ADAPTING AI WEATHER FOUNDATION MODELS TO A CHANGING CLIMATE

Anonymous authors
Paper under double-blind review

ABSTRACT

AI weather foundation models learn from historical archives, while the climate distribution encountered during deployment continues to evolve. This growing model–climate mismatch motivates continual adaptation without repeated full-model retraining, but the structure of weather forecast errors determines how such adaptation should be performed. We develop an online adaptation framework that exploits two complementary error structures: geographically persistent biases and flow-dependent errors that vary with the atmospheric state. Analysis of Aurora-1.3B reveals strongly concentrated forecast-loss sensitivity in its decoder, motivating flow-dependent corrections within fixed low-dimensional sensitivity subspaces. A residual memory tracks persistent errors remaining after this correction, while a variable-wise online gate controls the contribution of the persistent component. We evaluate Aurora-1.3B using ERA5 from 1959–1968 for offline initialization and a strictly chronological 1969–2022 stream for online evaluation, with updates restricted to previously available verification targets. For 2-metre temperature, distributional departure from the historical reference, measured directly from ERA5, is associated with larger frozen-model errors and greater adaptation gains, while the adapted model maintains more stable skill across evaluation periods. Across 69 output fields and five lead times from 6 to 72 hours, our method reduces MSE and MAE by 4.17% and 2.76%, respectively, relative to δ-Adapter, the strongest of six online adaptation baselines. Evolving climate distributions create the need for continual adaptation, while weather-specific forecast-error structure guides how that adaptation should be performed.

1 INTRODUCTION

AI-based weather forecasting has advanced rapidly, with models such as Pangu-Weather, GraphCast, and Aurora achieving competitive global medium-range forecasting skill (Bi et al., 2023; Lam et al., 2023; Bodnar et al., 2025). Long-term deployment, however, poses a different challenge: these models learn from finite historical archives, whereas the distribution of weather continues to evolve as the climate changes (Sippel et al., 2020). As deployment conditions depart from those represented during training, a frozen model can face increasing model–climate mismatch, while repeatedly retraining billion-parameter weather foundation models is computationally expensive. Related work on climate emulation has similarly highlighted the difficulty of generalizing across changing climate distributions (Stanley-Clamp et al., 2026). This motivates continual adaptation from newly verified forecasts while restricting the scope and cost of each update.

Evolving climate distributions create the need for continual adaptation, while weather-specific forecast-error structure guides how that adaptation should be performed. Existing online time-series methods track changing or recurring patterns under sequentially revealed supervision (Gama et al., 2014; Pham et al., 2022; Zhao & Shen, 2025), but weather forecasting involves high-dimensional, spatially coupled fields with characteristic error structures. Persistent regional biases, particularly in near-surface variables such as 2-metre temperature (T2M), can remain coherent across successive forecast cycles (Libonati et al., 2008; Haiden et al., 2018), whereas other errors vary strongly with the instantaneous atmospheric circulation (Gelaro et al., 1998; Leutbecher & Palmer, 2008; Ferranti et al., 2015). These complementary structures suggest retaining verified residual information where errors persist while learning state-dependent corrections where errors vary with the flow. Because

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.

their relative importance differs across atmospheric variables and forecast conditions, an effective adaptation mechanism should also adjust their contribution online.

The main difficulty is making flow-dependent adaptation efficient for a billion-parameter model operating on high-dimensional atmospheric states. Classical predictability studies show that forecast-error growth can concentrate within relatively low-dimensional dynamically sensitive directions (Gelaro et al., 1998; Patil et al., 2001; Trevisan & Uboldi, 2004). We observe an analogous structure in Aurora: historical forecast-loss gradients exhibit strongly concentrated and asymmetric decoder sensitivity, with rank-32 input-side subspaces in the later decoder stage capturing more than 90% of the gradient trace at the median across layers (Figure 2). Moreover, although these sensitivity bases are estimated offline and kept fixed during deployment, their activation changes with the evolving atmospheric circulation (Figure 3). This suggests that online adaptation can be confined to historically error-sensitive directions without losing responsiveness to the current atmospheric state.

Guided by these observations, we develop a structure-aware online adaptation framework for weather foundation models. Flow-dependent corrections are learned by updating coefficients within fixed low-dimensional forecast-sensitivity subspaces while the pretrained backbone remains frozen. An exponentially weighted spatial residual memory tracks persistent errors remaining after the flow-dependent correction, and a variable-wise online gate controls how much of this persistent correction is added to the forecast. The gate is updated from sequentially revealed verification losses, and we establish a regret guarantee against the better fixed correction choice in hindsight. All updates use only information available at each point in the chronological stream, preserving a strictly causal online protocol.

We evaluate Aurora-1.3B using ERA5 reanalysis (Hersbach et al., 2020), with 1959–1968 used for offline initialization and sensitivity estimation and 1969–2022 forming a strictly chronological 54-year online evaluation stream. To characterize changing deployment conditions independently of forecast performance, we measure the departure of the T2M distribution from the historical reference directly from ERA5 using Jensen–Shannon distance. Larger departures are associated with higher frozen-model errors and greater adaptation gains, while the adapted model maintains more stable T2M skill across evaluation periods (Figure 5). Across 69 output fields and lead times of 6, 12, 24, 48, and 72 hours, our method reduces MSE and MAE by 4.17% and 2.76%, respectively, relative to δ-Adapter (Liang et al., 2026), the strongest of six online adaptation baselines. Together, these results support continual adaptation as deployment climates evolve and show that forecast-error structure provides a principled basis for how that adaptation can be carried out efficiently.

2 Related Works

AI weather forecasting models. Data-driven global weather forecasting has advanced rapidly through models such as FourCastNet (Kurth et al., 2023), Pangu-Weather (Bi et al., 2023), GraphCast (Lam et al., 2023), FuXi (Chen et al., 2023), Aurora (Bodnar et al., 2025), GenCast (Price et al., 2025), and AIFS (Lang et al., 2024). These systems achieve competitive deterministic or probabilistic medium-range forecasting, often at substantially lower inference cost than conventional numerical weather prediction. Most, however, are trained offline and deployed with fixed parameters. Continual adaptation from sequentially verified forecasts as deployment distributions evolve remains comparatively underexplored.

Online forecasting under distribution shift. Online time-series methods adapt predictors as new targets become available under non-stationary dynamics. Representative approaches include FSNet (Pham et al., 2022), OneNet (Zhang et al., 2023), Proceed (Zhao & Shen, 2025), δ-Adapter (Liang et al., 2026), ADAPT-Z (Huang et al., 2026), and ELF (Lee et al., 2025). These methods address distribution shift through parameter updates, lightweight corrections, feature adaptation, memory, or online ensembling. Weather forecasting, however, exhibits persistent spatial biases and rapidly varying flow-dependent errors that call for different forms of correction. Our approach explicitly models these complementary structures and adapts their relative contribution online.

Weather forecast-error structure. Persistent and flow-dependent forecast errors have long been studied in numerical weather prediction. Sequential bias-correction methods use previously verified residuals to reduce systematic regional errors (Libonati et al., 2008; Haiden et al., 2018), while pre-

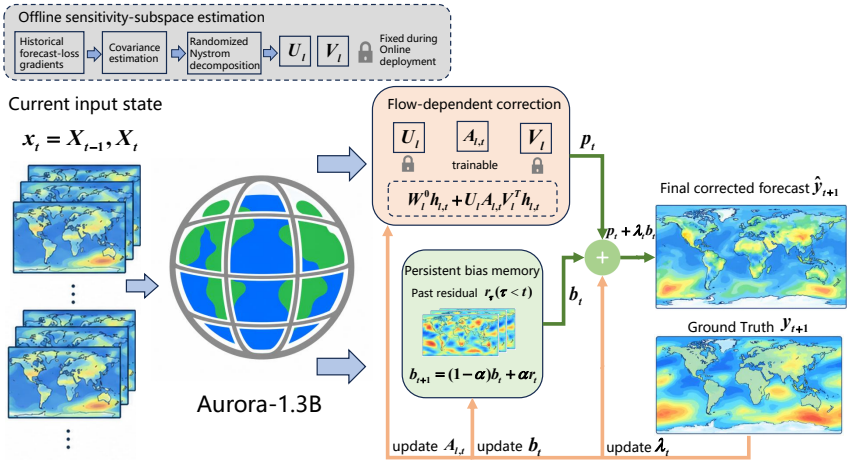


Figure 1: **Overview of the proposed online adaptation framework.** The Aurora-1.3B backbone remains frozen, while flow-dependent corrections are learned within fixed forecast-sensitivity subspaces and persistent residual errors are tracked by an online residual memory. A variable-wise gate adaptively controls the contribution of the persistent residual correction, and all updates use only previously revealed verification.

dictability studies show that forecast-error growth depends on the instantaneous circulation and can concentrate along dynamically sensitive directions (Gelaro et al., 1998; Patil et al., 2001; Leutbecher & Palmer, 2008; Ferranti et al., 2015; Trevisan & Uboldi, 2004). Neural weather forecasting has also begun to explore perturbation generation and model sensitivity (Scher & Messori, 2021; Tian et al., 2026). We build on these insights by using forecast-error structure not merely for diagnosis or ensemble generation, but to determine how a weather foundation model should adapt online.

3 METHODS

We consider supervised online adaptation of an AI weather foundation model (AI-WFM) under a strictly chronological data stream. We adopt Aurora (Bodnar et al., 2025), an Earth-system forecasting model with approximately 1.3 billion parameters, as the backbone. Our framework is guided by the structure of forecast errors encountered during deployment. We first characterize persistent and flow-dependent error structures that motivate complementary forms of correction. We then estimate low-dimensional forecast-sensitivity subspaces from historical data to enable efficient flow-dependent adaptation. Finally, a spatial residual memory tracks persistent errors remaining after the flow-dependent correction, while a variable-wise online gate controls the contribution of this persistent correction. Figure 1 provides an overview of the framework.

3.1 STRUCTURED FORECAST ERRORS IN LONG-TERM DEPLOYMENT

Forecast errors encountered during long-term deployment exhibit complementary temporal and spatial structures. Some errors are geographically coherent and persist across successive forecast cycles, particularly for near-surface variables (Libonati et al., 2008). Other errors depend strongly on the instantaneous atmospheric circulation, with their magnitude and spatial structure varying with the evolving flow (Leutbecher & Palmer, 2008; Ferranti et al., 2015). These structures motivate different forms of online correction: persistent errors can be tracked using previously verified residuals, whereas flow-dependent errors require corrections that respond to the current atmospheric state.

Consider online forecasting round t. Let $x_t = (X_{t-1}, X_t)$ denote the multivariate atmospheric input state, $y_{t+1,c}(s)$ the verifying target, and $f_{\theta_0,c}(x_t; s)$ the corresponding prediction of the frozen foundation model. Here, θ_0 denotes the fixed pretrained parameters, c indexes an atmospheric variable, and s denotes its spatial coordinate. We operationally represent the forecast error as

$$e_{t,c}(s) = y_{t+1,c}(s) - f_{\theta_0,c}(x_t; s) = b_{t,c}(s) + d_{t,c}(x_t; s) + \varepsilon_{t,c}(s). \tag{1}$$

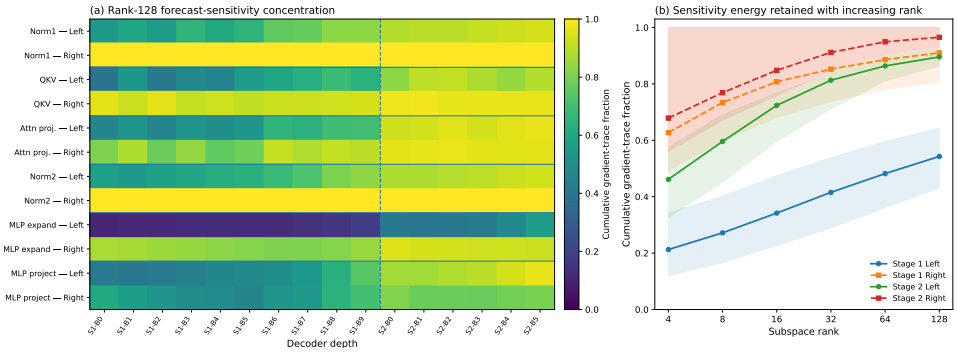


Figure 2: **Low-dimensional structure of forecast-loss sensitivity in the Aurora decoder. (a)** Cumulative gradient-trace fraction captured by the leading rank-128 left and right sensitivity subspaces across decoder layers. **(b)** Cumulative gradient-trace fraction versus subspace rank r; lines show the median across layers and shaded regions denote the interquartile range.

The term $b_{t,c}(s)$ represents a temporally persistent and spatially coherent component, while $d_{t,c}(x_t; s)$ denotes a state-dependent, flow-responsive component. The remaining term $\varepsilon_{t,c}(s)$ collects forecast errors that cannot be reliably captured from the available online information. Equation 1 is an operational decomposition for adaptation rather than a causal decomposition of climate-change effects, and the components need not be strictly orthogonal. We next construct a compact model-space representation for the flow-dependent component.

3.2 Forecast-Sensitivity Subspaces for Flow-Dependent Adaptation

NWP predictability studies have shown that, despite the high dimensionality of atmospheric state space, dominant finite-time forecast-error growth can concentrate within a small number of dynamically sensitive directions (Gelaro et al., 1998; Patil et al., 2001; Annan, 2004). Motivated by this structure, we examine whether forecast-loss sensitivity within a weather foundation model is similarly concentrated in a low-dimensional subspace.

For each selected linear layer ℓ, we characterize historical forecast-loss sensitivity using per-sample weight gradients

$$G_{\ell,i} = \nabla_{W_\ell} \mathcal{L}\big(f_{\theta_0}(x_i), y_i\big).$$

The corresponding output- and input-side second-moment matrices are

$$C_\ell^{\text{out}} = \sum\nolimits_i G_{\ell,i} G_{\ell,i}^\top, \qquad C_\ell^{\text{in}} = \sum\nolimits_i G_{\ell,i}^\top G_{\ell,i}. \tag{2}$$

Randomized Nyström subspace estimation. Let $P_{\mathcal{S}}$ denote the orthogonal projector onto subspace $\mathcal{S}$. For each side $q \in \{\text{out}, \text{in}\}$, we seek the rank-$r$ subspace that captures the largest amount of historical forecast-loss sensitivity:

$$\mathcal{S}_{\ell,r}^q = \underset{\dim(\mathcal{S})=r}{\arg\max}\ \operatorname{tr}(P_{\mathcal{S}} C_\ell^q)\,. \tag{3}$$

By the Rayleigh–Ritz principle (Horn & Johnson, 2012), the solution is spanned by the leading r eigenvectors of C_ℓ^q. Because these second-moment matrices are too large to construct explicitly, we estimate their dominant eigenspaces using a randomized Nyström sketch.

Specifically, we draw an orthonormalized Gaussian matrix $\Omega_\ell^q \in \mathbb{R}^{d_\ell^q \times k}$, where $k = r + p$, and form $Y_\ell^q = C_\ell^q \Omega_\ell^q$. The sketches can be accumulated directly from per-sample gradients without materializing C_ℓ^q:

$$Y_\ell^{\text{out}} = \sum_i G_{\ell,i} \left(G_{\ell,i}^\top \Omega_\ell^{\text{out}}\right), \qquad Y_\ell^{\text{in}} = \sum_i G_{\ell,i}^\top \left(G_{\ell,i} \Omega_\ell^{\text{in}}\right). \tag{4}$$

We then recover the leading eigendirections from the Nyström approximation

$$\widehat{C}_\ell^q = Y_\ell^q \left[(\Omega_\ell^q)^\top Y_\ell^q\right]^\dagger (Y_\ell^q)^\top. \tag{5}$$

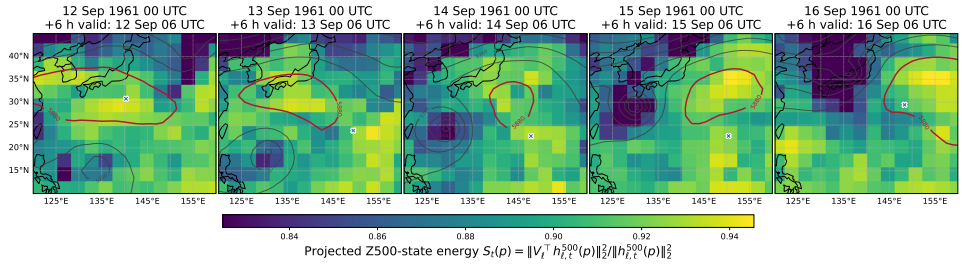


Figure 3: **Flow-dependent activation of the Z500 forecast-sensitivity subspace.** We visualize five consecutive 00 UTC forecasts from 12–16 September 1961. Colors show the fraction of local Z500 latent-state energy captured by the offline rank-128 sensitivity subspace, $S_t(p)$. Gray contours denote the verifying 500-hPa geopotential height, the red contour marks the 5880-gpm boundary conventionally associated with the western North Pacific subtropical high, and the white cross indicates the energy-weighted centroid of the top-5% sensitivity scores.

where † denotes the Moore–Penrose pseudoinverse. Its leading r eigenvectors define the output- and input-side sensitivity bases U_ℓ and V_ℓ, respectively. Complete factorization and implementation details are provided in Appendix C.2.

Figure 2 reveals a pronounced low-dimensional and asymmetric sensitivity structure. The right subspaces are highly concentrated, with the Stage 2 median capturing over 90% of the gradient trace at rank 32 and about 97% at rank 128. In contrast, the Stage 1 left subspace captures only about 55% at rank 128. Forecast-loss sensitivity is therefore substantially more compressible on the input side and becomes increasingly concentrated toward the later decoder stages.

Although the sensitivity bases are estimated offline and remain fixed during deployment, their activation depends on the current atmospheric state. Figure 3 shows that regions of high projected Z500-state energy move with the evolving 5880-gpm subtropical-ridge boundary, consistent with NWP studies locating forecast sensitivity around tropical cyclones, their environmental steering flows, and weaknesses in the subtropical ridge (Peng & Reynolds, 2006; Reynolds et al., 2009; Torn et al., 2025). This state-dependent activation allows a fixed low-dimensional basis to support flow-responsive online corrections.

3.3 Flow-Dependent Adaptation with Persistent Residual Gating

With the sensitivity bases established, we translate the operational error structure in Eq. 1 into two complementary online adaptation mechanisms. The flow-dependent correction is learned within the fixed sensitivity subspaces, while persistent errors remaining after this correction are tracked through spatial residual memory. The flow-dependent correction is always applied; a variable-wise gate controls only the additional contribution of the persistent residual correction.

Flow-dependent correction. The forecast-sensitivity bases U_ℓ and V_ℓ estimated during the offline period remain fixed throughout deployment. For a selected linear layer ℓ, let $h_{\ell,t}$ denote its current input activation and W_ℓ^0 its frozen pretrained weight. We introduce an online coefficient matrix $A_{\ell,t} \in \mathbb{R}^{r_\ell^{\text{out}} \times r_\ell^{\text{in}}}$ and modify the layer output as

$$\widetilde{a}_{\ell,t} = W_\ell^0 h_{\ell,t} + U_\ell A_{\ell,t} V_\ell^\top h_{\ell,t}. \tag{6}$$

Here, $V_\ell^\top h_{\ell,t}$ extracts coordinates along historically error-sensitive input directions, $A_{\ell,t}$ maps these coordinates to correction coefficients, and U_ℓ maps them into the output-side sensitivity subspace. Because the projected coordinates depend on the current activation $h_{\ell,t}$, the resulting correction remains state-dependent even though the sensitivity bases themselves are fixed. We denote the resulting forecast by

$$p_{t,c}(x_t; s) = f_{\theta_0,c}(x_t; s) + d_{t,c}(x_t; s).$$

Persistent residual memory. To track geographically structured errors that persist after flow-dependent adaptation, we maintain a spatial residual memory $b_{t,c}^{(i)}(s)$ for each atmospheric variable c and initialization-time stream $i \in \{00\,\text{UTC}, 12\,\text{UTC}\}$. Separate states allow the two forecast cycles

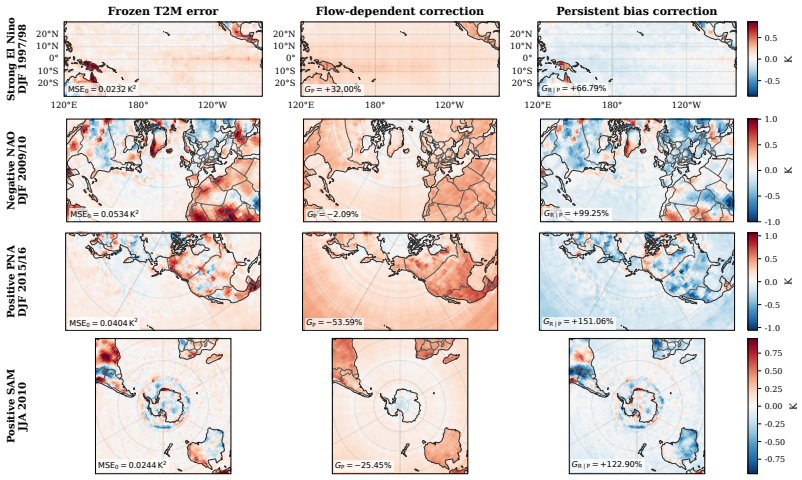


Figure 4: **Decomposition of T2M forecast-error correction across large-scale circulation regimes.** Rows correspond to a strong El Niño (DJF 1997/98), negative NAO (DJF 2009/10), positive PNA (DJF 2015/16), and positive SAM (JJA 2010). The frozen-model MSE and the relative contributions of the flow-dependent (G_{P}) and persistent ($G_{\mathrm{R|P}}$) corrections are reported in each panel.

to capture systematic differences between initialization times. We parameterize the update rate α by a half-life H, $\alpha = 1-2^{-\Delta t/H}$, where Δt is the interval between successive updates within the same stream. Once the corresponding verification becomes available, the residual memory is updated as

$$r_{t,c}(s) = y_{t+1,c}(s) - p_{t,c}(x_t; s), \qquad b^{(i)}_{t+1,c}(s) = (1-\alpha) b^{(i)}_{t,c}(s) + \alpha r_{t,c}(s). \tag{7}$$

Thus, the memory tracks persistent residual errors remaining after the flow-dependent correction rather than biases of the frozen backbone alone.

Variable-wise residual gating. For each atmospheric variable c, we maintain a gating weight $q_{t,c} \in [0,1]$ that controls the usefulness of adding the persistent residual correction on top of the flow-corrected forecast:

$$\widehat{y}_{t+1,c}(s) = f_{\theta_0,c}(x_t; s) + d_{t,c}(x_t; s) + q_{t,c} b^{(i)}_{t,c}(s). \tag{8}$$

Both the residual memory and the gating weight used at round t depend only on previously revealed verification. After the verifying target $y_{t+1}(s)$ becomes available, we update the flow-dependent adapter coefficients using the forecast loss:

$$A_{\ell,t+1} = A_{\ell,t} - \eta \nabla \mathcal{L}(\widehat{y}_{t+1}(s), y_{t+1}(s)), \tag{9}$$

where η denotes the online learning rate. After observing $y_{t+1,c}(s)$, we evaluate two endpoint forecasts: the flow-corrected forecast alone and the same forecast with the full persistent residual correction. Their losses are

$$\begin{aligned} \ell_{t,c,e_1} &= \|p_{t,c}(x_t; s) - y_{t+1,c}(s)\|_2^2, \\ \ell_{t,c,e_2} &= \left\| p_{t,c}(x_t; s) + b^{(i)}_{t,c}(s) - y_{t+1,c}(s) \right\|_2^2. \end{aligned} \tag{10}$$

Because endpoint losses have strongly field-dependent scales, let $\widetilde{\ell}_{t,c,e} \in [0,1]$ denote the scale-normalized endpoint loss. We initialize $q_{1,c} = 1/2$ and update it by exponential weighting:

$$q_{t+1,c} = \frac{q_{t,c} \exp\left(-\eta_q \widetilde{\ell}_{t,c,e_2}\right)}{(1-q_{t,c}) \exp\left(-\eta_q \widetilde{\ell}_{t,c,e_1}\right) + q_{t,c} \exp\left(-\eta_q \widetilde{\ell}_{t,c,e_2}\right)}. \tag{11}$$

Here, $\eta_q > 0$ controls how rapidly the gate responds to differences between the two endpoint losses.

Theorem 3.1 (Regret against the best fixed correction choice)**.** *Fix an atmospheric variable c. Let $\widetilde{\ell}_{t,c,e_1}, \widetilde{\ell}_{t,c,e_2} \in [0,1]$ denote the normalized losses of the flow-corrected forecasts without and with*

Table 1: MSE for eight key output fields at a 6-hour lead time. Bold and underline indicate the lowest and second-lowest values, respectively. Gain denotes the relative MSE reduction of Ours over the second-best method for each field.

Field	δ-Adapter	Proceed	OneNet	FSNet	ADAPT-Z	ELF	Ours	Gain (%)
Z500	<u>676.367</u>	915.514	733.233	762.317	842.627	1409.867	**582.499**	**13.88**
T850	<u>0.12001</u>	0.12290	0.12083	0.12343	0.12196	0.12912	**0.11811**	**1.59**
U850	<u>0.35459</u>	0.35608	0.35605	0.36020	0.35656	0.35667	**0.35384**	**0.21**
V850	<u>0.36539</u>	0.36688	0.36686	0.37118	0.36732	0.36706	**0.36485**	**0.15**
T2M	<u>0.31019</u>	0.32012	0.32107	0.31981	0.32382	0.32694	**0.25162**	**18.88**
U10	<u>0.13960</u>	0.14051	0.14050	0.14195	0.14060	0.14063	**0.13853**	**0.77**
V10	<u>0.14258</u>	0.14367	0.14367	0.14528	0.14374	0.14374	**0.14190**	**0.48**
MSL	<u>1088.852</u>	1166.726	1162.141	1111.857	1194.673	1202.782	**976.540**	**10.31**

the full persistent residual correction, respectively. Initialize $q_{1,c} = 1/2$ and update it by exponential weighting. Then, for any $\eta_q > 0$,

$$\sum_{t=1}^{T} \left[(1 - q_{t,c})\widetilde{\ell}_{t,c,e_1} + q_{t,c}\widetilde{\ell}_{t,c,e_2} \right] - \min_{e \in \{e_1, e_2\}} \sum_{t=1}^{T} \widetilde{\ell}_{t,c,e} \leq \frac{\log 2}{\eta_q} + \frac{\eta_q T}{8}. \tag{12}$$

In particular, choosing $\eta_q = \sqrt{8 \log 2/T}$ gives an upper bound of $\sqrt{T \log 2/2}$.

Thus, the average regret of the endpoint-weighting objective vanishes as $O(T^{-1/2})$ relative to the better fixed endpoint in hindsight. The guarantee compares flow-dependent correction alone with flow-dependent correction plus the full persistent residual correction, rather than an arbitrary fixed mixture. The proof is provided in Appendix A. We next evaluate whether these complementary adaptation mechanisms improve forecasting skill during long-term deployment.

4 Experiments

4.1 Experimental Setup

Dataset and temporal split. All experiments use ERA5 reanalysis spanning 1959–2022 (64 years), regridded to a 1.5° equiangular grid of 121×240 points. The data are sampled at 6-hour intervals, corresponding to 00:00 UTC, 06:00 UTC, 12:00 UTC, and 18:00 UTC each day, while forecast evaluation is performed only for initializations at 00:00 UTC and 12:00 UTC. Following WeatherBench2 (Rasp et al., 2024), we use five pressure-level variables, each defined on 13 pressure levels: geopotential (Z), specific humidity (Q), temperature (T), and the zonal and meridional wind components (U and V), together with four surface variables: 10-m zonal and meridional wind (U10 and V10), 2-m temperature (T2M), and mean sea-level pressure (MSL). We use the first ten years (1959–1968) exclusively as the offline initialization set, while the remaining 54 years (1969–2022) form a strictly chronological online evaluation stream. Dataset details are provided in Appendix B.

Baselines. We compare our method with six recent online adaptation approaches: δ-Adapter (Liang et al., 2026), Proceed (Zhao & Shen, 2025), OneNet (Zhang et al., 2023), FSNet (Pham et al., 2022), ADAPT-Z (Huang et al., 2026), and ELF (Lee et al., 2025). Hyperparameters for all baselines follow the best-performing settings reported in the corresponding papers.

Metrics. Following WeatherBench (Rasp et al., 2024), we evaluate deterministic forecasts using latitude-area-weighted errors over the global grid. For an atmospheric field c, let $\hat{y}_{t,c}(s)$ and $y_{t,c}(s)$ denote the forecast and ERA5 verification at grid point s and valid time t, respectively. The latitude weight is $\bar{w}_s = \frac{\cos \phi_s}{\sum_{s' \in \Omega} \cos \phi_{s'}}$., where ϕ_s is the latitude of grid point s. Two commonly used metrics as defined below:

$$\mathrm{MSE} = \frac{1}{|\mathcal{T}_{\mathrm{test}}|} \sum_{t \in \mathcal{T}_{\mathrm{test}}} \sum_{s \in \Omega} \bar{w}_s \left(\hat{y}_{t,c}(s) - y_{t,c}(s) \right)^2, \mathrm{MAE} = \frac{1}{|\mathcal{T}_{\mathrm{test}}|} \sum_{t \in \mathcal{T}_{\mathrm{test}}} \sum_{s \in \Omega} \bar{w}_s \left| \hat{y}_{t,c}(s) - y_{t,c}(s) \right| \tag{13}$$

Implementation Details. Our default configuration uses an online learning rate of 2×10^{-3}, a sensitivity-subspace rank of 128, and $\eta_q = 1$. All experiments are conducted on an 8×NVIDIA

Table 2: MAE for eight key output fields at a 6-hour lead time. Bold and underline indicate the lowest and second-lowest values, respectively. Gain denotes the relative MAE reduction of Ours over the second-best method for each field.

Field	δ-Adapter	Proceed	OneNet	FSNet	ADAPT-Z	ELF	Ours	Gain (%)
Z500	<u>20.381</u>	23.656	21.208	21.561	22.565	29.298	**18.803**	**7.74**
T850	<u>0.24135</u>	0.24529	0.24178	0.24544	0.24353	0.25464	**0.23936**	**0.83**
U850	<u>0.43015</u>	0.43118	0.43114	0.43402	0.43155	0.43166	**0.42938**	**0.18**
V850	<u>0.43527</u>	0.43626	0.43624	0.43905	0.43662	0.43635	**0.43454**	**0.17**
T2M	<u>0.33130</u>	0.33356	0.33415	0.33422	0.33541	0.33800	**0.30100**	**9.15**
U10	<u>0.26774</u>	0.26856	0.26856	0.27020	0.26869	0.26873	**0.26647**	**0.47**
V10	<u>0.27248</u>	0.27349	0.27351	0.27521	0.27360	0.27358	**0.27147**	**0.37**
MSL	<u>24.726</u>	25.687	25.620	24.971	25.992	26.136	**23.308**	**5.73**

A800 GPU cluster. Offline training of Aurora-1.3B takes approximately 576 GPU-hours, while our online adaptation takes approximately 57 GPU-hours. Hyperparameter selection and sensitivity analyses are provided in Appendix C.3 and C.4.

4.2 Online Forecast Performance

Overall comparison. We evaluate forecasts at lead times of 6, 12, 24, 48, and 72 hours. Tables 1 and 2 report the 6-hour forecast performance for eight key atmospheric variables, measured by mean squared error (MSE) and mean absolute error (MAE). Averaging the variable-wise relative reductions over these eight variables, our method reduces MSE and MAE by 5.78% and 3.08%, respectively, compared with the strongest baseline, δ-Adapter. Averaged over all five lead times, our method outperforms δ-Adapter by 4.17% in MSE and 2.76% in MAE across the 69 variables. Among the eight key variables, the largest 6-hour MSE gains over the second-best method occur for T2M, Z500, and MSL, with reductions of 18.88%, 13.88%, and 10.31%, respectively. Improvements for the wind fields are smaller but remain consistent, indicating that the proposed adaptation is particularly effective for thermodynamic and mass-related variables while retaining gains across the representative fields. Detailed results for all 69 variables at lead times of 6, 12, 24, 48, and 72 hours are provided in Appendix D.1.

Temporal drift analysis. Figure 5 examines how T2M forecast performance evolves as the deployment data depart from the 1959–1968 reference period. Across successive evaluation periods, Frozen Aurora's MSE rises from approximately 0.30 to $0.37\,\mathrm{K}^2$, whereas Ours remains comparatively stable near $0.25\,\mathrm{K}^2$. MAE exhibits a similar pattern, resulting in a widening performance gap between the frozen and adapted models. To quantify distributional change independently of forecast performance, we compute the annual Jensen–Shannon distance between ERA5 T2M distributions and the historical reference. Larger distances are associated with higher frozen-model MSE (Pearson $r = 0.84$) and greater relative MSE reduction from online adaptation ($r = 0.79$). Thus, the benefit of adaptation is most pronounced in years whose T2M distribution differs more strongly from the training-period climate. Appendix D.4 extends the analysis to seven additional atmospheric variables and shows that the strength of this relationship varies substantially across variables.

Gating weight evolution. Appendix D.2 further reveals clear variable-dependent gating behavior. The persistent-bias endpoint receives dominant weight ($q_t > 0.5$) for Z500, T2M, and MSL, with T2M gradually approaching $q_t = 1$, whereas the gate suppresses persistent correction for the wind fields ($q_t < 0.5$). T850 remains close to an equal mixture. The similar 00:00 and 12:00 UTC trajectories indicate that these preferences are robust across initialization times, supporting the need for variable-specific adaptive gating.

4.3 Ablation and Sensitivity Analysis

Ablation of the two correction components. Figure D.12 in Appendix D.5 reveals clear variable-dependent roles of the two corrections. Persistent bias correction substantially improves Z500, T2M, and MSL, with the strongest effect for T2M, whereas flow-dependent correction is more effective for T850 and the wind fields. The case studies further demonstrate that this balance varies with the atmospheric state. For Z500 during Typhoons Haiyan and Hagibis (Fig. D.4 in Appendix D.3),

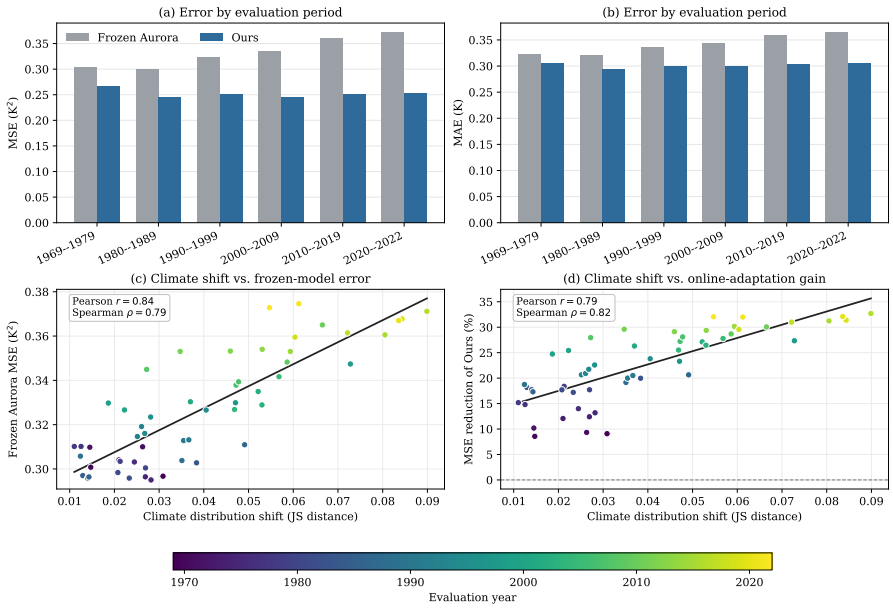


Figure 5: T2M forecast error under climate distribution shift. Panels (a,b) compare Frozen Aurora and Ours across evaluation periods. Panels (c,d) show annual climate shift versus frozen-model MSE and the MSE reduction of Ours. Climate shift is measured relative to the 1959–1968 distribution using Jensen–Shannon distance.

flow-dependent correction contributes 65.71% and 30.50%, respectively, while persistent correction provides limited or negative additional benefit. In contrast, persistent correction consistently contributes positively to T2M across four large-scale climate regimes (Fig. 4), and compensates for unfavorable flow-dependent corrections in several regimes. These results demonstrate that the relative importance of persistent and flow-dependent errors varies across both variables and atmospheric states, motivating their adaptive combination.

Effect of residual-memory timescale. Table C.4 in Appendix C.4 examines the sensitivity of our method to the residual-memory half-life. The 7-day configuration provides the most robust overall performance, achieving the lowest MSE on 29 fields and the lowest MAE on 27 fields. Shortening the half-life to one day makes the residual estimate overly sensitive to transient forecast errors, whereas longer half-lives delay its response to evolving biases. Accordingly, the 1, 30, and 90 day configurations all exhibit lower macro-averaged skill than the 7-day setting, supporting an intermediate residual-memory timescale.

Effect of online gating. Figure D.12 in Appendix D.5 shows that the online gate consistently improves over both fixed correction strategies by adapting the contribution of persistent bias correction over time. For Z500, T2M, and MSL, it largely retains the beneficial persistent component, whereas for T850 and the wind fields it suppresses this component when it becomes detrimental. The gated prediction therefore achieves lower cumulative regret across all variables, indicating that the optimal balance between persistent and flow-dependent corrections is both variable- and time-dependent.

5 Conclusion

We introduced an online adaptation framework for AI weather foundation models under evolving climate distributions. It combines low-dimensional sensitivity updates for flow-dependent errors, residual memory for persistent biases, and a variable-wise gate. On a 54-year chronological Aurora-1.3B evaluation, it consistently outperforms recent online adaptation baselines, with larger climate shifts associated with higher frozen-model errors and greater adaptation gains. These results demonstrate the value of weather-specific error structure for sustaining forecast skill under distribution shift.

AI USE STATEMENT

Generative AI tools were used to assist with methodological discussion, code development and debugging, literature search. They were also used to improve the clarity, grammar, and organization of the manuscript. All AI-assisted methodological suggestions were reviewed and validated by the authors, all generated or modified code was tested by the authors, and cited literature was checked against the original sources. Generative AI tools were not used to generate experimental data or replace the authors' scientific judgment. The authors take full responsibility for the final content of this work, including all text, claims, results, and artifacts produced with the aid of generative AI.

REFERENCES

JD Annan. On the orthogonality of bred vectors. *Monthly weather review*, 132(3):843–849, 2004.

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.

Cristian Bodnar, Wessel P Bruinsma, Ana Lucic, Megan Stanley, Anna Allen, Johannes Brandstetter, Patrick Garvan, Maik Riechert, Jonathan A Weyn, Haiyu Dong, et al. A foundation model for the earth system. *Nature*, 641(8065):1180–1187, 2025.

Lei Chen, Xiaohui Zhong, Feng Zhang, Yuan Cheng, Yinghui Xu, Yuan Qi, and Hao Li. Fuxi: A cascade machine learning forecasting system for 15-day global weather forecast. *npj climate and atmospheric science*, 6(1):190, 2023.

Laura Ferranti, Susanna Corti, and Martin Janousek. Flow-dependent verification of the ecmwf ensemble over the euro-atlantic sector. *Quarterly Journal of the Royal Meteorological Society*, 141(688):916–924, 2015.

João Gama, Indrė Žliobaitė, Albert Bifet, Mykola Pechenizkiy, and Abdelhamid Bouchachia. A survey on concept drift adaptation. *ACM computing surveys (CSUR)*, 46(4):1–37, 2014.

R Gelaro, R Buizza, TN Palmer, and E Klinker. Sensitivity analysis of forecast errors and the construction of optimal perturbations using singular vectors. *Journal of the Atmospheric Sciences*, 55(6):1012–1037, 1998.

Thomas Haiden, Irina Sandu, Gianpaolo Balsamo, Gabriele Arduini, and Anton Beljaars. Addressing biases in near-surface forecasts. *ECMWF Newsletter*, 157:20–25, 2018.

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.

Roger A Horn and Charles R Johnson. *Matrix analysis*. Cambridge university press, 2012.

Xiannan Huang, Shuhan Qiu, Jiayuan Du, and Chao Yang. Online time series prediction using feature adjustment. In *The Fourteenth International Conference on Learning Representations*, 2026. URL https://openreview.net/forum?id=s4U2FWEMTU.

Thorsten Kurth, Shashank Subramanian, Peter Harrington, Jaideep Pathak, Morteza Mardani, David Hall, Andrea Miele, Karthik Kashinath, and Anima Anandkumar. Fourcastnet: Accelerating global high-resolution weather forecasting using adaptive fourier neural operators. In *Proceedings of the platform for advanced scientific computing conference*, pp. 1–11, 2023.

Remi Lam, Alvaro Sanchez-Gonzalez, Matthew Willson, Peter Wirnsberger, Meire Fortunato, Ferran Alet, Suman Ravuri, Timo Ewalds, Zach Eaton-Rosen, Weihua Hu, et al. Learning skillful medium-range global weather forecasting. *Science*, 382(6677):1416–1421, 2023.

Simon Lang, Mihai Alexe, Matthew Chantry, Jesper Dramsch, Florian Pinault, Baudouin Raoult, Mariana CA Clare, Christian Lessig, Michael Maier-Gerber, Linus Magnusson, et al. Aifs–ecmwf's data-driven forecasting system. *arXiv preprint arXiv:2406.01465*, 2024.

Thomas L. Lee, William Toner, Rajkarn Singh, Artjom Joosen, and Martin Asenov. Lightweight online adaption for time series foundation model forecasts. In *Proceedings of the 42nd International Conference on Machine Learning*, volume 267 of *Proceedings of Machine Learning Research*, pp. 33736–33764. PMLR, 2025. URL https://proceedings.mlr.press/v267/lee25ag.html.

Martin Leutbecher and Tim N Palmer. Ensemble forecasting. *Journal of computational physics*, 227 (7):3515–3539, 2008.

Daojun Liang, Qi Li, Yinglong Wang, Jing Chen, Hu Zhang, Xiaoxiao Cui, Qizheng Wang, and Shuo Li. The forecast after the forecast: A post-processing shift in time series. In *The Fourteenth International Conference on Learning Representations*, 2026. URL https://openreview.net/forum?id=syfWdclGE1.

Renata Libonati, Isabel Trigo, and Carlos C DaCamara. Correction of 2 m-temperature forecasts using kalman filtering technique. *Atmospheric Research*, 87(2):183–197, 2008.

D. J. Patil, Brian R. Hunt, Eugenia Kalnay, James A. Yorke, and Edward Ott. Local low dimensionality of atmospheric dynamics. *Physical Review Letters*, 86(26):5878–5881, 2001. doi: 10.1103/PhysRevLett.86.5878.

Melinda S Peng and Carolyn A Reynolds. Sensitivity of tropical cyclone forecasts as revealed by singular vectors. *Journal of the atmospheric sciences*, 63(10):2508–2528, 2006.

Quang Pham, Chenghao Liu, Doyen Sahoo, and Steven CH Hoi. Learning fast and slow for online time series forecasting. *arXiv preprint arXiv:2202.11672*, 2022.

Ilan Price, Alvaro Sanchez-Gonzalez, Ferran Alet, Tom R Andersson, Andrew El-Kadi, Dominic Masters, Timo Ewalds, Jacklynn Stott, Shakir Mohamed, Peter Battaglia, et al. Probabilistic weather forecasting with machine learning. *Nature*, 637(8044):84–90, 2025.

Stephan Rasp, Stephan Hoyer, Alexander Merose, Ian Langmore, Peter Battaglia, Tyler Russell, Alvaro Sanchez-Gonzalez, Vivian Yang, Rob Carver, Shreya Agrawal, et al. Weatherbench 2: A benchmark for the next generation of data-driven global weather models. *Journal of Advances in Modeling Earth Systems*, 16(6):e2023MS004019, 2024.

Carolyn A Reynolds, Melinda S Peng, and Jan-Huey Chen. Recurving tropical cyclones: Singular vector sensitivity and downstream impacts. *Monthly Weather Review*, 137(4):1320–1337, 2009.

Sebastian Scher and Gabriele Messori. Ensemble methods for neural network-based weather forecasts. *Journal of Advances in Modeling Earth Systems*, 13(2), 2021.

Sebastian Sippel, Nicolai Meinshausen, Erich M Fischer, Enikő Székely, and Reto Knutti. Climate change now detectable from any single day of weather at global scale. *Nature climate change*, 10 (1):35–41, 2020.

Bradley Stanley-Clamp, Anson Lei, Hannah M Christensen, and Ingmar Posner. No epoch like the present: Robust climate emulation requires out-of-distribution generalisation. *arXiv preprint arXiv:2605.22248*, 2026.

Xiaoxu Tian, Daniel Holdaway, and Daryl Kleist. Evaluating machine learning weather models for data assimilation: Fundamental limitations in tangent linear and adjoint properties. *Geophysical Research Letters*, 53(2):e2025GL119402, 2026.

Ryan D Torn, Michael J Brennan, and Jason P Dunion. Application of ensemble sensitivity for hurricane track forecast sensitivity and flight planning. *Weather and Forecasting*, 40(3):411–424, 2025.

Anna Trevisan and Francesco Uboldi. Assimilation of standard and targeted observations within the unstable subspace of the observation–analysis–forecast cycle system. *Journal of the atmospheric sciences*, 61(1):103–113, 2004.

Yi-Fan Zhang, Qingsong Wen, Xue Wang, Weiqi Chen, Liang Sun, Zhang Zhang, Liang Wang, Rong Jin, and Tieniu Tan. OneNet: Enhancing time series forecasting models under concept drift by online ensembling. In *Advances in Neural Information Processing Systems*, volume 36, 2023. URL https://papers.neurips.cc/paper_files/paper/2023/hash/dd6a47bc0aad6f34aa5e77706d90cdc4-Abstract-Conference.html.

Lifan Zhao and Yanyan Shen. Proactive model adaptation against concept drift for online time series forecasting. In *Proceedings of the 31st ACM SIGKDD Conference on Knowledge Discovery and Data Mining V. 1*, pp. 2020–2031, 2025.

A PROOF OF THE ONLINE REGRET BOUND

Proof. Fix an atmospheric variable c and suppress the index c throughout the proof. Write $\ell_{t,j} := \widetilde{\ell}_{t,c,e_j} \in [0,1]$ for $j \in \{1,2\}$, and define $\pi_{t,1} = 1 - q_{t,c}, \pi_{t,2} = q_{t,c}$. The weighted loss incurred at round t is

$$m_t := \sum_{j=1}^{2} \pi_{t,j}\ell_{t,j} = (1 - q_{t,c})\ell_{t,1} + q_{t,c}\ell_{t,2}.$$

Let $\eta := \eta_q > 0$. Introduce unnormalized weights initialized as $w_{1,1} = w_{1,2} = 1/2$ and updated according to $w_{t+1,j} = w_{t,j}\exp(-\eta\ell_{t,j}), j \in \{1,2\}$. Define the potential $Z_t := w_{t,1} + w_{t,2}, Z_1 = 1$. Normalizing the weights gives $\pi_{t,j} = w_{t,j}/Z_t$. In particular,

$$\pi_{t+1,2} = \frac{\pi_{t,2}\exp(-\eta\ell_{t,2})}{\pi_{t,1}\exp(-\eta\ell_{t,1}) + \pi_{t,2}\exp(-\eta\ell_{t,2})},$$

which is precisely the stated exponential-weighting update for $q_{t+1,c}$. For the cumulative endpoint loss

$$L_{t,j} := \sum_{s=1}^{t} \ell_{s,j}, \qquad L_{0,j} := 0,$$

the weight recursion implies

$$w_{t+1,j} = \frac{1}{2}\exp(-\eta L_{t,j}).$$

At each round,

$$\frac{Z_{t+1}}{Z_t} = \sum_{j=1}^{2} \pi_{t,j}\exp(-\eta\ell_{t,j}).$$

To bound this quantity, condition on the realized history and the current endpoint losses, and introduce an auxiliary random variable X_t satisfying $\Pr(X_t = \ell_{t,j}) = \pi_{t,j}$. Then $X_t \in [0,1]$ and $\mathbb{E}[X_t] = m_t$. Hoeffding's lemma states that, for any random variable $X \in [a,b]$ and any $\lambda \in \mathbb{R}$,

$$\log \mathbb{E}\left[\exp\big(\lambda(X - \mathbb{E}[X])\big)\right] \le \frac{\lambda^2(b-a)^2}{8}.$$

Applying this inequality with $\lambda = -\eta$, $a = 0$, and $b = 1$ yields

$$\log\frac{Z_{t+1}}{Z_t} = \log \mathbb{E}\left[\exp(-\eta X_t)\right] = -\eta m_t + \log \mathbb{E}\left[\exp\big(-\eta(X_t - m_t)\big)\right] \le -\eta m_t + \frac{\eta^2}{8}.$$

Summing over $t = 1, \ldots, T$ telescopes the log-potential ratios:

$$\log\frac{Z_{T+1}}{Z_1} \le -\eta\sum_{t=1}^{T} m_t + \frac{\eta^2 T}{8}.$$

Since $Z_1 = 1$, we obtain

$$\log Z_{T+1} \le -\eta\sum_{t=1}^{T} m_t + \frac{\eta^2 T}{8}. \tag{A.1}$$

Let $j^\star \in \arg\min_{j\in\{1,2\}} L_{T,j}$. The potential includes the weight of this endpoint, so

$$Z_{T+1} = \frac{1}{2}\sum_{j=1}^{2}\exp(-\eta L_{T,j}) \ge \frac{1}{2}\exp(-\eta L_{T,j^\star}).$$

Taking logarithms gives

$$\log Z_{T+1} \ge -\log 2 - \eta \min_{j\in\{1,2\}} L_{T,j}. \tag{A.2}$$

Combining Eqs. equation A.1 and equation A.2 gives

$$-\log 2 - \eta \min_{j\in\{1,2\}} L_{T,j} \le -\eta\sum_{t=1}^{T} m_t + \frac{\eta^2 T}{8}.$$

Rearranging and dividing by $\eta > 0$ yields

$$\sum_{t=1}^{T} m_t - \min_{j\in\{1,2\}} L_{T,j} \le \frac{\log 2}{\eta} + \frac{\eta T}{8}.$$

Restoring the original notation proves

$$\sum_{t=1}^{T} \left[(1-q_{t,c})\widetilde{\ell}_{t,c,e_1} + q_{t,c}\widetilde{\ell}_{t,c,e_2}\right] - \min_{e\in\{e_1,e_2\}} \sum_{t=1}^{T} \widetilde{\ell}_{t,c,e} \le \frac{\log 2}{\eta_q} + \frac{\eta_q T}{8}.$$

The upper bound

$$B(\eta) = \frac{\log 2}{\eta} + \frac{\eta T}{8}$$

is minimized over $\eta > 0$ when

$$B'(\eta) = -\frac{\log 2}{\eta^2} + \frac{T}{8} = 0,$$

which gives

$$\eta^\star = \sqrt{\frac{8\log 2}{T}}.$$

Substituting this value yields

$$B(\eta^\star) = \sqrt{\frac{T\log 2}{8}} + \sqrt{\frac{T\log 2}{8}} = \sqrt{\frac{T\log 2}{2}}.$$

This completes the proof. □

B DATASET

In this section, we are going to introduce the dataset we used in this study detailedly. For the global forecasting, we conduct experiments on the WeatherBench2 (Rasp et al., 2024) benchmark, a subset of ERA5 reanalysis data. The WeatherBench2 benchmark we used is the version of 1.5°resolution (121 × 240), which spans from 1959 to 2022. This subset contains 5 variables (Z, Q, T, U, V) with 13 pressure levels (50 hPa, 100 hPa, 150 hPa, 200 hPa, 250 hPa, 300 hPa, 400 hPa, 500 hPa, 600 hPa, 700 hPa, 850 hPa, 925 hPa and 1,000 hPa) and 4 variables (U10M, V10M, T2M, MSLP) with surface level. All the data we used are shown in Table B.1.

C EXPERIMENTS DETAILS

C.1 OFFLINE INITIALIZATION

Forecasting Model. We use Aurora-1.3B as the forecasting backbone. The model takes two consecutive atmospheric states as input and predicts the next 6-h state. It operates on ERA5 fields regridded to a 1.5° grid, using five upper-air variables (Z, Q, T, U, V) on 13 pressure levels together with four surface variables (T2M, U10, V10, MSL). The backbone uses an embedding dimension of 512, a patch size of 2, encoder depths $(6, 10, 8)$ and decoder depths $(8, 10, 6)$, with attention heads $(8, 16, 32)$ across the encoder stages. Local attention uses a three-dimensional window of $(2, 6, 12)$. The Perceiver components use 16 attention heads and four latent levels. Both attention dropout and drop-path are set to zero. The first ten years of ERA5 (1959–1968) are used exclusively for offline initialization, while forecasts are initialized at 00 and 12 UTC and verified at the corresponding +6-h valid times.

Training Objective and Hyperparameters. The forecasting model is trained using a latitude-weighted multivariate ℓ_1 objective over all upper-air and surface variables. The upper-air variables (Z, Q, T, U, V) are assigned weights $(3.00, 0.60, 1.50, 0.77, 0.54)$, while the surface variables (T2M, U10, V10, MSL) use $(3.00, 0.77, 0.66, 1.50)$, respectively. The aggregated upper-air and surface losses are further weighted by 1.0 and 0.25.

Table B.1: ECMWF variables we used in datasets. The "Type" column indicates whether the variable represents a static property, a time-varying single-level property (e.g., surface variables are included), or a time-varying atmospheric property. The "Long name" and "Short name" columns are ECMWF's labels. The "level" column represents whether the variable is a surface variable or a upper variable, and indicates how many pressure levels were selected for the upper variables. The "Role" column indicates whether the variable is something Aurora takes as input and predicts, or only uses as input context.

Type	Long name	Short name	Levels	Role
Atmospheric	Geopotential	Z	13 levels	Input/Predicted
Atmospheric	Temperature	T	13 levels	Input/Predicted
Atmospheric	Specific humidity	Q	13 levels	Input/Predicted
Atmospheric	U component of wind	U	13 levels	Input/Predicted
Atmospheric	V component of wind	V	13 levels	Input/Predicted
Single	2m temperature	T2M	Singe level	Input/Predicted
Single	10m u component of wind	U10	Singe level	Input/Predicted
Single	10m v component of wind	V10	Singe level	Input/Predicted
Single	Mean sea level pressure	MSL	Singe level	Input/Predicted
Static	Land binary mask	LSM	Single level	Input
Static	Soil type	SLT	Single level	Input
Static	Orography	OR	Single level	Input

We employ a hybrid Muon–AdamW optimizer: Muon is applied to trainable two-dimensional linear weight matrices, while AdamW is used for the remaining parameters. Both optimizers use an initial learning rate of 3×10^{-4} and weight decay of 5×10^{-6}. Muon uses momentum 0.95 with Nesterov acceleration and five Newton–Schulz iterations, while AdamW uses $\beta = (0.9, 0.95)$ and $\epsilon = 10^{-8}$. The learning rate follows a cosine decay schedule after an optional warm-up stage. Training is performed with distributed data parallelism on eight GPUs, with a per-GPU batch size of 1.

C.2 Forecast-Sensitivity Subspace Estimation

Gradient Collection. We estimate forecast-sensitivity directions from the offline period 1959–1968 using forecasts initialized at 00 and 12 UTC with a 6-h lead time. The pretrained Aurora backbone is kept fixed, and gradients are enabled only temporarily for the two-dimensional weights of all linear layers in the final two decoder stages and the surface and atmospheric output heads. For each selected layer ℓ and offline sample i, we compute

$$G_{\ell,i} = \nabla_{W_\ell} \mathcal{L}_{\mathrm{off}}\big(f_{\theta_0}(x_i), y_i\big), \tag{C.1}$$

where $\mathcal{L}_{\mathrm{off}}$ is the same normalized, latitude-weighted multivariate ℓ_1 loss used for forecasting. Gradients are computed independently for each sample with batch size one and are used only to estimate the sensitivity statistics; no model parameter is updated during this stage. By default, every eighth offline sample is retained for gradient estimation.

Subspace Estimation and Hyperparameters. For each selected weight matrix $W_\ell \in \mathbb{R}^{d_\ell^{\mathrm{out}} \times d_\ell^{\mathrm{in}}}$, the per-sample gradient $G_{\ell,i}$ has the same shape as W_ℓ. We characterize its output- and input-side forecast-loss sensitivities using the gradient second-moment matrices

$$C_\ell^{\mathrm{out}} = \sum_i G_{\ell,i} G_{\ell,i}^\top, \qquad C_\ell^{\mathrm{in}} = \sum_i G_{\ell,i}^\top G_{\ell,i}. \tag{C.2}$$

Here, $C_\ell^{\mathrm{out}} \in \mathbb{R}^{d_\ell^{\mathrm{out}} \times d_\ell^{\mathrm{out}}}$ and $C_\ell^{\mathrm{in}} \in \mathbb{R}^{d_\ell^{\mathrm{in}} \times d_\ell^{\mathrm{in}}}$ describe sensitivity along the output and input dimensions of the layer, respectively.

Because these matrices are too large to construct explicitly, we estimate their dominant eigenspaces using a one-pass randomized Nyström sketch. For each side $q \in \{\mathrm{out}, \mathrm{in}\}$, we draw a Gaussian probe matrix $\Omega_\ell^q \in \mathbb{R}^{d_\ell^q \times k}$ and orthonormalize its columns, where $k = r_{\max} + p$. We then form

$$Y_\ell^q = C_\ell^q \Omega_\ell^q. \tag{C.3}$$

Importantly, the sketches can be accumulated directly as each per-sample gradient is observed:

$$Y_\ell^{\text{out}} = \sum_i G_{\ell,i} \left(G_{\ell,i}^\top \Omega_\ell^{\text{out}} \right), \qquad Y_\ell^{\text{in}} = \sum_i G_{\ell,i}^\top \left(G_{\ell,i} \Omega_\ell^{\text{in}} \right). \tag{C.4}$$

Thus, neither gradient second-moment matrix needs to be materialized or stored.

After the one-pass accumulation, we construct the small core matrix

$$B_\ell^q = (\Omega_\ell^q)^\top Y_\ell^q = (\Omega_\ell^q)^\top C_\ell^q \Omega_\ell^q. \tag{C.5}$$

For numerical stability, we symmetrize the core matrix as $B_\ell^q \leftarrow \frac{1}{2}\left(B_\ell^q + (B_\ell^q)^\top\right)$ before computing its eigendecomposition

$$B_\ell^q = S_\ell^q \Lambda_\ell^q (S_\ell^q)^\top. \tag{C.6}$$

Eigenvalues below the relative pseudoinverse tolerance $\tau_{\text{pinv}} \lambda_{\max}(B_\ell^q)$ are discarded. Let $S_{\ell,+}^q$ and $\Lambda_{\ell,+}^q$ denote the retained eigenvectors and positive eigenvalues. The Moore–Penrose pseudoinverse is then computed as

$$(B_\ell^q)^\dagger = S_{\ell,+}^q (\Lambda_{\ell,+}^q)^{-1} (S_{\ell,+}^q)^\top. \tag{C.7}$$

The corresponding Nyström approximation is

$$\widehat{C}_\ell^q = Y_\ell^q (B_\ell^q)^\dagger (Y_\ell^q)^\top. \tag{C.8}$$

To recover its dominant eigendirections without explicitly constructing $\widehat{C}_\ell^q$, we form

$$F_\ell^q = Y_\ell^q S_{\ell,+}^q (\Lambda_{\ell,+}^q)^{-1/2}, \tag{C.9}$$

which satisfies $\widehat{C}_\ell^q = F_\ell^q (F_\ell^q)^\top$. A compact singular-value decomposition

$$F_\ell^q = \widetilde{U}_\ell^q \Sigma_\ell^q (Q_\ell^q)^\top \tag{C.10}$$

therefore gives

$$\widehat{C}_\ell^q = \widetilde{U}_\ell^q (\Sigma_\ell^q)^2 (\widetilde{U}_\ell^q)^\top. \tag{C.11}$$

The leading columns of $\widetilde{U}_\ell^q$ approximate the dominant eigenvectors of C_ℓ^q. We consequently define the output- and input-side sensitivity bases as

$$U_\ell = \widetilde{U}_\ell^{\text{out}}[:, 1:r], \qquad V_\ell = \widetilde{U}_\ell^{\text{in}}[:, 1:r]. \tag{C.12}$$

We estimate a single maximum-rank basis with $r_{\max} = 128$ and use $r \in \{4, 8, 16, 32, 64, 128\}$ as nested prefixes of this basis. The Nyström oversampling dimension is $p = 16$, giving a sketch size of $k = 144$, and the random seed is fixed to 2026. Unless otherwise specified, all experiments use $r = 128$. Once estimated from the offline period, U_ℓ and V_ℓ remain fixed throughout online adaptation; only the projected adapter coefficients are updated.

C.3 ONLINE LEARNING

Hyperparameter During online adaptation, all pretrained Aurora parameters remain frozen and only the low-dimensional projected coefficients are updated. We use SGD with a learning rate of 2×10^{-3}, zero momentum, and no weight decay. The forecast-sensitivity adapters are applied to the final two decoder stages and output heads with rank $r = 128$. The adaptive gate uses $\eta_q = 1.0$, a 7-day timescale for its running states, and a maximum normalized loss of 1. Online adaptation proceeds sequentially with batch size one over 1969–2022, with separate residual and routing states for 00 and 12 UTC. All updates are performed only after the corresponding 6-h verification becomes available, ensuring a strictly causal evaluation protocol.

Table C.1: Sensitivity to the forecast-sensitivity subspace rank, with $\eta_A = 2 \times 10^{-3}$ and $\eta_q = 1$ fixed. Changes are relative to the default rank of 128.

Rank	Average MSE	ΔMSE (%)	Average MAE	ΔMAE (%)
16	205.2233	+1.07	5.03718	+0.52
32	204.5598	+0.75	5.02971	+0.37
64	203.8227	+0.38	5.02055	+0.19
128 (default)	**203.0428**	–	**5.01111**	–

Table C.2: Sensitivity to the online adaptation learning rate η, with rank 128 and $\eta_q = 1$ fixed. Changes are relative to the default $\eta = 2 \times 10^{-3}$; positive values indicate higher error.

η_A	Average MSE	ΔMSE (%)	Average MAE	ΔMAE (%)
5×10^{-4}	204.4093	+0.67	5.02476	+0.27
$\mathbf{1 \times 10^{-3}}$	**202.6691**	−0.18	**5.00347**	−0.15
2×10^{-3} (default)	203.0428	–	5.01111	–
3×10^{-3}	227.4027	+12.00	5.30524	+5.87
5×10^{-3}	475.6694	+134.27	7.56256	+50.92

C.4 Hyperparameter Sensitivity

Subspace rank. Table C.1 shows a monotonic decrease in average MSE and MAE as the rank increases from 16 to 128. Rank 128 also improves both metrics for all 69 fields relative to rank 16. The modest aggregate differences indicate limited sensitivity to the precise rank.

Online learning rate. Table C.2 shows that $\eta = 10^{-3}$ achieves the lowest average MSE and MAE, improving on the default 2×10^{-3} by 0.18% and 0.15%, respectively. However, the default performs better on 62 of the 69 fields. Larger learning rates substantially degrade geopotential-height forecasts, despite improving several other fields.

Gate learning rate. Table C.3 shows that the default $\eta_q = 1$ achieves the lowest average MSE and MAE. Both smaller and larger values increase aggregate error, although the changes around the default are relatively small.

Table C.3: Sensitivity to the router learning rate η_q, with rank 128 and $\eta_A = 2 \times 10^{-3}$ fixed. Changes are relative to the default $\eta_q = 1$.

η_q	Average MSE	ΔMSE (%)	Average MAE	ΔMAE (%)
0.25	205.5877	+1.25	5.04191	+0.61
0.5	204.0712	+0.51	5.02413	+0.26
1 (default)	**203.0428**	–	**5.01111**	–
2	203.5988	+0.27	5.01742	+0.13
4	203.9941	+0.47	5.02265	+0.23

C.5 Computational Efficiency Analysis

We further evaluate the computational efficiency of the proposed method and several mainstream online adaptation baselines on the full 1969–2022 evaluation stream. Table C.5 reports the comparison in terms of total wall-clock runtime, GPU memory consumption, and theoretical FP32 computational overhead.

As shown in the table C.5, lightweight adaptation methods including OneNet and ELF achieve the best and second-best resource efficiency. Specifically, OneNet yields the lowest runtime (56.16 h), GPU memory usage (9.43 GiB), and FP32 operations (3.942 EFLOPs), while ELF exhibits nearly identical efficiency with only tiny increments in all metrics. Our method maintains moderate computational cost over the evaluation period. Although it does not attain the minimal resource footprint, it attains a good trade-off between computational efficiency and forecasting accuracy, avoiding the

Table C.4: Sensitivity to the residual-memory half-life. ΔMSE and ΔMAE denote the macro-averaged relative improvement over the 7-day configuration; higher is better. Best-field counts indicate the number of the 69 atmospheric variables on which each configuration achieves the lowest error.

Half-life	ΔMSE (%) ↑	ΔMAE (%) ↑	MSE wins ↑	MAE wins ↑
1 day	−0.900	−0.418	4	9
7 days	**0.000**	**0.000**	**29**	**27**
30 days	<u>−0.621</u>	<u>−0.253</u>	<u>22</u>	<u>21</u>
90 days	−1.494	−0.678	14	12

Table C.5: Computational efficiency over the complete 1969–2022 evaluation stream. Runtime denotes total wall-clock time, GPU memory denotes the observed device-memory usage, and theoretical FP32 operations are estimated by multiplying runtime by the 19.5-TFLOP/s peak FP32 throughput of one NVIDIA A800 GPU. Bold and underline indicate the lowest and second-lowest values.

Method	Runtime (h)	GPU Memory (GiB)	FP32 Ops. (EFLOPs)
δ-Adapter	57.43	12.71	4.032
Proceed	56.61	22.98	3.974
OneNet	**56.16**	**9.43**	**3.942**
FSNet	57.24	15.09	4.018
Adapt-Z	81.08	13.70	5.692
ELF	<u>56.21</u>	<u>9.45</u>	<u>3.946</u>
Ours	57.49	20.09	4.036

notable performance drop observed in ultra-lightweight baselines. By contrast, Adapt-Z consumes substantially more runtime and computation, showing inferior efficiency for long sequential online weather adaptation.

D DETAILED RESULTS

D.1 NUMERICAL RESULTS

Table D.1: MSE for all 69 output fields at a 6-hour lead time. Bold and underline indicate the lowest and second-lowest values, respectively. Gain is the relative MSE reduction of Ours over the best baseline for each field; negative values indicate that Ours has a larger error.

Field	δ-Adapter	Proceed	OneNet	FSNet	ADAPT-Z	ELF	Ours	Gain (%)
Z50	4209.465	5437.505	4607.417	4504.681	4948.431	6066.945	**3585.281**	**14.83**
Z100	1823.899	2982.668	2091.184	2185.517	2518.817	4014.980	**1547.844**	**15.14**
Z150	1377.309	2397.649	1504.826	1583.329	1944.765	3980.225	**1147.644**	**16.67**
Z200	1159.681	1990.212	1231.392	1359.458	1586.408	3593.339	**973.499**	**16.05**
Z250	1038.302	1719.849	1102.080	1215.049	1404.534	3116.652	**881.763**	**15.08**
Z300	953.204	1507.493	1012.962	1131.921	1252.109	2691.851	**815.082**	**14.49**
Z400	796.685	1161.953	859.637	904.398	1027.819	1928.172	**686.996**	**13.77**
Z500	676.367	915.514	733.233	762.317	842.627	1409.867	**582.499**	**13.88**
Z600	620.398	787.330	672.440	687.435	742.707	1101.385	**537.065**	**13.43**
Z700	599.239	706.275	645.591	636.727	684.050	886.478	**522.776**	**12.76**
Z850	610.595	677.551	654.190	631.953	669.119	753.958	**540.112**	**11.54**
Z925	637.072	693.680	682.624	654.907	692.967	744.814	**567.275**	**10.96**
Z1000	697.274	747.520	745.524	716.886	755.092	781.247	**625.111**	**10.35**
Q50	3.447e-16	4.171e-16	3.687e-16	3.824e-16	4.132e-16	4.411e-16	**3.206e-16**	**6.97**
Q100	**6.291e-15**	6.470e-15	6.469e-15	6.415e-15	6.491e-15	6.484e-15	6.399e-15	**-1.72**
Q150	**4.496e-13**	4.568e-13	4.559e-13	4.572e-13	4.577e-13	4.589e-13	4.548e-13	**-1.14**
Q200	**1.599e-11**	1.623e-11	1.619e-11	1.637e-11	1.623e-11	1.624e-11	1.623e-11	**-1.50**
Q250	**1.630e-10**	1.650e-10	1.648e-10	1.663e-10	1.650e-10	1.651e-10	1.652e-10	**-1.36**
Q300	**6.320e-10**	6.376e-10	6.368e-10	6.457e-10	6.373e-10	6.389e-10	6.387e-10	**-1.07**
Q400	**4.508e-09**	4.541e-09	4.537e-09	4.602e-09	4.543e-09	4.548e-09	4.556e-09	**-1.06**
Q500	**1.405e-08**	1.414e-08	1.412e-08	1.435e-08	1.413e-08	1.418e-08	1.416e-08	**-0.76**
Q600	**3.146e-08**	3.159e-08	3.156e-08	3.205e-08	3.161e-08	3.161e-08	3.158e-08	**-0.37**
Q700	**5.689e-08**	5.705e-08	5.701e-08	5.780e-08	5.717e-08	5.708e-08	5.689e-08	**-0.00**
Q850	1.106e-07	1.110e-07	1.109e-07	1.123e-07	1.113e-07	1.111e-07	**1.102e-07**	**0.36**
Q925	9.249e-08	9.295e-08	9.271e-08	9.401e-08	9.302e-08	9.330e-08	**9.078e-08**	**1.85**
Q1000	5.574e-08	5.599e-08	5.573e-08	5.690e-08	5.621e-08	5.643e-08	**5.204e-08**	**6.63**
T50	0.41468	0.45817	0.43139	0.43957	0.45597	0.49767	**0.40596**	**2.10**
T100	0.20679	0.21194	0.20861	0.20947	0.21259	0.22333	**0.20468**	**1.02**
T150	**0.11128**	0.11406	0.11372	0.11368	0.11514	0.11701	0.11154	**-0.23**
T200	**0.09175**	0.09439	0.09413	0.09402	0.09415	0.09615	0.09230	**-0.60**
T250	**0.06869**	0.07118	0.06943	0.07052	0.07013	0.07427	0.06906	**-0.53**
T300	**0.05849**	0.06162	0.05879	0.05975	0.06009	0.06646	0.05863	**-0.24**
T400	0.06154	0.06530	0.06181	0.06300	0.06350	0.07258	**0.06146**	**0.13**
T500	0.05575	0.05895	0.05594	0.05703	0.05720	0.06621	**0.05548**	**0.48**
T600	0.06233	0.06571	0.06251	0.06357	0.06447	0.07369	**0.06141**	**1.47**
T700	0.07358	0.07699	0.07412	0.07660	0.07562	0.08558	**0.07252**	**1.44**
T850	0.12001	0.12290	0.12083	0.12343	0.12196	0.12912	**0.11811**	**1.59**
T925	0.11166	0.11432	0.11241	0.11443	0.11337	0.12008	**0.10688**	**4.28**
T1000	0.12309	0.12604	0.12367	0.12643	0.12489	0.13273	**0.11358**	**7.73**
U50	0.97374	1.00614	0.99920	0.99241	1.01118	1.01938	**0.96532**	**0.86**
U100	**0.87208**	0.88841	0.88598	0.88530	0.88873	0.89204	0.87575	**-0.42**
U150	**1.01394**	1.02766	1.02513	1.03337	1.02689	1.03161	1.02196	**-0.79**
U200	**1.06337**	1.07640	1.07576	1.08766	1.07745	1.08216	1.07330	**-0.93**
U250	**1.05802**	1.06826	1.06766	1.08611	1.06864	1.07239	1.06782	**-0.93**
U300	**0.98975**	0.99619	0.99601	1.01239	0.99701	0.99928	0.99689	**-0.72**
U400	**0.78506**	0.78969	0.78938	0.80343	0.79031	0.79195	0.79076	**-0.73**
U500	**0.52347**	0.52554	0.52544	0.53361	0.52599	0.52677	0.52564	**-0.41**
U600	**0.45040**	0.45221	0.45217	0.45899	0.45264	0.45335	0.45146	**-0.24**
U700	**0.41799**	0.41941	0.41944	0.42375	0.42008	0.42000	0.41818	**-0.05**
U850	0.35459	0.35608	0.35605	0.36020	0.35656	0.35667	**0.35384**	**0.21**
U925	0.28734	0.28875	0.28869	0.29255	0.28942	0.28917	**0.28612**	**0.43**
U1000	0.17799	0.17908	0.17898	0.18142	0.17959	0.17930	**0.17678**	**0.68**
V50	**1.03290**	1.07863	1.06644	1.06782	1.07911	1.08113	1.03600	**-0.30**
V100	**0.85322**	0.87012	0.86834	0.86835	0.87096	0.87112	0.85983	**-0.78**
V150	**0.99631**	1.01024	1.00917	1.01417	1.01096	1.01197	1.00611	**-0.98**
V200	**1.06793**	1.08005	1.07986	1.09256	1.08152	1.08115	1.07869	**-1.01**
V250	**1.09066**	1.10005	1.10000	1.11665	1.10194	1.10063	1.10114	**-0.96**
V300	**1.04446**	1.05028	1.05002	1.06984	1.05210	1.05049	1.05296	**-0.81**
V400	**0.83510**	0.83910	0.83919	0.85310	0.84017	0.83967	0.84133	**-0.75**
V500	**0.54855**	0.55039	0.55030	0.55930	0.55097	0.55069	0.55111	**-0.47**
V600	**0.46586**	0.46749	0.46745	0.47459	0.46799	0.46767	0.46724	**-0.30**
V700	**0.43149**	0.43297	0.43295	0.43844	0.43346	0.43311	0.43189	**-0.09**
V850	0.36539	0.36688	0.36686	0.37118	0.36732	0.36706	**0.36485**	**0.15**
V925	0.30288	0.30460	0.30454	0.30881	0.30507	0.30479	**0.30178**	**0.36**
V1000	0.18392	0.18523	0.18515	0.18856	0.18575	0.18533	**0.18301**	**0.50**
T2M	0.31019	0.32012	0.32107	0.31981	0.32382	0.32694	**0.25162**	**18.88**
U10	0.13960	0.14051	0.14050	0.14195	0.14060	0.14063	**0.13853**	**0.77**
V10	0.14258	0.14367	0.14367	0.14528	0.14374	0.14374	**0.14190**	**0.48**
MSL	1088.852	1166.726	1162.141	1111.857	1194.673	1202.782	**976.540**	**10.31**

Table D.2: MAE for all 69 output fields at a 6-hour lead time. Bold and underline indicate the lowest and second-lowest values, respectively. Gain is the relative MAE reduction of Ours over the best baseline for each field; negative values indicate that Ours has a larger error.

Field	δ-Adapter	Proceed	OneNet	FSNet	ADAPT-Z	ELF	Ours	Gain (%)
Z50	50.485	57.351	52.790	52.133	54.616	60.939	**46.383**	**8.12**
Z100	33.516	42.085	35.773	36.259	38.567	48.570	**30.720**	**8.34**
Z150	29.126	37.653	30.315	30.900	33.752	48.033	**26.405**	**9.34**
Z200	26.711	34.440	27.462	28.618	30.560	45.727	**24.292**	**9.06**
Z250	25.252	32.101	25.966	27.066	28.789	42.761	**23.096**	**8.54**
Z300	24.179	30.128	24.873	26.134	27.246	39.958	**22.186**	**8.24**
Z400	22.092	26.565	22.920	23.409	24.789	34.087	**20.370**	**7.79**
Z500	20.381	23.656	21.208	21.561	22.565	29.298	**18.803**	**7.74**
Z600	19.511	21.977	20.317	20.516	21.277	26.017	**18.062**	**7.42**
Z700	19.105	20.783	19.854	19.696	20.420	23.355	**17.775**	**6.96**
Z850	19.084	20.165	19.797	19.424	20.031	21.350	**17.897**	**6.22**
Z925	19.354	20.259	20.091	19.627	20.253	21.058	**18.211**	**5.91**
Z1000	20.056	20.831	20.805	20.340	20.959	21.343	**18.932**	**5.61**
Q50	1.127e-08	1.274e-08	1.134e-08	1.198e-08	1.270e-08	1.308e-08	**1.051e-08**	**6.73**
Q100	**4.235e-08**	4.269e-08	4.274e-08	4.284e-08	4.288e-08	4.302e-08	4.250e-08	**-0.37**
Q150	3.592e-07	3.595e-07	**3.590e-07**	3.603e-07	3.603e-07	3.610e-07	3.593e-07	**-0.07**
Q200	2.104e-06	2.102e-06	**2.100e-06**	2.114e-06	2.103e-06	2.105e-06	2.107e-06	**-0.29**
Q250	6.759e-06	6.745e-06	**6.742e-06**	6.777e-06	6.746e-06	6.751e-06	6.769e-06	**-0.40**
Q300	1.382e-05	1.380e-05	**1.379e-05**	1.390e-05	1.380e-05	1.382e-05	1.386e-05	**-0.51**
Q400	3.898e-05	3.883e-05	**3.882e-05**	3.914e-05	3.885e-05	3.886e-05	3.902e-05	**-0.51**
Q500	6.921e-05	6.908e-05	**6.903e-05**	6.974e-05	6.907e-05	6.920e-05	6.938e-05	**-0.51**
Q600	1.100e-04	1.098e-04	**1.098e-04**	1.110e-04	1.099e-04	1.099e-04	1.101e-04	**-0.31**
Q700	1.541e-04	1.539e-04	**1.538e-04**	1.553e-04	1.541e-04	1.540e-04	1.540e-04	**-0.09**
Q850	2.190e-04	2.191e-04	2.190e-04	2.208e-04	2.194e-04	2.194e-04	**2.183e-04**	**0.30**
Q925	1.971e-04	1.972e-04	1.968e-04	1.988e-04	1.971e-04	1.981e-04	**1.951e-04**	**0.87**
Q1000	1.556e-04	1.558e-04	1.552e-04	1.574e-04	1.565e-04	1.568e-04	**1.517e-04**	**2.30**
T50	0.48224	0.50434	0.49214	0.49499	0.50412	0.52609	**0.47634**	**1.22**
T100	0.33812	0.34246	0.33953	0.34063	0.34346	0.35294	**0.33590**	**0.66**
T150	**0.24448**	0.24805	0.24725	0.24751	0.24946	0.25167	0.24458	**-0.04**
T200	**0.21729**	0.22134	0.22045	0.22025	0.22070	0.22421	0.21772	**-0.20**
T250	**0.19002**	0.19464	0.19100	0.19282	0.19241	0.20025	0.19034	**-0.17**
T300	0.17753	0.18362	0.17790	0.17935	0.18043	0.19233	**0.17751**	**0.01**
T400	0.18571	0.19212	0.18608	0.18791	0.18888	0.20366	**0.18539**	**0.17**
T500	0.17616	0.18194	0.17646	0.17823	0.17867	0.19407	**0.17563**	**0.30**
T600	0.18338	0.18928	0.18358	0.18506	0.18680	0.20241	**0.18235**	**0.56**
T700	0.19677	0.20266	0.19720	0.20168	0.19989	0.21713	**0.19541**	**0.69**
T850	0.24135	0.24529	0.24178	0.24544	0.24353	0.25464	**0.23936**	**0.83**
T925	0.22529	0.22904	0.22553	0.22815	0.22714	0.23880	**0.22078**	**2.00**
T1000	0.23207	0.23642	0.23210	0.23506	0.23406	0.24798	**0.22306**	**3.88**
U50	0.75451	0.76711	0.76526	0.76229	0.76960	0.77203	**0.75047**	**0.53**
U100	**0.69206**	0.69861	0.69785	0.69798	0.69894	0.70013	0.69299	**-0.13**
U150	**0.73231**	0.73765	0.73669	0.74095	0.73738	0.73919	0.73503	**-0.37**
U200	**0.75194**	0.75688	0.75674	0.76167	0.75735	0.75915	0.75543	**-0.46**
U250	**0.74709**	0.75119	0.75096	0.75845	0.75134	0.75289	0.75052	**-0.46**
U300	**0.72253**	0.72523	0.72520	0.73139	0.72556	0.72675	0.72501	**-0.34**
U400	**0.64095**	0.64315	0.64297	0.64946	0.64338	0.64445	0.64317	**-0.35**
U500	**0.51744**	0.51860	0.51853	0.52312	0.51879	0.51950	0.51838	**-0.18**
U600	**0.48040**	0.48152	0.48149	0.48575	0.48174	0.48240	0.48084	**-0.09**
U700	0.46444	0.46536	0.46537	0.46820	0.46580	0.46581	**0.46436**	**0.02**
U850	0.43015	0.43118	0.43114	0.43402	0.43155	0.43166	**0.42938**	**0.18**
U925	0.38204	0.38302	0.38296	0.38592	0.38359	0.38339	**0.38083**	**0.32**
U1000	0.30279	0.30373	0.30365	0.30617	0.30427	0.30401	**0.30137**	**0.47**
V50	**0.78328**	0.79991	0.79621	0.79621	0.80014	0.80080	0.78346	**-0.02**
V100	**0.69246**	0.69937	0.69871	0.69937	0.69975	0.69983	0.69465	**-0.32**
V150	**0.73223**	0.73730	0.73700	0.73971	0.73768	0.73798	0.73558	**-0.46**
V200	**0.75954**	0.76411	0.76408	0.76956	0.76478	0.76447	0.76338	**-0.51**
V250	**0.76522**	0.76882	0.76874	0.77516	0.76963	0.76892	0.76890	**-0.48**
V300	**0.74767**	0.74994	0.74976	0.75745	0.75071	0.74989	0.75063	**-0.40**
V400	**0.66387**	0.66560	0.66562	0.67156	0.66609	0.66581	0.66623	**-0.35**
V500	**0.53270**	0.53367	0.53360	0.53857	0.53404	0.53378	0.53384	**-0.21**
V600	**0.49083**	0.49178	0.49175	0.49614	0.49212	0.49184	0.49140	**-0.12**
V700	0.47322	0.47411	0.47408	0.47764	0.47445	0.47415	**0.47318**	**0.01**
V850	0.43527	0.43626	0.43624	0.43905	0.43662	0.43635	**0.43454**	**0.17**
V925	0.38993	0.39119	0.39112	0.39405	0.39160	0.39129	**0.38871**	**0.31**
V1000	0.30912	0.31026	0.31018	0.31350	0.31082	0.31036	**0.30783**	**0.42**
T2M	0.33130	0.33356	0.33415	0.33422	0.33541	0.33800	**0.30100**	**9.15**
U10	0.26774	0.26856	0.26856	0.27020	0.26869	0.26873	**0.26647**	**0.47**
V10	0.27248	0.27349	0.27351	0.27521	0.27360	0.27358	**0.27147**	**0.37**
MSL	24.726	25.687	25.620	24.971	25.992	26.136	**23.308**	**5.73**

Table D.3: MSE for all 69 output fields at a 12-hour lead time. Bold and underline indicate the lowest and second-lowest values, respectively. Gain is the relative MSE reduction of Ours over the best baseline for each field; negative values indicate that Ours has a larger error.

Field	δ-Adapter	Proceed	OneNet	FSNet	ADAPT-Z	ELF	Ours	Gain (%)
Z50	14365.431	17979.110	15225.157	15752.742	16317.057	20452.004	**9734.218**	**32.24**
Z100	8748.692	12367.749	9460.087	10411.313	10883.791	15764.031	**6154.921**	**29.65**
Z150	8577.524	12002.753	8979.619	9983.670	10614.123	17284.912	**6143.984**	**28.37**
Z200	8828.636	11628.980	9082.756	10302.973	10409.702	16520.377	**6574.119**	**25.54**
Z250	8807.777	11028.791	9011.525	9995.923	10113.434	15105.278	**6926.823**	**21.36**
Z300	8136.609	9904.915	8323.254	9205.649	9183.314	13219.004	**6615.908**	**18.69**
Z400	5980.176	7065.809	6150.743	6582.951	6715.863	9097.567	**5051.664**	**15.53**
Z500	4479.378	5146.016	4624.836	4888.316	4965.632	6393.002	**3876.690**	**13.45**
Z600	3775.349	4207.243	3896.713	4040.716	4095.322	4955.815	**3339.922**	**11.53**
Z700	3521.298	3757.807	3612.450	3627.865	3693.463	4159.154	**3168.191**	**10.03**
Z850	3578.299	3692.262	3646.650	3593.742	3664.602	3831.073	**3262.542**	**8.82**
Z925	3761.948	3843.465	3826.983	3759.401	3834.540	3925.982	**3438.348**	**8.54**
Z1000	4068.438	4136.328	4133.910	4061.871	4138.477	4190.080	**3719.447**	**8.43**
Q50	7.223e-16	9.169e-16	8.171e-16	8.589e-16	9.045e-16	1.006e-15	**6.710e-16**	**7.10**
Q100	**1.191e-14**	1.236e-14	1.237e-14	1.223e-14	1.240e-14	1.260e-14	1.211e-14	**-1.70**
Q150	**8.256e-13**	8.479e-13	8.419e-13	8.471e-13	8.492e-13	8.587e-13	8.309e-13	**-0.64**
Q200	**3.062e-11**	3.131e-11	3.110e-11	3.180e-11	3.129e-11	3.137e-11	3.104e-11	**-1.37**
Q250	**3.408e-10**	3.486e-10	3.472e-10	3.506e-10	3.489e-10	3.495e-10	3.447e-10	**-1.12**
Q300	**1.475e-09**	1.496e-09	1.491e-09	1.522e-09	1.494e-09	1.501e-09	1.482e-09	**-0.42**
Q400	1.138e-08	1.150e-08	1.146e-08	1.166e-08	1.147e-08	1.155e-08	**1.136e-08**	**0.24**
Q500	3.859e-08	3.919e-08	3.893e-08	3.979e-08	3.899e-08	3.945e-08	**3.822e-08**	**0.96**
Q600	8.394e-08	8.433e-08	8.425e-08	8.445e-08	8.426e-08	8.475e-08	**8.221e-08**	**2.07**
Q700	1.533e-07	1.548e-07	1.549e-07	1.561e-07	1.550e-07	1.554e-07	**1.504e-07**	**1.94**
Q850	2.647e-07	2.667e-07	2.662e-07	2.673e-07	2.678e-07	2.654e-07	**2.606e-07**	**1.56**
Q925	1.877e-07	1.894e-07	1.882e-07	1.907e-07	1.896e-07	1.881e-07	**1.793e-07**	**4.47**
Q1000	1.166e-07	1.178e-07	1.158e-07	1.162e-07	1.188e-07	1.144e-07	**9.989e-08**	**12.70**
T50	0.76570	0.89049	0.81231	0.85165	0.88635	1.01712	**0.71378**	**6.78**
T100	0.45723	0.47227	0.46039	0.46477	0.47131	0.53150	**0.42096**	**7.93**
T150	0.30672	0.31561	0.31259	0.30901	0.31768	0.33794	**0.28241**	**7.93**
T200	0.31925	0.33019	0.32910	0.32415	0.33019	0.33968	**0.30280**	**5.15**
T250	0.26817	0.27962	0.27131	0.27482	0.27511	0.29400	**0.25264**	**5.79**
T300	0.22177	0.23523	0.22345	0.22540	0.22993	0.25562	**0.20827**	**6.09**
T400	0.23492	0.25034	0.23650	0.24263	0.24382	0.27566	**0.21787**	**7.26**
T500	0.22779	0.23990	0.22910	0.23272	0.23386	0.26279	**0.21203**	**6.92**
T600	0.23179	0.24610	0.23388	0.23788	0.24321	0.27160	**0.21495**	**7.27**
T700	0.25072	0.26506	0.25465	0.26919	0.26150	0.29181	**0.23441**	**6.51**
T850	0.37970	0.39123	0.38414	0.39385	0.38876	0.41359	**0.35407**	**6.75**
T925	0.33277	0.34350	0.33602	0.34165	0.34003	0.36046	**0.30610**	**8.02**
T1000	0.34046	0.35323	0.34441	0.34925	0.35107	0.37533	**0.29720**	**12.71**
U50	2.40112	2.52507	2.48765	2.48370	2.54946	2.60119	**2.26492**	**5.67**
U100	2.61415	2.67030	2.64655	2.63341	2.66724	2.70483	**2.51181**	**3.91**
U150	3.53407	3.58318	3.56010	3.55543	3.57274	3.60797	**3.39184**	**4.02**
U200	4.03633	4.09380	4.07726	4.08373	4.08281	4.11487	**3.94107**	**2.36**
U250	4.13495	4.18775	4.17725	4.17949	4.18007	4.21330	**4.05331**	**1.97**
U300	3.78359	3.82343	3.81694	3.81621	3.81815	3.84036	**3.71785**	**1.74**
U400	2.80912	2.83742	2.83136	2.83125	2.83502	2.85214	**2.77439**	**1.24**
U500	2.06742	2.08460	2.08131	2.07451	2.08251	2.09292	**2.03467**	**1.58**
U600	1.74725	1.76150	1.75899	1.76061	1.76132	1.76667	**1.71165**	**2.04**
U700	1.60583	1.61580	1.61465	1.61091	1.61604	1.62089	**1.56633**	**2.46**
U850	1.36485	1.37463	1.37323	1.37145	1.37596	1.37921	**1.32834**	**2.68**
U925	1.18227	1.19034	1.18869	1.18754	1.19330	1.19252	**1.14611**	**3.06**
U1000	0.70416	0.71192	0.71069	0.71038	0.71360	0.71393	**0.68234**	**3.10**
V50	2.30638	2.39871	2.36756	2.37764	2.40043	2.40075	**2.24887**	**2.49**
V100	2.49264	2.52943	2.51665	2.51234	2.53340	2.55818	**2.42734**	**2.62**
V150	3.48179	3.51702	3.50478	3.50325	3.51607	3.54741	**3.38674**	**2.73**
V200	4.03378	4.07479	4.06642	4.07917	4.07243	4.08155	**3.95346**	**1.99**
V250	4.22930	4.26622	4.26040	4.26150	4.26444	4.27381	**4.15495**	**1.76**
V300	3.94827	3.97868	3.97367	3.98988	3.97911	3.98484	**3.89456**	**1.36**
V400	2.95926	2.98261	2.97930	2.97733	2.98534	2.98704	**2.93772**	**0.73**
V500	2.15842	2.17233	2.17079	2.16673	2.17342	2.17587	**2.13773**	**0.96**
V600	1.81618	1.82889	1.82693	1.82293	1.83006	1.83183	**1.78863**	**1.52**
V700	1.66414	1.67377	1.67271	1.67155	1.67499	1.67658	**1.62846**	**2.14**
V850	1.43192	1.43986	1.43803	1.43985	1.44244	1.44255	**1.38659**	**3.17**
V925	1.27045	1.27949	1.27819	1.27677	1.28213	1.28212	**1.23075**	**3.12**
V1000	0.75336	0.76018	0.75947	0.75944	0.76209	0.76189	**0.73072**	**3.01**
T2M	0.42578	0.42802	0.42968	0.42920	0.43248	0.44118	**0.32318**	**24.10**
U10	0.56511	0.57107	0.57048	0.56912	0.57196	0.57260	**0.54705**	**3.19**
V10	0.60308	0.60850	0.60800	0.60686	0.60993	0.60969	**0.58464**	**3.06**
MSL	6548.522	6654.903	6657.919	6547.125	6749.201	6726.893	**5971.179**	**8.80**

Table D.4: MAE for all 69 output fields at a 12-hour lead time. Bold and underline indicate the lowest and second-lowest values, respectively. Gain is the relative MAE reduction of Ours over the best baseline for each field; negative values indicate that Ours has a larger error.

Field	δ-Adapter	Proceed	OneNet	FSNet	ADAPT-Z	ELF	Ours	Gain (%)
Z50	90.390	102.544	92.985	95.301	96.925	110.463	**73.010**	**19.23**
Z100	70.222	84.863	73.360	77.346	78.755	95.140	**58.131**	**17.22**
Z150	68.098	82.814	69.932	74.813	76.753	100.329	**56.762**	**16.65**
Z200	67.695	80.703	68.932	75.294	74.980	99.348	**57.160**	**15.56**
Z250	66.354	77.603	67.315	73.047	72.872	94.526	**57.310**	**13.63**
Z300	63.079	72.891	63.994	69.600	68.715	88.044	**55.254**	**12.41**
Z400	53.739	61.082	54.776	58.232	58.513	72.624	**48.079**	**10.53**
Z500	46.709	51.972	47.719	50.267	50.384	60.542	**42.400**	**9.22**
Z600	42.878	46.587	43.751	45.421	45.538	52.531	**39.445**	**8.01**
Z700	41.070	43.118	41.688	42.109	42.491	46.692	**38.179**	**7.04**
Z850	40.692	41.494	41.050	40.691	41.187	42.694	**38.162**	**6.22**
Z925	41.384	41.831	41.676	41.179	41.690	42.434	**38.883**	**5.58**
Z1000	42.841	43.121	43.089	42.608	43.067	43.390	**40.290**	**5.44**
Q50	1.617e-08	1.859e-08	1.680e-08	1.786e-08	1.853e-08	1.974e-08	**1.500e-08**	**7.24**
Q100	**6.011e-08**	6.088e-08	6.106e-08	6.109e-08	6.130e-08	6.202e-08	6.021e-08	**-0.15**
Q150	4.953e-07	4.972e-07	4.955e-07	4.984e-07	4.987e-07	5.016e-07	**4.929e-07**	**0.48**
Q200	**2.934e-06**	2.944e-06	2.935e-06	2.970e-06	2.944e-06	2.953e-06	2.935e-06	**-0.01**
Q250	**9.761e-06**	9.787e-06	9.767e-06	9.833e-06	9.789e-06	9.811e-06	9.778e-06	**-0.18**
Q300	2.093e-05	2.095e-05	**2.091e-05**	2.114e-05	2.094e-05	2.100e-05	2.097e-05	**-0.25**
Q400	6.148e-05	6.135e-05	**6.126e-05**	6.176e-05	6.130e-05	6.148e-05	6.137e-05	**-0.18**
Q500	1.154e-04	1.157e-04	1.154e-04	1.167e-04	1.155e-04	1.161e-04	**1.152e-04**	**0.15**
Q600	1.812e-04	1.810e-04	1.809e-04	1.818e-04	1.810e-04	1.815e-04	**1.797e-04**	**0.69**
Q700	2.558e-04	2.562e-04	2.563e-04	2.579e-04	2.564e-04	2.568e-04	**2.535e-04**	**0.91**
Q850	3.427e-04	3.435e-04	3.431e-04	3.447e-04	3.441e-04	3.430e-04	**3.395e-04**	**0.93**
Q925	2.847e-04	2.859e-04	2.847e-04	2.873e-04	2.857e-04	2.857e-04	**2.780e-04**	**2.34**
Q1000	2.326e-04	2.338e-04	2.313e-04	2.318e-04	2.349e-04	2.301e-04	**2.160e-04**	**6.11**
T50	0.65690	0.70306	0.67486	0.68893	0.70165	0.75524	**0.63224**	**3.75**
T100	0.50680	0.51590	0.50848	0.51240	0.51625	0.55390	**0.48448**	**4.40**
T150	0.40865	0.41523	0.41181	0.40985	0.41654	0.43183	**0.38921**	**4.76**
T200	0.39452	0.40488	0.40278	0.39960	0.40414	0.41393	**0.38290**	**2.95**
T250	0.36452	0.37664	0.36702	0.37205	0.37158	0.39164	**0.35267**	**3.25**
T300	0.34024	0.35500	0.34155	0.34381	0.34910	0.37506	**0.32856**	**3.43**
T400	0.35517	0.37060	0.35650	0.36288	0.36394	0.39363	**0.34140**	**3.88**
T500	0.34742	0.36000	0.34852	0.35256	0.35334	0.38189	**0.33467**	**3.67**
T600	0.35130	0.36564	0.35310	0.35710	0.36205	0.38955	**0.33747**	**3.94**
T700	0.36193	0.37620	0.36489	0.38044	0.37182	0.40161	**0.34890**	**3.60**
T850	0.43088	0.44069	0.43351	0.44413	0.43779	0.46078	**0.41403**	**3.91**
T925	0.39135	0.40144	0.39339	0.39924	0.39744	0.41889	**0.37212**	**4.91**
T1000	0.39878	0.41036	0.40136	0.40562	0.40756	0.43156	**0.36784**	**7.76**
U50	1.18980	1.21970	1.21109	1.21070	1.22590	1.23692	**1.15493**	**2.93**
U100	1.20963	1.22332	1.21781	1.21466	1.22244	1.23106	**1.18461**	**2.07**
U150	1.38895	1.40002	1.39505	1.39382	1.39754	1.40510	**1.35992**	**2.09**
U200	1.48824	1.49968	1.49640	1.49717	1.49751	1.50370	**1.46939**	**1.27**
U250	1.50956	1.52034	1.51825	1.51780	1.51881	1.52520	**1.49327**	**1.08**
U300	1.44303	1.45170	1.45031	1.44901	1.45049	1.45510	**1.42923**	**0.96**
U400	1.24422	1.25167	1.25009	1.24960	1.25097	1.25576	**1.23530**	**0.72**
U500	1.06666	1.07185	1.07092	1.06851	1.07119	1.07445	**1.05674**	**0.93**
U600	0.98384	0.98848	0.98775	0.98816	0.98843	0.99012	**0.97190**	**1.21**
U700	0.94469	0.94797	0.94773	0.94627	0.94802	0.94970	**0.93076**	**1.47**
U850	0.85855	0.86206	0.86168	0.86107	0.86242	0.86381	**0.84409**	**1.68**
U925	0.77349	0.77646	0.77589	0.77604	0.77752	0.77734	**0.75787**	**2.02**
U1000	0.59487	0.59850	0.59800	0.59848	0.59911	0.59961	**0.58282**	**2.03**
V50	1.17243	1.19476	1.18793	1.18925	1.19513	1.19518	**1.15663**	**1.35**
V100	1.18858	1.19749	1.19450	1.19363	1.19855	1.20341	**1.17203**	**1.39**
V150	1.38137	1.38820	1.38569	1.38609	1.38815	1.39363	**1.36131**	**1.45**
V200	1.49214	1.50002	1.49828	1.50065	1.49980	1.50092	**1.47606**	**1.08**
V250	1.53301	1.54025	1.53901	1.53786	1.53999	1.54132	**1.51732**	**1.02**
V300	1.47938	1.48543	1.48432	1.48719	1.48553	1.48622	**1.46657**	**0.87**
V400	1.27752	1.28301	1.28220	1.28086	1.28369	1.28375	**1.27135**	**0.48**
V500	1.08933	1.09309	1.09274	1.09105	1.09345	1.09385	**1.08274**	**0.60**
V600	1.00196	1.00571	1.00518	1.00353	1.00613	1.00639	**0.99235**	**0.96**
V700	0.95995	0.96288	0.96262	0.96228	0.96330	0.96358	**0.94695**	**1.35**
V850	0.87693	0.87937	0.87888	0.87960	0.88019	0.88019	**0.85879**	**2.07**
V925	0.79804	0.80087	0.80055	0.80045	0.80173	0.80168	**0.78047**	**2.20**
V1000	0.61570	0.61850	0.61826	0.61901	0.61938	0.61926	**0.60335**	**2.00**
T2M	0.39767	0.39550	0.39607	0.39667	0.39759	0.40680	**0.34750**	**12.14**
U10	0.53199	0.53506	0.53482	0.53445	0.53561	0.53598	**0.52092**	**2.08**
V10	0.55156	0.55402	0.55381	0.55359	0.55479	0.55461	**0.54035**	**2.03**
MSL	53.381	53.667	53.663	53.179	53.964	53.939	**50.116**	**5.76**

Table D.5: MSE for all 69 output fields at a 24-hour lead time. Bold and underline indicate the lowest and second-lowest values, respectively. Gain is the relative MSE reduction of Ours over the best baseline for each field; negative values indicate that Ours has a larger error.

Field	δ-Adapter	Proceed	OneNet	FSNet	ADAPT-Z	ELF	Ours	Gain (%)
Z50	26384.444	35494.019	29055.287	30083.238	31629.178	43581.490	**20660.919**	**21.69**
Z100	16474.766	24327.944	18396.249	19935.711	21189.427	34017.417	**12972.019**	**21.26**
Z150	17020.844	24122.492	18115.992	19857.270	21265.656	37503.782	**13772.419**	**19.08**
Z200	18855.467	24490.470	19590.327	21580.535	22028.481	34481.412	**15841.694**	**15.98**
Z250	20104.820	24402.604	20700.654	22090.788	22650.018	32376.043	**17707.925**	**11.92**
Z300	19388.488	22716.411	19957.603	21041.784	21442.618	28922.522	**17547.251**	**9.50**
Z400	14892.992	16872.329	15389.543	15762.731	16319.090	20502.533	**13792.311**	**7.39**
Z500	11389.439	12592.577	11796.738	11940.985	12350.254	14763.485	**10613.034**	**6.82**
Z600	9632.399	10411.206	9959.954	9973.811	10297.083	11717.354	**9003.781**	**6.53**
Z700	8952.254	9399.254	9197.243	9073.906	9353.905	10124.612	**8390.842**	**6.27**
Z850	9033.452	9255.214	9210.851	9018.586	9284.647	9535.843	**8490.247**	**5.86**
Z925	9501.206	9664.218	9666.334	9461.995	9723.223	9859.536	**8936.572**	**5.55**
Z1000	10268.858	10408.201	10429.438	10217.099	10483.659	10564.622	**9654.424**	**5.51**
Q50	1.572e-15	2.083e-15	1.846e-15	1.981e-15	2.045e-15	2.349e-15	**1.424e-15**	**9.38**
Q100	**2.329e-14**	2.449e-14	2.444e-14	2.417e-14	2.465e-14	2.569e-14	2.357e-14	**-1.22**
Q150	1.579e-12	1.627e-12	1.608e-12	1.623e-12	1.631e-12	1.675e-12	**1.566e-12**	**0.84**
Q200	**6.356e-11**	6.534e-11	6.468e-11	6.648e-11	6.530e-11	6.548e-11	6.460e-11	**-1.63**
Q250	**8.002e-10**	8.206e-10	8.157e-10	8.250e-10	8.202e-10	8.194e-10	8.150e-10	**-1.85**
Q300	**3.644e-09**	3.716e-09	3.697e-09	3.780e-09	3.706e-09	3.719e-09	3.673e-09	**-0.77**
Q400	**2.680e-08**	2.736e-08	2.725e-08	2.761e-08	2.732e-08	2.745e-08	2.689e-08	**-0.31**
Q500	9.139e-08	9.334e-08	9.276e-08	9.460e-08	9.299e-08	9.370e-08	**9.106e-08**	**0.36**
Q600	1.891e-07	1.910e-07	1.908e-07	1.902e-07	1.912e-07	1.913e-07	**1.864e-07**	**1.43**
Q700	3.414e-07	3.487e-07	3.494e-07	3.506e-07	3.507e-07	3.498e-07	**3.393e-07**	**0.61**
Q850	5.719e-07	5.802e-07	5.796e-07	5.780e-07	5.850e-07	5.751e-07	**5.696e-07**	**0.40**
Q925	3.452e-07	3.496e-07	3.476e-07	3.525e-07	3.514e-07	3.457e-07	**3.317e-07**	**3.91**
Q1000	2.087e-07	2.126e-07	2.076e-07	2.081e-07	2.132e-07	2.050e-07	**1.821e-07**	**11.17**
T50	1.20334	1.49485	1.34037	1.39513	1.49073	1.86209	**1.11532**	**7.31**
T100	0.78426	0.82711	0.80386	0.80629	0.83001	1.00704	**0.71550**	**8.77**
T150	0.57952	0.60717	0.60058	0.58780	0.61575	0.68579	**0.52698**	**9.07**
T200	0.64766	0.68093	0.67703	0.66674	0.68177	0.72033	**0.61589**	**4.91**
T250	0.57286	0.60476	0.58204	0.59508	0.59183	0.64802	**0.54605**	**4.68**
T300	0.46077	0.49277	0.46550	0.47124	0.48085	0.54628	**0.43762**	**5.02**
T400	0.46684	0.50030	0.47293	0.48670	0.48642	0.55625	**0.44584**	**4.50**
T500	0.45878	0.48361	0.46443	0.46947	0.47245	0.53054	**0.44200**	**3.66**
T600	0.45741	0.48388	0.46447	0.46786	0.48023	0.52939	**0.43931**	**3.96**
T700	0.48274	0.51093	0.49548	0.52085	0.50598	0.55752	**0.46335**	**4.02**
T850	0.73089	0.75175	0.74399	0.75395	0.74886	0.78723	**0.69462**	**4.96**
T925	0.65212	0.66921	0.66025	0.66678	0.66523	0.69172	**0.61113**	**6.29**
T1000	0.62419	0.63802	0.62985	0.63601	0.63879	0.66457	**0.55099**	**11.73**
U50	4.10629	4.42928	4.36888	4.36230	4.53122	4.65283	**3.80467**	**7.35**
U100	4.57341	4.71367	4.67359	4.64892	4.74279	4.89919	**4.35609**	**4.75**
U150	6.45472	6.56209	6.50650	6.50120	6.54287	6.70119	**6.18181**	**4.23**
U200	7.99150	8.13487	8.09697	8.08515	8.11263	8.19327	**7.81375**	**2.22**
U250	8.76784	8.91294	8.88132	8.86755	8.89621	8.98164	**8.64127**	**1.44**
U300	8.28182	8.40140	8.37837	8.37717	8.39226	8.44844	**8.19075**	**1.10**
U400	6.08743	6.17035	6.15478	6.13734	6.16809	6.21342	**6.04427**	**0.71**
U500	4.36562	4.41217	4.40704	4.38684	4.41287	4.43383	**4.32002**	**1.04**
U600	3.53051	3.56569	3.56217	3.55707	3.57094	3.57857	**3.47577**	**1.55**
U700	3.13502	3.15939	3.15885	3.14681	3.16601	3.17206	**3.06724**	**2.16**
U850	2.78330	2.80872	2.80559	2.79456	2.81707	2.81953	**2.71490**	**2.46**
U925	2.55334	2.57528	2.57066	2.55863	2.58336	2.58151	**2.49010**	**2.48**
U1000	1.53717	1.55677	1.55274	1.54662	1.56079	1.56127	**1.50014**	**2.41**
V50	3.64068	3.77781	3.72227	3.74122	3.78367	3.81925	**3.53164**	**3.00**
V100	4.37908	4.45684	4.43380	4.41624	4.47364	4.62481	**4.23252**	**3.35**
V150	6.47042	6.54004	6.51800	6.50680	6.54276	6.70279	**6.28905**	**2.80**
V200	8.12935	8.23681	8.21175	8.22245	8.23216	8.27404	**7.98515**	**1.77**
V250	9.13264	9.24815	9.21668	9.20290	9.24198	9.28066	**9.00471**	**1.40**
V300	8.79479	8.90478	8.87463	8.88369	8.90916	8.94278	**8.71598**	**0.90**
V400	6.53775	6.61663	6.60082	6.57392	6.62507	6.64832	**6.51284**	**0.38**
V500	4.65198	4.70008	4.69369	4.67572	4.70657	4.71956	**4.62343**	**0.61**
V600	3.71717	3.75429	3.74875	3.73889	3.76214	3.76795	**3.67659**	**1.09**
V700	3.24592	3.27496	3.27220	3.25908	3.28361	3.28597	**3.19551**	**1.55**
V850	2.89806	2.92135	2.91684	2.90510	2.93175	2.92954	**2.83032**	**2.34**
V925	2.74389	2.77261	2.76748	2.74933	2.78301	2.78249	**2.68689**	**2.08**
V1000	1.65877	1.67950	1.67636	1.66798	1.68593	1.68569	**1.62507**	**2.03**
T2M	0.65710	0.65849	0.65932	0.66176	0.66478	0.68897	**0.51926**	**20.98**
U10	1.25447	1.26972	1.26714	1.26138	1.27241	1.27337	**1.22405**	**2.43**
V10	1.34583	1.36203	1.35984	1.35203	1.36671	1.36687	**1.31830**	**2.05**
MSL	16575.148	16810.660	16854.626	16518.112	17152.954	17033.964	**15575.616**	**5.71**

Table D.6: MAE for all 69 output fields at a 24-hour lead time. Bold and underline indicate the lowest and second-lowest values, respectively. Gain is the relative MAE reduction of Ours over the best baseline for each field; negative values indicate that Ours has a larger error.

Field	δ-Adapter	Proceed	OneNet	FSNet	ADAPT-Z	ELF	Ours	Gain (%)
Z50	118.743	141.393	124.028	128.305	131.246	159.717	**101.165**	**14.80**
Z100	95.434	118.877	101.260	106.527	109.220	135.536	**82.462**	**13.59**
Z150	95.775	117.858	99.352	105.778	108.911	144.822	**83.760**	**12.54**
Z200	98.123	117.053	100.677	108.599	108.811	143.980	**87.047**	**11.29**
Z250	98.184	114.108	100.103	106.729	107.439	137.687	**89.151**	**9.20**
Z300	94.338	107.955	96.205	102.119	102.269	128.460	**86.859**	**7.93**
Z400	81.093	90.956	83.029	86.104	87.522	106.177	**75.871**	**6.44**
Z500	70.940	77.824	72.701	74.606	75.807	88.890	**66.748**	**5.91**
Z600	65.307	70.013	66.724	67.732	68.810	77.637	**61.642**	**5.61**
Z700	62.705	65.228	63.627	63.448	64.519	69.811	**59.306**	**5.42**
Z850	62.308	63.095	62.697	61.881	62.851	64.572	**58.977**	**4.69**
Z925	63.574	63.891	63.827	63.016	63.835	64.661	**60.204**	**4.46**
Z1000	65.970	66.077	66.121	65.413	66.102	66.459	**62.503**	**4.45**
Q50	2.385e-08	2.797e-08	2.524e-08	2.736e-08	2.789e-08	3.045e-08	**2.168e-08**	**9.11**
Q100	8.781e-08	8.955e-08	8.991e-08	8.981e-08	9.060e-08	9.287e-08	**8.761e-08**	**0.22**
Q150	7.024e-07	7.061e-07	7.022e-07	7.081e-07	7.091e-07	7.181e-07	**6.928e-07**	**1.34**
Q200	**4.275e-06**	4.308e-06	4.287e-06	4.345e-06	4.307e-06	4.328e-06	4.278e-06	**-0.07**
Q250	**1.492e-05**	1.501e-05	1.496e-05	1.509e-05	1.500e-05	1.504e-05	1.502e-05	**-0.62**
Q300	**3.259e-05**	3.274e-05	3.266e-05	3.308e-05	3.270e-05	3.279e-05	3.275e-05	**-0.47**
Q400	**9.383e-05**	9.418e-05	9.400e-05	9.461e-05	9.416e-05	9.435e-05	9.408e-05	**-0.26**
Q500	**1.786e-04**	1.797e-04	1.793e-04	1.811e-04	1.794e-04	1.801e-04	1.789e-04	**-0.21**
Q600	2.708e-04	2.713e-04	2.713e-04	2.716e-04	2.714e-04	2.717e-04	**2.694e-04**	**0.52**
Q700	3.782e-04	3.810e-04	3.812e-04	3.828e-04	3.818e-04	3.818e-04	**3.770e-04**	**0.34**
Q850	4.924e-04	4.956e-04	4.946e-04	4.957e-04	4.972e-04	4.939e-04	**4.906e-04**	**0.36**
Q925	3.821e-04	3.844e-04	3.826e-04	3.863e-04	3.847e-04	3.833e-04	**3.738e-04**	**2.19**
Q1000	3.144e-04	3.175e-04	3.129e-04	3.133e-04	3.174e-04	3.112e-04	**2.929e-04**	**5.89**
T50	0.82454	0.91214	0.86557	0.88251	0.91007	1.02818	**0.79100**	**4.07**
T100	0.66941	0.68925	0.67759	0.68202	0.69305	0.77459	**0.63582**	**5.02**
T150	0.56519	0.58028	0.57394	0.56911	0.58430	0.61951	**0.53379**	**5.56**
T200	0.56495	0.58699	0.58189	0.57870	0.58568	0.60925	**0.54715**	**3.15**
T250	0.53206	0.55481	0.53701	0.54895	0.54515	0.58422	**0.51658**	**2.91**
T300	0.48820	0.51245	0.49081	0.49530	0.50343	0.54756	**0.47358**	**2.99**
T400	0.49541	0.51979	0.49906	0.51008	0.50937	0.55590	**0.48214**	**2.68**
T500	0.48575	0.50428	0.48907	0.49408	0.49482	0.53650	**0.47485**	**2.24**
T600	0.48822	0.50743	0.49235	0.49525	0.50330	0.53904	**0.47577**	**2.55**
T700	0.49517	0.51530	0.50186	0.52293	0.50981	0.54842	**0.48284**	**2.49**
T850	0.58619	0.59841	0.59103	0.60171	0.59528	0.62238	**0.56895**	**2.94**
T925	0.53713	0.54798	0.54023	0.54528	0.54428	0.56565	**0.51492**	**4.14**
T1000	0.52997	0.53807	0.53137	0.53567	0.53805	0.55868	**0.49279**	**7.01**
U50	1.55366	1.61059	1.59954	1.59862	1.62858	1.64484	**1.49574**	**3.73**
U100	1.60792	1.63377	1.62685	1.62306	1.63839	1.65839	**1.56864**	**2.44**
U150	1.88050	1.89890	1.89030	1.88876	1.89538	1.91496	**1.84090**	**2.11**
U200	2.09626	2.11620	2.11096	2.10883	2.11337	2.12335	**2.07071**	**1.22**
U250	2.19722	2.21681	2.21286	2.21039	2.21466	2.22496	**2.17887**	**0.84**
U300	2.12196	2.13882	2.13578	2.13534	2.13717	2.14452	**2.10775**	**0.67**
U400	1.80983	1.82418	1.82153	1.81922	1.82336	1.83211	**1.80190**	**0.44**
U500	1.53905	1.54835	1.54749	1.54367	1.54835	1.55278	**1.52908**	**0.65**
U600	1.39161	1.39942	1.39875	1.39802	1.40055	1.40208	**1.37813**	**0.97**
U700	1.31193	1.31748	1.31760	1.31493	1.31898	1.32040	**1.29411**	**1.36**
U850	1.20599	1.21228	1.21162	1.20970	1.21414	1.21498	**1.18649**	**1.62**
U925	1.10205	1.10730	1.10614	1.10466	1.10931	1.10883	**1.08252**	**1.77**
U1000	0.85001	0.85601	0.85488	0.85417	0.85715	0.85763	**0.83576**	**1.68**
V50	1.47456	1.50055	1.48978	1.49293	1.50149	1.50688	**1.45129**	**1.58**
V100	1.58095	1.59462	1.59036	1.58796	1.59759	1.61553	**1.55392**	**1.71**
V150	1.88432	1.89409	1.89052	1.88988	1.89472	1.91136	**1.85737**	**1.43**
V200	2.11573	2.12996	2.12649	2.12837	2.12980	2.13406	**2.09553**	**0.95**
V250	2.24063	2.25517	2.25119	2.24839	2.25440	2.25808	**2.22166**	**0.85**
V300	2.17830	2.19215	2.18833	2.19016	2.19256	2.19586	**2.16464**	**0.63**
V400	1.85951	1.87116	1.86891	1.86483	1.87244	1.87534	**1.85390**	**0.30**
V500	1.57487	1.58327	1.58233	1.57940	1.58458	1.58622	**1.56805**	**0.43**
V600	1.41673	1.42406	1.42311	1.42140	1.42575	1.42625	**1.40611**	**0.75**
V700	1.32359	1.32965	1.32926	1.32665	1.33159	1.33158	**1.30982**	**1.04**
V850	1.21885	1.22373	1.22284	1.22078	1.22606	1.22531	**1.19957**	**1.58**
V925	1.13046	1.13632	1.13544	1.13208	1.13882	1.13821	**1.11230**	**1.61**
V1000	0.88062	0.88622	0.88554	0.88416	0.88838	0.88790	**0.86766**	**1.47**
T2M	0.51618	0.51277	0.51295	0.51424	0.51581	0.53480	**0.45774**	**10.73**
U10	0.76400	0.76917	0.76838	0.76727	0.77027	0.77064	**0.75109**	**1.69**
V10	0.79349	0.79841	0.79785	0.79591	0.80016	0.79994	**0.78177**	**1.48**
MSL	82.467	82.544	82.637	81.950	83.109	82.967	**78.121**	**4.67**

Table D.7: MSE for all 69 output fields at a 48-hour lead time. Bold and underline indicate the lowest and second-lowest values, respectively. Gain is the relative MSE reduction of Ours over the best baseline for each field; negative values indicate that Ours has a larger error.

Field	δ-Adapter	Proceed	OneNet	FSNet	ADAPT-Z	ELF	Ours	Gain (%)
Z50	68812.226	89114.694	75692.650	77697.253	83445.627	112711.138	**58733.177**	**14.65**
Z100	44757.219	59551.570	49718.579	51682.857	55013.873	86097.088	**37863.339**	**15.40**
Z150	50015.601	62731.708	52945.670	55696.941	58191.989	93854.958	**43563.154**	**12.90**
Z200	60880.093	70775.338	62955.439	65752.393	66834.789	89626.985	**54928.430**	**9.78**
Z250	70460.119	77784.912	72221.423	73771.876	75202.398	92636.280	**65853.003**	**6.54**
Z300	70995.989	76544.662	72711.675	73530.367	74920.230	87697.973	**67557.670**	**4.84**
Z400	55518.409	58743.690	56986.463	56736.475	58309.483	64965.811	**53531.662**	**3.58**
Z500	41063.557	42963.359	42205.014	41778.181	43002.050	46638.255	**39626.369**	**3.50**
Z600	32727.101	33893.064	33578.390	33139.009	34112.140	36124.515	**31514.127**	**3.71**
Z700	28730.041	29363.253	29325.317	28851.468	29622.726	30673.522	**27593.673**	**3.96**
Z850	27638.927	27866.404	28016.502	27559.140	28220.694	28456.837	**26466.883**	**3.96**
Z925	28977.157	29103.189	29308.003	28847.425	29486.128	29561.387	**27729.805**	**3.87**
Z1000	31284.857	31365.268	31583.131	31106.134	31760.887	31763.873	**29899.705**	**3.88**
Q50	3.633e-15	4.887e-15	4.240e-15	4.625e-15	4.776e-15	5.672e-15	**3.226e-15**	**11.20**
Q100	**4.698e-14**	5.010e-14	4.988e-14	4.961e-14	5.086e-14	5.456e-14	4.733e-14	**-0.75**
Q150	2.808e-12	2.865e-12	2.832e-12	2.861e-12	2.878e-12	2.968e-12	**2.732e-12**	**2.72**
Q200	**1.143e-10**	1.176e-10	1.164e-10	1.195e-10	1.178e-10	1.173e-10	1.164e-10	**-1.79**
Q250	**1.538e-09**	1.576e-09	1.569e-09	1.595e-09	1.575e-09	1.566e-09	1.578e-09	**-2.60**
Q300	**7.755e-09**	7.912e-09	7.879e-09	8.055e-09	7.886e-09	7.883e-09	7.869e-09	**-1.47**
Q400	**6.079e-08**	6.264e-08	6.242e-08	6.292e-08	6.272e-08	6.266e-08	6.158e-08	**-1.29**
Q500	**2.195e-07**	2.247e-07	2.244e-07	2.273e-07	2.246e-07	2.249e-07	2.205e-07	**-0.45**
Q600	**4.645e-07**	4.740e-07	4.741e-07	4.726e-07	4.765e-07	4.723e-07	4.657e-07	**-0.27**
Q700	**8.325e-07**	8.582e-07	8.620e-07	8.610e-07	8.662e-07	8.553e-07	8.450e-07	**-1.50**
Q850	**1.323e-06**	1.349e-06	1.348e-06	1.333e-06	1.369e-06	1.329e-06	1.327e-06	**-0.31**
Q925	7.304e-07	7.365e-07	7.339e-07	7.416e-07	7.451e-07	7.251e-07	**7.063e-07**	**2.59**
Q1000	4.602e-07	4.698e-07	4.639e-07	4.646e-07	4.721e-07	4.565e-07	**4.280e-07**	**6.25**
T50	2.06680	2.62365	2.42336	2.40449	2.62761	3.58658	**1.90697**	**7.73**
T100	1.50660	1.62076	1.57643	1.57322	1.65034	2.11847	**1.36902**	**9.13**
T150	1.30040	1.37498	1.35849	1.32805	1.41133	1.61907	**1.18916**	**8.55**
T200	1.67585	1.76547	1.75982	1.74092	1.78197	1.89389	**1.60937**	**3.97**
T250	1.54468	1.62121	1.57489	1.60330	1.59807	1.72815	**1.48619**	**3.79**
T300	1.17791	1.24644	1.19360	1.20622	1.22601	1.36727	**1.12721**	**4.30**
T400	1.22289	1.29096	1.24651	1.27383	1.26697	1.40233	**1.18274**	**3.28**
T500	1.25560	1.30642	1.27882	1.28192	1.28874	1.40037	**1.22683**	**2.29**
T600	1.22752	1.27552	1.25421	1.25176	1.27270	1.35083	**1.19861**	**2.36**
T700	1.25455	1.30684	1.29448	1.32396	1.30357	1.37851	**1.22442**	**2.40**
T850	1.75888	1.79555	1.79378	1.79254	1.79606	1.84367	**1.70681**	**2.96**
T925	1.61674	1.64482	1.63802	1.63978	1.64367	1.67234	**1.54721**	**4.30**
T1000	1.39961	1.41277	1.40693	1.41192	1.41775	1.44655	**1.27201**	**9.12**
U50	7.61314	8.22787	8.11565	8.17798	8.51020	9.15729	**6.83007**	**10.29**
U100	8.75427	9.00953	9.01198	8.96343	9.19080	9.87435	**8.28851**	**5.32**
U150	13.44799	13.62240	13.54321	13.58255	13.63093	14.06197	**12.97544**	**3.51**
U200	18.58910	18.90435	18.84742	18.83957	18.90482	19.06500	**18.31288**	**1.49**
U250	22.63548	23.00023	22.94908	22.88135	23.01762	23.15301	**22.45295**	**0.81**
U300	22.85982	23.19109	23.13994	23.05726	23.22330	23.30520	**22.71278**	**0.64**
U400	16.91908	17.15712	17.13447	16.99316	17.18035	17.25698	**16.83828**	**0.48**
U500	11.61351	11.74288	11.74644	11.65164	11.76571	11.78984	**11.51946**	**0.81**
U600	8.84159	8.93171	8.94205	8.89730	8.96171	8.95579	**8.74163**	**1.13**
U700	7.48917	7.55343	7.56722	7.52721	7.58950	7.57435	**7.37900**	**1.47**
U850	6.83261	6.88241	6.88224	6.84097	6.91410	6.89890	**6.71895**	**1.66**
U925	6.64748	6.69404	6.68909	6.64101	6.71725	6.70563	**6.54679**	**1.42**
U1000	3.97541	4.02096	4.01376	3.98611	4.03242	4.02750	**3.92545**	**1.26**
V50	5.65080	5.83526	5.74247	5.76584	5.85104	6.05829	**5.56440**	**1.53**
V100	7.86176	7.96068	7.92442	7.87165	7.98901	8.56362	**7.67318**	**2.40**
V150	13.27880	13.39713	13.35983	13.32590	13.43862	13.80829	**13.04190**	**1.78**
V200	19.12469	19.45513	19.37561	19.36961	19.49479	19.57979	**18.98304**	**0.74**
V250	23.94097	24.36892	24.25995	24.19508	24.41528	24.51879	**23.83064**	**0.46**
V300	24.56815	24.98805	24.88015	24.77864	25.05234	25.15289	**24.49655**	**0.29**
V400	18.29262	18.57907	18.53261	18.34072	18.62389	18.69706	**18.27263**	**0.11**
V500	12.38053	12.54420	12.53310	12.42002	12.57585	12.61550	**12.34299**	**0.30**
V600	9.23650	9.34457	9.34394	9.28162	9.37727	9.38747	**9.18493**	**0.56**
V700	7.63837	7.71607	7.72545	7.66485	7.74958	7.74688	**7.57661**	**0.81**
V850	6.97020	7.03050	7.03462	6.97858	7.05954	7.05184	**6.89500**	**1.08**
V925	7.09876	7.17444	7.17400	7.10667	7.20055	7.20367	**7.03765**	**0.86**
V1000	4.30717	4.36405	4.36312	4.32791	4.38046	4.38100	**4.27196**	**0.82**
T2M	1.33653	1.33024	1.33187	1.34218	1.33819	1.37966	**1.12154**	**15.69**
U10	3.28445	3.31963	3.31483	3.29307	3.32668	3.32529	**3.24490**	**1.20**
V10	3.51962	3.56280	3.56281	3.53448	3.57359	3.57538	**3.49257**	**0.77**
MSL	50415.835	50607.795	50966.782	50194.391	51624.555	51195.142	**48204.632**	**3.96**

Table D.8: MAE for all 69 output fields at a 48-hour lead time. Bold and underline indicate the lowest and second-lowest values, respectively. Gain is the relative MAE reduction of Ours over the best baseline for each field; negative values indicate that Ours has a larger error.

Field	δ-Adapter	Proceed	OneNet	FSNet	ADAPT-Z	ELF	Ours	Gain (%)
Z50	178.619	212.370	186.219	193.602	200.091	247.914	**155.990**	**12.67**
Z100	150.908	181.499	159.593	166.625	170.897	206.615	**133.830**	**11.32**
Z150	158.488	185.849	164.471	172.926	176.003	222.310	**143.903**	**9.20**
Z200	168.741	191.621	173.226	182.484	182.655	225.647	**156.106**	**7.49**
Z250	174.021	192.761	177.413	184.698	185.745	221.968	**164.372**	**5.54**
Z300	169.648	185.352	172.909	178.663	179.595	210.225	**162.010**	**4.50**
Z400	146.370	157.395	149.531	151.712	154.116	175.585	**141.189**	**3.54**
Z500	126.361	133.788	129.064	129.773	132.118	146.970	**122.008**	**3.45**
Z600	114.000	118.776	115.982	115.843	118.047	127.855	**109.926**	**3.57**
Z700	107.489	109.784	108.609	107.655	109.439	115.287	**103.408**	**3.80**
Z850	105.269	105.444	105.435	104.439	105.559	107.262	**100.822**	**3.46**
Z925	107.530	107.150	107.438	106.644	107.432	108.168	**102.911**	**3.50**
Z1000	111.754	111.109	111.441	110.852	111.439	111.628	**106.875**	**3.59**
Q50	3.684e-08	4.356e-08	3.922e-08	4.289e-08	4.348e-08	4.820e-08	**3.265e-08**	**11.37**
Q100	1.306e-07	1.344e-07	1.350e-07	1.351e-07	1.368e-07	1.426e-07	**1.300e-07**	**0.47**
Q150	9.745e-07	9.755e-07	9.699e-07	9.802e-07	9.807e-07	9.939e-07	**9.506e-07**	**1.99**
Q200	**5.989e-06**	6.046e-06	6.014e-06	6.085e-06	6.048e-06	6.057e-06	5.999e-06	**-0.17**
Q250	**2.162e-05**	2.179e-05	2.171e-05	2.194e-05	2.177e-05	2.176e-05	2.185e-05	**-1.07**
Q300	**4.915e-05**	4.948e-05	4.938e-05	5.002e-05	4.940e-05	4.944e-05	4.960e-05	**-0.92**
Q400	**1.439e-04**	1.455e-04	1.452e-04	1.458e-04	1.456e-04	1.455e-04	1.453e-04	**-0.95**
Q500	**2.826e-04**	2.854e-04	2.853e-04	2.873e-04	2.854e-04	2.856e-04	2.849e-04	**-0.80**
Q600	**4.273e-04**	4.306e-04	4.308e-04	4.306e-04	4.314e-04	4.300e-04	4.292e-04	**-0.44**
Q700	**5.880e-04**	5.958e-04	5.969e-04	5.974e-04	5.982e-04	5.955e-04	5.929e-04	**-0.83**
Q850	**7.404e-04**	7.475e-04	7.462e-04	7.452e-04	7.521e-04	7.432e-04	7.422e-04	**-0.24**
Q925	5.517e-04	5.543e-04	5.523e-04	5.567e-04	5.565e-04	5.513e-04	**5.419e-04**	**1.70**
Q1000	4.613e-04	4.662e-04	4.622e-04	4.622e-04	4.666e-04	4.583e-04	**4.420e-04**	**3.55**
T50	1.07565	1.21239	1.15509	1.15941	1.20938	1.43698	**1.02488**	**4.72**
T100	0.93514	0.97347	0.95695	0.96178	0.98771	1.13432	**0.88483**	**5.38**
T150	0.84338	0.87040	0.85938	0.85292	0.88142	0.94554	**0.79570**	**5.65**
T200	0.88843	0.92682	0.91748	0.91881	0.92808	0.96890	**0.86216**	**2.96**
T250	0.84603	0.88198	0.85612	0.87761	0.86927	0.92920	**0.82481**	**2.51**
T300	0.76269	0.79669	0.76837	0.77733	0.78624	0.84784	**0.74305**	**2.58**
T400	0.77176	0.80454	0.78041	0.79791	0.79219	0.85166	**0.75631**	**2.00**
T500	0.76504	0.78897	0.77339	0.77952	0.77863	0.83081	**0.75378**	**1.47**
T600	0.76340	0.78601	0.77281	0.77504	0.78287	0.82222	**0.75031**	**1.71**
T700	0.76480	0.78925	0.77826	0.80127	0.78528	0.82608	**0.75199**	**1.67**
T850	0.88271	0.89555	0.89056	0.89917	0.89446	0.92021	**0.86606**	**1.89**
T925	0.81701	0.82697	0.82174	0.82500	0.82585	0.84403	**0.79254**	**2.99**
T1000	0.76596	0.76815	0.76376	0.76878	0.77111	0.78995	**0.72244**	**5.41**
U50	2.09843	2.17764	2.16154	2.17167	2.21560	2.25818	**1.99054**	**5.14**
U100	2.23121	2.26600	2.26595	2.26084	2.28786	2.32390	**2.17150**	**2.68**
U150	2.73353	2.75562	2.74759	2.74997	2.75627	2.78543	**2.68597**	**1.74**
U200	3.19943	3.22821	3.22374	3.22254	3.22875	3.24095	**3.17242**	**0.84**
U250	3.49520	3.52594	3.52277	3.51853	3.52721	3.53762	**3.47762**	**0.50**
U300	3.45194	3.48031	3.47656	3.47422	3.48190	3.48962	**3.43781**	**0.41**
U400	2.92282	2.94733	2.94513	2.93647	2.94878	2.95998	**2.91476**	**0.28**
U500	2.44719	2.46283	2.46368	2.45556	2.46558	2.46916	**2.43472**	**0.51**
U600	2.15926	2.17240	2.17389	2.17030	2.17662	2.17543	**2.14400**	**0.71**
U700	1.98691	1.99690	1.99899	1.99454	2.00264	1.99969	**1.96816**	**0.94**
U850	1.82997	1.83808	1.83756	1.83314	1.84324	1.84046	**1.80905**	**1.14**
U925	1.69820	1.70486	1.70343	1.69874	1.70904	1.70630	**1.67823**	**1.18**
U1000	1.30550	1.31394	1.31231	1.30904	1.31637	1.31534	**1.29336**	**0.93**
V50	1.82422	1.85203	1.83592	1.84026	1.85408	1.87538	**1.80745**	**0.92**
V100	2.12827	2.14110	2.13608	2.13082	2.14497	2.18805	**2.10229**	**1.22**
V150	2.71559	2.72861	2.72439	2.72253	2.73305	2.75741	**2.69146**	**0.89**
V200	3.22775	3.25639	3.24990	3.25029	3.26026	3.26606	**3.21485**	**0.40**
V250	3.54955	3.58290	3.57500	3.57088	3.58633	3.59235	**3.53975**	**0.28**
V300	3.51349	3.54624	3.53801	3.53551	3.55034	3.55688	**3.50760**	**0.17**
V400	2.96624	2.99187	2.98770	2.97559	2.99580	3.00182	**2.96534**	**0.03**
V500	2.46627	2.48403	2.48287	2.47467	2.48799	2.49118	**2.46242**	**0.16**
V600	2.16018	2.17431	2.17411	2.16868	2.17895	2.17876	**2.15299**	**0.33**
V700	1.96568	1.97700	1.97826	1.97129	1.98223	1.98050	**1.95597**	**0.49**
V850	1.81065	1.81977	1.81989	1.81367	1.82459	1.82230	**1.79796**	**0.70**
V925	1.72000	1.73045	1.73029	1.72247	1.73496	1.73418	**1.70838**	**0.68**
V1000	1.34729	1.35705	1.35698	1.35224	1.36096	1.35984	**1.33912**	**0.61**
T2M	0.73183	0.72249	0.72379	0.72784	0.72734	0.75205	**0.66444**	**8.04**
U10	1.17683	1.18420	1.18295	1.17994	1.18616	1.18558	**1.16638**	**0.89**
V10	1.21760	1.22617	1.22612	1.22146	1.22917	1.22871	**1.21065**	**0.57**
MSL	139.990	139.174	139.617	139.146	140.113	139.759	**134.003**	**3.70**

Table D.9: MSE for all 69 output fields at a 72-hour lead time. Bold and underline indicate the lowest and second-lowest values, respectively. Gain is the relative MSE reduction of Ours over the best baseline for each field; negative values indicate that Ours has a larger error.

Field	δ-Adapter	Proceed	OneNet	FSNet	ADAPT-Z	ELF	Ours	Gain (%)
Z50	127630.585	157641.646	139631.404	140705.027	154219.067	194965.406	**114226.842**	**10.50**
Z100	88607.526	108506.637	97391.056	98428.520	104697.355	151564.655	**79194.299**	**10.62**
Z150	105769.800	122059.215	110973.460	113858.677	117057.361	168633.847	**96760.022**	**8.52**
Z200	136764.726	149241.663	140463.066	143324.937	144575.951	174454.667	**128267.043**	**6.21**
Z250	165035.910	174096.839	168342.360	169374.636	171109.688	194112.624	**158221.614**	**4.13**
Z300	168433.735	175121.396	171627.696	171483.450	173396.785	189888.639	**163163.223**	**3.13**
Z400	130863.149	134552.518	133469.039	132087.359	134346.056	142558.114	**127744.759**	**2.38**
Z500	94418.273	96396.175	96326.891	95054.908	96828.131	101070.648	**92146.142**	**2.41**
Z600	72480.915	73513.951	73812.466	72797.576	74193.219	76327.426	**70598.729**	**2.60**
Z700	61139.090	61519.730	61990.840	61168.920	62209.226	63195.099	**59415.836**	**2.82**
Z850	56649.655	56497.206	57096.222	56452.867	57278.752	57297.905	**54903.006**	**2.75**
Z925	59234.209	58958.331	59595.768	58958.662	59749.964	59595.146	**57359.833**	**2.71**
Z1000	63880.485	63533.439	64168.021	63509.760	64316.188	64093.046	**61770.988**	**2.74**
Q50	5.832e-15	7.911e-15	6.735e-15	7.444e-15	7.714e-15	9.290e-15	**5.118e-15**	**12.24**
Q100	6.958e-14	7.408e-14	7.362e-14	7.406e-14	7.566e-14	8.189e-14	**6.945e-14**	**0.18**
Q150	3.802e-12	3.831e-12	3.793e-12	3.828e-12	3.847e-12	3.930e-12	**3.630e-12**	**4.29**
Q200	**1.511e-10**	1.549e-10	1.535e-10	1.576e-10	1.556e-10	1.537e-10	1.534e-10	**-1.51**
Q250	**2.077e-09**	2.122e-09	2.115e-09	2.158e-09	2.124e-09	2.104e-09	2.126e-09	**-2.39**
Q300	**1.112e-08**	1.132e-08	1.129e-08	1.158e-08	1.130e-08	1.126e-08	1.129e-08	**-1.53**
Q400	**9.396e-08**	9.694e-08	9.670e-08	9.734e-08	9.727e-08	9.686e-08	9.555e-08	**-1.70**
Q500	**3.523e-07**	3.607e-07	3.611e-07	3.649e-07	3.611e-07	3.608e-07	3.554e-07	**-0.88**
Q600	**7.674e-07**	7.857e-07	7.865e-07	7.839e-07	7.914e-07	7.814e-07	7.751e-07	**-1.01**
Q700	**1.361e-06**	1.404e-06	1.412e-06	1.409e-06	1.419e-06	1.394e-06	1.392e-06	**-2.26**
Q850	**2.067e-06**	2.106e-06	2.105e-06	2.078e-06	2.144e-06	2.071e-06	2.069e-06	**-0.08**
Q925	1.136e-06	1.141e-06	1.139e-06	1.149e-06	1.159e-06	1.123e-06	**1.099e-06**	**2.18**
Q1000	7.790e-07	7.944e-07	7.901e-07	7.919e-07	8.009e-07	7.753e-07	**7.494e-07**	**3.34**
T50	2.90211	3.62793	3.51035	3.30476	3.65298	5.06522	**2.66920**	**8.03**
T100	2.34523	2.52353	2.45752	2.45173	2.58776	3.35917	**2.14464**	**8.55**
T150	2.28163	2.39949	2.37056	2.32649	2.47541	2.81440	**2.12157**	**7.02**
T200	3.25353	3.39821	3.40382	3.36980	3.45103	3.61290	**3.15386**	**3.06**
T250	2.91773	3.03071	2.97982	3.00203	3.01120	3.18168	**2.82814**	**3.07**
T300	2.11757	2.21766	2.15005	2.16068	2.19588	2.39245	**2.04242**	**3.55**
T400	2.35764	2.45795	2.40784	2.44162	2.42666	2.61125	**2.30280**	**2.33**
T500	2.48922	2.56503	2.53735	2.53340	2.53934	2.70371	**2.44910**	**1.61**
T600	2.41644	2.48341	2.47073	2.45911	2.47726	2.58165	**2.37705**	**1.63**
T700	2.42109	2.49055	2.49517	2.51035	2.48575	2.57838	**2.37855**	**1.76**
T850	3.17144	3.21729	3.22994	3.20866	3.22191	3.26888	**3.09818**	**2.31**
T925	2.92411	2.95436	2.95586	2.94624	2.95556	2.98035	**2.82213**	**3.49**
T1000	2.39374	2.39781	2.39948	2.39610	2.40353	2.43240	**2.21637**	**7.41**
U50	11.03250	11.84339	11.71563	11.79570	12.32634	13.53785	**9.83968**	**10.81**
U100	13.53727	13.81244	13.92218	13.83320	14.21210	15.45337	**12.79053**	**5.52**
U150	22.30717	22.45561	22.40940	22.54106	22.52762	23.11099	**21.60090**	**3.17**
U200	33.15385	33.56901	33.52859	33.59432	33.63162	33.79565	**32.75436**	**1.20**
U250	42.51323	43.03856	43.03693	42.92896	43.13520	43.21130	**42.21980**	**0.69**
U300	43.76355	44.26589	44.24840	44.04981	44.36844	44.35046	**43.49393**	**0.62**
U400	32.06457	32.43237	32.44376	32.14324	32.48634	32.51887	**31.90296**	**0.50**
U500	21.54208	21.73467	21.77371	21.58777	21.78301	21.77095	**21.36780**	**0.81**
U600	15.99065	16.11722	16.16757	16.06882	16.17312	16.12846	**15.82602**	**1.03**
U700	13.26648	13.35812	13.40670	13.33581	13.42657	13.37101	**13.10600**	**1.21**
U850	12.12437	12.17629	12.19511	12.12289	12.22564	12.18646	**11.96558**	**1.30**
U925	12.11921	12.17315	12.18269	12.09369	12.20308	12.17863	**11.98984**	**0.86**
U1000	7.20972	7.27277	7.27304	7.22095	7.29001	7.27609	**7.15388**	**0.77**
V50	7.69285	7.90130	7.77322	7.79335	7.92332	8.24515	**7.58965**	**1.34**
V100	11.83043	11.90321	11.88265	11.81035	11.96160	12.87497	**11.61879**	**1.62**
V150	22.11713	22.20185	22.17446	22.18359	22.29513	22.65495	**21.80277**	**1.42**
V200	34.97853	35.50271	35.40088	35.42003	35.62080	35.61859	**34.79561**	**0.52**
V250	46.15950	46.90562	46.79925	46.65352	47.04992	47.07571	**46.02083**	**0.30**
V300	47.97438	48.70294	48.60219	48.32908	48.85478	48.88685	**47.86474**	**0.23**
V400	35.00195	35.49229	35.47064	35.06309	35.57015	35.62241	**34.95459**	**0.14**
V500	23.08485	23.35130	23.37413	23.13610	23.40162	23.43702	**23.00809**	**0.33**
V600	16.72291	16.88487	16.92217	16.78395	16.93378	16.92696	**16.63893**	**0.50**
V700	13.47266	13.57959	13.62745	13.51284	13.62897	13.60267	**13.38688**	**0.64**
V850	12.23704	12.30452	12.33940	12.24323	12.33527	12.31612	**12.15260**	**0.69**
V925	12.89034	12.97775	13.00467	12.89169	12.99815	12.99973	**12.82380**	**0.52**
V1000	7.84872	7.92497	7.94184	7.88254	7.94293	7.93872	**7.81624**	**0.41**
T2M	2.19598	2.17321	2.18246	2.19267	2.18067	2.22968	**1.92042**	**11.63**
U10	5.95073	5.99814	5.99881	5.96021	6.00611	6.00079	**5.91091**	**0.67**
V10	6.39210	6.44803	6.46178	6.41535	6.45564	6.45554	**6.37177**	**0.32**
MSL	102714.972	102272.805	103303.174	102208.550	103938.146	103126.643	**99366.202**	**2.78**

Table D.10: MAE for all 69 output fields at a 72-hour lead time. Bold and underline indicate the lowest and second-lowest values, respectively. Gain is the relative MAE reduction of Ours over the best baseline for each field; negative values indicate that Ours has a larger error.

Field	δ-Adapter	Proceed	OneNet	FSNet	ADAPT-Z	ELF	Ours	Gain (%)
Z50	230.963	269.933	240.091	248.546	259.298	316.863	**205.915**	**10.85**
Z100	205.115	237.709	215.530	223.400	229.036	267.376	**186.958**	**8.85**
Z150	223.690	251.878	231.169	240.753	243.103	291.668	**209.195**	**6.48**
Z200	244.447	267.750	250.120	259.968	259.493	303.395	**232.414**	**4.92**
Z250	256.870	275.788	261.302	268.739	269.383	306.413	**247.945**	**3.47**
Z300	252.148	267.820	256.338	261.763	262.672	293.768	**245.182**	**2.76**
Z400	217.321	228.068	221.183	222.663	225.325	247.076	**212.673**	**2.14**
Z500	185.546	192.539	188.721	188.621	191.408	206.438	**181.575**	**2.14**
Z600	164.716	168.907	166.899	166.032	168.712	178.485	**160.916**	**2.31**
Z700	152.784	154.432	153.857	152.552	154.539	160.263	**148.859**	**2.42**
Z850	147.400	146.760	147.201	146.403	147.282	148.742	**142.832**	**2.44**
Z925	150.626	149.398	150.068	149.594	150.058	150.536	**145.755**	**2.44**
Z1000	156.686	155.158	155.821	155.617	155.833	155.711	**151.383**	**2.43**
Q50	4.742e-08	5.614e-08	5.031e-08	5.543e-08	5.601e-08	6.230e-08	**4.129e-08**	**12.93**
Q100	1.633e-07	1.683e-07	1.689e-07	1.700e-07	1.720e-07	1.804e-07	**1.617e-07**	**0.96**
Q150	1.166e-06	1.159e-06	1.153e-06	1.166e-06	1.164e-06	1.175e-06	**1.125e-06**	**2.41**
Q200	**7.136e-06**	7.192e-06	7.158e-06	7.234e-06	7.197e-06	7.174e-06	7.141e-06	**-0.06**
Q250	**2.603e-05**	2.622e-05	2.615e-05	2.645e-05	2.621e-05	2.613e-05	2.632e-05	**-1.09**
Q300	**6.070e-05**	6.110e-05	6.104e-05	6.188e-05	6.104e-05	6.094e-05	6.132e-05	**-1.02**
Q400	**1.827e-04**	1.850e-04	1.848e-04	1.855e-04	1.854e-04	1.848e-04	1.849e-04	**-1.21**
Q500	**3.658e-04**	3.697e-04	3.701e-04	3.721e-04	3.700e-04	3.696e-04	3.698e-04	**-1.09**
Q600	**5.583e-04**	5.641e-04	5.646e-04	5.643e-04	5.657e-04	5.624e-04	5.637e-04	**-0.97**
Q700	**7.609e-04**	7.726e-04	7.746e-04	7.743e-04	7.762e-04	7.702e-04	7.717e-04	**-1.42**
Q850	**9.348e-04**	9.438e-04	9.426e-04	9.401e-04	9.510e-04	9.372e-04	9.375e-04	**-0.28**
Q925	6.918e-04	6.940e-04	6.922e-04	6.970e-04	6.981e-04	6.898e-04	**6.807e-04**	**1.33**
Q1000	5.941e-04	6.001e-04	5.974e-04	5.974e-04	6.018e-04	5.912e-04	**5.796e-04**	**1.97**
T50	1.26387	1.42320	1.37398	1.35316	1.42027	1.71815	**1.19930**	**5.11**
T100	1.16548	1.21419	1.19516	1.20014	1.23613	1.42328	**1.10687**	**5.03**
T150	1.10214	1.13496	1.12083	1.11440	1.15180	1.23021	**1.04752**	**4.96**
T200	1.20406	1.25023	1.24000	1.24554	1.25660	1.30177	**1.17392**	**2.50**
T250	1.13560	1.17697	1.14988	1.17504	1.16567	1.23048	**1.11156**	**2.12**
T300	1.00936	1.04747	1.01800	1.02946	1.03811	1.10460	**0.98810**	**2.11**
T400	1.04238	1.07876	1.05566	1.07681	1.06679	1.12805	**1.02834**	**1.35**
T500	1.04423	1.07044	1.05701	1.06375	1.06057	1.11645	**1.03433**	**0.95**
T600	1.03887	1.06258	1.05297	1.05576	1.06016	1.09910	**1.02689**	**1.15**
T700	1.03460	1.05969	1.05326	1.07412	1.05702	1.09607	**1.02203**	**1.22**
T850	1.16964	1.18101	1.17905	1.18532	1.18190	1.20408	**1.15234**	**1.48**
T925	1.08505	1.09203	1.08965	1.09161	1.09304	1.10734	**1.05950**	**2.35**
T1000	0.98289	0.97892	0.97692	0.98264	0.98356	0.99960	**0.93581**	**4.21**
U50	2.50889	2.59607	2.58046	2.59523	2.65418	2.70667	**2.36952**	**5.56**
U100	2.78646	2.81785	2.82872	2.82128	2.85835	2.89477	**2.71087**	**2.71**
U150	3.54572	3.56231	3.55899	3.56826	3.56851	3.59456	**3.49081**	**1.55**
U200	4.26697	4.29612	4.29469	4.29897	4.30096	4.31027	**4.23688**	**0.71**
U250	4.74758	4.78103	4.78240	4.77992	4.78692	4.79308	**4.72717**	**0.43**
U300	4.72536	4.75729	4.75647	4.75476	4.76297	4.76536	**4.70741**	**0.38**
U400	3.98252	4.01033	4.01073	4.00039	4.01472	4.02249	**3.97154**	**0.28**
U500	3.30303	3.32089	3.32452	3.31504	3.32646	3.32625	**3.28691**	**0.49**
U600	2.88077	2.89589	2.90088	2.89570	2.90255	2.89742	**2.86331**	**0.61**
U700	2.62192	2.63397	2.63907	2.63394	2.64308	2.63568	**2.60214**	**0.75**
U850	2.40640	2.41389	2.41474	2.40921	2.42112	2.41554	**2.38410**	**0.93**
U925	2.26099	2.26740	2.26698	2.26055	2.27267	2.26845	**2.24197**	**0.82**
U1000	1.73299	1.74229	1.74145	1.73669	1.74569	1.74324	**1.72304**	**0.57**
V50	2.09357	2.12075	2.09960	2.10609	2.12316	2.15060	**2.07591**	**0.84**
V100	2.61623	2.62401	2.62135	2.61560	2.63046	2.68085	**2.59251**	**0.88**
V150	3.51686	3.52625	3.52377	3.52570	3.53406	3.55103	**3.49256**	**0.69**
V200	4.32350	4.35876	4.35269	4.35399	4.36692	4.36660	**4.31284**	**0.25**
V250	4.83525	4.88003	4.87412	4.86834	4.88860	4.88939	**4.82995**	**0.11**
V300	4.81024	4.85414	4.84709	4.84187	4.86257	4.86478	**4.80840**	**0.04**
V400	**4.02400**	4.05841	4.05552	4.03869	4.06509	4.06866	4.02551	**-0.04**
V500	3.30518	3.32847	3.32934	3.31709	3.33456	3.33657	**3.30189**	**0.10**
V600	2.85626	2.87408	2.87672	2.86776	2.88042	2.87802	**2.85018**	**0.21**
V700	2.56582	2.57938	2.58380	2.57436	2.58659	2.58183	**2.55773**	**0.32**
V850	2.35467	2.36416	2.36677	2.35912	2.36993	2.36560	**2.34501**	**0.41**
V925	2.27804	2.28871	2.29089	2.28142	2.29352	2.29182	**2.26999**	**0.35**
V1000	1.78979	1.80023	1.80218	1.79615	1.80472	1.80235	**1.78469**	**0.28**
T2M	0.92252	0.90641	0.91005	0.91574	0.91280	0.93746	**0.85188**	**6.02**
U10	1.56282	1.57094	1.57027	1.56629	1.57336	1.57186	**1.55502**	**0.50**
V10	1.61719	1.62632	1.62792	1.62220	1.62945	1.62808	**1.61357**	**0.22**
MSL	196.520	194.649	195.499	195.510	195.883	195.275	**190.096**	**2.34**

D.2 GATING WEIGHT RESULTS



Figure D.1: Temporal evolution of the routing weights for geopotential (Z50–Z1000) and upper-level specific humidity (Q50–Q850). Blue and orange curves show the 365-day rolling means for the 00 and 12 UTC initialization streams, respectively.



Figure D.2: Temporal evolution of the routing weights for lower-level specific humidity (Q925 and Q1000), temperature (T50–T1000), and upper-level zonal wind (U50–U600). Blue and orange curves show the 365-day rolling means for the 00 and 12 UTC initialization streams, respectively.

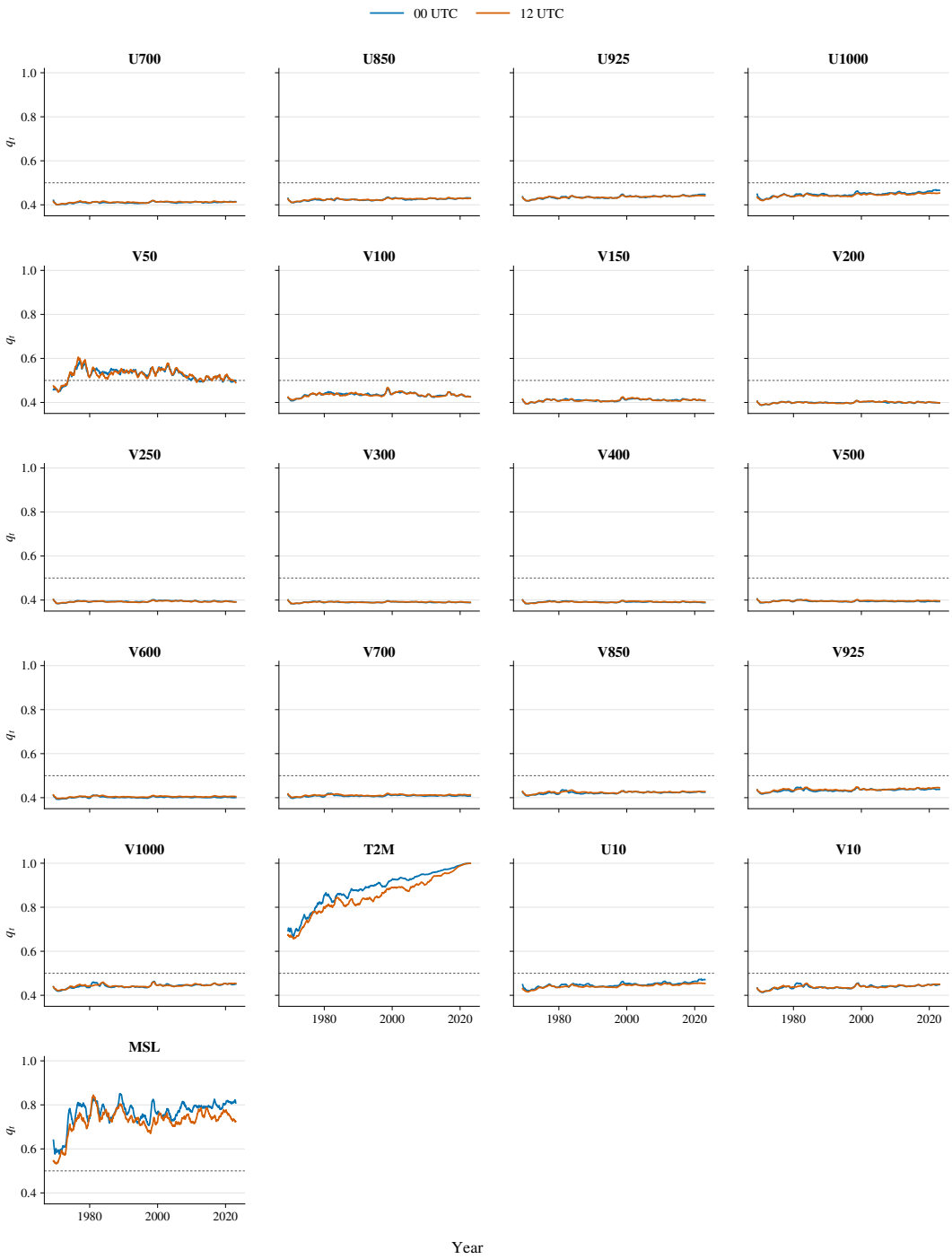


Figure D.3: Temporal evolution of the routing weights for lower-level zonal wind (U700–U1000), meridional wind (V50–V1000), and surface fields (T2M, U10, V10, and MSL). Blue and orange curves show the 365-day rolling means for the 00 and 12 UTC initialization streams, respectively.

D.3 CASE-STUDY VISUALIZATIONS

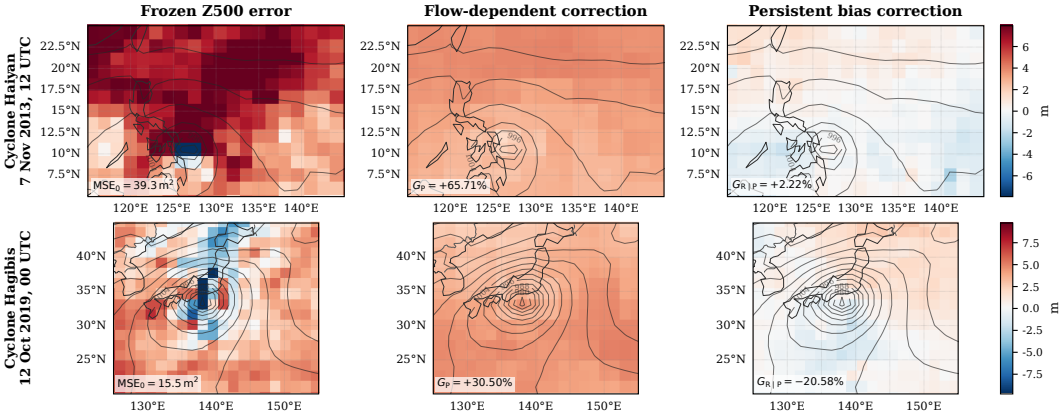


Figure D.4: **Decomposition of Z500 forecast-error correction for two tropical-cyclone cases.** Rows show forecasts initialized during Typhoon Haiyan (7 November 2013, 12 UTC) and Typhoon Hagibis (12 October 2019, 00 UTC). The frozen-model MSE and the relative contributions of the flow-dependent (G_{P}) and persistent ($G_{\mathrm{R|P}}$) corrections are reported in each panel.

D.4 VARIABLE-WISE CLIMATE-SHIFT AND ADAPTATION ANALYSIS

To examine whether the temporal-drift behavior observed for T2M extends to other atmospheric variables, we repeat the analysis for Z500, T850, U850, V850, U10, V10, and MSL.

Figures D.5 and D.6 show the clearest temporal degradation. For Z500, climate shift is strongly associated with both the MSE of Frozen Aurora ($r = 0.87$) and the MSE reduction achieved by our method ($r = 0.96$). T850 exhibits similarly strong relationships, with $r = 0.89$ for frozen-model error and $r = 0.92$ for adaptation gain. The period-averaged results further show that the errors of Frozen Aurora increase markedly in the later evaluation periods, whereas online adaptation substantially limits this degradation.

The wind variables exhibit weaker relationships between climate shift and frozen-model error, but more consistent relationships between shift and adaptation gain. As shown in Figures D.8, D.7, D.11, and D.9, the correlations with frozen-model MSE range from 0.40 to 0.58, while the correlations with adaptation gain range from 0.68 to 0.83. Although the absolute improvements for the wind fields are smaller than those for Z500 and T850, the increasing adaptation gains indicate that online updating becomes more useful as their deployment distributions move farther from the training climate.

MSL presents a different pattern (Figure D.10). Its climate distribution shift has almost no association with frozen-model error ($r = 0.06$), although it remains moderately associated with adaptation gain ($r = 0.55$). Unlike temperature and geopotential, MSL errors depend strongly on the spatial organization and displacement of synoptic pressure systems. Consequently, a global marginal-distribution distance may not fully capture the circulation-related changes that govern MSL forecast errors. This result highlights that the relationship between climate shift and forecast degradation is variable dependent and that no single marginal drift measure is expected to describe every atmospheric field equally well.

Overall, the additional variables support the broader conclusion that forecast degradation and the benefit of online adaptation generally increase under stronger distribution shift. However, the magnitude of this relationship depends on the physical characteristics of each variable and on whether the selected drift measure captures its relevant spatial and dynamical structure.

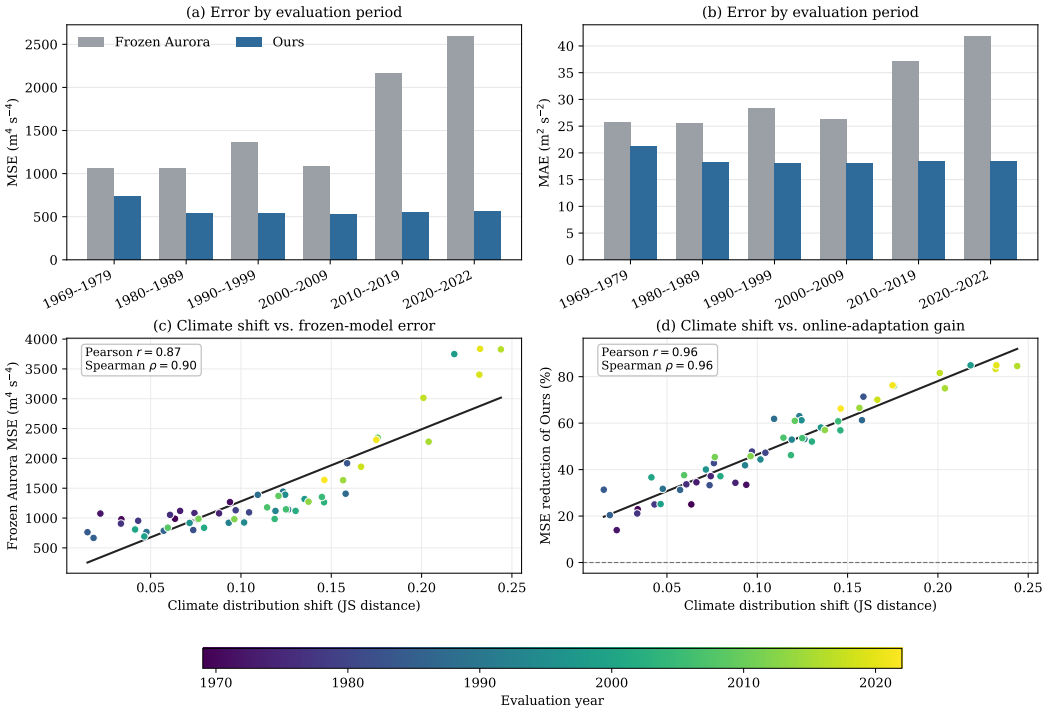


Figure D.5: Climate-shift analysis for Z500.

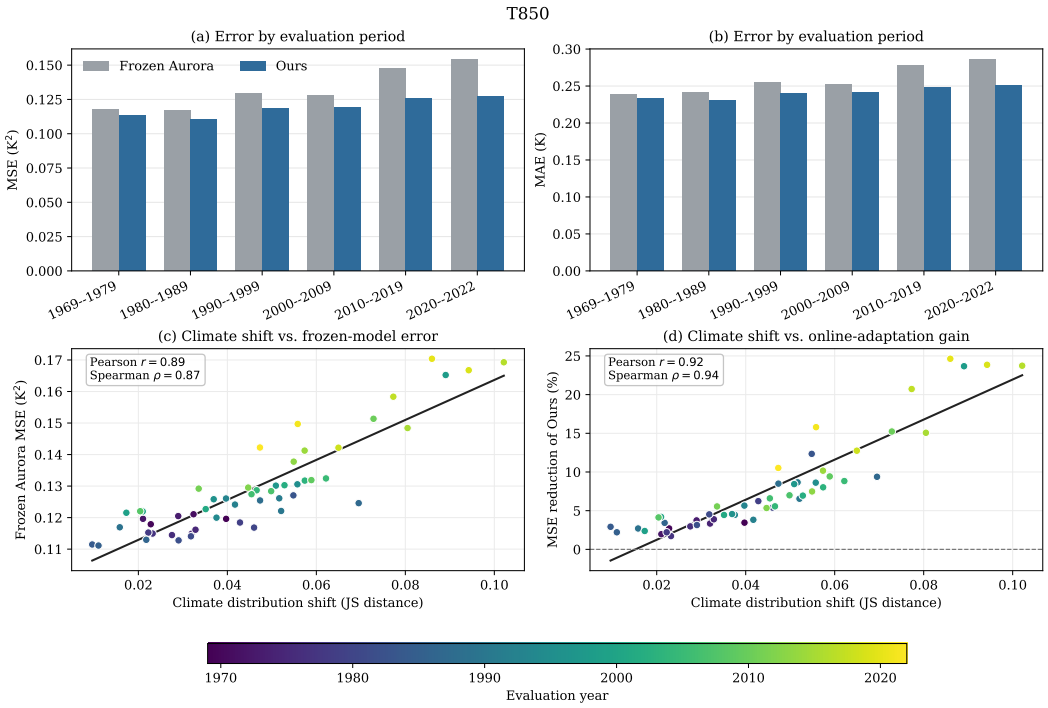


Figure D.6: Climate-shift analysis for T850.

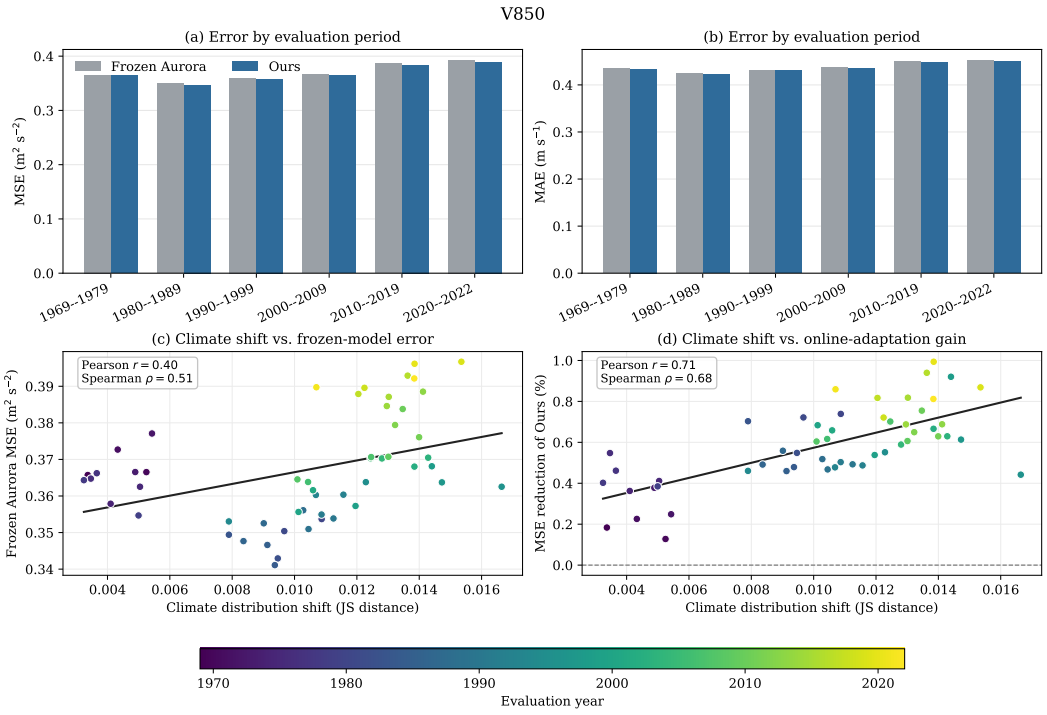


Figure D.7: Climate-shift analysis for V850.

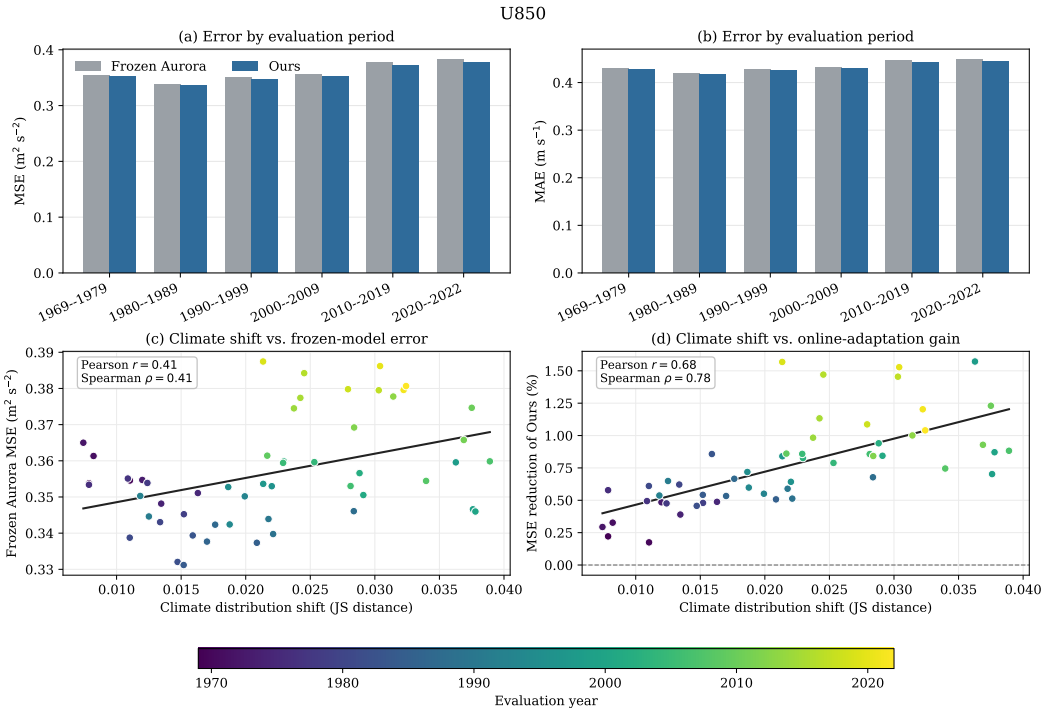


Figure D.8: Climate-shift analysis for U850.

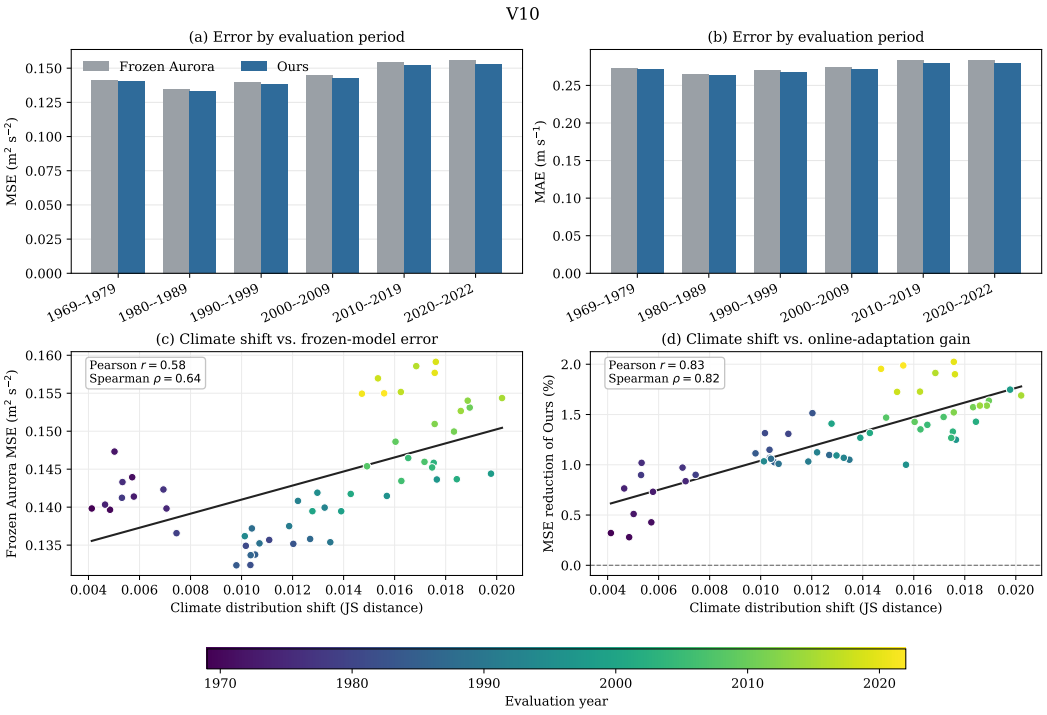


Figure D.9: Climate-shift analysis for V10.

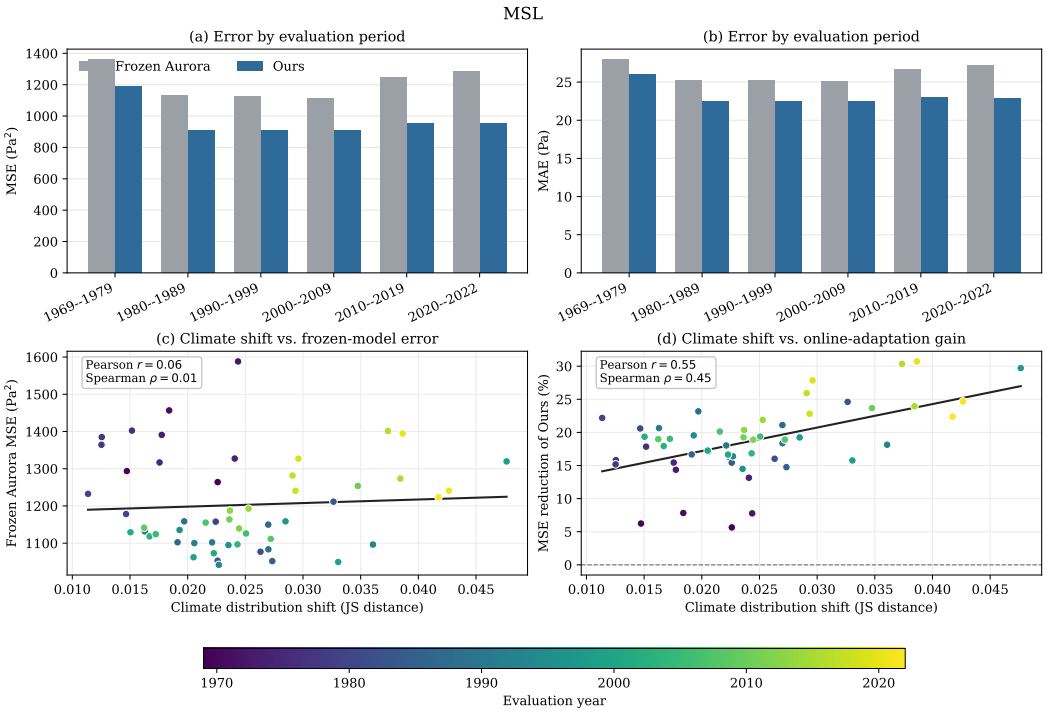


Figure D.10: Climate-shift analysis for MSL.

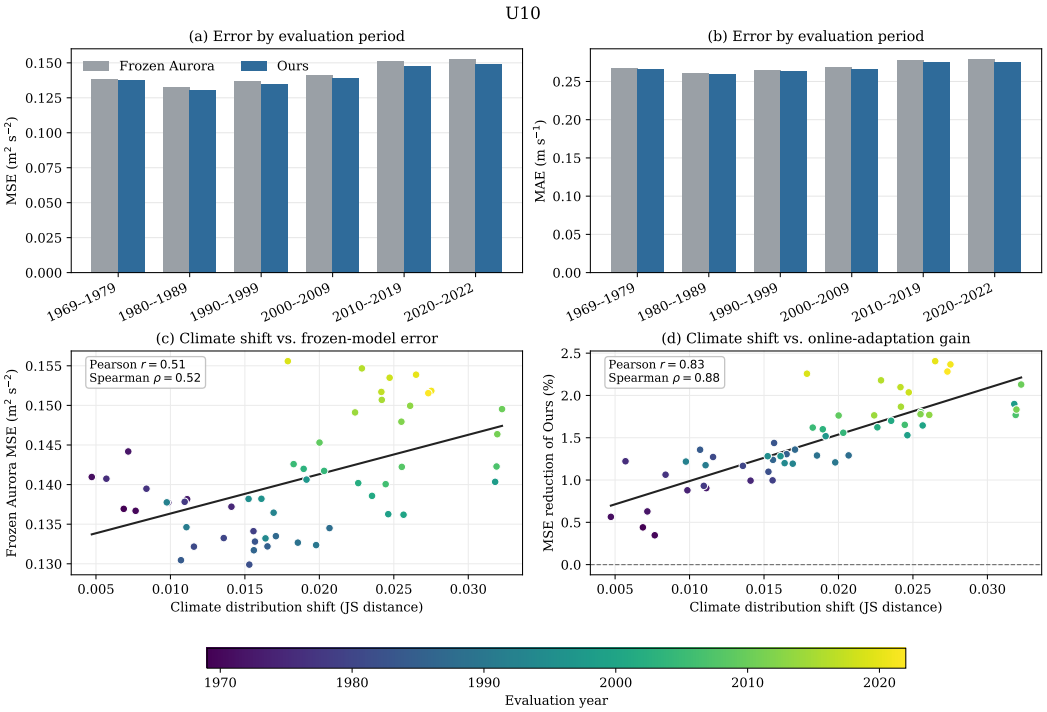


Figure D.11: Climate-shift analysis for U10.

D.5 Ablation of the two correction components

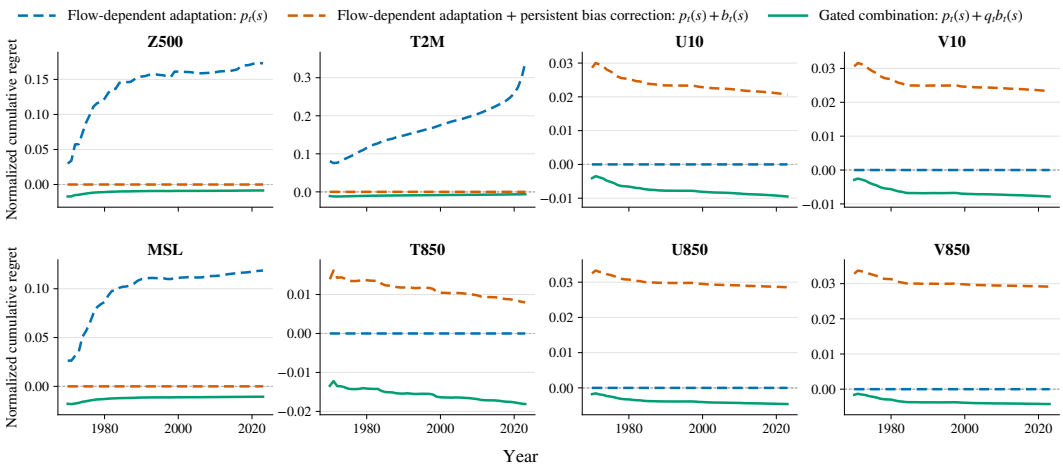


Figure D.12: Normalized cumulative regret of the adaptation endpoints and routed forecast relative to the best fixed endpoint in hindsight. Negative regret indicates that the routed mixture outperforms the best fixed endpoint over the corresponding period.