DATA-ANCHORED RESPONSE SPACES FOR CLOSED-LOOP NEURAL COMPONENTS IN NON-DIFFERENTIABLE NUMERICAL SIMULATORS

Anonymous authors
Paper under double-blind review

ABSTRACT

Learning unresolved physics with neural networks can improve climate and other scientific models, but good offline accuracy does not guarantee reliable closed-loop behavior. Existing approaches use physical constraints or differentiable simulators to improve long-horizon performance, yet the former depends on predefined constraints and the latter is difficult to apply to legacy non-differentiable solvers. We introduce DARS, a data-anchored response representation that anchors learned outputs to mixtures of multivariate response prototypes recovered from training trajectories. A learnable router determines the state-dependent combination, while short differentiable surrogate rollouts updates only the router through short differentiable surrogate rollouts, without gradients through the original simulator. Despite comparable offline accuracy, before adaptation DARS completes the full 72 h rollout across three matched training seeds, whereas unrestricted readouts fail after 15.0–24.3 h and no componentwise limiter reaches the horizon. This shows that limiting individual amplitudes is insufficient to reproduce the benefit of preserving joint response structure. Separately, an adapted, frozen deployment schedule reduces full-trajectory mean (0–72 h) specific-humidity and temperature RMSE relative to its offline parent across 22 unseen initial conditions by $31.4\% \pm 2.0\%$ and $61.2\% \pm 7.2\%$, respectively. These results highlight learned-response representation as a key factor in reliable long-horizon solver–AI modeling.

1 INTRODUCTION

Learned components embedded in numerical simulators face a recurring mismatch between offline prediction and coupled deployment. Once deployed, their outputs modify the simulated state and therefore the inputs encountered later. This feedback can shift the model away from its training distribution and produce biases or instabilities that are not apparent from offline accuracy alone. Such discrepancies have been reported in learned atmospheric parameterizations (Brenowitz & Bretherton, 2018; 2019; Brenowitz et al., 2020). Our experiments illustrate the same issue: models with similar offline accuracy and short-horizon adaptation success can behave very differently over longer coupled integrations.

Existing approaches address different parts of this problem. Physical constraints restrict undesirable outputs, coupled or multistep training exposes models to deployment-induced states, and differentiable hybrid models propagate feedback through the numerical dynamics (Yuval & O'Gorman, 2020; Rasp, 2020; Kochkov et al., 2024). When the production solver cannot be differentiated end to end, learned surrogates can instead provide short-horizon gradients while the original simulator remains responsible for coupled evaluation (Frezat et al., 2026). These approaches improve physical consistency, deployment exposure, or access to optimization signals. Less explicit is how the *multivariate response space itself* should be represented while the learned component adapts under feedback.

This distinction matters because a conventional neural readout produces a free multivariate response vector. Componentwise bounds can suppress extreme amplitudes but do not preserve relationships among variables and vertical levels. The solver, however, repeatedly receives and reacts to the complete response profile. We therefore treat *joint response structure as a distinct design problem*, complementary to amplitude control and closed-loop optimization.

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.

To address this problem, we introduce the **data-anchored response space (DARS)**. Using training trajectories and known resolved dynamics, DARS constructs transition-consistent multivariate response prototypes. In our atmospheric instantiation, each prototype represents a joint temperature and moisture tendency profile. Predictions combine training-derived responses; sequential adaptation holds them fixed and updates their state-dependent weights. DARS thus separates *which responses are available* from *how they are selected*. Short differentiable surrogate rollouts adapt this selection rule, while the original simulator remains authoritative for coupled evaluation.

We compare the three offline readouts at $k = 0$, then examine sequential adaptation with seed-dependent limiter completion. At one initialization, DARS and unrestricted readouts have similar offline accuracy. Before adaptation, DARS completes the full 72 h rollout across all three matched training seeds, while unrestricted readouts fail after 15.0–24.3 h and no componentwise limiter reaches the horizon. This separation indicates that controlling individual amplitudes is not sufficient to reproduce the long-horizon behavior obtained by preserving joint response structure. Separately, a frozen deployment schedule reduces full-trajectory mean specific-humidity and temperature field RMSE across 22 initializations not used for adaptation by $31.4\% \pm 2.0\%$ and $61.2\% \pm 7.2\%$ relative to its offline parent. Together, these results formulate response-space construction as an underexplored component of closed-loop solver–AI design, provide DARS as a concrete realization of this idea, and show that structured response spaces can improve long-horizon behavior while retaining adaptation and transfer beyond the conditions used for tuning.

2 Related Work

Closed-loop learning in hybrid scientific models. Learned components can behave differently after deployment because their outputs modify the simulated state and therefore future inputs. Yuval and O'Gorman (Yuval & O'Gorman, 2020) demonstrated stable coupled atmospheric simulations with a random-forest parameterization, while Rasp (Rasp, 2020) used coupled online learning to expose a model to deployment-induced states. Kochkov et al. (Kochkov et al., 2024) instead optimize learned physics within a differentiable atmospheric core. These works address how learned components are trained or constrained under feedback; DARS elevates the multivariate response space from an implicit consequence of the readout to an explicit object of design in closed-loop adaptation.

Surrogate-assisted adaptation for non-differentiable simulators. When the production simulator is not differentiable, learned dynamics can provide short virtual rollouts for optimization. SLBO and MBPO use this model-based strategy while controlling accumulated model error (Luo et al., 2019; Janner et al., 2019). Frezat et al. (Frezat et al., 2026) similarly use a differentiable emulator to optimize a learned parameterization while retaining the original scientific solver for evaluation. Our surrogate-assisted adaptation follows this established pattern; DARS instead changes the set of responses over which such adaptation operates.

Data-anchored and structured response representations. Sample-based predictors such as random forests and kernel methods construct outputs from training-associated responses (Breiman, 2001; Nadaraya, 1964), while RBF networks and mixture-of-experts models provide smooth state-dependent combinations of localized components (Schwenker et al., 2001; Jacobs et al., 1991). These ideas motivate data anchoring and structured combination, but do not explicitly address how a multivariate response space should be preserved during repeated adaptation inside a numerical solver. Smooth routing combines training-derived response atoms according to the current state; in the sequential adaptation study, updates change only the routing rule and keep the response basis fixed.

3 Methodology

DARS is designed to provide a structured response space within which repeated adaptation can proceed without unrestricted output extrapolation (Figure 1). Using reference transitions and the fixed GRIST backbone (Zhang et al., 2020), we construct complete joint temperature and moisture response atoms. Smooth routing selects mixtures of these atoms; in the matched-readout study, adaptation updates routing while preserving the response basis.

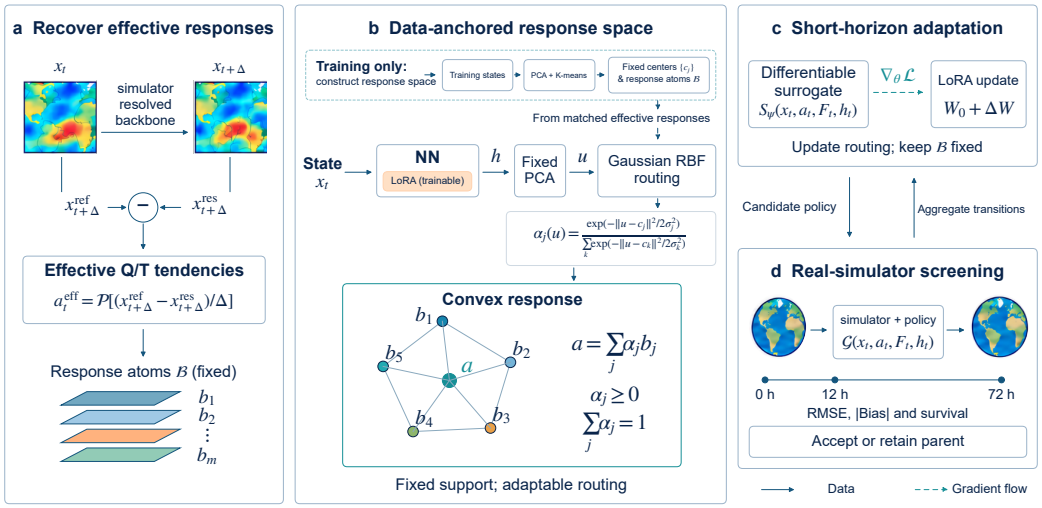

Figure 1: **DARS workflow.** (a) Reference transitions and the resolved GRIST backbone define transition-consistent moisture and temperature responses. (b) A learned representation and fixed input anchors define Gaussian routing over complete response atoms; the evaluated DARS readout uses anchor-specific anisotropic metrics. (c) Short differentiable surrogate rollouts guide low-rank encoder updates while the response dictionary remains fixed. (d) Exported checkpoints are evaluated through uninterrupted GRIST integration. Policy-update windows and surrogate replay have different time horizons (Appendix O).

3.1 Closed-Loop Hybrid Simulation

DARS leaves the numerical simulator unchanged and adapts only its learned physical component. Let $\mathbf{a}_t$ denote the response, π_θ the state-dependent response model, $\mathbf{z}_t$ the packed model input, $\mathbf{x}_t$ the simulator state, $\mathbf{F}_t$ the forcing, and $\mathbf{h}_t$ the solver history. The coupled system evolves as

$$\mathbf{a}_t = \pi_\theta(\mathbf{z}_t), \qquad \mathbf{x}_{t+\Delta} = G(\mathbf{x}_t, \mathbf{a}_t, \mathbf{F}_t, \mathbf{h}_t), \tag{1}$$

where G is the original, non-differentiable simulator. In the atmospheric model instance, $\mathbf{z}_t \in \mathbb{R}^{151}$ and $\mathbf{a}_t \in \mathbb{R}^{60}$ contains 30 specific-humidity tendencies and 30 temperature tendencies. The coupling interval is $\Delta = 1200\,\mathrm{s}$. Evaluation targets are ERA5 reanalysis states (Hersbach et al., 2020), denoted "Truth" in diagnostic figures rather than direct observations.

During coupled evaluation, candidates advance from their own states; reference states are never written back. The simulator supplies authoritative trajectories and coupled evaluation. Further input and evaluation details are in Appendix B.

3.2 Transition-Consistent Effective Responses

Reference trajectories contain successive simulator states but do not directly label the unresolved physical response applied between them. Let $D_\Delta(\mathbf{x}_t, \mathbf{F}_t, \mathbf{h}_t)$ denote the resolved-state advance over one coupling interval when the learned response is omitted. For a reference transition $\mathbf{x}_t^{\mathrm{ref}} \to \mathbf{x}_{t+\Delta}^{\mathrm{ref}}$, we construct

$$\mathbf{a}_t^{\mathrm{eff}} = \mathcal{P}\left[\frac{\mathbf{x}_{t+\Delta}^{\mathrm{ref}} - D_\Delta(\mathbf{x}_t^{\mathrm{ref}}, \mathbf{F}_t, \mathbf{h}_t)}{\Delta}\right], \tag{2}$$

where $\mathcal{P}$ selects and converts the residual into the 30-level moisture and temperature tendencies represented by the learned component. The same forcing and solver-history inputs used by the reference transition are supplied to D_Δ.

We call $\mathbf{a}_t^{\mathrm{eff}}$ a *transition-consistent response*, rather than a uniquely observed physical action. It is the response that agrees with the observed transition under the resolved backbone used in the construction. Only the training partition is used to construct the response dictionary; the dataset definition and replay-check protocol are given in Appendix B.

3.3 Data-Anchored Response Space

3.3.1 State representation and response dictionary

The policy maps each packed input through a residual encoder E_θ to a 128-dimensional representation,

$$\mathbf{h}_i = E_\theta(\mathbf{z}_i). \tag{3}$$

The evaluated DARS implementation uses $K = 4096$ training-derived input anchors $\{\mathbf{c}_j\}_{j=1}^{K}$, each paired with a complete 60-dimensional response atom $\mathbf{b}_j$. The response dictionary is

$$\mathcal{B} = \{\mathbf{b}_1, \ldots, \mathbf{b}_K\}. \tag{4}$$

During sequential adaptation, the input anchors and response atoms remain fixed. Queries and anchors are encoded by the same E_θ, so an encoder update changes both their embedding coordinates. No 64-dimensional PCA projection is used in this deployed routing path. Network settings are given in Appendix N; diagnostic PCA projections are separate from the deployed readout.

3.3.2 Normalized Anisotropic Gaussian Routing

For the evaluated DARS readout, let $\boldsymbol{\delta}_j = E_\theta(\mathbf{z}) - E_\theta(\mathbf{c}_j)$. Anchor j receives the Gaussian response

$$\phi_j(\mathbf{z}) = \exp\Big[-\tfrac{1}{2}\boldsymbol{\delta}_j^\top M_j \boldsymbol{\delta}_j\Big], \quad M_j = D_j + U_j U_j^\top, \tag{5}$$

where D_j is a positive diagonal inverse-width matrix and $U_j \in \mathbb{R}^{128\times 16}$. These metric tensors are frozen during adaptation. The ideal normalized weights and response are

$$\alpha_j(\mathbf{z}) = \frac{\phi_j(\mathbf{z})}{\sum_{m=1}^{K}\phi_m(\mathbf{z})}, \qquad \pi_\theta(\mathbf{z}) = \sum_{j=1}^{K}\alpha_j(\mathbf{z})\mathbf{b}_j, \tag{6}$$

with $\alpha_j(\mathbf{z}) \geq 0$ and $\sum_j \alpha_j(\mathbf{z}) = 1$. Consequently, under exact normalization,

$$\pi_\theta(\mathbf{z}) \in \mathcal{H}, \qquad \mathcal{H} = \mathrm{conv}(\mathcal{B}). \tag{7}$$

The dictionary determines which complete responses are available, while the router determines their state-dependent mixture. With $S = \sum_j \phi_j$ and $\epsilon = 10^{-12}$, the matched-readout study implements $\widetilde{\alpha}_j = \phi_j/(S+\epsilon)$ rather than $\alpha_j = \phi_j/S$. For $S > 0$, its standardized response is $\widetilde{\pi} = [S/(S+\epsilon)]\pi$, a contraction of the ideal mixture toward the standardized zero response. It therefore belongs to $\mathrm{conv}(\mathcal{B} \cup \{0\})$, not necessarily $\mathrm{conv}(\mathcal{B})$. After output de-standardization, this extra vertex is the output normalization mean, not zero physical tendency. If all kernels underflow to zero, the implementation returns that mean. Shrinkage is small only when $S \gg \epsilon$; this is not assumed uniformly over coupled states. The rank-16 metric factor is distinct from the rank-8 adaptation rank.

3.4 Representation-Level Properties

The fixed response hull provides a round-independent bound on the available outputs: ideal routing changes their mixture, not the response class. With epsilon normalization, the corresponding bound uses the hull augmented by the standardized zero response (Appendix J). This is the mechanism we test as a foundation for stable long-horizon coupling and successive refinement. It bounds the response representation, not the complete simulator's trajectory; coupled stability and monotonic metric improvement remain empirical results.

3.5 Surrogate-Assisted Closed-Loop Adaptation

Because G is non-differentiable, a learned surrogate S_ψ supplies short-rollout gradients. Only the initial state comes from an archive; subsequent virtual states are generated by S_ψ. The matched-readout study uses six 20-minute virtual leads and optimizes scaled Q/T mean-square errors over the first four leads, with a squared adapter-update penalty. The surrogate is frozen during this optimization; the full schedule is in Appendix O.

In the matched-readout study, parent weights are frozen and rank-8 LoRA adapters (Hu et al., 2022) update three encoder linear layers (22,712 parameters) in DARS and the unrestricted arm. DARS

input anchors, response atoms, and metric tensors remain fixed; their encoded coordinates change with the encoder. Limiter adaptation coverage is seed-dependent: seed 1117 stops in round 1, seed 2027 in round 3, and seed 3037 completes all three rounds. The limiter is therefore not a complete matched three-seed adaptation arm.

3.6 SIMULATOR-IN-THE-LOOP REFINEMENT

Starting from the offline checkpoint $k = 0$, each adaptation round refines an arm-specific surrogate, updates encoder adapters through short virtual rollouts, and evaluates the exported checkpoint in an uninterrupted original-simulator rollout. We evaluate two distinct outcomes: whether every stage retains 72 h survival, and whether a fixed full-horizon error metric decreases between successive stages. The launcher accumulates valid rollout prefixes rather than acceptance-filtered replay. Policy-update diagnostics are confined to the first 12 h, whereas surrogate replay in the sequential adaptation study may include later states; the separate 22-IC deployment uses the configuration below.

The simulator remains fixed and supplies the deployment trajectories; checkpoints are frozen during evaluation. The matched-readout study refines its surrogate between rounds and updates only the selected encoder adapters during response-model optimization. Numerical survival and accuracy are evaluated in the original simulator, not inferred from virtual-rollout loss.

The separate 22-IC transfer study uses adaptation data only from the first 12 h of the 30 September 1988 initialization; longer trajectories are not used for tuning or deployment selection. Unlike the sequential adaptation study's router-only updates, this system includes response-basis adapters and a prescribed uniform temperature-tendency correction. The resulting request-indexed deployment schedule is frozen and reused on 22 other initializations without per-case retraining. Its pre-exported model sequence covers the first 12 h; a 20-minute transition removes the added uniform temperature tendency, after which a fixed continuation model performs state-dependent neural inference. All coupled rollouts remain uninterrupted, with no reference-state replacement. Here, 12 h denotes the adaptation-data window, not training compute time or the extent of offline pretraining. Transfer across initializations and improvement beyond this window are evaluated separately from the matched-readout comparison.

4 EXPERIMENTS AND RESULTS

4.1 EXPERIMENTAL DESIGN AND EVIDENCE HIERARCHY

The experiments follow a stability-to-adaptation progression. First, the matched offline $k = 0$ comparison tests whether readout structure supplies a stable long-horizon starting point before fine-tuning. Second, $k = 1$–3 test whether DARS retains that survival while successive updates improve accuracy. Third, a separately selected 12 h-adapted deployment tests cross-initialization transfer and improvement beyond its adaptation window. These complementary tests distinguish the foundation supplied by the readout from subsequent adaptation gains; the 22-IC study is not a repeated three-readout ablation. The Native-physics baseline uses the Weather Physics (PhysW) suite (Li et al., 2023); vertical and spectral diagnostics assess residual physical fidelity, not the causal mechanism of survival differences. The tuning initialization is 00:00 UTC on 30 September 1988; a 72-hour run has 216 twenty-minute transitions and 217 states. The matched-readout study uses training seeds 1117, 2027, and 3037. Stage $k = 0$ is the offline checkpoint; $k = 1, 2, 3$ are sequential updates within each seed, not additional independent trials. Each rollout freezes its checkpoint. DARS and unrestricted adaptation share the prescribed budget and adapter locations but refine separate surrogates. Limiter coverage differs by seed: 1117 completes no update, 2027 completes rounds 1–2, and 3037 completes rounds 1–3.

The long-range configuration applies a pre-specified post-12-hour policy schedule without restarting the simulator. The prognostic state, solver history, and integration clock remain continuous. The exact tail-policy definition and its use in each comparison arm are recorded in the experiment manifest.

For Q and T, native-grid area-weighted RMSE is the primary field-error measure, alongside signed bias; spectrally filtered quantities are used only to diagnose residual structure. Exact metric definitions are in Appendix C.

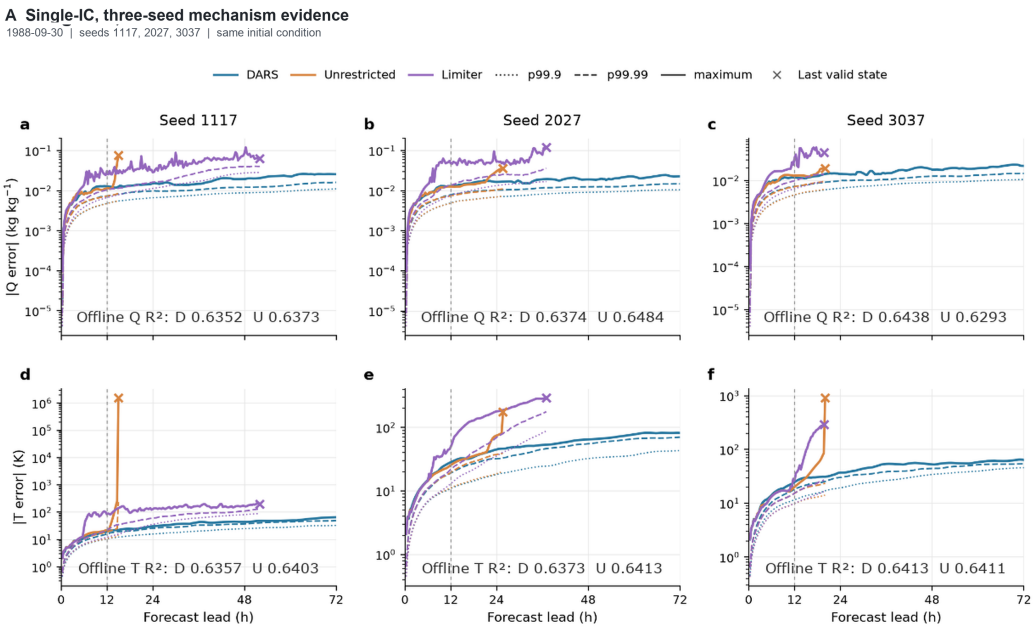

Figure 2: Single-initialization, three-seed trajectory diagnostics. Crosses mark the last archived valid state. Displayed checkpoints are not all budget matched; Table 1 reports the strict pre-adaptation $k = 0$ comparison. Adapted endpoints and limiter variants are descriptive, not three completed matched adaptation histories. The full figure is in Appendix 5.

4.2 Stable coupling as a foundation for successive adaptation

Table 1: Strictly matched $k = 0$ gate-survival time (h).

Seed	DARS	Unrestricted	Limiter
1117	72.0	15.0	52.3
2027	72.0	24.3	40.3
3037	72.0	19.3	19.3

Each arm uses its offline-trained $k = 0$ model, frozen throughout integration and not online-adapted. The limiter adds a post-hoc clamp to the corresponding unrestricted model. Survival is the archived valid-row count times the coupling interval, capped at 72 h, not an inferred exact crash time.

Offline validation shows similar DARS and unrestricted accuracy: the largest seed-level moisture or temperature R^2 difference is 0.0145 (Appendix C). The strictly matched survival comparison below uses the $k = 0$ offline models, not checkpoints obtained by passing an adaptation gate.

At $k = 0$, DARS completes 72 h for all three seeds, whereas unrestricted survival is 15.0, 24.3, and 19.3 h and limiter survival is 52.3, 40.3, and 19.3 h (Table 1). The advantage is already present before online updates: the structured readout provides an empirically stable starting point, rather than stability being created by fine-tuning. DARS then retains 72 h survival at every $k = 1$–3 checkpoint across the same three seeds (12 completed rollouts including $k = 0$). Unrestricted checkpoints never exceed 26 h. Limiter adaptation is incomplete for two seeds, and no available limiter rollout reaches 72 h. These stages are repeated measurements within seed histories.

Stable survival permits full-horizon refinement. For DARS seed 1117, full-trajectory (0–72 h) mean Q RMSE decreases monotonically across $k = 0, 1, 2, 3$: $1.1518, 1.1402, 1.1268, 1.1175$ g kg^{-1}. Corresponding T RMSE is $3.5548, 3.5187, 3.4854, 3.4564$ K; mean absolute signed biases also decrease at each stage. The other seeds show lower final full-horizon T RMSE (2027: $3.790 \rightarrow 3.675$ K; 3037: $4.118 \rightarrow 3.947$ K). Short-window improvement alone is insufficient: unrestricted seed 1117 improves its 0–12 h metrics while still terminating near 15 h.

4.3 AMPLITUDE BOUNDS VERSUS STRUCTURED RESPONSE GEOMETRY

Figure 3 separates output constraints (panel a) from learned representations of the same physical samples (panel b). The unrestricted and limiter arms share an embedding and nonlinear readout, whereas DARS combines fixed complete response atoms with state-dependent routing.

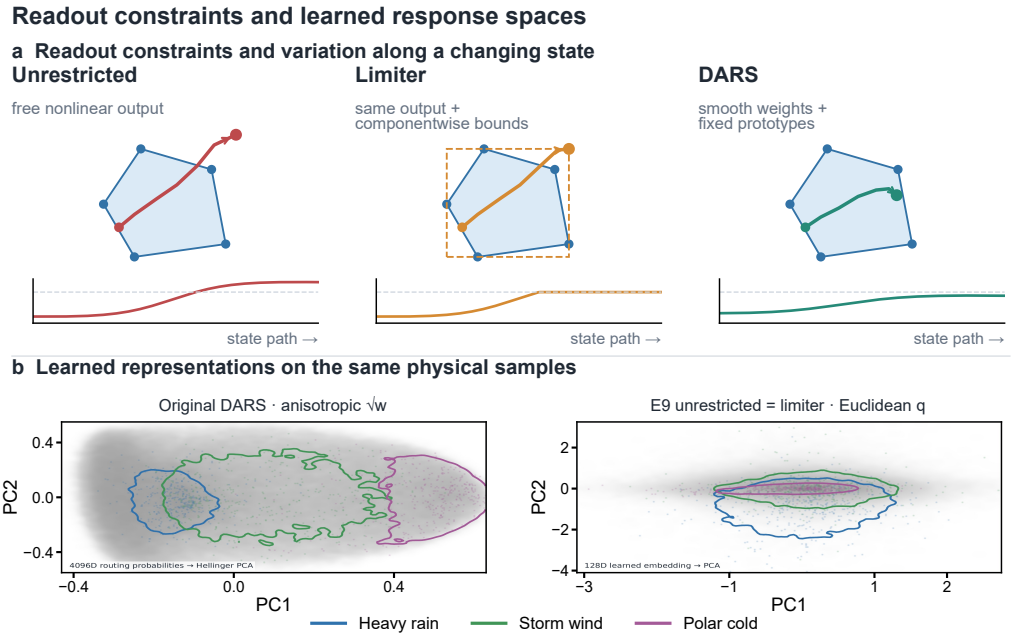

Figure 3: Readout constraints and learned response representations. Panel a illustrates ideal normalized mixtures: DARS stays in the prototype hull, whereas unrestricted and componentwise-limited outputs can leave it. The epsilon implementation requires the augmented-hull qualification in Section 3.3. The embedding panels use different PCA bases and are descriptive, not a matched causal comparison or a stability guarantee.

Panel a highlights why the response class can support this progression. The unrestricted readout can leave the joint-response hull; clipping bounds individual channels but permits unsupported joint profiles and introduces threshold kinks. DARS instead changes smooth mixtures of complete atoms, preserving its bounded response class during routing adaptation. The schematic depicts ideal normalization; the epsilon implementation uses the augmented hull specified in Method. This supplies a response-level constraint that is consistent with the observed survival, not a stability theorem for the coupled solver.

Panel b displays DARS through anchor-indexed routing weights (square-root/Hellinger PCA), whereas the shared unrestricted/limiter embedding uses a separate Euclidean PCA. The DARS regime separation thus describes the combined router–anchor representation, not the encoder alone. These projections visualize representations rather than defining the deployed routing metric. Because the coordinates and PCA bases differ, apparent distances across panels are not directly comparable.

With the DARS response atoms fixed, encoder updates alter how particular states mix the available physical responses rather than introducing arbitrary new output profiles. Panel b illustrates the joint router–anchor organization that makes such rerouting possible, but does not itself measure adaptation locality or cross-regime interference. The survival histories and fixed-horizon metrics, rather than embedding separation alone, establish the empirical stability-to-improvement progression.

4.4 TRANSFER ACROSS INITIAL CONDITIONS

The separate transfer study adapts a DARS deployment using only the first 12 h of the 30 September 1988 initialization, then freezes its request-indexed schedule for 22 initializations not used for adaptation. It does not fine-tune each case or repeat the matched three-readout comparison. Figure 4 compares the deployment with its offline parent and independently rerun Native physics. Appendix Figure 5A shows the separate single-initialization readout diagnostics; panel B gives three transfer examples. Deployment settings are in Appendix Q.

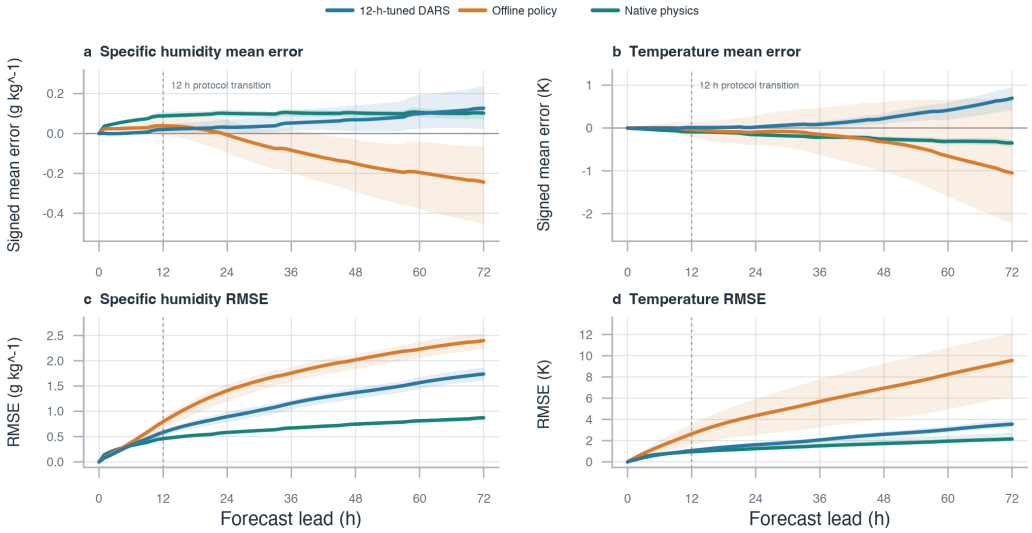

Figure 4: Time evolution across 22 initializations not used for adaptation for the frozen 12 h-tuned DARS deployment, its offline parent, and Native physics. Curves are across-case means; shading spans the full range. Top: signed mean error. Bottom: native-grid RMSE. After 12 h, a 20-minute transition removes the added uniform temperature tendency while neural inference continues.

Figure 4 shows a lead-time-dependent bias advantage. The underlying per-case metrics show smaller absolute signed Q and T bias for DARS than for Native physics in all 22 cases at every integer hour from 24 through 36 h. At 48 h, the across-case mean advantage remains for both variables, but the case-level advantage has narrowed to 21/22 for Q and 15/22 for T. Thus the low-bias behavior of the 12 h-tuned deployment persists beyond its tuning window, robustly to 36 h and in the mean to about 48 h.

After 48 h, this advantage erodes: mean absolute T bias exceeds Native's from about 51 h, and mean absolute Q bias from about 62 h. By 72 h, DARS has larger mean absolute bias for both variables. These bias comparisons do not imply lower field RMSE than Native.

The RMSE comparison has a different baseline and time profile. DARS improves on its offline parent at every post-initial integer hour for both Q and T in all 22 cases, including at 72 h; Native physics retains lower field RMSE at that endpoint. Averaging hourly RMSE over each full 0–72 h trajectory, the mean case-level reductions relative to the offline parent are $31.4\% \pm 2.0\%$ for Q and $61.2\% \pm 7.2\%$ for T (sample standard deviation across 22 cases). Further lead-time and minimum-case details are in Appendix C.

4.5 PHYSICAL STRUCTURE OF THE REMAINING ERROR

Across the 22 cases, DARS closes $52.2\% \pm 3.1\%$ of the Q RMSE gap and $83.8\% \pm 5.3\%$ of the T RMSE gap to native physics, while native physics retains lower field RMSE in every case. The gap definition, aggregate statistics, and illustrative diagnostics are provided in Appendix D. Appendix Figure 7 shows that, for the illustrative case, the remaining difference is structured across the vertical column rather than explained by a simple global bias: DARS keeps signed biases closer to zero

over much of the column, while retaining higher per-level RMSE at several levels. Thus, DARS substantially reduces systematic drift and narrows the gap to native physics, but does not yet match its spatially resolved accuracy.

5 CONCLUSION

DARS provides a bounded joint-response class that empirically supports stable coupling and serves as a foundation for successive adaptation. Before fine-tuning, all three DARS seeds complete 72 h at $k = 0$, while unrestricted and componentwise-limited readouts terminate earlier despite comparable offline accuracy. DARS retains full-horizon survival through $k = 1$–3, and seed 1117 improves full-horizon Q/T RMSE and absolute biases monotonically across the three updates; the other seeds also improve final temperature RMSE. This is evidence under the evaluated protocol, not a guarantee of simulator stability or monotonic optimization.

A DARS deployment adapted using only the first 12 h of the 30 September 1988 initialization also transfers across initializations and forward in time. Its frozen deployment schedule is reused without retraining on 22 initializations not used for adaptation, reducing full-trajectory (0–72 h) mean specific-humidity and temperature field RMSE by $31.4\% \pm 2.0\%$ and $61.2\% \pm 7.2\%$ relative to its offline parent, and closing $52.2\% \pm 3.1\%$ and $83.8\% \pm 5.3\%$ of the corresponding gaps to native physics. At every hourly lead from 24 to 36 h, absolute Q/T bias is lower than native physics in all 22 cases, demonstrating improvement beyond the 12 h adaptation window. At 48 h, the advantage persists in 21/22 Q cases and 15/22 T cases before subsequently eroding. These findings concern the complete adapted deployment, not isolated router updates; native physics retains lower full-trajectory field RMSE.

Together, the offline comparison, sequential adaptation, and frozen transfer results support the progression from structured responses to reliable coupled refinement and transferable accuracy gains. Continuous checkpoints are stages within three seed histories, not independent replicates. The transfer study evaluates the complete adapted system and does not establish that the three-readout survival separation holds at every initialization.

Our methodological contribution is a hybrid learning framework that combines data-anchored response-space design, surrogate-assisted adaptation, and evaluation in the original non-differentiable simulator. We do not claim predictive accuracy beyond existing baselines: DARS improves on its offline parent, while native physics retains lower full-trajectory field RMSE in the evaluated transfer cases.

Reproducibility Statement

The Method section specifies the response construction, routing readout, and surrogate-assisted adaptation procedure. Appendix B provides additional construction and evaluation details, and Appendix J states the normalization assumptions underlying the response-space bound, including the implemented denominator offset. Appendices N and O document the network architectures, adapter budgets, optimization settings, and sequential training schedule. Appendix Q separately documents the first-12-hour adaptation and frozen deployment schedule used for the 22-initialization transfer evaluation. These records distinguish offline readout comparisons, sequential adaptation checkpoints, and frozen transfer tests so that their training information and evaluation scope are explicit.

AI Use Statement

Generative AI tools assisted with manuscript drafting and language editing, LaTeX formatting, literature searches, feedback on experimental methodology, interpretation and presentation of results, and the development and review of analysis and figure-generation scripts. AI-assisted suggestions and outputs were reviewed against the experimental records and source materials during manuscript preparation. The authors are responsible for checking the accuracy of the scientific claims, numerical results, citations, code, and manuscript text, and take full responsibility for the content of this submission.

References

Leo Breiman. Random forests. *Machine Learning*, 45:5–32, 2001.

Noah D. Brenowitz and Christopher S. Bretherton. Prognostic validation of a neural network unified physics parameterization. *Geophysical Research Letters*, 45(12):6289–6298, 2018. doi: 10.1029/2018GL078510.

Noah D. Brenowitz and Christopher S. Bretherton. Spatially extended tests of a neural network parametrization trained by coarse-graining. *Journal of Advances in Modeling Earth Systems*, 11 (8):2728–2744, 2019.

Noah D. Brenowitz, Benjamin Henn, Jeremy McGibbon, Spencer K. Clark, Anna Kwa, William A. Perkins, Oliver Watt-Meyer, and Christopher S. Bretherton. Machine learning climate model dynamics: Offline versus online performance. *arXiv preprint arXiv:2011.03081*, 2020.

Hugo Frezat, Ronan Fablet, Guillaume Balarac, and Julien Le Sommer. Gradient-free online learning of subgrid-scale dynamics with neural emulators. *Journal of Advances in Modeling Earth Systems*, 18:e2023MS004140, 2026.

Scott Fujimoto, David Meger, and Doina Precup. Off-policy deep reinforcement learning without exploration. In *Proceedings of the 36th International Conference on Machine Learning*, volume 97, pp. 2052–2062, 2019.

Hans Hersbach, Bill Bell, Paul Berrisford, et al. The ERA5 global reanalysis. *Quarterly Journal of the Royal Meteorological Society*, 146(730):1999–2049, 2020. doi: 10.1002/qj.3803.

Edward J. Hu, Yelong Shen, Phillip Wallis, Zeyuan Allen-Zhu, Yuanzhi Li, Shean Wang, Lu Wang, and Weizhu Chen. LoRA: Low-rank adaptation of large language models. In *International Conference on Learning Representations*, 2022. URL https://openreview.net/forum?id=nZeVKeeFYf9.

Robert A. Jacobs, Michael I. Jordan, Steven J. Nowlan, and Geoffrey E. Hinton. Adaptive mixtures of local experts. *Neural Computation*, 3(1):79–87, 1991. doi: 10.1162/neco.1991.3.1.79.

Michael Janner, Justin Fu, Marvin Zhang, and Sergey Levine. When to trust your model: Model-based policy optimization. In *Advances in Neural Information Processing Systems*, volume 32, 2019.

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

Aviral Kumar, Justin Fu, George Tucker, and Sergey Levine. Stabilizing off-policy q-learning via bootstrapping error reduction. In *Advances in Neural Information Processing Systems*, volume 32, 2019.

X. Li, Y. Zhang, X. Peng, B. Zhou, J. Li, and Y. Wang. Intercomparison of the weather and climate physics suites of a unified forecast–climate model system (GRIST-A22.7.28) based on single-column modeling. *Geoscientific Model Development*, 16(10):2975–2993, 2023. doi: 10.5194/gmd-16-2975-2023. URL https://gmd.copernicus.org/articles/16/2975/2023/.

Yuping Luo, Huazhe Xu, Yuanzhi Li, Yuandong Tian, Trevor Darrell, and Tengyu Ma. Algorithmic framework for model-based deep reinforcement learning with theoretical guarantees. In *International Conference on Learning Representations*, 2019.

E. A. Nadaraya. On estimating regression. *Theory of Probability and Its Applications*, 9(1):141–142, 1964. doi: 10.1137/1109020.

Paul A. O'Gorman and John G. Dwyer. Using machine learning to parameterize moist convection: Potential for modeling of climate, climate change, and extreme events. *Journal of Advances in Modeling Earth Systems*, 10(10):2548–2563, 2018.

Stephan Rasp. Coupled online learning as a way to tackle instabilities and biases in neural network parameterizations: General algorithms and lorenz 96 case study (v1.0). *Geoscientific Model Development*, 13(5):2185–2196, 2020.

Stéphane Ross, Geoffrey J. Gordon, and J. Andrew Bagnell. A reduction of imitation learning and structured prediction to no-regret online learning. In *Proceedings of the Fourteenth International Conference on Artificial Intelligence and Statistics*, pp. 627–635, 2011.

Friedhelm Schwenker, Hans A. Kestler, and Günther Palm. Three learning phases for radial-basis-function networks. *Neural Networks*, 14(4–5):439–458, 2001.

Geoffrey S. Watson. Smooth regression analysis. *Sankhyā: The Indian Journal of Statistics, Series A*, 26:359–372, 1964.

Tianhe Yu, Garrett Thomas, Lantao Yu, Stefano Ermon, James Y. Zou, Sergey Levine, Chelsea Finn, and Tengyu Ma. Mopo: Model-based offline policy optimization. In *Advances in Neural Information Processing Systems*, volume 33, 2020.

Janni Yuval and Paul A. O'Gorman. Stable machine-learning parameterization of subgrid processes for climate modeling at a range of resolutions. *Nature Communications*, 11:3295, 2020. doi: 10.1038/s41467-020-17142-3.

Janni Yuval, Paul A. O'Gorman, and Christopher N. Hill. Use of neural networks for stable, accurate and physically consistent parameterization of subgrid atmospheric processes with good performance at reduced precision. *Geophysical Research Letters*, 48:e2020GL091363, 2021.

Yi Zhang, Jian Li, Rucong Yu, Zhuang Liu, Yu Zhou, Xiaohan Li, and Xiaomeng Huang. A multiscale dynamical model in a dry-mass coordinate for weather and climate modeling: Moist dynamics and its coupling to physics. *Monthly Weather Review*, 148(7):2671–2699, 2020. doi: 10.1175/MWR-D-19-0305.1.

A Extended Motivation and Related Work

A.1 Extended comparison with prior work

Closed-loop learning in hybrid scientific models. Previous work has addressed this mismatch through several complementary mechanisms. Physical constraints and stability-oriented designs restrict harmful outputs or encode known conservation structures (Yuval & O'Gorman, 2020; Yuval et al., 2021). Coupled online learning exposes learned components to states induced by their own deployment (Rasp, 2020), while multistep training and online evaluation test behavior under extended feedback (Brenowitz et al., 2020; Rasp, 2020). When host dynamics are differentiable, differentiable hybrid models can incorporate feedback directly into end-to-end training; NeuralGCM demonstrates this regime with a differentiable atmospheric dynamical core (Kochkov et al., 2024).

Existing approaches modify training objectives, exposure to deployment states, or physical admissibility. DARS addresses the joint response representation. Our strict offline $k = 0$ comparison precedes online updates; sequential DARS/unrestricted adaptation and the separate frozen transfer study supply distinct evidence. The limiter is not claimed to complete the same three-round adaptation procedure.

Surrogate-assisted adaptation for non-differentiable simulators. When a production simulator is difficult or impossible to differentiate end to end, a common strategy is to train a surrogate that provides gradient or sensitivity information for policy or readout updates. This setting is closely related to model-based reinforcement learning and iterative system identification. SLBO alternates among model fitting, virtual policy optimization, and data collection from the original system, while MBPO limits model-generated rollouts to short horizons to reduce accumulated model error (Luo et al., 2019; Janner et al., 2019). Because learned surrogates are imperfect, optimization through them can move policies toward regions with little data support or unreliable dynamics. Model-based methods therefore use uncertainty estimates, trust regions, or behavior-support constraints to control such excursions (Fujimoto et al., 2019; Kumar et al., 2019; Yu et al., 2020). Visited-state aggregation is also related to DAgger-style distribution matching (Ross et al., 2011). Unlike standard policy-learning settings with observed actions, the reference trajectories in our setting provide state transitions but no direct labels for the unresolved physical responses. The resolved dynamics are therefore used to recover effective responses before the policy is adapted.

The work of Frezat et al. (2026) is particularly close to our setting. Their differentiable neural emulator provides gradients for a learned parameterization coupled to a non-differentiable solver, while the original simulator remains responsible for final evaluation (Frezat et al., 2026). We use the same general surrogate-assisted pattern: short rollouts provide optimization gradients, visited coupled transitions refine the surrogate, and exported readouts are evaluated through uninterrupted integration in the original simulator. Replay in the sequential adaptation study accumulates valid trajectory prefixes, not acceptance-filtered rollouts. DARS contributes a different element within this pipeline. It constructs a fixed dictionary of joint response prototypes from transition-consistent temperature and moisture tendencies, and, under ideal normalization, restricts predictions to nonnegative convex combinations; the epsilon implementation instead requires the augmented-hull qualification in Method. Low-rank updates modify the routing rule without redefining the dictionary.

Data-anchored and structured response representations. Sample-based predictors provide a closer design precedent for DARS. Random forests average labels associated with training samples and can inherit useful range or conservation properties from the data (Breiman, 2001; O'Gorman & Dwyer, 2018; Yuval & O'Gorman, 2020). Tree ensembles rely on hard partitions, while neighborhood methods use fixed metrics or retrieve nearby samples. Kernel predictors provide smooth distance-weighted averages, but their underlying geometry is usually specified in advance (Nadaraya, 1964; Watson, 1964). DARS instead uses anchor-specific anisotropic Gaussian distances in a learned embedding for response routing, while keeping each response component anchored to complete multivariate structure derived from training transitions.

DARS is also structurally related to radial-basis-function networks and mixture-of-experts models. Classical RBF networks combine local radial units with learned output coefficients, while mixture-of-experts models jointly learn expert predictors and state-dependent gating rules (Schwenker et al., 2001; Jacobs et al., 1991). The key distinction lies in the response components. DARS does not learn free expert functions or an unrestricted output vector. Each component is anchored to a complete

multivariate response derived from a training transition. Ideal unit-sum routing gives prototype-hull membership and smooth state-dependent selection; the implemented epsilon denominator requires the augmented-hull qualification in Method.

The fixed routing geometry should be distinguished from the low-rank adaptation used during closed-loop fine-tuning. The former defines how states are compared and routed to response prototypes through anchor-specific Gaussian metrics. The latter provides a parameter-efficient update to the routing model. In the matched adaptation study, the response dictionary remains fixed while low-rank routing updates change state-dependent selection. DARS is also complementary to support regularization in offline reinforcement learning: BCQ and BEAR constrain policies toward behavior data, while MOPO penalizes regions in which the learned dynamics model is uncertain (Fujimoto et al., 2019; Kumar et al., 2019; Yu et al., 2020). DARS places the corresponding structural control directly in the response representation.

Overall, the closest prior work provides deployment-aware optimization, sample-based prediction, or smooth state-dependent mixing. DARS combines these ideas through a different division of responsibilities. Complete multivariate responses derived from training transitions define the response dictionary, while learning is concentrated on smooth, state-dependent selection within that dictionary. The offline $k = 0$ comparison matches all three readouts. Subsequent rounds prescribe the same adapter budget but use arm-specific surrogates; limiter completion differs by seed (Appendix O.4). Much of the existing work has focused on how to optimize under feedback. DARS complements this focus by asking which response space that optimization should use.

B Additional Method Details

DARS uses a 151-component policy input and emits 60 tendencies: 30 specific-humidity and 30 temperature components every $\Delta = 1200\,\mathrm{s}$. Only the training partition is used to construct

$$\mathcal{D}_{\mathrm{resp}} = \left\{(\mathbf{z}_i, \mathbf{a}_i^{\mathrm{eff}})\right\}_{i=1}^{N_{\mathrm{eff}}}. \tag{8}$$

The replay-check definition appears in Appendix G; the routing-displacement bound and its proof appear in Appendix J. Table 2 summarizes the division of roles in the matched study.

Direct differentiation through G is unavailable. We therefore train a differentiable state-transition surrogate S_ψ. During a virtual rollout,

$$\widehat{\mathbf{x}}_{t+\Delta} = S_\psi(\widehat{\mathbf{x}}_t, \mathbf{a}_t, \mathbf{F}_t, \mathbf{h}_t), \qquad \widehat{\mathbf{z}}_t = H(\widehat{\mathbf{x}}_t, \mathbf{F}_t, \mathbf{h}_t), \qquad \mathbf{a}_t = \pi_\theta(\widehat{\mathbf{z}}_t). \tag{9}$$

Only the initial state is taken from an archived trajectory; subsequent states are generated by S_ψ. The verified sequential adaptation objective and optimizer settings are specified in Appendix O.2.

Adaptation proceeds in four steps:

1. **Fit or refine the surrogate.** For the matched-readout study, train S_ψ on offline training transitions and accumulated valid coupled rollout prefixes.
2. **Optimize through virtual rollouts.** Update the routing adapters through short differentiable surrogate rollouts while keeping $\mathcal{B}$ fixed.
3. **Evaluate in the original simulator.** Deploy the candidate in an uninterrupted rollout of G; no reference state is written back during candidate integration.
4. **Evaluate and aggregate.** Record numerical survival and Q/T errors in the original simulator. The sequential adaptation launcher accumulates generated valid rollout prefixes for the next round; this is not claimed to implement acceptance-filtered replay.

The surrogate supplies optimization gradients, while the original simulator remains the final authority for deployment evidence. Acceptance-filtered replay is a possible protocol, not an established property of the sequential adaptation launcher; the distinction is recorded in Appendix K.

C Additional Evaluation Details

The separate 22-IC transfer test freezes a deployment schedule adapted using only the first 12 h of the 30 September 1988 initialization. Evaluation cases span autumn 1988, winter 1997/98, and spring

Component	Construction/source	During matched adaptation
Numerical simulator $\mathcal{G}$	Existing validated scientific code	Fixed and authoritative
Input normalization and anchors	Training partition only	Fixed; anchor embeddings change with the encoder
Response dictionary B	Training transition-consistent responses	Fixed in the matched ablation
Base policy at adapted layers	Offline policy	Parent weights frozen
Selected routing layers	Parent routing model	Low-rank adapters updated
Transition surrogate S_ψ	Reference and visited transitions	Refined between rounds
Candidate simulator state	Archived initial state	Evolves freely under $\mathcal{G}$

Table 2: Separation of fixed and adapted components in the matched closed-loop study.

2005 and are not used for adaptation or selection. The protocol is described in Appendix Q; it is not the sequential adaptation experiment.

For variable $v \in \{Q, T\}$, the evaluation reports area-weighted native-grid RMSE and the time average of the absolute signed bias. Let $\omega_{i\ell}$ denote the normalized weights over horizontal cells i and levels ℓ, with $\sum_{i,\ell} \omega_{i\ell} = 1$. The hourly metrics are

$$B_v(t) = \sum_{i,\ell} \omega_{i\ell} \left(\widehat{v}_{i\ell t} - v^*_{i\ell t} \right), \tag{10}$$

$$R_v(t) = \left[\sum_{i,\ell} \omega_{i\ell} \left(\widehat{v}_{i\ell t} - v^*_{i\ell t} \right)^2 \right]^{1/2}. \tag{11}$$

Native-grid RMSE is the primary field-error measure. Spectrally filtered quantities are used only to diagnose the structure of residual errors.

The $k = 0$ offline models have similar validation accuracy. Unrestricted/DARS moisture R^2 values are 0.6373/0.6352 (seed 1117), 0.6484/0.6374 (2027), and 0.6293/0.6438 (3037); temperature values are 0.6403/0.6357, 0.6413/0.6373, and 0.6411/0.6413. The largest difference is 0.0145. These values describe offline validation and do not make the $k = 0$ survival separation an effect of subsequent adaptation.

Figure 4 reports the across-case mean with the full case range. At 72 hours, the mean DARS specific-humidity RMSE is approximately $1.7\,\mathrm{g\,kg^{-1}}$, compared with approximately $2.4\,\mathrm{g\,kg^{-1}}$ for the offline parent. The corresponding temperature RMSE values are approximately $3.5\,\mathrm{K}$ and $9.5\,\mathrm{K}$. Native physics remains more accurate, with approximate specific-humidity and temperature RMSE values of $0.9\,\mathrm{g\,kg^{-1}}$ and $2.2\,\mathrm{K}$, respectively.

The full-trajectory case-level reductions reported in the main text have minima of 26.7% for specific humidity and 42.0% for temperature; both variables improve in all 22 cases. Improvement relative to the offline parent remains positive at every post-initial integer hour through 72 h in every case. Hour 0 is the shared zero-error initial state, for which a percentage reduction is undefined.

Figure 6 expands the representation diagnostic in Figure 3b to eight weather regimes. It uses identical physical samples in both rows, but not a matched checkpoint pair: the upper row is the archived original DARS model, while the lower row is the unrestricted model from the matched-readout study at seed 1117, stage $k = 0$; the post-hoc limiter shares that embedding. The upper projection uses square-root routing weights and Hellinger PCA, incorporating the original DARS checkpoint's anchor-specific anisotropic metric. The lower projection applies Euclidean PCA to the unrestricted/limiter learned embedding. Each PCA is fitted independently, so separation may be compared qualitatively within a row, not by distance between rows. The main Method specifies the evaluated anisotropic DARS readout; this diagnostic uses a different DARS checkpoint and does not isolate the effect of the embedding, anchors, or metric on coupled survival or adaptation locality.

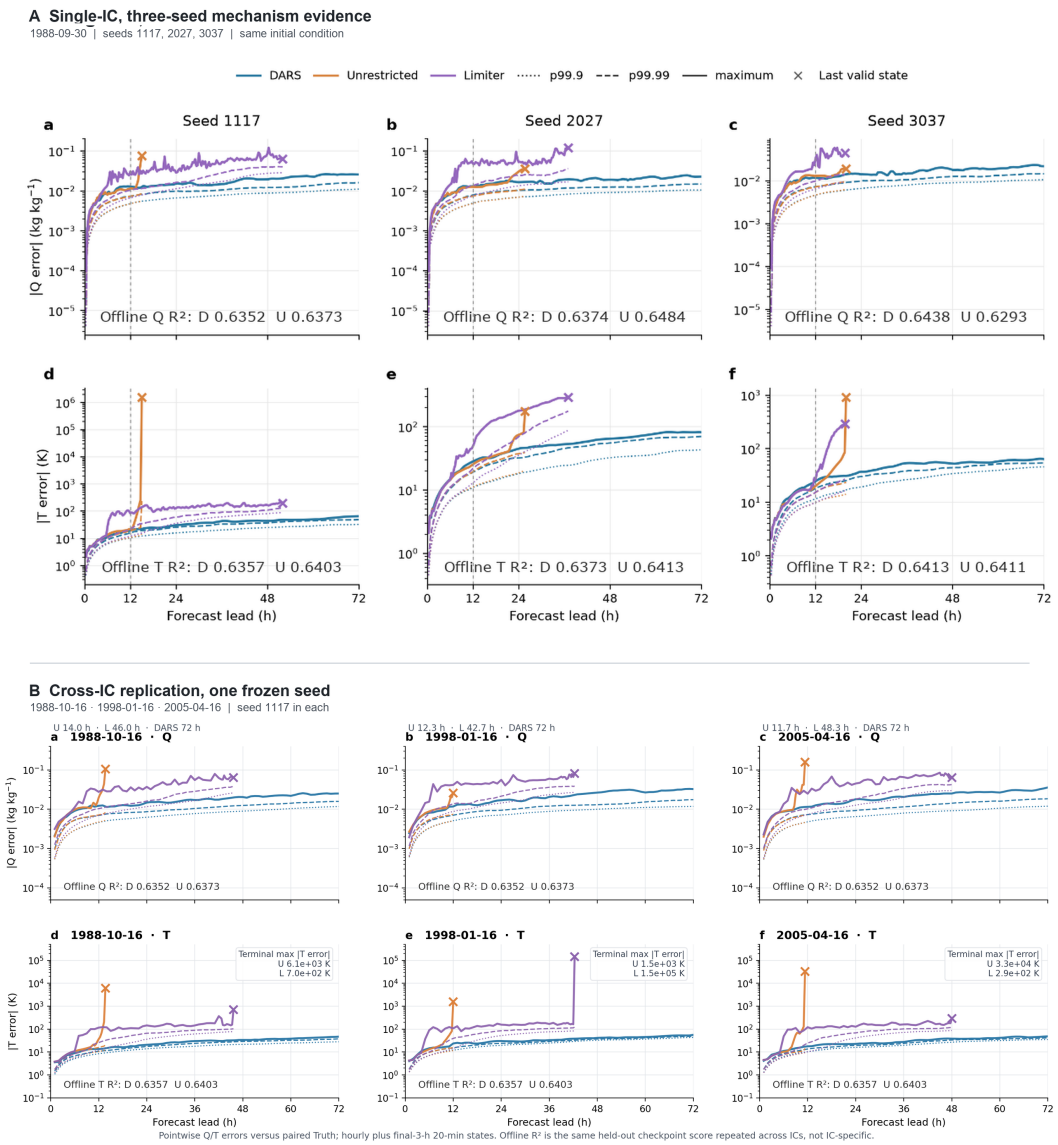

Figure 5: Long-horizon endpoint and trajectory diagnostics under three response parameterizations. Curves show DARS, an unrestricted readout, and a componentwise limiter for three training seeds; crosses indicate the last archived valid simulator state. Dotted/dashed/solid summaries show upper-tail and maximum absolute errors. The displayed DARS/unrestricted endpoints and limiter checkpoints are not all strictly budget matched, so this figure does not replace the strict pre-adaptation k0 comparison in Table 1.

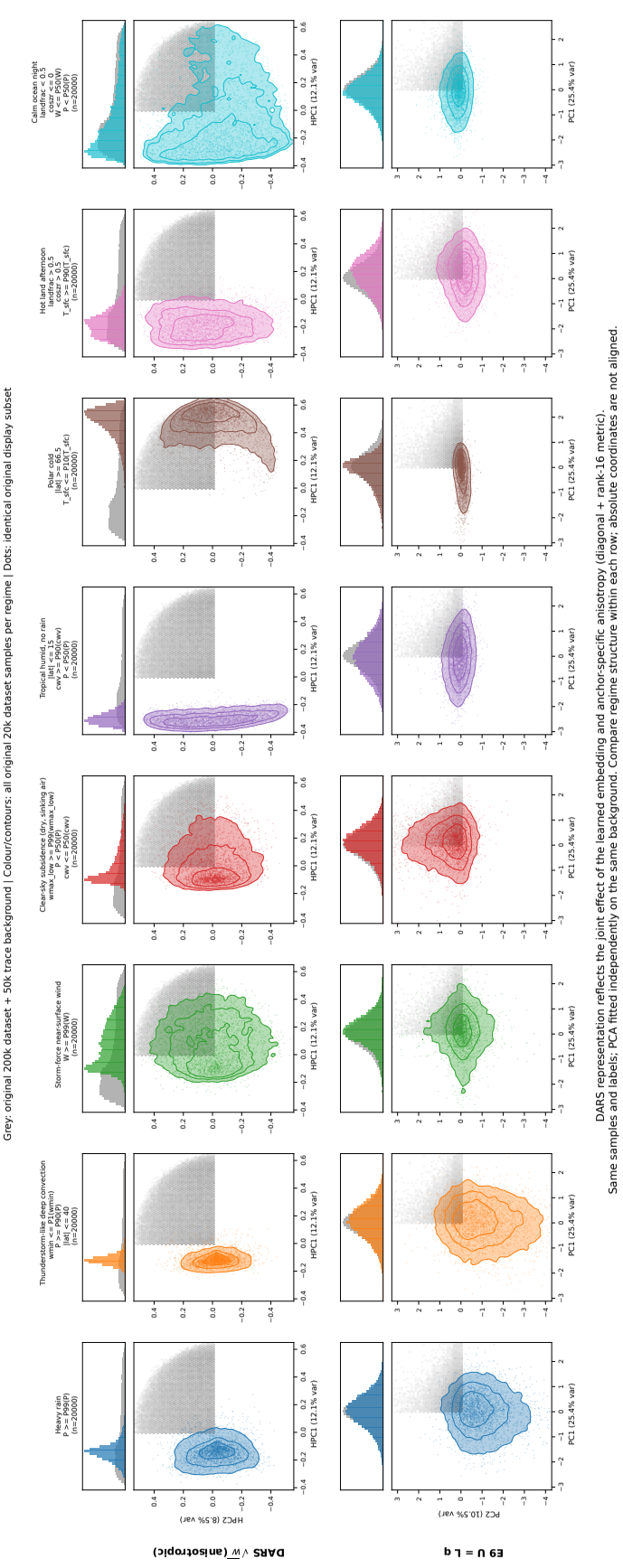

Figure 6: **Weather-regime structure in two diagnostic response representations.** The same atmospheric-column samples and regime labels are shown in both rows across eight weather regimes. Top: archived original DARS square-root routing probabilities over 4,096 anchors, projected by Hellinger PCA; this checkpoint uses anchor-specific anisotropic distances. Bottom: the unrestricted model from the matched-readout study (seed 1117, stage $k = 0$) 128-dimensional learned embedding projected by Euclidean PCA; the post-hoc limiter shares this embedding. The PCA bases and model checkpoints differ. The panels are descriptive and do not constitute a matched causal comparison or establish the exact configuration of the selected transfer deployment.

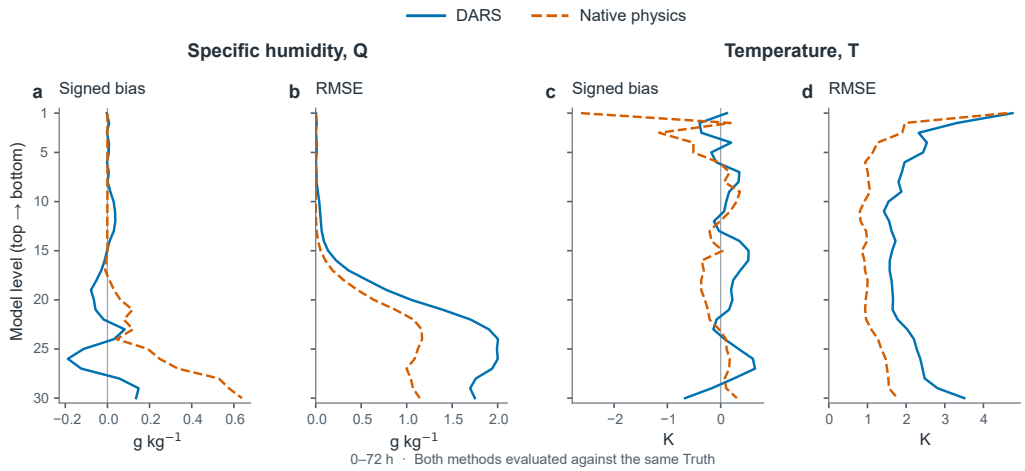

Figure 7: **Vertical signed-bias and RMSE profiles for both specific humidity and temperature.** For the illustrative transfer initial condition 19881004, profiles aggregate the 0–72 h DARS and independently rerun Native-physics errors relative to the same Truth trajectory. Signed bias can cancel across levels and cells, whereas RMSE retains error magnitude.

D Additional Physical Diagnostics

For variable v, define the gap closure relative to native physics as

$$\mathrm{GapClosure}_v = \frac{R_v^{\mathrm{offline}} - R_v^{\mathrm{DARS}}}{R_v^{\mathrm{offline}} - R_v^{\mathrm{native}}}. \tag{12}$$

Across the 22 cases, DARS closes $52.2\% \pm 3.1\%$ of the specific-humidity gap and $83.8\% \pm 5.3\%$ of the temperature gap, using the same case-level definition. Native physics nevertheless retains lower field RMSE in all 22 cases. DARS therefore narrows the gap to native physics without matching it across all metrics.

The illustrative vertical profiles in Figure 7 describe residual physical fidelity, not the causal mechanism of the matched survival difference. Spectral and geographic diagnostics appear in Figure 10.

In the illustrative case, DARS keeps signed bias comparatively small but has larger per-level RMSE than native physics across much of the column. Thus, lower bias does not imply native-physics-level field accuracy; the remaining error has vertical and spectral structure rather than being a single global offset.

Table 3: **Twenty-two-case 0–72 h field-RMSE summary.** Values are across-case means ± sample standard deviations where shown. Gap closure is $100(\mathrm{RMSE}_{\mathrm{PV0}} - \mathrm{RMSE}_{\mathrm{DARS}})/(\mathrm{RMSE}_{\mathrm{PV0}} - \mathrm{RMSE}_{\mathrm{Native}})$.

Comparison	Q	T
DARS reduction versus PV0	31.43% (22/22)	61.22% (22/22)
Parent-to-Native gap closed	$52.24\% \pm 3.06$	$83.79\% \pm 5.32$
Native lower field RMSE	22/22	22/22

Figure 7 uses initial condition 19881004 as an illustrative case. Aggregate claims are based on all 22 initial conditions; this case is shown to make the vertical patterns readable. For specific humidity, the signed bias remains close to zero for both DARS and Native physics over most levels. The lower levels show a larger positive Native-physics bias in this case, while DARS remains closer to zero. The RMSE profile shows a different pattern: DARS has larger errors in the lower and middle levels, reaching approximately $1.8\,\mathrm{g\,kg^{-1}}$, whereas Native physics remains mostly between 1.0 and $1.2\,\mathrm{g\,kg^{-1}}$.

Temperature shows the same distinction between bias and field error. In this illustrative case, native physics has a negative bias across most levels, while DARS has a smaller positive bias. DARS temperature RMSE remains higher than native physics at most levels, particularly in the middle and upper parts of the profile. The remaining error is therefore structured across levels and scales rather than explained by a single global offset. DARS controls the large-scale drift of the offline parent, but its vertical structure does not yet reach the accuracy of native physics.

E Cross-Initialization Lead-Time Consistency

Table 4 summarizes the same 22 newly evaluated initial conditions in six 12 h lead-time windows. It excludes the prior 12 October case and the 30 September tuning archive. Every case-level reduction is positive in every window for both variables. Repeated windows within a case and nearby initialization dates are not independent statistical replicates.

Table 4: **Mean RMSE reduction (%) by lead-time window over 22 unseen cases.** Parentheses give the minimum case-level reduction; all 22 cases are positive in every cell.

	0–12 h	12–24 h	24–36 h	36–48 h	48–60 h	60–72 h
Q	19.87 (12.88)	33.14 (29.20)	36.04 (30.75)	32.84 (28.08)	31.29 (25.95)	28.51 (22.86)
T	53.94 (29.77)	60.27 (38.91)	62.04 (42.16)	61.38 (40.54)	61.79 (42.42)	62.02 (45.62)

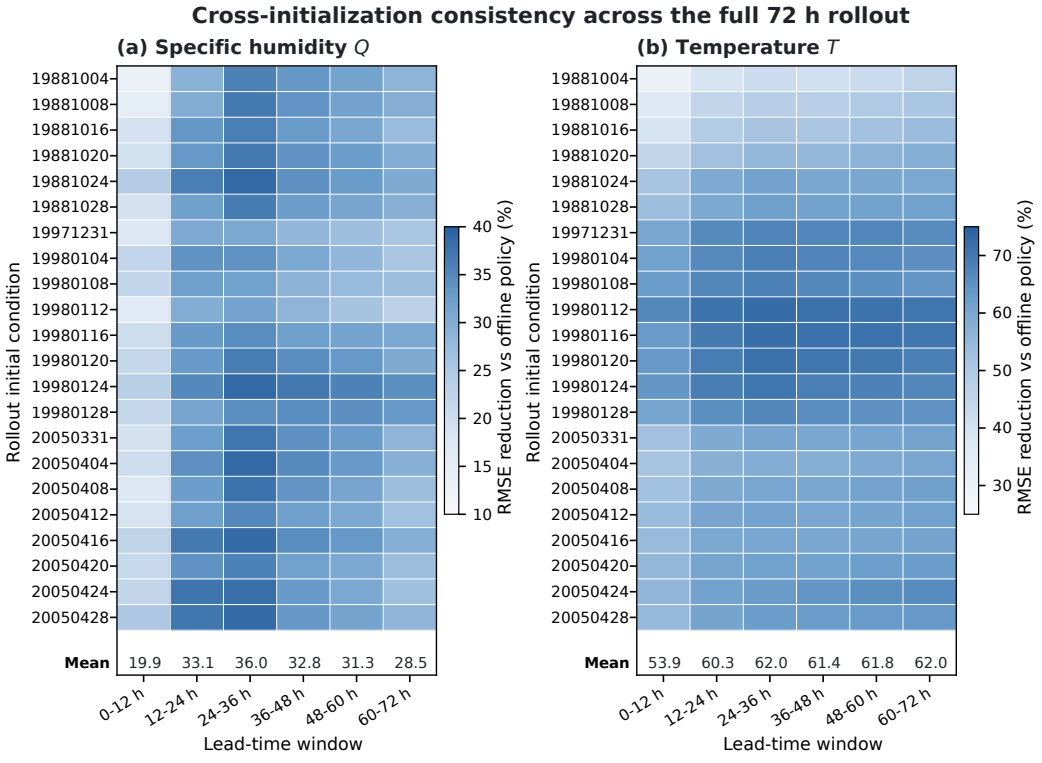

Figure 8: **Case-level persistence of the selected 12 h-tuned DARS deployment across lead-time windows.** Rows are the same 22 rollout initial conditions used in Figure 4; columns are six prespecified 12 h windows. Each cell reports the percent reduction in window-mean native-grid RMSE relative to the matched offline policy, with positive values favoring DARS. Values below each heatmap are across-case means. All 264 variable–case–window comparisons are positive. The visualization exposes case heterogeneity but does not turn correlated windows or nearby initialization dates into independent replicates.

F Diagnostic PCA and Input Anchors

The matched-readout study routes directly in a 128-dimensional learned embedding with 4,096 fixed input anchors and anchor-specific metrics (Appendix N.2). Its readout does not use a 64-dimensional PCA projection. Stored time rows and atmospheric-column samples are different sampling units: the 1,944 offline time rows must not be compared directly with the anchor count to infer a clustering inconsistency. The offline anchor-construction and response-aggregation procedure requires its own construction record; it is not inferred here from the exported tensors.

Figure 9 displays a separate, descriptive diagnostic of encoded states and routing. Its two-component PCA was fitted only to 4,096 archived training-derived RBF input centers after the frozen 128-dimensional encoder, not to the evaluation case. This PCA is a visualization transform, not an deployed routing projection. The panels use 40,962 paired native-grid columns at each shown lead time. They show a distributional change within one initial condition, not a causal test of stability.

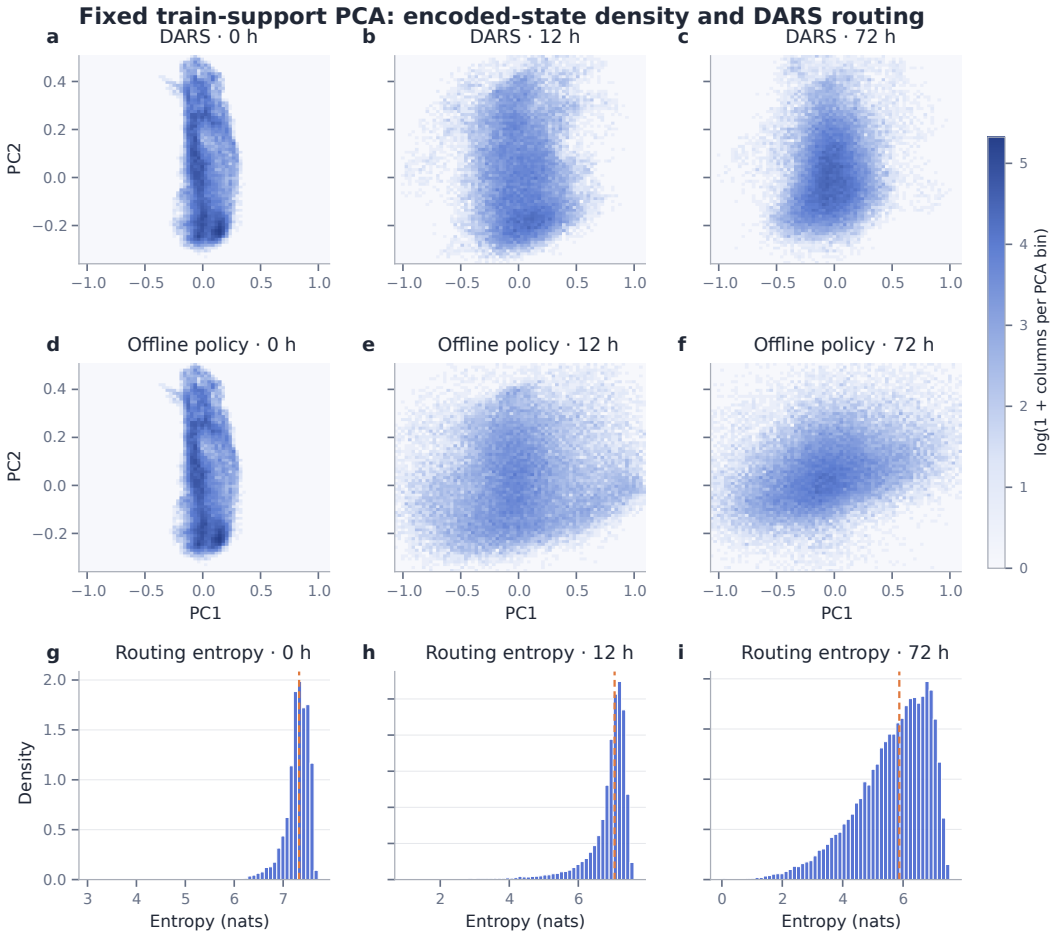

Figure 9: **Encoded-state density and DARS routing along one continuous 72-h case.** DARS and the offline policy are projected through the same fixed two-component PCA fitted to 4,096 archived training-derived RBF input centers. Columns show 0, 12, and 72 h using identical axes; the lower row shows DARS routing entropy. Each method–hour panel includes 40,962 paired native-grid columns, without area weighting. Native physics has no documented common policy encoder and is therefore excluded. This visualization describes state occupancy and routing, without establishing that their changes caused the observed survival difference.

G Effective-Response Replay Check

The resolved advance used in the response extraction must use the same forcing files, executable, coupling interval, and solver-history convention as the corresponding reference transition. For each validation transition, reconstruct a_t^{eff}, replay one simulator step,

$$\widehat{x}_{t+\Delta}^{\mathrm{ref}} = \mathcal{G}(x_t^{\mathrm{ref}}, a_t^{\mathrm{eff}}, F_t, h_t),$$

and report a reconstruction metric

$$e_{\mathrm{replay}}(t) = \mathrm{metric}\big(\widehat{x}_{t+\Delta}^{\mathrm{ref}}, x_{t+\Delta}^{\mathrm{ref}}\big)\,.$$

Large replay error would indicate that the extracted residual absorbs backbone, forcing, history, or numerical mismatch. The available archive does not provide numerical replay-error summaries, so this subsection specifies the check without claiming a measured replay accuracy.

H Surrogate and Policy Loss Details

The surrogate is trained on state/response/forcing/history transitions. Any normalization statistics must be fitted on the training partition and frozen. During a virtual rollout,

$$\widehat{x}_{k+1} = S_\psi(\widehat{x}_k, \pi_\theta(H(\widehat{x}_k, F_k, h_k)), F_k, h_k)\,.$$

For variable v, let $\mathrm{RMSE}_v^{(k)}$ and $\mathrm{Bias}_v^{(k)}$ denote the weighted field error and signed bias at lead k. A representative policy objective is

$$\mathcal{L}_{\mathrm{policy}} = \frac{1}{H_s}\sum_{k=1}^{H_s} \lambda_k \left(\ell_Q^{(k)} + \ell_T^{(k)}\right) + \lambda_{\mathrm{anc}} R_{\mathrm{anc}},$$

with

$$\ell_v^{(k)} = c_{\mathrm{RMSE},v}\,\mathrm{RMSE}_v^{(k)} + c_{\mathrm{bias},v}\,|\mathrm{Bias}_v^{(k)}|.$$

This equation is a generic schematic, not the sequential adaptation loss. The matched-readout study uses scaled Q/T mean-square errors over the first four leads of a six-lead rollout, with a squared adapter-update penalty of coefficient 0.1; no explicit bias-loss term is established for the matched-readout study. Its exact optimizer budget is given in Appendix O.2.

I Low-Rank Routing Adaptation

For a selected linear layer W, the parameter-efficient update has the usual form

$$W' = W + \frac{\alpha}{r} B_L A_L, \qquad r \ll \min(\dim W).$$

The parent weights remain frozen. The matched-readout study uses rank-8, alpha-16 adapters at three encoder linear layers: the input projection, the third residual block's second linear layer, and the embedding projection. Their total trainable budget is 22,712 parameters. Input anchors, response atoms, and metric tensors remain fixed, while query and anchor embeddings change together. Full settings are given in Appendix O; the separate transfer deployment uses the settings in Appendix Q.

J Proof of the Routing-Displacement Bound

The following proof assumes ideal nonnegative weights satisfying $\sum_j w_j = \sum_j w'_j = 1$. It is a property of response mixtures, not a guarantee for coupled simulator trajectories. Let $\delta_j = w'_j - w_j$ and

$$p = \sum_{\delta_j > 0} \delta_j = \frac{1}{2}\|\delta\|_1.$$

The positive and negative parts of δ/p define two points $b_+, b_- \in \mathcal{H} = \mathrm{conv}(B)$. Therefore

$$\pi_{w'}(z) - \pi_w(z) = p(b_+ - b_-),$$

which gives

$$\|\pi_{w'}(z) - \pi_w(z)\|_A \le \frac{\operatorname{diam}(\mathcal{H})}{2}\|w'(z) - w(z)\|_1.$$

Convex-hull membership and this bound follow under ideal unit-sum normalization. For the implementation, introduce a zero atom in standardized output coordinates with weight $\widetilde{w}_0 = \epsilon/(S+\epsilon)$ and use $\widetilde{w}_j = \phi_j/(S+\epsilon)$ for the original atoms. These augmented weights sum to one, so the same argument holds with $\mathcal{H}_0 = \operatorname{conv}(B \cup \{0\})$ and the augmented weight vector, not necessarily the original hull. The deviation from the ideal mixture for $S > 0$ is $\|\widetilde{\pi} - \pi\|_A = [\epsilon/(S+\epsilon)]\|\pi\|_A$. De-standardization maps the zero atom to the output normalization mean. No uniform negligible-shrinkage assertion is made for states with small kernel sums.

For an optional local smoothness statement under ideal normalization, the following isotropic Gaussian logits provide an illustrative special case; the gradient condition below also applies to the anisotropic logits in Method. Write

$$s_j(u) = -\frac{\|u - c_j\|_2^2}{2\sigma_j^2}, \qquad \alpha(u) = \operatorname{softmax}(s(u)).$$

On a convex domain Ω, if every logit is differentiable and $\|\nabla s_j(u)\|_{U,*} \le G$, then differentiating the mixture and integrating along the line segment in Ω yields

$$\|\pi(u') - \pi(u)\|_A \le G\ \operatorname{diam}(\mathcal{H})\ \|u' - u\|_U.$$

This is a conditional local sensitivity bound for the representation, not a theorem about long-horizon stability of the numerical simulator.

K Real-Simulator Gate and Reproducibility Contract

An acceptance-filtered protocol would require a prespecified numerical-survival horizon and parent-relative moisture/temperature tolerances. The archived sequential adaptation launcher instead accumulates generated valid rollout prefixes; those records do not establish filtering by such acceptance criteria. No reference state is written back during evaluation. We therefore distinguish reported survival and accuracy from candidate acceptance, and do not infer numerical acceptance tolerances from the manuscript summary.

The matched-readout reproducibility record comprises the offline training split and normalization statistics; the 4,096 input anchors, response dictionary, and anchor-specific metric tensors; the encoder and surrogate checkpoints; replay members and sampling settings; adapter configuration; and simulator executable, forcing, initialization, and coupling interval. Network and schedule settings are in Appendices N and O. Coupled evaluation records should retain valid-prefix lengths and error metrics, rather than imply acceptance-filtered trajectories or unspecified gate tolerances. The separate transfer record is given in Appendix Q.

L VERTICAL, SPECTRAL, AND GEOGRAPHIC ERROR DIAGNOSTICS

The following figures diagnose the selected 12 h-tuned DARS deployment and independently rerun Native physics relative to the same Truth trajectories over 0–72 h. They are separate from the matched response-space ablation. Global signed bias can cancel across cells and levels, whereas RMSE retains local error magnitude; spectral coherence measures phase-aligned agreement, and a prediction-to-Truth power ratio of one indicates matched spectral amplitude.

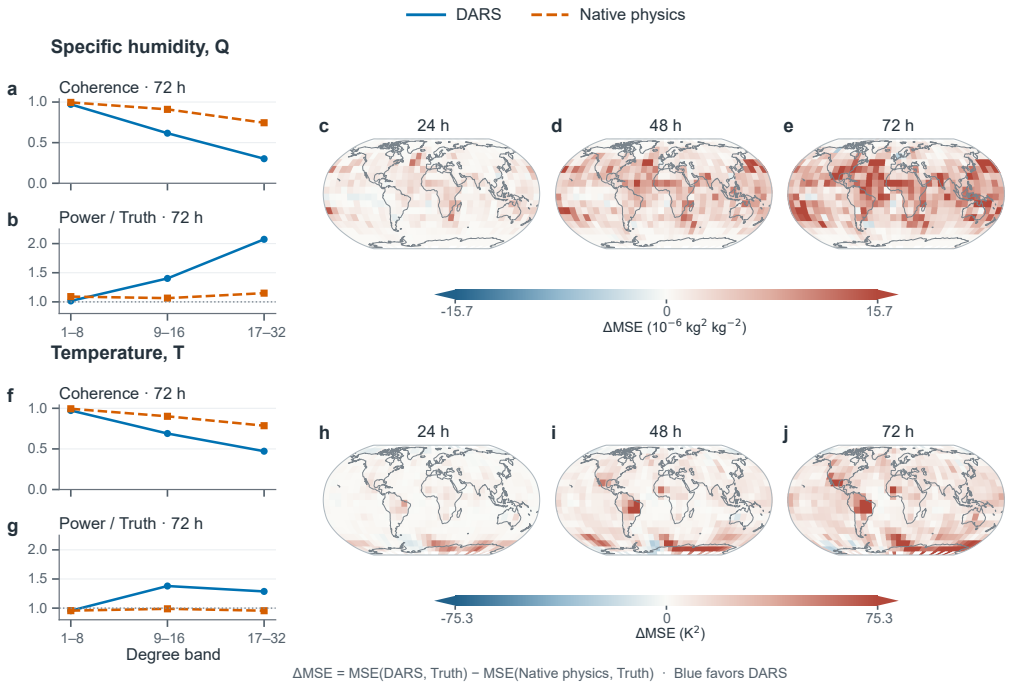

Figure 10: **Spectral and geographic error for both specific humidity and temperature.** For the illustrative transfer initial condition 19881004, panels compare 12 h-tuned DARS with independently rerun Native physics against the same Truth trajectory. The left panels show 72 h spectral coherence and prediction-to-Truth power ratios; maps show DARS-minus-Native mean-squared-error differences at 24, 48, and 72 h. Blue favors DARS and red favors Native physics. Vertical bias and RMSE profiles are shown separately in Figure 7.

M NATIVE-GRID FORECAST ERRORS AND POLICY RESPONSES

Figures 11–14 show local errors and DARS responses at two endpoints that complement the vertical profiles in Figure 7 and the spectral/geographic summary in Figure 10: 12 h marks the adaptation window, and 72 h marks the deployment horizon. Each field derives from 40,962 paired G6 cells against the same Truth; no remapping enters the reported metrics. Hexagonal longitude–latitude aggregation is for display only. Error scales are symmetric and shared by DARS and Native physics within each variable, level, and time. The DARS response is evaluated on its archived real-simulator parent state, not a surrogate trajectory.

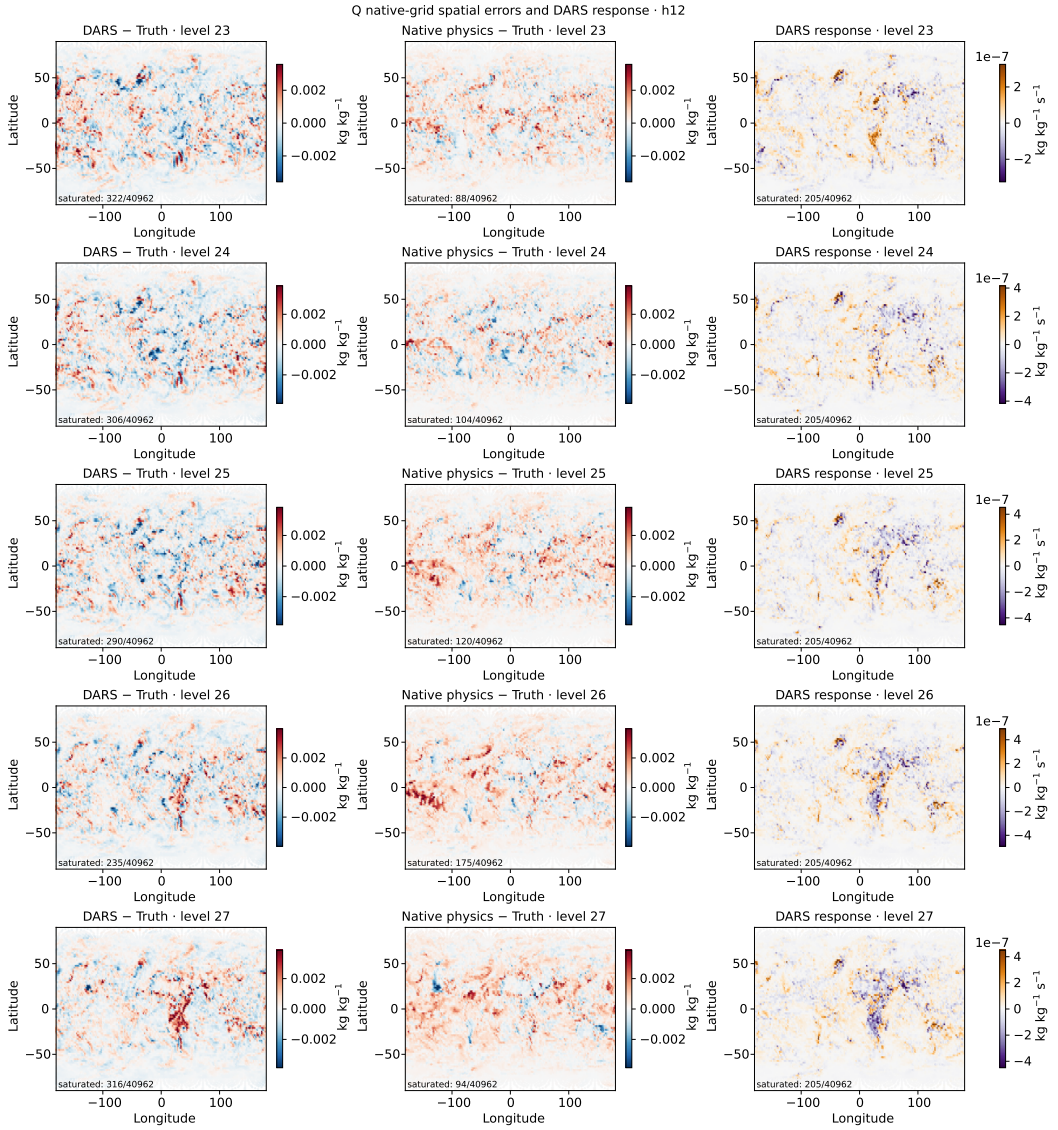

Figure 11: **Specific-humidity errors and DARS responses at 12 h.** Rows are levels 23–27 (top to bottom); columns show DARS-minus-Truth error, native-physics-minus-Truth error, and the DARS response evaluated on the matching DARS parent state. Native-grid values are displayed through equal-angle hexbin means only; quantitative metrics retain the original 40,962 cells and official area weights. For each level and time, the two error panels share a zero-centered symmetric scale set by their combined 99.5th percentile; printed saturation counts report values outside that display range. The response uses a separate tendency scale.

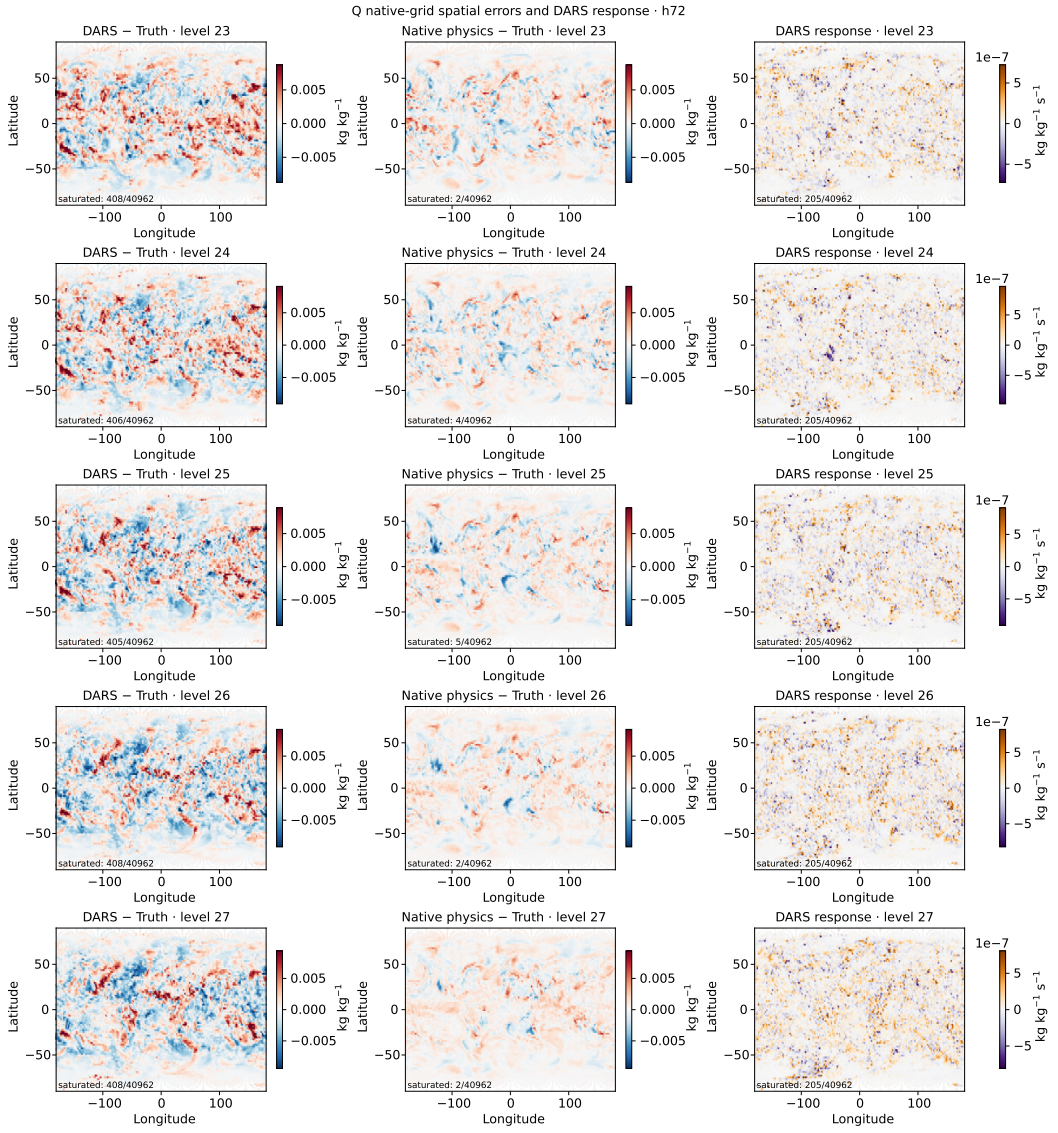

Figure 12: **Specific-humidity errors and DARS responses at 72 h.** Layout, levels, Truth reference, hexbin display, percentile clipping, and saturation-count convention follow Figure 11. Color magnitude should not be compared directly across lead times because each level–time pair has its own shared method scale.

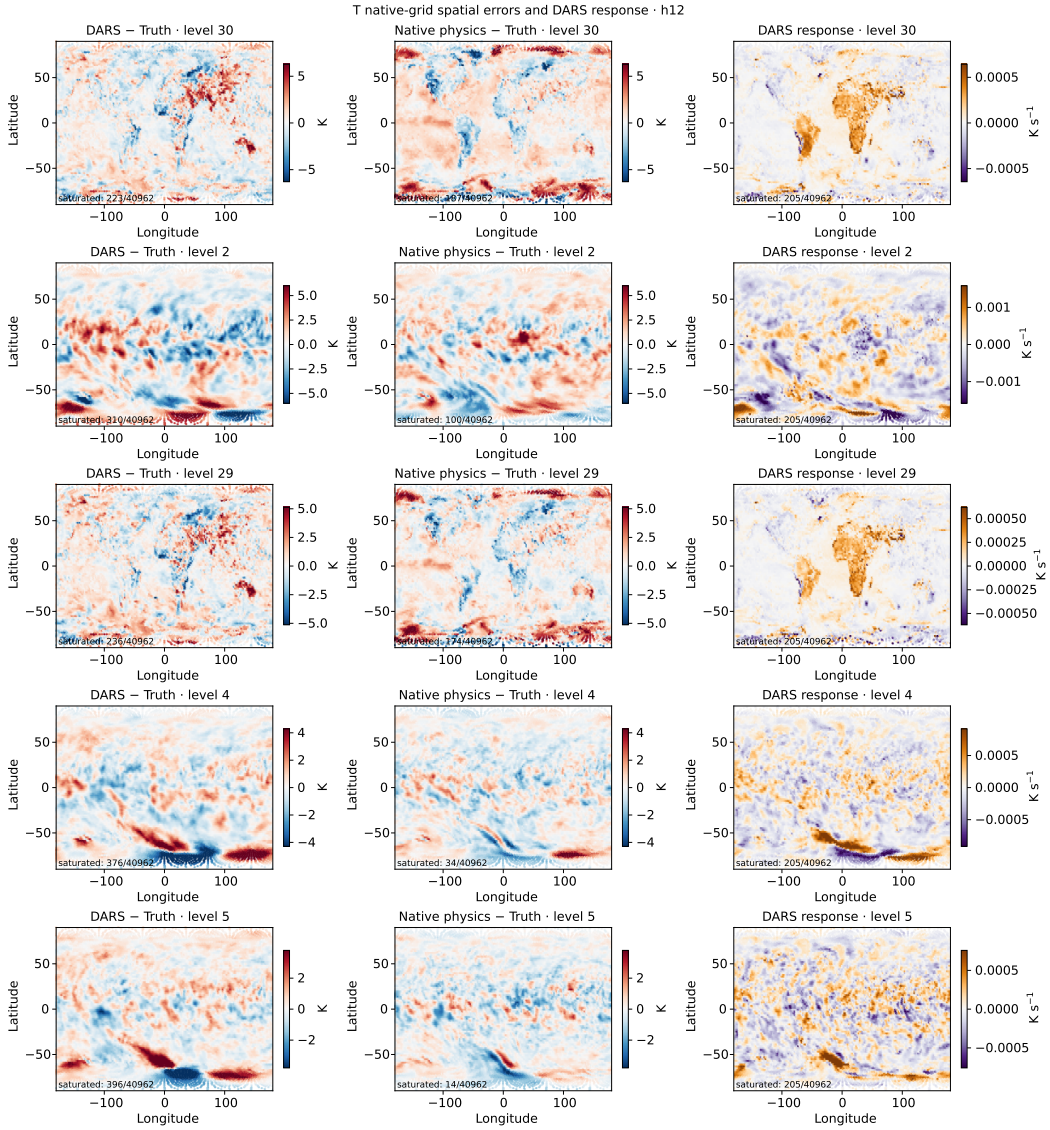

Figure 13: **Temperature errors and DARS responses at 12 h.** Rows are levels 30, 2, 29, 4, and 5; columns show DARS-minus-Truth error, native-physics-minus-Truth error, and the DARS response evaluated on the matching DARS parent state. Native-grid values are displayed through equal-angle hexbin means only; quantitative metrics retain the original 40,962 cells and official area weights. For each level and time, the two error panels share a zero-centered symmetric scale set by their combined 99.5th percentile; printed saturation counts report values outside that display range. The response uses a separate tendency scale.

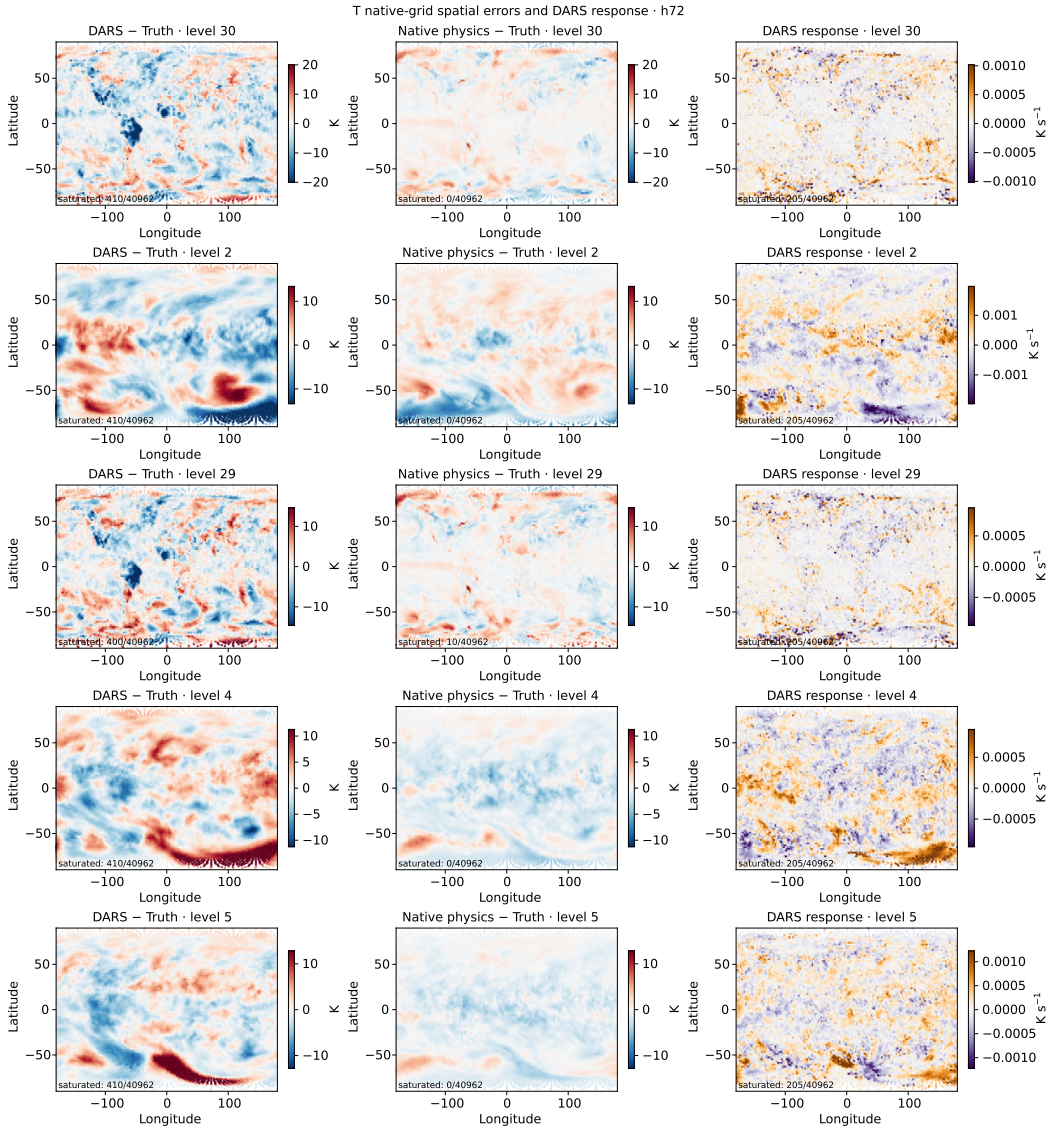

Figure 14: **Temperature errors and DARS responses at 72 h.** Layout, levels, Truth reference, hexbin display, percentile clipping, and saturation-count convention follow Figure 13. Color magnitude should not be compared directly across lead times because each level–time pair has its own shared method scale.

N NETWORK ARCHITECTURES

N.1 SHARED INPUTS, OUTPUTS, AND FEATURE EXTRACTOR

The matched-readout comparison uses a common atmospheric-column input and output interface for all three response parameterizations. The packed input contains 151 features: the 30-level profiles of specific humidity, temperature, resolved humidity tendency, resolved temperature tendency, and pressure, together with the cosine of the solar zenith angle. Inputs and outputs are standardized using statistics computed from the offline training split. The model predicts a joint 60-dimensional response consisting of 30 humidity physical tendencies and 30 temperature physical tendencies. The deployed wrapper selects the required features from the fixed simulator API and converts the standardized predictions back to physical units.

All three arms use the same feature-extractor architecture, but the separately trained DARS and unrestricted models need not have identical learned weights. The limiter uses the unrestricted model's weights with an additional output clamp.

Table 5: Shared feature-extractor configuration.

Component	Configuration
Input projection	Linear, $151 \to 512$
Residual stack	Six residual MLP blocks
Each residual block	LayerNorm(512) $\to$ Linear(512, 1024) $\to$ GELU $\to$ Linear(1024, 512), with a residual connection
Residual scaling	Configured scale 1.0, together with the block's learned residual multiplier
Embedding projection	Linear, $512 \to 128$
Embedding dimension	128
Dropout	0
Response dimension	60, jointly covering Q and T

N.2 DARS READOUT IN THE MATCHED-READOUT STUDY

The DARS arm in the matched-readout study uses 4,096 input anchors and a corresponding dictionary of 60-dimensional response atoms. The query and the fixed input anchors are passed through the same feature extractor. Consequently, adapting the feature extractor changes both the query coordinates and the encoded anchor coordinates, while the original input anchors and response dictionary remain fixed during sequential adaptation.

For query embedding $h(x)$ and encoded input anchor $h(c_j)$, the DARS routing distance is

$$d_j(x) = \delta_j(x)^\top M_j \delta_j(x), \qquad \delta_j(x) = h(x) - h(c_j), \tag{13}$$

$$M_j = \mathrm{diag}[\exp(-2s_j)] + U_j U_j^\top, \qquad U_j \in \mathbb{R}^{128\times 16}. \tag{14}$$

Each anchor has a 128-dimensional log-width vector and a rank-16 metric factor; diagonal inverse-width terms are capped at 10^6. The readout uses $w_j = \exp[-d_j/2]/(\sum_m \exp[-d_m/2] + 10^{-12})$. These weights are nonnegative but sum to less than one. Exact convex-hull membership of the original atoms describes ideal normalization; the implemented standardized prediction belongs to the hull augmented by zero (Appendix J). No freely learned nonlinear decoder follows the mixture.

Table 6: DARS readout configuration during sequential adaptation.

DARS readout component	Configuration during sequential adaptation
Input anchors	$4{,}096 \times 151$; fixed
Response dictionary	$4{,}096 \times 60$; fixed
Query/anchor embedding	Shared 128-dimensional feature extractor
Gaussian metric	Anchor-specific diagonal plus rank-16 positive-semidefinite term
Log-width construction setting	Initial value -1.5; checkpoint values retained during adaptation
Metric parameters	Frozen during adaptation
Output	Normalized Gaussian-weighted dictionary mixture

The rank-16 metric factor is distinct from the rank-8 LoRA adapters used for fine-tuning. It specifies the routing geometry rather than the adaptation budget.

N.3 Unrestricted and componentwise-limited readouts

The unrestricted arm replaces the dictionary readout with a nonlinear decoder,

$$\widehat{a}(x) = W_2\,\mathrm{GELU}(W_1 h(x) + b_1) + b_2, \tag{15}$$

with dimensions $128 \to 48{,}459 \to 60$. The unusually wide hidden layer is used to closely match the DARS model's registered parameter count; it is not the feature-extractor width.

The componentwise limiter uses the same unrestricted network followed by a parameter-free clamp in physical output units,

$$a_i(x) = \min\{u_i, \max\{\ell_i, \widehat{a}_i(x)\}\}. \tag{16}$$

The 60 lower and upper bounds are the per-channel minima and maxima over the July 1988 offline training block, row IDs 73–2016. No validation or evaluation rows are used to determine these bounds. The bounds are reordered to match the deployed vertical-level convention before application. Limiting enforces individual channel ranges, but does not impose membership in the joint response dictionary's convex hull. The adaptation implementation uses the actual clamp gradient, which is zero in saturated regions, rather than a straight-through estimator.

Table 7: Registered and trainable parameter counts for the three evaluated DARS readouts.

Quantity	DARS	Unrestricted	Limiter
Registered model parameters in the exported base bundle	15,608,966	15,609,121	15,609,121
Trainable response-model parameters per adaptation round	22,712	22,712	22,712
Additional limiter parameters	—	—	0

The limiter count of 22,712 applies if the policy-update stage is reached. Registered totals exclude buffers and include tensors frozen during adaptation, including DARS dictionary and metric tensors. They are not the number of parameters optimized during fine-tuning. The unrestricted and DARS registered totals differ by 155 parameters.

O Three-Round Adaptation Schedule

O.1 Stage definitions and experimental scope

The matched-readout study uses offline training seeds 1117, 2027, and 3037. Stage $k = 0$ is the exported offline model before any online adaptation. Successful DARS/unrestricted rounds $r = 1, 2, 3$ yield sequential checkpoints $k = 1, 2, 3$; this does not imply that the limiter completes them. These stages are repeated measurements within three seed histories, not independent replicates. All use the 30 September 1988 initialization; each evaluation freezes one checkpoint during integration.

The generic schedule alternates surrogate refinement, response-model adaptation, model export, and an original-simulator rollout. Each arm maintains its own on-policy replay and separately refined surrogate; the arms do not share a fitted surrogate. The surrogate is initialized from the same locked base dynamics checkpoint at the start of each round, whereas the response model is initialized from the preceding round's deployed checkpoint. The updated response-model adapters are merged into that checkpoint before export.

O.2 Response-model adaptation

Only three feature-extractor linear layers receive LoRA updates: the $151 \to 512$ input projection, the $1024 \to 512$ second linear layer of the third residual block, and the $512 \to 128$ embedding projection. The same locations and adapter dimensions are used across the response parameterizations.

Table 8: Response-model adaptation settings.

Setting	Value
LoRA rank	8
LoRA alpha	16; update scaling alpha/rank $= 2$
LoRA dropout	0
LoRA initialization	A: Kaiming uniform; B: zero
Trainable parameter count	$8(151 + 512) + 8(1024 + 512) + 8(512 + 128) = 22{,}712$
Optimizer	Adam
Learning rate	1×10^{-4}, constant
Optimizer steps per round	4
Windows per optimizer step	4
Window draws per round	16
Virtual rollout	Six 20-minute leads
State-loss leads	First four leads of the six-lead rollout
State loss	Equal weighting of scaled Q and T mean-square errors
Adapter regularization	0.1 times the sum of squared LoRA weight updates
Gradient clipping	Global norm 1.0
Trust-region/backtracking option	Disabled in the generic schedule
Reference-surrogate window screen	Disabled in the frozen generic protocol

The loss differentiates through the fixed, round-specific surrogate, while only the response-model adapters are updated. The DARS dictionary and metric tensors, the unrestricted decoder, and the limiter bounds remain unchanged. Policy-update windows require the archived deployed-input diagnostics, which are available for the first 36 simulator steps, corresponding to the initial 12 hours. These windows are therefore short-horizon adaptation data even when the parent checkpoint's evaluation rollout continues beyond 12 hours.

O.3 Surrogate refinement and replay

The surrogate uses rank-8, alpha-16, zero-dropout LoRA updates, AdamW with learning rate 10^{-4} and zero weight decay, and gradient-norm clipping at 1.0. Each round contains 12 optimizer steps. At every optimizer step, six windows are sampled *from each replay member*, and the member losses are combined using the prescribed replay weights. Thus, "six windows per step" is not a total batch size of six windows across the whole replay pool.

Each window contains eight consecutive stored states for a six-lead, 20-minute-step surrogate rollout. Offline replay comes from the July training block, containing 1,944 stored rows. On-policy replay uses the arm's accumulated coupled trajectories, using the available valid prefix when integration terminates early. Window sampling uses shuffled queues without replacement until a queue is exhausted.

Table 9: Replay masses and window draws by adaptation round.

Round	Offline replay mass	On-policy replay members and masses	Surrogate steps	Surrogate window draws	Response steps / draws
1	0.50	$k0$: 0.50	12	$12 \times 6 \times 2 = 144$	4 / 16
2	0.50	$k0$: 0.25; $k1$: 0.25	12	$12 \times 6 \times 3 = 216$	4 / 16
3	0.50	$k0$: 0.125; $k1$: 0.125; $k2$: 0.25	12	$12 \times 6 \times 4 = 288$	4 / 16

The draws in this table are window evaluations, not necessarily distinct windows. In the archived seed-1117 DARS records, each on-policy replay member contains 217 stored states spanning the 72-hour rollout. In the corresponding unrestricted records, each member contains 45 valid stored states. Accordingly, the archived surrogate refinement in the sequential adaptation study is *not restricted to the first 12 hours*, even though the response-model update uses the short diagnostic-supported window. The two data horizons must be reported separately.

O.4 Limiter completion and evaluation

The frozen generic schedule was prescribed for the limiter as well, but its seed-1117 first-round surrogate refinement failed with a nonphysical virtual pushforward state, before response-model adaptation. This failure was reproduced in two attempts. Completion differs by seed: 1117 completes no response-model update, 2027 completes rounds 1–2 before failure in round 3, and 3037 completes all three rounds. Thus the limiter does not provide three complete, budget-matched adaptation histories; no available limiter rollout reaches 72 h. Its $k = 0$ rollout remains a valid offline-readout comparison; screened or rescue variants must be identified separately rather than pooled with the frozen generic protocol.

After each successfully exported checkpoint, the original simulator evaluates an uninterrupted rollout toward the 72-hour horizon using the same initialization. Checkpoint stages and seeds are reported separately. The strict $k = 0$ comparison, the sequential $k = 1$–3 adaptation results, and any limiter rescue experiments are distinct experimental records.

P Scope of the Matched-Readout Configuration

The main Method now describes the evaluated DARS readout and adaptation settings detailed above. Its fixed quantities are the input anchors, response atoms, metric tensors, and base weights; the encoded query and anchor coordinates both change when the encoder adapters are updated. Policy optimization uses short diagnostic-supported windows, while surrogate replay can extend beyond 12 h. The launcher accumulates valid generated rollout prefixes rather than establishing acceptance-filtered replay. The separate 22-initialization transfer deployment includes response-basis adaptation and prescribed temperature tendencies; its architecture, training windows, and frozen deployment schedule are documented in Appendix Q, rather than inferred from the matched-readout study.

Q ADAPTATION SETTINGS FOR THE 22-INITIALIZATION TRANSFER STUDY

Q.1 SCOPE AND MODEL CONFIGURATION

The transfer experiment is separate from the matched-readout study. Adaptation and deployment selection use only the first 12 h of the 30 September 1988 initialization, not the subsequent long-range trajectories. The selected deployment is reused on all 22 evaluation initializations without per-case retraining. These initializations are distinct from the adaptation initialization; they are not all held out from offline pretraining. The 12 h restriction concerns adaptation data, not the duration of optimization or the offline training-data window.

The exported DARS model has a 151-dimensional input, a width-512 encoder with six residual blocks (LayerNorm, $512 \to 1024 \to 512$, with GELU), and a 128-dimensional embedding. It uses 4,096 input anchors, a $4{,}096 \times 60$ response basis, and anchor-specific diagonal-plus-rank-16 routing metrics. The exported bundle contains 15,608,966 registered parameters; this is not the fine-tuning budget. Unlike the sequential adaptation study's fixed-basis encoder adaptation, the transfer system includes response-basis updates and prescribed temperature-tendency corrections, merged into frozen deployment models.

Q.2 TRAINING STAGES AND FINAL TEMPERATURE ADAPTER

The selected training lineage includes an early dynamics-surrogate stage, a surrogate-assisted temperature update, subsequent calibration, and a final directly supervised temperature-basis adapter. The early surrogate replay loader reads coupled rows 2–9 (approximately 0.7–3 h at a 20-minute interval). The subsequent surrogate-assisted temperature update reads state rows 18–25 (6–8.3 h) and diagnostic steps 19–24. These are short windows within the adaptation trajectory, not longer-range transfer trajectories. The final temperature stage does not use the dynamics surrogate.

Table 10: Final supervised temperature-basis adaptation in the selected transfer lineage. These settings are not the encoder-adapter configuration in the sequential adaptation study.

Setting	Value
Adaptation diagnostics	Steps 3–36, corresponding to 1–12 h
Inputs and targets	Equal numbers of reference-state and coupled-state inputs; reference physical-temperature tendencies as targets
Updated factors	$A_T \in \mathbb{R}^{4096\times 16}$, $B_T \in \mathbb{R}^{16\times 30}$
Rank / alpha	16 / 16
Trainable parameters	66,016 in this T stage; inherited Q factors and parent weights frozen
Optimizer / learning rate	Adam / 3×10^{-4}
Epochs / batch size	5 / 2,048 atmospheric columns
Gradient clipping	Global gradient norm capped at 1
Random seed	117
Selected loss mode	Raw, centered, and bias terms; no spectral term

For each diagnostic step and vertical level, the temperature-tendency error is divided by the reference target's root-mean-square magnitude, floored at 2×10^{-6} K s^{-1}. Let $e_{i\ell}$ denote this normalized error, let $\langle\cdot\rangle_A$ denote the cell-area-weighted batch mean, and define $\mu_\ell = \langle e_{i\ell}\rangle_A$. The selected loss is

$$\mathcal{L}_T = \lambda(t)\left[\langle e^2\rangle_{A,\ell} + \tfrac{1}{2}\langle (e-\mu)^2\rangle_{A,\ell} + \tfrac{1}{2}\langle \mu^2\rangle_\ell\right], \tag{17}$$

where $\lambda(t) = 1$ for $t \le 4$ h, 1.5 for $4 < t \le 6$ h, and 2 thereafter within the 12 h window. Diagnostic steps and columns are shuffled each epoch. The 66,016-parameter count applies only to this final temperature stage, not to the cumulative training lineage. Saved humidity factors are inherited and frozen in this stage.

Q.3 SELECTION AND FROZEN DEPLOYMENT

Selection uses hourly 1–12 h raw temperature field RMSE on the adaptation initialization, subject to humidity and temperature screening thresholds. The selected configuration is verified in two uninterrupted 12 h simulator rollouts. Neither the 22 transfer initializations nor longer-range trajectories are used for adaptation or deployment selection.

The selected configuration includes a spatially and vertically uniform added temperature tendency, not a direct temperature-state offset. Its hourly nodes from 0 through 12 h are, in K s^{-1},

$$(3.4 \times 10^{-6},\ 3.4 \times 10^{-6},\ 5.95 \times 10^{-7},\ 4.25 \times 10^{-7},\ 2.55 \times 10^{-7},$$
$$-\ 8.5 \times 10^{-8},\ -4.25 \times 10^{-7},\ -4.25 \times 10^{-7},\ -2.55 \times 10^{-7},\ 0,$$
$$2.55 \times 10^{-7},\ -2.55 \times 10^{-6},\ -1.7 \times 10^{-6}).$$

These nodes are linearly interpolated at the 20-minute coupling interval and merged into the response basis. Requests 1–36 use the pre-exported prefix models. Request 37 uses a 20-minute transition with added tendency -8.5×10^{-7} K s^{-1}; requests 38–217 use the fixed continuation model with the additional uniform tendency removed. State-dependent DARS inference continues throughout. The schedule is fixed before each transfer rollout: there is no runtime optimization, simulator restart, or reference-state injection. Thus the experiment evaluates a frozen deployment schedule rather than one unchanged checkpoint across the entire horizon.