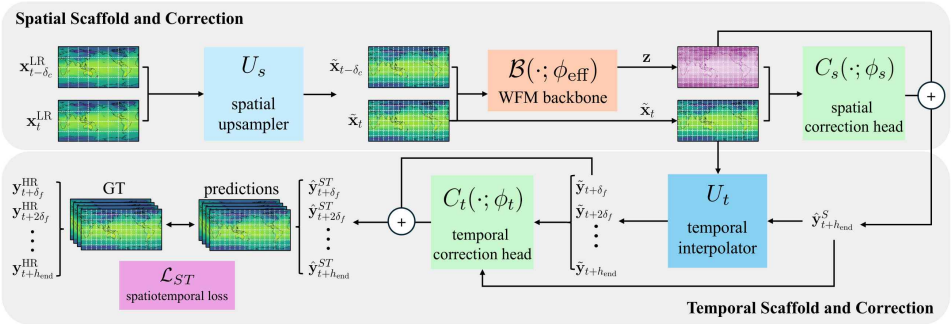
1 **Prediction pipeline**
How do we predict a fine-grained future sequence?

2 **Rank allocation**
How do we allocate a fixed rank budget?

3 **Update scaling**
How strongly should each component be updated?

2. Methodology

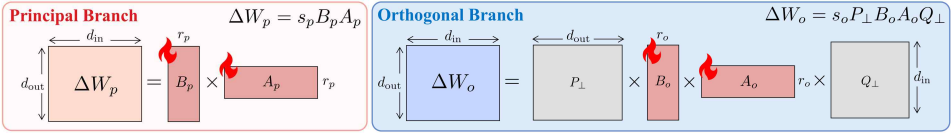
Scaffold-and-correct pipeline



Spatial endpoint correction anchors temporal refinement.

2. Methodology

Budget allocation via dual-branch low-rank adaptation



Principal branch

Initialized from leading singular directions; factors remain trainable

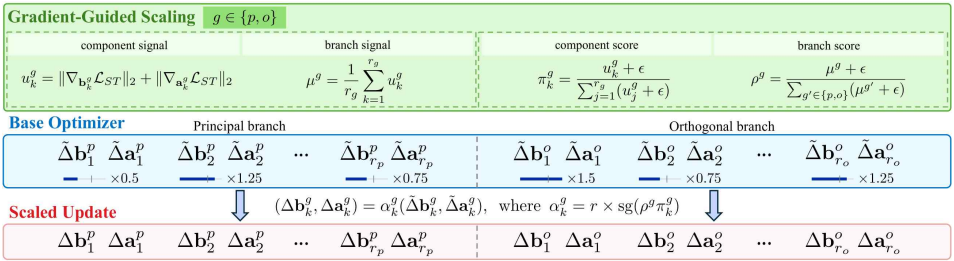
Orthogonal branch

Corrections outside the initial principal subspaces

Fixed total rank · function-preserving initialization

2. Methodology

Gradient-guided scaling of optimizer updates

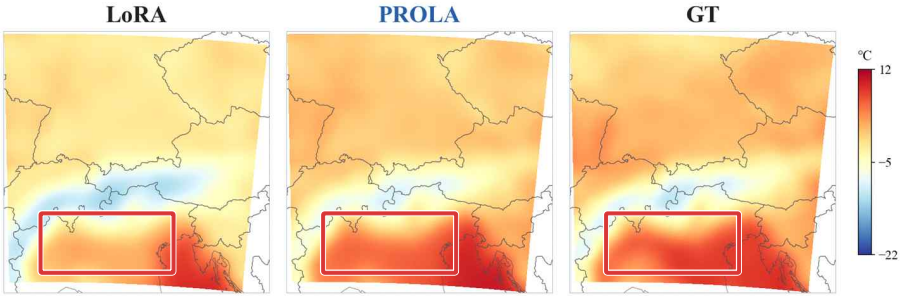


Gradient signals guide update strength without increasing trainable rank.

3. Results

Regional temperature patterns over time

ERA5-to-CERRA · T2M · 6h forecasts



Closer recovery of regional temperature contrasts

3. Results

Consistent gains across global PSWD variables

6h latitude-weighted RMSE · mean ± std over three seeds

Method	Trainable Params	Z500	T850	T2M	U10	V10
LoRA	3.92M	79.17 ± 0.33	0.994 ± 0.003	1.280 ± 0.012	1.314 ± 0.005	1.306 ± 0.002
DoRA	4.16M	78.40 ± 0.30	0.989 ± 0.003	1.272 ± 0.013	1.306 ± 0.006	1.298 ± 0.001
AdaLoRA	5.35M	77.51 ± 0.94	0.983 ± 0.006	1.251 ± 0.002	1.282 ± 0.004	1.279 ± 0.002
PiSSA	3.92M	78.43 ± 0.93	0.989 ± 0.004	1.251 ± 0.009	1.298 ± 0.004	1.289 ± 0.003
OPLoRA	3.92M	79.60 ± 0.38	0.996 ± 0.003	1.281 ± 0.012	1.319 ± 0.005	1.311 ± 0.002
PROLA	3.92M	**72.35** ± 0.64	**0.956** ± 0.002	**1.212** ± 0.006	**1.240** ± 0.005	**1.234** ± 0.002

6.7%

lower Z500 RMSE

vs. strongest PEFT (AdaLoRA)

77.51 → **72.35**

<0.4%

of backbone size

3.92M trainable parameters
1.26B backbone parameters

3. Results

Generalization to regional weather tasks

ERA5-to-CERRA downscaling

Latitude-weighted RMSE at 3h

Method	Z500	T850	T2M	U10	V10
LoRA	72.14	1.076	2.150	1.370	1.327
PiSSA	71.80	1.072	2.126	1.354	1.311
OPLoRA	72.56	1.081	2.132	1.371	1.332
PROLA	**68.44**	**1.028**	**2.021**	**1.316**	**1.274**

4.9% lower T2M RMSE than PiSSA

Regional precipitation forecasting

12h lead · 6h totals · RMSE in cm

Method	SEEPS ↓	RMSE ↓	ACC ↑
LoRA	0.608	0.212	0.660
PiSSA	0.612	0.213	0.656
OPLoRA	0.605	0.212	0.661
PROLA	**0.590**	**0.208**	**0.676**

Lower SEEPS and RMSE; higher ACC

The gains extend to regional downscaling and precipitation forecasting.

3. Results

Both the prediction pipeline and PROLA design matter

Scaffold-and-correct pipeline

3h RMSE ↓ · PROLA in all variants

Method / Variant	Z500	T850	T2M	U10	V10
Direct prediction	53.13	0.898	1.165	1.211	1.177
w/o C_s, C_t	87.73	0.959	1.307	1.277	1.270
w/o C_t	84.53	0.938	1.290	1.259	1.254
w/o C_s	53.35	0.880	1.138	1.155	1.133
Full pipeline	**50.30**	**0.866**	**1.128**	**1.140**	**1.119**

84.53 → **50.30**

PROLA branch and scaling ablation

3h RMSE ↓ · fixed total rank 8

Configuration	r_p	r_o	Scaling	Z500	T850	T2M
Principal	8	0	–	53.66	0.894	1.193
Orthogonal	0	8	–	54.54	0.899	1.197
Both branches	4	4	–	51.36	0.875	1.143
PROLA	4	4	✓	**50.30**	**0.866**	**1.128**

51.36 → **50.30**

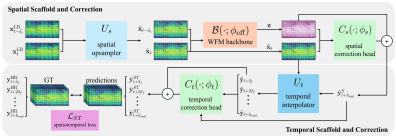
NEURAL INFORMATION PROCESSING SYSTEMS NeurIPS 2026

PROLA: Principal-Orthogonal Low-rank Adaptation for Predictive Spatiotemporal Weather Downscaling

Predictive spatiotemporal weather downscaling
Future prediction with spatial and temporal refinement.

Principal–orthogonal adaptation
A fixed rank budget for complementary adaptation roles.

Gradient-guided update scaling
Better adaptation across global and regional weather tasks.



Thank you

Read the paper

arXiv:2611.11318

MLV Machine Learning & Vision Lab KAIST