WHEN MEMORY UPDATES HELP PREDICTION BUT HURT CONTROL: A FLIGHT DYNAMICS STUDY

Anonymous authors
Paper under double-blind review

ABSTRACT

Robots use recent observations to predict how they will respond to future actions. During planning, however, imagined actions supply no new observations. Should a dynamics model keep the memory inferred from observed history fixed, or update it along its own predicted trajectory? We study this question in quadrotor simulation using matched, separately trained Frozen and Updated memory designs under a common planner. With the original logged-training recipe and no layer normalization (LN), Updated improves logged prediction, including physical position, yet produces worse closed-loop tracking. Replacing logged actions with planner candidates reverses the mean prediction ranking even when histories and initial physical states are held fixed. We further show that training recipe and optimization budget can reverse which design controls better. At the long budget, Mixed training with candidate actions improves Updated's tracking but worsens Frozen's relative to Re-simulated training with logged actions. A prospective test on five new datasets confirms these opposing effects at both extrapolation winds. These findings show that memory design, training recipe and optimization budget jointly shape the link between prediction and control. They motivate assessing robot dynamics models on logged prediction, planner queries and closed-loop control before choosing how memory should evolve during planning.

1 INTRODUCTION

Before choosing an action, a robot must predict motion under conditions it may not directly observe. Wind changes a drone's response to thrust; recent motion provides evidence about that response. A world model can encode this experience into a context vector and condition its predictions on it. The planner then asks a different question: what would happen under actions the robot has not yet taken? Using experience to answer such counterfactual questions connects history-based adaptation to model-based control.

This connection introduces a design decision that is easy to hide inside a recurrent network. As the model rolls forward a physical state, should it also roll forward the memory inferred from past observations? Both choices are reasonable. Holding context fixed preserves a common reference across candidate futures. Updating memory can represent useful features of an evolving trajectory. But an imagined transition brings no new external measurement: a changed memory is now partly a consequence of the model's own prediction. The architecture must therefore determine how observed history and imagined evolution interact.

We compare these choices using a shared decoder architecture with matched inputs, parameters, state integration and selection rules. A secondary factor crosses updating with memory-input LN (Ba et al., 2016). Each architecture is trained in its own rollout mode; the comparison concerns learned designs, not a test-time switch that isolates one internal mechanism.

Our central finding is that the preferred memory design depends jointly on the actions evaluated, the training recipe and the optimization budget. The starting point is a paired prediction–control reversal under the original logged-training recipe. With no LN, Updated has the lowest mean logged prediction error, but Frozen provides better mean tracking at nonzero wind under the same planner. At 6.1 m/s, prediction error favors Updated (0.03905 versus 0.04092), whereas tracking RMSE favors Frozen (0.210 versus 0.323 m). Updating also reduces logged physical-position error, so its prediction advantage is not solely a consequence of the normalized state-error metric. To test transfer to

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.

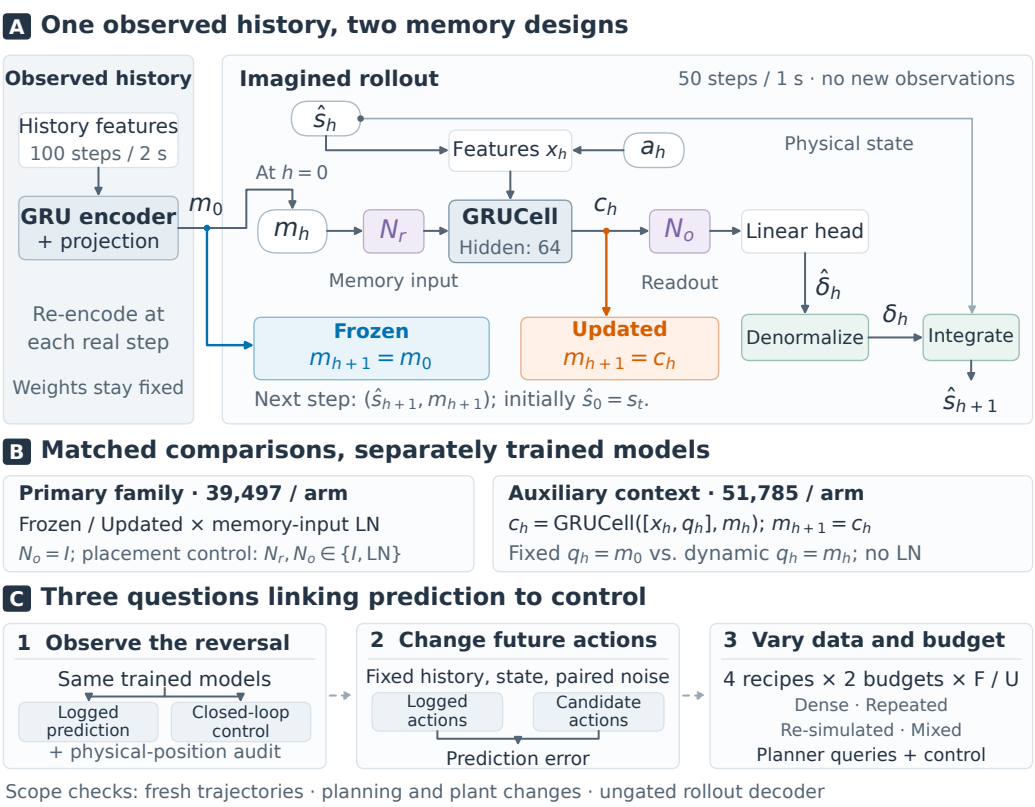

Figure 1: Controlled architecture and evidence sequence. **A:** One rollout step, repeated for $h = 0, \ldots, 49$. History initializes m_0 once per real step; blue and orange routes assign next-step memory under alternative memory designs. Transition features x_h contain state, body-frame velocity and action, normalized with training statistics. The full GRUCell output c_h branches before readout normalization; $\widehat{\delta}_h$ is de-normalized before physical integration. **B:** The primary family changes memory design; a separate no-LN family gives both arms dynamic state and varies auxiliary context access. Parameter counts match within each family, not across them. Normalization is a secondary control; A depicts the primary cell rather than the augmented cell. **C:** Research proceeds from reversal to action source to training; the latter includes a separate new-data confirmation. Dashed arrows do not indicate causality or checkpoint transfer. F/U denotes Frozen/Updated.

planning queries, we hold observed histories and physical states fixed and replace logged actions with planner candidates. This reverses the mean prediction ranking.

Prediction and control objectives can disagree, as value-aware learning and objective-mismatch studies establish (Farahmand et al., 2017; Lambert et al., 2020). Our contribution is to establish how future action source, training recipe and optimization budget change the preferred memory design in this flight-dynamics setting. The contribution has three linked parts. **We establish a paired reversal:** the same trained models favor Updated on logged prediction and Frozen on closed-loop tracking, with replication on fresh trajectories. **We locate an action-source boundary:** replacing logged actions with planner candidates reverses the mean prediction ranking at fixed observed history, initial state and paired physical noise. **We map when the preferred memory design changes:** four data recipes crossed with two budgets show that longer logged-action training can improve prediction yet hurt control; at the long budget, Mixed versus Re-simulated training improves Updated while worsening Frozen at both extrapolation winds, confirmed prospectively on five new datasets. Together, these contrasts show why neither logged prediction error nor a memory design considered alone determines the better controller in this study.

2 MEMORY BETWEEN OBSERVATION AND IMAGINATION

At real time t, observed history initializes m_0. Frozen reuses it throughout a candidate future; Updated carries each new decoder output. After executing one action, both controllers re-encode fresh observations. Thus "Frozen" refers only to imagination, not to adaptation between real steps; weights remain fixed and context inference involves no online gradient training. The encoder need not identify wind or a static physical parameter. We test a computational design choice by crossing updating with normalization and following all four arms through logged prediction, planner queries and closed-loop tracking. This compares two information paths: direct access to m_0 at every imagined step versus retaining context through the evolving m_h. It does not isolate recurrence independently of context access. A follow-up therefore gives both arms a dynamic rollout state and varies an auxiliary context input (Appendix N). This narrows the context-access question, while learned representations and optimization remain possible mediators.

3 CONTROLLED MODELS AND EVALUATION

Task and common transition. A model receives $K = 100$ observed steps (2 s at 50 Hz) and predicts $H = 50$ steps (1 s), given future actions. State comprises world velocity v, rotation R and body angular rate ω. The encoder receives state, body-frame velocity, actions and observed state increments. The decoder receives predicted state, body-frame velocity and the supplied action. Future observed states are never rollout inputs. Offline evaluation uses logged actions; the planner supplies candidate actions in closed-loop control. Input normalization and metric scales use training data only.

A 64-dimensional GRU encoder and linear projection produce m_0. Write $\hat{s}_h$ for the state predicted at real step $t + h$, with $\hat{s}_0 = s_t$, and a_h for the action at that step. At step h, all arms use the same GRU transition and linear prediction head:

$$c_h = \mathrm{GRUCell}(x_h, N_r(m_h)), \qquad \widehat{\delta}_h = WN_o(c_h) + b, \tag{1}$$

$$m_{h+1} = \begin{cases} m_0 & \text{Frozen,} \\ c_h & \text{Updated.} \end{cases} \tag{2}$$

Here x_h contains normalized transition features and $\widehat{\delta}_h$ is the normalized nine-dimensional head output. Before integration we recover $\delta_h = \sigma_\delta \odot \widehat{\delta}_h + \mu_\delta$ using training statistics. The four primary arms cross memory updating with $N_r \in \{I, \mathrm{LN}\}$, holding $N_o = I$. All have exactly 39,497 trainable parameters. LN is non-affine with $\epsilon = 10^{-5}$; it acts on the full hidden-state argument of the GRUCell, including its gated carry branch. It differs from the original LN GRU formulation; conclusions about this intervention do not generalize to that architecture. The placement study also crosses readout normalization $N_o \in \{I, \mathrm{LN}\}$. Updated arms always carry the raw cell output c_h, including when the readout is normalized.

Both memory designs compute a new cell output at every step and integrate predicted increments into the changing physical state: $v_{h+1} = v_h + \delta v_h$, $R_{h+1} = R_h \exp([\delta\theta_h]_\times)$ and $\omega_{h+1} = \omega_h + \delta\omega_h$. At identical weights and normalization settings, the first prediction is identical; subsequent predictions can differ through memory carry-over. Each arm is trained in its own rollout mode, rather than switching an already trained model to a different mode at test time. Training combines teacher-forced increment and open-loop rollout losses. Full feature dimensions, optimization settings and checkpoint rules appear in Appendix A.

Real flights and the target comparison. Public Neural-Fly logs (O'Connell et al., 2022) provide real-flight prediction evidence before controlled simulation. Held-out conditions include interpolation and extrapolation; wind and flight session remain confounded, and training seeds are not independent physical trials. Main models receive no wind metadata. Appendix B specifies temporal and whole-flight splits and the separate figure-8 stress test.

Prediction and control outcomes. Prediction error averages velocity, rotation and angular-rate errors over windows and horizon:

$$E = \frac{1}{NH} \sum_{i=1}^{N} \sum_{h=1}^{H} \frac{e_{v,i,h} + e_{R,i,h} + e_{\omega,i,h}}{3}. \tag{3}$$

Velocity and rate terms are mean absolute coordinate errors divided by coordinate-wise training scales; rotation uses geodesic error divided by its training scale. Scales are standard deviations of training H-step changes, clamped below at 10^{-6}. This convention does not make persistence equal to one. Displacement is a secondary outcome, excluded from E to avoid counting velocity twice. Closed-loop performance is measured separately as position tracking RMSE in meters.

Selection and uncertainty. Checkpoints use ID validation. Fixed-log comparisons pair ten training seeds, averaging overlapping windows within each seed. Two-sided Wilcoxon tests use Holm correction within each prospectively specified study family; paired t 95% intervals are descriptive, not simultaneous. Appendix C retains all 24 primary contrasts, including null results. Inspected OOD results motivated later questions, so separate follow-up families do not constitute one untouched confirmation test. In simulation, the unit is an independently generated training dataset ($n = 5$ per cohort); model seeds and trajectories are averaged within it. We report descriptive intervals and no simulation p-values.

Original simulation datasets. Five independently generated training datasets use shared quadrotor physics with aerodynamic drag, gusts and actuation noise. Each has 16 training flights, eight ID validation flights and ten disjoint test trajectories per condition. Training winds match the real-log range; static tests use 3.7, 6.1 and 8.5 m/s, with additional changing-wind tests. Three model seeds per dataset and four primary arms give 60 checkpoints. Normalizers and error scales use each dataset's training data only. These datasets vary training experience, not the underlying physics. Appendix G reports all prediction means; normalization and selection controls are collected in Appendix M.

4 When the prediction ranking reverses in control

4.1 The same predictors, inside the same planner

We place all 60 simulation checkpoints into a common MPPI controller (Williams et al., 2017), retaining the complete roster without OOD selection. Learned models, Nominal-physics MPPI and True-mean-wind MPPI use the same 64 action candidates, 50-step horizon, action-noise settings, cost regularization, trajectory preview and planning-seed mapping. Nominal-physics MPPI assumes zero wind; True-mean-wind MPPI receives the static mean wind but not future gusts. PD and Preview PD provide additional references. The goal is to measure the consequences of the model substitution under a common planning budget, not to compare separately optimized control systems.

Each episode shares initial conditions and environmental seeds across controllers, uses a two-second PD warmup and scores the next 30 seconds. Four static winds (0, 4.9, 6.1 and 8.5 m/s) span an accurate nominal-model condition, the training wind boundary and two extrapolation conditions. The study contains 2,400 learned episodes (5 datasets $\times 3$ model seeds $\times 4$ arms $\times 4$ winds $\times 10$ episodes) and 160 unique reference episodes. References are reused in paired comparisons, not copied into artificial independent replicates. All models re-encode the latest two seconds of observations at each real control step, following Figure 1.

The primary metric is failure-aware position RMSE: a divergent or nonfinite episode would receive the prespecified capped-error score rather than be dropped. All 2,560 episodes passed acceptance and none met failure criteria. We retain all outcomes; the absence of observed failures does not establish safety or a failure-rate advantage. Dataset-level uncertainty is conditional on ten common evaluation episodes per wind, distinct from prediction-test trajectories. Appendix H specifies caps, thresholds and the full protocol.

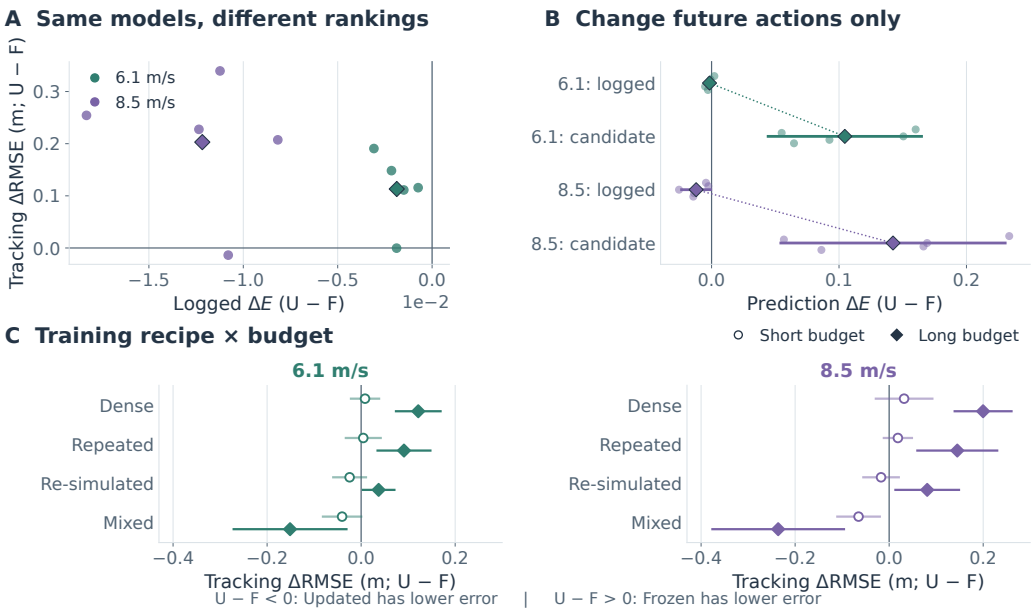

Figure 2: Evidence from ranking reversal to a training-data intervention. **A:** Original no-LN models: Updated minus Frozen (U–F). Four of five training datasets improve logged prediction but worsen tracking at each wind; prediction and control use distinct trajectories. **B:** At fixed logged histories, physical states and paired noise, replace logged actions with a new MPPI candidate bank. The U–F prediction gap shifts adversely at both winds. **C:** U–F tracking gaps across four training recipes and two budgets (short: circles; long: diamonds). The 240 no-LN models use matched parameters and within-budget optimization steps; 60 short-budget models are reused from the first coverage study. Candidate exposure changes the long-budget ranking. A/B use the original models. A/B dots are dataset scores and diamonds means; C markers are means. B/C bars are descriptive paired 95% intervals across five datasets.

4.2 Prediction and control favor different memory designs

Figure 2A pairs the no-LN prediction and control differences. Updated minus Frozen prediction error is -0.00187 ($[-0.00294, -0.00080]$) at 6.1 m/s and -0.01217 ($[-0.01683, -0.00751]$) at 8.5 m/s. Four of five datasets jointly improve prediction and worsen control at each wind; the same dataset is the exception in both. Complete four-arm means remain in Appendices G and H.

Updated minus Frozen tracking RMSE is 0.11311 m at 6.1 m/s and 0.20298 m at 8.5 m/s, with paired dataset-level intervals $[0.02502, 0.20120]$ and $[0.04009, 0.36588]$ m. Four of five datasets favor Frozen at every nonzero wind; the fifth slightly favors Updated and is retained. At zero wind, the 0.03043 m mean difference has an interval spanning zero. These descriptive intervals condition on the common evaluation episodes. Appendix H retains all 32 planned contrasts.

A locked follow-up evaluates the same 30 no-LN checkpoints on 20 new trajectories per shared wind (1,200 episodes). Updated minus Frozen RMSE is 0.112 m ($[0.022, 0.202]$) at 6.1 m/s and 0.200 m ($[0.050, 0.349]$) at 8.5 m/s, favoring Frozen in 4/5 and 5/5 datasets. All 20 episode-level contrasts are positive at each wind after averaging the trained roster (Appendix K). This extends the ordering to new draws from the same task distribution; it does not imply that better prediction causes worse control.

The offline advantage includes position. A full-window audit integrates predicted velocity into physical position and compares it with recorded positions, reproducing all accepted normalized error arrays exactly. Updated reduces horizon-averaged position error by 0.00611 m ($[0.00410, 0.00812]$) at 6.1 m/s and 0.05666 m ($[0.04653, 0.06678]$) at 8.5 m/s. Thus its offline advantage also holds for a control-relevant quantity; it is not explained solely by averaging different state components. Component and horizon decompositions retain early exceptions.

4.3 USEFUL WIND COMPENSATION HAS CLEAR BOUNDARIES

Both Frozen variants outperform Nominal-physics MPPI in all five datasets at 4.9 and 6.1 m/s. At 6.1 m/s, Frozen (no LN)/(LN) mean RMSE is $0.210/0.188$ m versus 0.315 m for Nominal-physics MPPI. At 8.5 m/s, Frozen (LN) improves in all five datasets; Frozen (no LN) improves in four, with an interval crossing zero. Context held fixed during imagination can thus support useful wind response through repeated inference from observations.

The boundary conditions matter. At zero wind, every learned variant loses to Nominal-physics MPPI in every dataset; both Updated variants also lose at 8.5 m/s. True-mean-wind MPPI has the lowest grand-mean MPPI error at every wind, but receives privileged wind information and is not an optimal-control bound. Figure 3 retains all controller means and the less certain closed-loop LN-by-memory interaction. Mean learned planning time is 24.14 ms on CUDA, above the 20 ms control period. Physics references run on CPU, so these timings establish neither an algorithmic speedup nor a real-time hardware-flight demonstration (Appendix H).

5 CHANGING THE FUTURE ACTIONS CHANGES THE PREDICTION RANKING

Offline logs and controller rollouts differ in both history and future actions. To isolate the action-source contrast, we fix 100 logged first-decision anchors (five datasets, ten flights, two winds), with the same history, initial state and eight paired physical-noise samples, and replace only the future action source: the logged sequence versus a freshly generated 64-sequence MPPI bank. On the logged sequence, Updated minus Frozen E is -0.00160 at 6.1 m/s (interval crosses zero) and -0.01210 at 8.5 m/s ($[-0.02352, -0.00068]$). On candidates it becomes $+0.10456$ and $+0.14237$, with intervals above zero. The paired candidate-minus-logged change is positive in all five datasets at each wind (Appendix L). This isolates action source in this query test; it neither explains the entire closed-loop gap nor identifies memory corruption as its internal cause.

5.1 ACTUAL PLANNER QUERIES AND RETURNED SOLUTIONS

We next examine the queries actually posed by MPPI, retaining all 60 GRU checkpoints. At the 40 shared first planning states (four winds, ten episodes), all arms see identical observed histories and the actual 64 candidate sequences. We compare their predicted trajectories and candidate-cost ordering against physical simulation. We also evaluate the complete softmin-weighted action sequence returned by each model, rather than substituting its best candidate. Eight common-noise physical rollouts construct the oracle costs; eight independent evaluation-noise rollouts score the solutions. This is an open-loop query test, not a decomposition of 30-second closed-loop tracking.

With no LN, Updated now has higher horizon-averaged prediction error E and poorer candidate-cost ordering at every wind: its logged-action advantage does not transfer to these planner proposals. Updated minus Frozen physical solution cost is $+0.361$, $+1.213$, $+1.723$ and $+2.632$ at winds 0, 4.9, 6.1 and 8.5 m/s, respectively, positive in all five training datasets in each case. Terminal component analysis localizes an angular-rate gap already at step 1; at the shared held-out winds, position differences become adverse by step 25. The gap also occurs in zero-perturbation candidates, so no large-action threshold is identified (Appendix J). On these first-decision queries, the logged-action advantage is lost in both trajectory prediction and the physical solution cost.

Later queries require preserving the planner's action-sequence warm start. We reconstruct 180 original Frozen/Updated donor episodes and evaluate their actual candidate banks at three fixed indices, yielding 540 anchors. All 225,180 replayed decisions reproduce the logged actions exactly. These comparisons are conditional on donor policy and time: all 18 strata are retained in Appendix I. Their prespecified balanced secondary differences in physical solution cost are $+0.047$ at zero wind, $+2.373$ at 6.1 m/s and $+3.472$ at 8.5 m/s, with intervals $[-0.874, 0.968]$, $[0.746, 3.999]$ and $[1.485, 5.460]$. Some late strata reverse; neither uniform harm nor a unique internal mechanism follows from the windy averages.

Table 1: Training recipe and budget change tracking. Failure-aware RMSE (m); learned models average five datasets, three seeds and ten episodes per wind. S/L: 5,760/21,360 steps. Physical references use the same ten episodes, without training. Figure 2C shows paired U–F uncertainty; Appendices P–Q retain all winds and contrasts.

		6.1 m/s		8.5 m/s	
Recipe	Budget	Frozen	Updated	Frozen	Updated
Dense	S	0.195	0.203	0.354	0.386
Dense	L	0.206	0.328	0.363	0.563
Repeated	S	0.213	0.218	0.371	0.390
Repeated	L	0.217	0.308	0.395	0.540
Re-simulated	S	0.209	0.185	0.351	0.333
Re-simulated	L	0.161	0.198	0.292	0.373
Mixed	S	0.178	0.138	0.306	0.241
Mixed	L	0.263	0.112	0.435	0.199
Nominal-physics MPPI		0.310		0.454	
True-mean-wind MPPI		0.067		0.132	

6 Training recipe and budget change the memory ranking

The first matched coverage study improves both architectures' mean tracking, but its Re-simulated arm already changes the original ordering (Appendix O). To examine that discrepancy, we cross four recipes with two budgets and Frozen/Updated: 240 no-LN models across five datasets and three seeds, reusing the 60 short-budget coverage models. Each logical epoch has 12,288 examples. *Dense* draws fresh windows from the original 45,600-window training pool; *Repeated* uses 1,536 original windows eight times each; *Re-simulated* retains those histories and logged actions but generates eight noisy futures; *Mixed* replaces half with MPPI candidate actions and their induced future states. Budgets are 5,760/21,360 steps, with complete learning-rate schedules and ID-only checkpoint selection. Budget therefore includes schedule duration and selection opportunities, not just compute. All models and 30 original no-LN reference models share 40 fresh episodes and common first-decision query banks, including zero wind. All 10,800 closed-loop episodes pass acceptance without failures (Appendix P).

More logged training need not improve control. At 6.1 m/s, longer Dense training reduces Updated's logged error E from 0.06797 to 0.03856, yet increases tracking RMSE from 0.203 to 0.328 m. The U–F tracking gap grows by 0.113 m ($[0.038, 0.189]$); the corresponding 8.5 m/s change is 0.168 m ($[0.059, 0.277]$). Both are positive in all five datasets. Dense and Repeated training favor Frozen at the long budget; their short-budget intervals cross zero at both primary winds. The Dense arm approaches the original mean ordering without reproducing its early-stop protocol.

Candidate exposure helps Updated but hurts Frozen at the long budget. At the long budget, Re-simulated training favors Frozen, whereas Mixed training favors Updated in all five datasets at both primary winds (Figure 2C). Mixed–Re-simulated changes in the U–F gap are -0.188 m ($[-0.302, -0.075]$) and -0.317 m ($[-0.422, -0.212]$). This combines absolute improvement for Updated with deterioration for Frozen: at 6.1 m/s their changes are -0.086 m ($[-0.128, -0.044]$) and $+0.102$ m ($[0.004, 0.201]$). The same opposing signs hold at 8.5 m/s. Both architectures instead improve in mean at the short budget, where the relative coverage effect remains uncertain at 6.1 m/s. The long–short changes in the coverage interaction are -0.172 m ($[-0.287, -0.058]$) and -0.269 m ($[-0.372, -0.166]$), negative in all five datasets.

At the long budget, Mixed training also reduces U–F gaps in physical solution cost from $1.254/2.574$ to $0.041/0.145$ at 6.1/8.5 m/s. The Mixed intervals cross zero and do not establish equivalence. Frozen's mean physical solution cost improves even as its closed-loop tracking worsens: query gains alone do not establish a control gain. Training-support and re-simulation contrasts are less decisive. The benefit of candidate exposure therefore depends on memory design and budget; these conditional changes do not establish coverage alone as a repair of the original reversal.

Table 2: Long-budget Mixed minus Re-simulated tracking RMSE (m). Each cohort contains five independently generated training datasets and uses its own shared evaluation roster. Cohorts are analyzed separately. Values are means with descriptive paired 95% intervals; negative values favor Mixed.

Wind (m/s)	Memory design	Original datasets	Confirmation datasets
6.1	Frozen	$+0.102$ $[0.004, 0.201]$	$+0.059$ $[0.023, 0.096]$
6.1	Updated	-0.086 $[-0.128, -0.044]$	-0.176 $[-0.294, -0.057]$
8.5	Frozen	$+0.144$ $[0.031, 0.257]$	$+0.085$ $[0.004, 0.165]$
8.5	Updated	-0.173 $[-0.244, -0.102]$	-0.286 $[-0.462, -0.111]$

On the training-bridge episodes, long-budget Mixed Updated has lower RMSE than Nominal-physics MPPI at every nonzero wind. It remains worse than True-mean-wind MPPI and loses to both references at zero wind (Table 1; Appendix Q).

The opposite effects recur on new training data. With a protocol fixed before new outcomes, we train 120 models on five new datasets, crossing Re-simulated/Mixed with both budgets and memory designs. All 4,800 fresh control episodes pass acceptance without failures. At the long budget, Mixed helps Updated and hurts Frozen in every new dataset at both primary winds; all four paired intervals exclude zero (Table 2). This confirms the opposite effects within the same physics family; cohorts are analyzed separately (Appendix R).

7 Robustness and limits

Fixed-protocol extensions use the original GRU checkpoints and five training datasets as the uncertainty unit. Three planning settings vary candidates (32/128 at $H = 50$) or horizon ($H = 25$, 64 candidates). Eight plant changes vary mass by $\pm 10\%$, drag by $\pm 20\%$, body-rate time constant to 0.06/0.10 s, or gust correlation time to 0.5/2 s, with nominal controller parameters. Across 27,920 episodes, all 44 setting–wind mean no-LN U–F RMSE differences remain positive. All 24 windy plant-shift intervals exclude zero, but every short-horizon and zero-wind interval crosses it (Appendix I). These findings do not test the coverage-trained models.

Replacing only the rollout GRUCell with a width-64 tanh RNNCell retains the GRU encoder, four-arm design, data and training recipe. The 60 models have 28,489 parameters each versus 39,497 for GRU; matching holds within families. With no LN, Updated improves prediction in all five datasets and all five offline conditions, yet has higher mean tracking error at every wind in 2,400 matched episodes. At 8.5 m/s the interval crosses zero despite 4/5 datasets favoring Frozen (Table 21). Thus the mean reversal extends beyond a gated rollout cell, with limited high-wind precision. All offline LN-by-memory interaction intervals cross zero. Control episodes in both extensions pass acceptance with zero numerical failures.

Normalization is a separate, narrower result. On real-flight logs, memory-input LN has a larger prediction penalty for Updated than Frozen: the interaction is 0.01105 (Holm $p = 0.02539$). It survives equal-budget ID learning-rate selection and is positive in all five simulation datasets; readout LN and Elman interactions remain inconclusive. This intervention differs from the original LN GRU formulation. Complete errors, selection results and exceptions remain in Appendix M. These findings motivate checking normalization jointly with memory design; they do not explain the no-LN prediction–control reversal.

8 Implications and scope

A practical evaluation procedure. Treat observation refresh, context access, imagined state evolution and training recipe as separate design choices. Compare memory designs at explicit optimization budgets. Before selecting a dynamics model by logged validation error, also test common candidate actions and the actual actions returned by the planner, then retain the complete model roster in closed-loop evaluation. Logged-action accuracy alone missed both the planner-query error reversal and the tracking disadvantage. Include conditions where Nominal-physics MPPI is already

adequate: all original learned variants lose at zero wind. Candidate-action training must be tested within each memory design and budget; its benefits can differ even in sign. This procedure does not prescribe freezing everywhere or assume latent coordinates represent physical quantities.

What remains unexplained and untested. The action-source test locates a transfer boundary; the training bridge tests conditional recipe changes, including action-induced future states, but excludes later policy-induced histories. Support and re-simulation are not fully crossed; longer budgets also change the learning-rate schedule and ID-selection opportunities. A separate dynamic-state architecture with persistent auxiliary context improves planner queries, with uncertain high-wind tracking benefit and worse logged position prediction (Appendix N). These tests do not identify representation geometry, gating or optimization as the cause. Later diagnostics remain conditional on two donor policies, with mixed strata. The ungated decoder retains a GRU encoder and changes total capacity. Two separate five-dataset cohorts vary training experience in one simulator; bounded plant shifts are evaluation perturbations, not independent-physics training. Real-flight logs supply prediction evidence, not hardware control validation. Fixed-log intervals measure seed variation; simulated intervals condition on the shared evaluation roster. OOD inspection motivated follow-ups, so complete protocols and rosters ensure traceability but do not remove adaptivity. The supported conclusion concerns matched memory designs under specified action, training and flight-simulation conditions.

9 Related work

History-based adaptation and world models. Hidden-parameter MDPs formalize dynamics conditioned on latent task properties (Doshi-Velez & Konidaris, 2016). CaDM separates history-context encoding from conditioned transition prediction and reuses context for future-step prediction (Lee et al., 2020). Recent context-aware neural ODE control also infers latent environmental information from state–action histories for MPC (Yu et al., 2026). Our question is how to carry this history information through imagined actions. We compare direct context access with recursive context retention under matched models, then test when action source, training recipe and budget change their prediction and control rankings. PEARL, rapid motor adaptation and Dreamer also connect experience to decisions (Rakelly et al., 2019; Kumar et al., 2021; Hafner et al., 2020); we do not reproduce these complete systems or rank their policies.

Dynamics learning for control. Ensemble planning (Chua et al., 2018) and short-rollout policy optimization (Janner et al., 2019) address how learned dynamics support decisions. Value-aware model learning incorporates value-function structure into its loss (Farahmand et al., 2017); objective-mismatch experiments show that one-step prediction likelihood need not correlate with control performance (Lambert et al., 2020). Our study connects this objective mismatch to a controlled memory-design comparison under a fixed planner and maps how training recipe and budget change the comparison. Neural-Lander and Neural-Fly connect learned aerodynamics to flight control (Shi et al., 2019; O'Connell et al., 2022). We use Neural-Fly logs for prediction and a simulator for action evaluation, without reproducing its adaptive controller or claiming state-of-the-art flight control.

10 Conclusion

In this flight-dynamics study, memory updating can improve logged prediction while worsening control. Fixed-anchor tests locate an action-source boundary; the training bridge shows that recipe and budget can reverse the preferred memory design. Compared with Re-simulated training, long-budget Mixed helps Updated but hurts Frozen at both extrapolation winds, confirmed on five new datasets. Evaluate robot dynamics models on logged prediction, planner queries and closed-loop control under explicit training conditions before adopting a memory design.

AI Use Disclosure. OpenAI tools were used for language editing and to assist with implementing portions of analysis and visualization code specified by the authors. All outputs were reviewed and validated by the authors.

Ethics statement

The study uses public flight logs and simulated trajectories. Closed-loop results are confined to simulation; the reported failure counts and timing do not certify safe autonomous operation. We retain unfavorable results and distinguish result-informed research from untouched-data confirmation.

Code availability

Anonymous code and accepted numerical records are available at `https://anonymous.4open.science/r/ICLR_NeuralFlyWAM-BF66/README.md`.

Reproducibility statement

Equations 1–3 define the controlled model and outcome. Appendices A–H specify preprocessing, training, statistical comparisons and simulation. Versioned manifests, training-only normalizers, per-run source hashes, evaluation artifacts and analysis scripts record the executed studies. All 24 planned fixed-log contrasts are included; simulation summaries retain all arms and the complete episode roster. The accompanying anonymous core supplement contains all 60 original, 30 auxiliary-context, 60 action-coverage, 180 training-bridge and 120 independent-confirmation GRU model tensor exports (450 unique models), resolved configurations, real-log and simulation split manifests, normalizers, accepted numerical records and scripts for the paired analyses and figures. Its README and hash manifest specify exact reproduction commands and distinguish portable exports from original accepted files. A self-contained replay reconstructs all 120 model evaluations at one first-decision query from exported tensors and simulator code, reproducing 360 prediction, position and physical solution-cost values exactly in the tested environment. Raw flight logs, full simulation/query traces and non-GRU checkpoints are outside this compact package; it supports auditing the core comparison, not complete end-to-end reproduction of every extension from the package alone.

REFERENCES

Jimmy Lei Ba, Jamie Ryan Kiros, and Geoffrey E. Hinton. Layer normalization. *arXiv preprint arXiv:1607.06450*, 2016. URL https://arxiv.org/abs/1607.06450.

Kurtland Chua, Roberto Calandra, Rowan McAllister, and Sergey Levine. Deep reinforcement learning in a handful of trials using probabilistic dynamics models. In *Advances in Neural Information Processing Systems*, 2018. URL https://arxiv.org/abs/1805.12114.

Finale Doshi-Velez and George Konidaris. Hidden parameter markov decision processes: A semiparametric regression approach for discovering latent task parametrizations. In *International Joint Conference on Artificial Intelligence*, 2016.

Amir-Massoud Farahmand, Andre Barreto, and Daniel Nikovski. Value-aware loss function for model-based reinforcement learning. In *Proceedings of the 20th International Conference on Artificial Intelligence and Statistics*, volume 54 of *Proceedings of Machine Learning Research*, pp. 1486–1494. PMLR, 2017. URL https://proceedings.mlr.press/v54/farahmand17a.html.

Akshat Gupta, Atahan Ozdemir, and Gopala Anumanchipalli. Geometric interpretation of layer normalization and a comparative analysis with RMSNorm. *arXiv preprint arXiv:2409.12951*, 2024.

Danijar Hafner, Timothy Lillicrap, Jimmy Ba, and Mohammad Norouzi. Dream to control: Learning behaviors by latent imagination. In *International Conference on Learning Representations*, 2020.

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*, 2019. URL https://arxiv.org/abs/1906.08253.

Ashish Kumar, Zipeng Fu, Deepak Pathak, and Jitendra Malik. RMA: Rapid motor adaptation for legged robots. In *Robotics: Science and Systems*, 2021.

Nathan Lambert, Brandon Amos, Omry Yadan, and Roberto Calandra. Objective mismatch in model-based reinforcement learning. In *Proceedings of the 2nd Conference on Learning for Dynamics and Control*, volume 120 of *Proceedings of Machine Learning Research*, pp. 761–770. PMLR, 2020. URL https://proceedings.mlr.press/v120/lambert20a.html.

Kimin Lee, Younggyo Seo, Seunghyun Lee, Honglak Lee, and Jinwoo Shin. Context-aware dynamics model for generalization in model-based reinforcement learning. In *Proceedings of the 37th International Conference on Machine Learning*, volume 119 of *Proceedings of Machine Learning Research*, pp. 5757–5766. PMLR, 2020. URL https://proceedings.mlr.press/v119/lee20g.html.

Clare Lyle, Zeyu Zheng, Khimya Khetarpal, James Martens, Hado van Hasselt, Razvan Pascanu, and Will Dabney. Normalization and effective learning rates in reinforcement learning. In *Advances in Neural Information Processing Systems*, 2024. URL https://arxiv.org/abs/2407.01800.

Michael O'Connell, Guanya Shi, Xichen Shi, Kamyar Azizzadenesheli, Anima Anandkumar, Yisong Yue, and Soon-Jo Chung. Neural-Fly enables rapid learning for agile flight in strong winds. *Science Robotics*, 7(66):eabm6597, 2022. doi: 10.1126/scirobotics.abm6597.

Simon Ouellette, Rolf Pfister, and Hansueli Jud. Counting and algorithmic generalization with transformers. *arXiv preprint arXiv:2310.08661*, 2023.

Kate Rakelly, Aurick Zhou, Deirdre Quillen, Chelsea Finn, and Sergey Levine. Efficient off-policy meta-reinforcement learning via probabilistic context variables. In *International Conference on Machine Learning*, 2019.

Guanya Shi, Xichen Shi, Michael O'Connell, Rose Yu, Kamyar Azizzadenesheli, Animashree Anandkumar, Yisong Yue, and Soon-Jo Chung. Neural lander: Stable drone landing control using learned dynamics. In *International Conference on Robotics and Automation*, 2019.

Grady Williams, Paul Drews, Brian Goldfain, James M. Rehg, and Evangelos A. Theodorou. Information theoretic model predictive control: Theory and applications to autonomous driving. *arXiv preprint arXiv:1707.02342*, 2017. URL https://arxiv.org/abs/1707.02342.

Shao-Yi Yu, Jen-Wei Wang, Maya Horii, Masayoshi Tomizuka, and Vikas Garg. Learning context-aware neural ODE dynamics for adaptive robotic control, 2026. URL https://arxiv.org/abs/2606.15469.

Biao Zhang and Rico Sennrich. Root mean square layer normalization. In *Advances in Neural Information Processing Systems*, 2019. URL https://arxiv.org/abs/1910.07467.

A TRAINING AND IMPLEMENTATION

Implementation. The state has 12 channels: velocity (3), a six-dimensional rotation representation (6), and body angular rate (3). The history has $12+3+5+9 = 29$ channels; the main-platform action is thrust plus an attitude quaternion. Transition features have $12 + 3 + 5 = 20$ channels. The history encoder applies dropout 0.1 to the final output of a 64-dimensional GRU, then a linear projection. The controlled head bias starts at zero and weights have standard deviation 10^{-3}. Shared parameters use the same initialization order across paired arms. The feature groups, memory equations and common integration are specified in Section 3.

Loss and optimization. The objective is the sum of two terms with unit weights. The first is mean squared error on the nine normalized next-step increments, with ground-truth states supplied to the transition. The second is the mean over windows and horizon of $(e_v^2 + e_R^2 + e_\omega^2)/3$, where e_v, e_R, e_ω are the component errors defined for Equation 3, evaluated on fully open-loop predictions. Horizon weights are uniform. The original logged-training recipe uses AdamW, weight decay 10^{-4}, one-cycle maximum learning rate 10^{-3} by default, batch size 256, gradient clipping 1.0, and at most 120 epochs. ID validation selects the checkpoint with early-stopping patience 20 and minimum improvement 10^{-4}. Training windows use stride 1; in-training ID validation uses stride 5, whereas exported evaluation uses stride 1. In-training validation batch size is 512; exported evaluation uses 256. All controlled studies reported here train and evaluate in FP32. GPU runs are seeded but not claimed bitwise deterministic across hardware.

Study records and seeds. Protocol names map to the experiments as follows: Round 9 is the initial factorial/information/causal-platform family; Round 10A is the metadata intervention; Round 10B is ID learning-rate selection; Round 11A is descriptive decomposition; Round 11B/C/D are simulation, placement and control. Fixed-log seed sets are respectively 20–29 for the initial factorial/information models, 0–9 for the second platform, 40–49 after learning-rate selection and 60–69 for placement. Simulation uses 70–72 within each of five datasets. Every planned seed is retained; pilots do not enter scientific comparisons.

B DATA AND CAUSAL PREPROCESSING

Real logs. The main platform uses the public Neural-Fly randomized-trajectory logs, with four training wind conditions and two whole held-out wind conditions. A checksummed split manifest specifies admissible window centers; training and validation blocks exclude windows that would overlap across the split. The separate figure-8 session is a combined trajectory/session/wind stress test. Static and sinusoidal flights with the same nominal wind label must remain separate when conditioning results on wind. Their earlier pooled archival comparisons are not used as evidence in this manuscript. Repeated windows and training seeds do not create independent sessions.

We use public Neural-Fly quadrotor logs (O'Connell et al., 2022). Training and ID validation occupy nonoverlapping temporal blocks at $\{0, 1.3, 2.5, 4.9\}$ m/s. The primary test, D2, aggregates whole held-out flights at 3.7 m/s (interpolation) and 6.1 m/s (extrapolation). Thus a held-out condition need not lie outside the training wind range. Main models receive no wind metadata. A separate figure-8 session is a combined trajectory, session and wind stress test. These logs motivate a realistic hidden-condition problem, but wind and flight session remain confounded; training seeds are not independent physical trials. Simulation subsequently tests independent trajectory generation under controlled wind schedules.

Second platform. The Intel-Aero logs have a reduced schema, including raw motor commands and no measured gyro. We decimate their 100 Hz records to 50 Hz. For $t \geq 1$ we use $\omega_t = \mathrm{Log}(R_{t-1}^{\mathsf{T}} R_t)/\Delta t$, a delayed interval-average rate over the preceding 20 ms, with $\omega_0 = 0$. It is an attitude-derived proxy, not an instantaneous gyro measurement. A property test changes future attitudes and checks that present history features are unchanged. Normalizers and metric scales are refitted to the causal representation. Training winds are $\{0, 1.3, 3.7\}$ m/s; held-out winds are 2.5 and 4.9 m/s. Each model is retrained on this platform; this is not zero-shot weight transfer. Older forward-difference preprocessing and mixed-precision runs remain archived but do not enter the corrected comparison.

Complementary models. The generic history model (M1) uses a width-64 GRU encoder and a width-72 recurrent decoder. The latent model (M2) uses a width-64 GRU encoder, a 16-dimensional projected context held fixed over the rollout, and a two-hidden-layer width-128 MLP transition. It applies affine LN to context unless a variant specifies otherwise. The width-64 latent variant narrows the transition to maintain a similar parameter budget. These models share history inputs and integration, but their decoder families differ. They are complementary information/platform controls; M1 versus M2 does not isolate the controlled memory-update effect.

Second-platform outcome. The causal Intel-Aero comparison favors the fixed-context MLP over the generic history GRU: M2 minus M1 is -0.02962 ($[-0.04297, -0.01627]$, 10/10 seeds lower, Holm $p = 0.02539$). Removing LN at latent widths 16 or 64 yields no detected advantage (Holm $p = 0.96094$ and 0.96680). These results bound a universal normalization claim but do not replicate the shared-decoder factorial.

C COMPLETE FIXED-LOG PRIMARY COMPARISONS

Each table reproduces one accepted, prospectively specified within-study family. All differences use ten paired training seeds. Intervals are descriptive paired t 95% intervals; familywise decisions use Holm-adjusted two-sided Wilcoxon p-values. The count column reports negative differences, not an interaction "win." Negative model differences favor the first condition named. Tables include every one of the 24 planned comparisons, regardless of outcome. The initial family includes 13 tests; the learning-rate and placement studies have three and four, respectively. A four-test metadata follow-up completes the 24 primary contrasts.

For memory design u, define the normalization effect $\Delta_u = E_{u,\mathrm{LN}} - E_{u,I}$ and interaction $\mathcal{I} = \Delta_{\mathrm{Updated}} - \Delta_{\mathrm{Frozen}}$. A positive interaction means a larger normalization penalty with updating. We also report absolute errors: a positive interaction does not imply that Frozen is the better predictor.

Table 3: Information, memory updating and causal transfer: all 13 primary contrasts, one within-study Holm family. Differences are in normalized static OOD error; intervals are descriptive, not simultaneous.

Contrast	Mean [95% CI]	$\Delta < 0$	Raw p	Holm p
Non-affine LN minus identity	$+0.00044$ $[-0.00377, +0.00464]$	4/10	0.76953	0.98438
RMS minus identity	-0.00168 $[-0.00597, +0.00261]$	5/10	0.49219	0.98438
Centering minus identity	-0.00154 $[-0.00414, +0.00106]$	7/10	0.19336	0.96680
LN + statistics minus LN	$+0.00130$ $[+0.00031, +0.00229]$	3/10	0.04883	0.34180
RMS + scale minus RMS	$+0.00185$ $[+0.00002, +0.00368]$	1/10	0.03711	0.29688
History + wind minus history (LN)	-0.01444 $[-0.01618, -0.01270]$	10/10	0.00195	0.02539
History + wind minus history (identity)	-0.01575 $[-0.01963, -0.01186]$	10/10	0.00195	0.02539
Frozen: LN minus identity	$+0.00205$ $[-0.00114, +0.00524]$	4/10	0.23242	0.96680
Updated: LN minus identity	$+0.01309$ $[+0.01085, +0.01534]$	0/10	0.00195	0.02539
LN-by-memory interaction	$+0.01105$ $[+0.00761, +0.01448]$	0/10	0.00195	0.02539
Second platform: M2 minus M1	-0.02962 $[-0.04297, -0.01627]$	10/10	0.00195	0.02539
Second platform: identity minus LN (16d)	$+0.00259$ $[-0.00196, +0.00714]$	3/10	0.16016	0.96094
Second platform: identity minus LN (64d)	$+0.00189$ $[-0.00249, +0.00628]$	3/10	0.19336	0.96680

Table 4: Fixed-checkpoint metadata interventions: all 4 primary contrasts, one within-study Holm family. Differences are in normalized static OOD error; intervals are descriptive, not simultaneous.

Contrast	Mean [95% CI]	$\Delta < 0$	Raw p	Holm p
LN: median minus original	$+0.03073$ $[+0.02805, +0.03341]$	0/10	0.00195	0.00781
LN: permuted minus original	$+0.01508$ $[+0.01261, +0.01756]$	0/10	0.00195	0.00781
Identity: median minus original	$+0.02840$ $[+0.02517, +0.03163]$	0/10	0.00195	0.00781
Identity: permuted minus original	$+0.01487$ $[+0.01282, +0.01692]$	0/10	0.00195	0.00781

Table 5: ID-selected learning-rate robustness: all 3 primary contrasts, one within-study Holm family. Differences are in normalized static OOD error; intervals are descriptive, not simultaneous.

Contrast	Mean [95% CI]	$\Delta < 0$	Raw p	Holm p
Frozen: LN minus identity	-0.00099 $[-0.00513, +0.00316]$	5/10	1.00000	1.00000
Updated: LN minus identity	$+0.01277$ $[+0.00979, +0.01575]$	0/10	0.00195	0.00586
LN-by-memory interaction	$+0.01375$ $[+0.00892, +0.01859]$	0/10	0.00195	0.00586

Table 6: Normalization location: all 4 primary contrasts, one within-study Holm family. Differences are in normalized static OOD error; intervals are descriptive, not simultaneous.

Contrast	Mean [95% CI]	$\Delta < 0$	Raw p	Holm p
Updated: memory-input LN effect	$+0.01504$ $[+0.01244, +0.01764]$	0/10	0.00195	0.00781
Updated: readout LN effect	-0.00212 $[-0.00581, +0.00156]$	7/10	0.27539	0.55078
Memory-input LN-by-memory interaction	$+0.01062$ $[+0.00589, +0.01536]$	1/10	0.00391	0.01172
Readout LN-by-memory interaction	-0.00041 $[-0.00640, +0.00559]$	5/10	0.92188	0.92188

D Normalization placement: absolute outcomes

Table 7 contains all eight trained arms. Memory-input LN and readout LN are non-affine. Memory-input LN acts on the memory input to the GRUCell; readout acts on the cell output immediately before the linear prediction head. The raw cell output is carried forward in Updated arms in every case. Both-norm arms are descriptive diagnostics; no additional primary tests were introduced for their nonadditivity.

Table 7: Placement study: static D2 error, mean and standard deviation over ten paired training seeds. All eight arms share 39,497 parameters.

Memory	Memory input	Readout	Mean	Seed SD
Frozen	I	I	0.13268	0.00326
Frozen	I	LN	0.13097	0.00378
Frozen	LN	I	0.13710	0.00233
Frozen	LN	LN	0.13536	0.00368
Updated	I	I	0.10962	0.00279
Updated	I	LN	0.10750	0.00523
Updated	LN	I	0.12466	0.00307
Updated	LN	LN	0.12567	0.00283

E Information controls and limits

For $z \in \mathbb{R}^d$, let $\mu = d^{-1}\sum_j z_j$, $s = \sqrt{d^{-1}\sum_j (z_j-\mu)^2 + \epsilon}$, and $r = \sqrt{d^{-1}\sum_j z_j^2 + \epsilon}$. The controls use identity, centering $z-\mu$, non-affine LN $(z-\mu)/s$, and RMS normalization z/r. Sufficient-statistic variants concatenate $[(z-\mu)/s, \mu, s]$ or $[z/r, r]$. These permit exact reconstruction through $s \cdot \mathrm{LN}(z) + \mu$ or $r \cdot (z/r)$; numerical and gradient tests check this property. LN statistics add 256 parameters to the first transition layer; RMS statistics and scalar wind each add 128. Auxiliary input weights start at zero, with the shared trunk initialized as in the corresponding baseline. This is an equal-trunk rather than exactly equal parameter-count comparison.

History-plus-wind controls retain the entire context pathway and append the wind label divided by 6 to the transition input. Main static OOD is the primary outcome because dynamic-flight labels are nominal, not instantaneous wind. Same-checkpoint interventions replace this channel with its training-window median or a global permutation, retaining all other inputs and weights. Ten fixed permutations (seeds 1000–1009) of the full canonical evaluation-window label vector are shared across checkpoints. Their outcomes are averaged within each training seed before the paired comparison; permutations are not extra independent observations. Their error increases show fitted

dependence on the paired metadata. They cannot separate wind physics from correlated session information or exclude a history–label support violation.

Adding wind metadata while retaining history reduces main-platform D2 error by 0.01444 with affine LN and 0.01575 with no LN (approximately 12–13%; both Holm $p = 0.02539$). At the same checkpoints, median replacement and global permutation increase error in every seed under both normalizations (all four Holm $p = 0.00781$). These outcomes demonstrate use of paired metadata, subject to the support and session qualifications above.

Reinserting LN sufficient statistics changes error by $+0.00130$ relative to non-affine LN (Holm $p = 0.34180$); reinserting RMS scale changes it by $+0.00185$ (Holm $p = 0.29688$). Neither contrast detects an improvement. Centering, non-affine LN and non-affine RMS versus identity are also inconclusive in this latent MLP family. These are not equivalence results and do not prove that context information is irrelevant. The original wind-only, probe, width and latent geometry studies motivated the follow-ups but are not used to support a unique mechanism here.

F CONDITION-SPECIFIC REAL-LOG DECOMPOSITION

These descriptive results reuse the accepted learning-rate follow-up, with ten paired seeds per flight and no new p-values. Table 8 separates all two static held-out flights and five figure-8 stress flights. The static 8.5 m/s and sinusoidal flights have the same nominal label but are separate rows. The 12.1 m/s stress reversal is part of the evidence, not a discarded outlier. Because these are different sessions and trajectories, they do not identify a wind-amplitude threshold at which the interaction changes sign.

Table 8: Accepted prediction decomposition by flight (wind in m/s; sinusoidal value is nominal). Effects average ten training seeds and use normalized error. Results are descriptive and no new tests are performed. Stress rows jointly shift session, trajectory and wind.

Flight	Δ_{Frozen}	Δ_{Updated}	$\mathcal{I}$
D2: static 3.7	-0.00258	+0.01938	+0.02195
D2: static 6.1	+0.00060	+0.00615	+0.00555
Stress: static 0	+0.00366	+0.02836	+0.02470
Stress: static 4.2	+0.00668	+0.01295	+0.00628
Stress: static 8.5	+0.00751	+0.01133	+0.00382
Stress: sinusoidal 8.5	+0.00309	+0.00700	+0.00391
Stress: static 12.1	+0.02365	-0.04252	-0.06617

G INDEPENDENT SIMULATED PREDICTION

Physics. The simulator uses rigid-body translation and rotation with a first-order inner attitude/rate loop. In body coordinates, airspeed is $u = R^{\mathsf{T}}(v - w)$ and drag is $f_b = -(d_1 + d_2|u|) \odot u$. World force is $Rf_b + TRe_3 + mg$. Parameters are mass 1 kg, maximum thrust 25 N, body-rate time constant 0.08 s, attitude gain 9, linear drag $(0.25, 0.25, 0.40)$ N s/m and quadratic drag $(0.020, 0.020, 0.035)$ N s^2/m^2. Ornstein–Uhlenbeck gusts have one-second correlation time and per-axis scale $0.08 \max(\|w_{\text{mean}}\|, 0.3)$; actuation noise scale is 0.01. Control runs at 50 Hz with ten Euler substeps. Position/velocity PD gains are 6 and 4. This is a controlled physics model at comparable scales to the logs, not a system identified from them, and all five datasets share these parameters.

Trajectories and independence. Each reference trajectory sums three sinusoids per coordinate, with frequencies sampled from 0.05–0.35 Hz and random phase. Horizontal range is normalized to 1.5 m and vertical excursion to 0.6 m around height 1.5 m. Every flight lasts 60 s. Within a dataset, four training trajectory identities are each crossed with the four training winds (16 flights); two separate ID identities give eight validation flights. Ten new static-test identities are crossed with three static winds; separate sets of ten trajectories cover each of the two dynamic families. Five datasets thus contain 370 files. The generation manifest fixes trajectory, noise and model seeds and

file hashes before training. Training, validation and test identity sets are disjoint, as are dataset replicates. Arms share trajectories and disturbances.

For each condition, the analysis averages ten trajectory scores and three model seeds within a dataset, then reports the five dataset values. It does not treat windows or the resulting 150 model/trajectory combinations as independent replicates. Each dataset has 45,600 admissible training centers, 22,800 validation centers, 85,500 static-test centers and 57,000 dynamic-test centers. All 60 runs and 180 split-level prediction artifacts passed acceptance before analysis. Table 9 reports absolute arm means; Table 31 reports effects and uncertainty.

Table 9: Simulated prediction: all four absolute normalized means. Each mean averages five dataset scores, each already averaged over three model seeds and ten trajectories. Scores use dataset-specific training scales.

Condition	Frozen (no LN)	Frozen (LN)	Updated (no LN)	Updated (LN)
Static 3.7 m/s	0.02860	0.02938	0.02768	0.03580
Static 6.1 m/s	0.04092	0.04164	0.03905	0.04985
Static 8.5 m/s	0.07540	0.07371	0.06323	0.08071
Dynamic $\nu = 0.25$	0.03486	0.03554	0.03268	0.04157
Dynamic $\nu = 1$	0.04385	0.04483	0.03827	0.04628

H MATCHED-BUDGET CONTROL AND COMPLETE OUTCOMES

Planning. MPPI uses 64 candidates over 50 steps, throttle noise 0.04, tilt noise 0.08, temperature coefficient 0.05, action regularization 0.05 and nominal-sequence decay 0.85. All MPPI variants use the same preview and candidate budget. The learned model observes the last 100 state/action steps. Nominal-physics MPPI and True-mean-wind MPPI use the same underlying physics predictor with respectively zero and true static mean wind assumptions; the latter does not observe future stochastic gusts. PD uses current position and velocity reference; Preview PD adds feedforward preview. For candidate k, cost is the horizon sum of squared position error plus 0.2 times squared velocity error, with regularizer $0.05H^{-1}\sum_h \|\delta a_{k,h}/\sigma\|^2$ on throttle/tilt perturbations. Softmin weights use $\exp[-(J_k - \min_j J_j)/(0.05(\mathrm{std}_j\, J_j + 10^{-9}))]$. Candidate noise is temporally smoothed with AR coefficient 0.7; two candidates are the pure nominal and shifted previous solution. The first action of the weighted perturbation sequence is executed. Physics predictors omit stochastic noise during candidate rollout. This is the specified MPPI implementation, not a claim that every MPPI variant has the same ranking.

Episode $e \in \{0, \ldots, 9\}$ uses environmental seed $310000+e$, trajectory seed $320000+e$ and planning seed $330000+e$, shared across controllers. Initial position is the first reference position, velocity and angular rate are zero, attitude is identity and initial gust is zero. A common 100-step PD warmup precedes 1,500 scored steps. All 60 checkpoints are crossed with these episode seeds at winds 0, 4.9, 6.1 and 8.5 m/s. Reference episodes are stored once; reusing their means in dataset-paired contrasts creates no extra observations. The same ten episodes are used across training datasets, so reported dataset intervals are conditional on this evaluation set.

Failure scoring and diagnostics. Controller failure includes a nonfinite state, action or prediction; tracking error above 5 m; position norm above 20 m; velocity norm above 30 m/s; or angular-rate norm above 60 rad/s. Pointwise position errors are capped at 5 m before computing RMSE, and all remaining steps after failure receive 5 m error. A warmup failure would receive 5 m throughout scoring. Software execution failures are recorded separately and resolved before acceptance. All 2,560 formal episodes passed roster, provenance, checkpoint-hash, scoring-trace and cross-controller warmup checks, with zero controller failures.

All controller RMSE means are in Table 10; all eight planned RMSE contrasts at each wind are in Tables 11 and 12. Failure rates and all failure-rate contrasts are zero. No p-values or seed-expansion gates are applied. Effort and latency are summarized in Table 13; the episode records retain the complete wind-specific values. Formal learned evaluation allocated 28.1075 GPU-hours, and unique references 4.2025 CPU-hours, excluding pilots. An initial pilot resource gate failed and

scheduling ceilings were amended before formal runs; controller budgets and the scientific roster were unchanged.

Original memory contrast and physical references. The prespecified no-LN control contrast between Updated and Frozen provides a paired view of this result. Updated minus Frozen RMSE is 0.07216 m at 4.9 m/s, 0.11311 m at 6.1 m/s and 0.20298 m at 8.5 m/s, with descriptive dataset-level intervals $[0.01060, 0.13373]$, $[0.02502, 0.20120]$ and $[0.04009, 0.36588]$ m. Four of five datasets favor Frozen at each nonzero wind. Thus the mean control ordering is supported by a planned within-dataset comparison, rather than only by comparing independently pooled averages. The fifth dataset is retained and favors Updated slightly; the result is not unanimity across training experience. At zero wind, the same contrast has a mean of 0.03043 m and an interval spanning zero. These are descriptive intervals conditional on the shared evaluation episodes, not additional hypothesis tests. The complete set of 32 planned control contrasts appears in Appendix H, including unfavorable and uncertain comparisons.

At 4.9 and 6.1 m/s, both Frozen variants outperform Nominal-physics MPPI in each of the five datasets. At 8.5 m/s, Frozen (LN) improves in all five; Frozen (no LN) improves in four, with a difference interval that crosses zero. Their grand-mean RMSEs are $0.157/0.135$ m at 4.9 m/s, $0.210/0.188$ m at 6.1 m/s and $0.368/0.347$ m at 8.5 m/s, compared with Nominal-physics MPPI at $0.250/0.315/0.460$ m. These conditions show that context held fixed during imagination can still support useful response to wind through repeated inference from real observations.

The favorable results do not extend to every operating condition. At zero wind, all learned variants lose to Nominal-physics MPPI in every dataset; their mean RMSEs range from 0.109 to 0.140 m versus 0.036 m for Nominal-physics MPPI. At 8.5 m/s, both Updated variants also lose to Nominal-physics MPPI in every dataset. True-mean-wind MPPI has the lowest grand-mean MPPI error at every wind, but its privileged input makes it a reference for missing environmental knowledge, not an optimal-control bound. Figure 3 retains these references and unfavorable conditions alongside the improvements.

The closed-loop LN-by-memory interaction is less certain than its prediction counterpart: its descriptive intervals cross zero at 0, 6.1 and 8.5 m/s. Mean learned planning time is 24.14 ms on CUDA, above the 20 ms control period; physics references run on CPU, so timings do not establish an algorithmic speedup. This is simulated closed-loop control, not a real-time hardware-flight demonstration. Appendix H retains all contrasts, effort and timing.

Table 10: All closed-loop tracking RMSE means (meters; lower is better). Learned scores average five datasets, three model seeds and ten episodes per wind; references use ten unique episodes per wind. Failure rates are zero for every row and wind.

Controller	0 m/s	4.9 m/s	6.1 m/s	8.5 m/s
Frozen (no LN)	0.10935	0.15702	0.21005	0.36820
Frozen (LN)	0.11439	0.13536	0.18790	0.34692
Updated (no LN)	0.13978	0.22918	0.32316	0.57118
Updated (LN)	0.11444	0.27142	0.36816	0.63551
Nominal-physics MPPI	0.03597	0.24966	0.31477	0.45955
True-mean-wind MPPI	0.03597	0.06545	0.07953	0.14414
PD	0.29453	0.41797	0.49105	0.68341
Preview PD	0.05082	0.30034	0.39526	0.61675

Table 11: Complete planned closed-loop RMSE contrasts (meters). Descriptive paired 95% intervals across five training datasets are conditional on the same ten episodes and are not simultaneous. Negative effects favor the first named condition; negative interactions are not model wins.

Wind	Contrast	Mean	95% interval	$\Delta < 0$
0	Frozen: LN effect	$+0.0050$	$[-0.0407, +0.0507]$	3/5
0	Updated: LN effect	-0.0253	$[-0.0657, +0.0150]$	4/5
0	LN-by-memory interaction	-0.0304	$[-0.1003, +0.0396]$	4/5
0	Updated (no LN) minus Frozen (no LN)	$+0.0304$	$[-0.0445, +0.1054]$	2/5
0	Frozen (no LN) minus Nominal-physics MPPI	$+0.0734$	$[+0.0197, +0.1271]$	0/5
0	Frozen (LN) minus Nominal-physics MPPI	$+0.0784$	$[+0.0176, +0.1392]$	0/5
0	Updated (no LN) minus Nominal-physics MPPI	$+0.1038$	$[+0.0484, +0.1592]$	0/5
0	Updated (LN) minus Nominal-physics MPPI	$+0.0785$	$[+0.0503, +0.1067]$	0/5
4.9	Frozen: LN effect	-0.0217	$[-0.0555, +0.0122]$	4/5
4.9	Updated: LN effect	$+0.0422$	$[-0.0144, +0.0989]$	1/5
4.9	LN-by-memory interaction	$+0.0639$	$[+0.0008, +0.1271]$	1/5
4.9	Updated (no LN) minus Frozen (no LN)	$+0.0722$	$[+0.0106, +0.1337]$	1/5
4.9	Frozen (no LN) minus Nominal-physics MPPI	-0.0926	$[-0.1345, -0.0508]$	5/5
4.9	Frozen (LN) minus Nominal-physics MPPI	-0.1143	$[-0.1470, -0.0816]$	5/5
4.9	Updated (no LN) minus Nominal-physics MPPI	-0.0205	$[-0.0858, +0.0448]$	3/5
4.9	Updated (LN) minus Nominal-physics MPPI	$+0.0218$	$[-0.0657, +0.1093]$	1/5

Table 12: Complete planned closed-loop RMSE contrasts (meters). Descriptive paired 95% intervals across five training datasets are conditional on the same ten episodes and are not simultaneous. Negative effects favor the first named condition; negative interactions are not model wins.

Wind	Contrast	Mean	95% interval	$\Delta < 0$
6.1	Frozen: LN effect	-0.0221	$[-0.0614, +0.0171]$	4/5
6.1	Updated: LN effect	$+0.0450$	$[-0.0262, +0.1162]$	1/5
6.1	LN-by-memory interaction	$+0.0671$	$[-0.0038, +0.1381]$	1/5
6.1	Updated (no LN) minus Frozen (no LN)	$+0.1131$	$[+0.0250, +0.2012]$	1/5
6.1	Frozen (no LN) minus Nominal-physics MPPI	-0.1047	$[-0.1608, -0.0487]$	5/5
6.1	Frozen (LN) minus Nominal-physics MPPI	-0.1269	$[-0.1744, -0.0793]$	5/5
6.1	Updated (no LN) minus Nominal-physics MPPI	$+0.0084$	$[-0.0795, +0.0963]$	3/5
6.1	Updated (LN) minus Nominal-physics MPPI	$+0.0534$	$[-0.0602, +0.1670]$	1/5
8.5	Frozen: LN effect	-0.0213	$[-0.0915, +0.0489]$	3/5
8.5	Updated: LN effect	$+0.0643$	$[-0.0632, +0.1919]$	2/5
8.5	LN-by-memory interaction	$+0.0856$	$[-0.0484, +0.2196]$	1/5
8.5	Updated (no LN) minus Frozen (no LN)	$+0.2030$	$[+0.0401, +0.3659]$	1/5
8.5	Frozen (no LN) minus Nominal-physics MPPI	-0.0913	$[-0.1887, +0.0060]$	4/5
8.5	Frozen (LN) minus Nominal-physics MPPI	-0.1126	$[-0.1961, -0.0291]$	5/5
8.5	Updated (no LN) minus Nominal-physics MPPI	$+0.1116$	$[-0.0227, +0.2460]$	0/5
8.5	Updated (LN) minus Nominal-physics MPPI	$+0.1760$	$[+0.0208, +0.3311]$	0/5

Table 13: Descriptive effort and timing, averaged over all planned winds and episodes (600 per learned arm; 40 per reference). Throttle RMS is relative to hover throttle 0.3924; tilt RMS is in radians. Time columns average per-episode mean and per-episode 95th-percentile planning times, respectively, in milliseconds; the latter is not a pooled percentile. Dashes indicate unrecorded planning times for PD. Learned models run on CUDA, references on CPU, so timings are not hardware-matched speedups.

Controller	Throttle RMS	Tilt RMS	Mean ms	P95 ms
Frozen (no LN)	0.0616	0.3225	22.93	23.72
Frozen (LN)	0.0639	0.3268	25.15	25.93
Updated (no LN)	0.0770	0.3399	23.00	23.77
Updated (LN)	0.0805	0.3521	25.48	26.20
Nominal-physics MPPI	0.0524	0.3059	106.45	110.36
True-mean-wind MPPI	0.0491	0.3013	110.31	113.16
PD	0.0477	0.3230	—	—
Preview PD	0.0406	0.2942	6.14	6.51

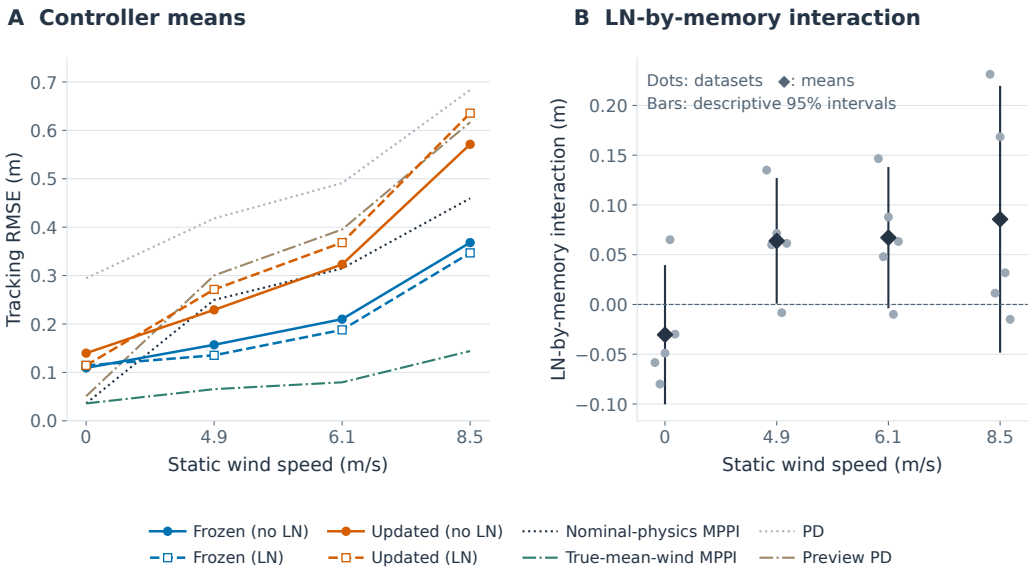

Figure 3: Closed-loop benefits and limits for the original checkpoints and evaluation roster. Left: all controller grand means. Right: LN-by-memory interaction, with five dataset dots, mean diamonds and descriptive paired 95% intervals. Learned values average three model seeds and ten episodes within a dataset; references use ten unique episodes per wind. Frozen models have lower mean tracking RMSE than Nominal-physics MPPI at nonzero wind, but all learned models lose at zero wind. Control interaction intervals are wide; they do not establish uniform normalization harm. All 2,560 episodes are included.

I PLANNER-QUERY DIAGNOSTICS AND BOUNDED EXTENSIONS

These studies were specified after inspecting the original prediction/control results. They are result-informed extensions on existing datasets, not untouched confirmation. All contrasts first average model seeds and paired episodes or anchors within each of five training datasets. Intervals are descriptive paired t intervals with four degrees of freedom, conditional on the evaluation roster; they are not simultaneous, new hypothesis tests or equivalence claims.

Common-query diagnostics. The first-decision study uses 40 common PD-warmup histories, 60 GRU checkpoints and 2,400 model–anchor observations. Later queries use 180 donor episodes (five datasets, three seeds, two no-LN donor policies, three winds and two episodes), with indices 350, 850 and 1350, giving 540 anchors and 2,160 model–anchor observations. Each model is scored on the same state, history and actual candidate bank within its paired comparison. Reconstructing the planner warm start reproduces 225,180 original donor decisions with zero action discrepancy; no tolerance was relaxed. Eight physical noise draws construct oracle candidate costs and eight independent common-noise draws evaluate each full weighted solution. Oracle and evaluation streams are separate. Physical solution cost is the original MPPI objective, not tracking RMSE. Its difference from the oracle-softmin solution can be negative: it is not nonnegative regret or an optimal-control bound. Monte Carlo uncertainty is distinct from five-dataset uncertainty. Later donor/time strata are primary; balanced donor/time averages in the main text are secondary. Complete metric, arm, dataset and Monte Carlo traces are retained in the analysis artifacts.

Here ρ is Spearman rank correlation across candidates between model-predicted cost and the mean physical cost over eight evaluation-noise draws (average ranks for ties; undefined for a constant vector). Query E averages the three normalized state errors in Eq. 3 uniformly over candidates, physical noise draws and all steps through the stated horizon; terminal errors use only its final step. The oracle uses mean candidate costs before softmin; each resulting weighted action sequence is evaluated on independent physical draws. These finite-sample oracle costs are Monte Carlo estimates, not exact expected costs. Neither query errors nor their correlations identify mediation of the closed-loop effect.

Table 14: All first-decision winds: no-LN Updated minus Frozen. Positive error/cost and negative correlation favor Frozen.

Wind	ΔE	$\Delta\rho$	Δ cost	Cost 95% interval
0.0	0.0642	-0.1579	0.3610	[0.2118, 0.5103]
4.9	0.1293	-0.2408	1.2132	[0.1495, 2.2769]
6.1	0.1556	-0.2611	1.7228	[0.2277, 3.2180]
8.5	0.2087	-0.2853	2.6318	[0.6900, 4.5735]

Table 15: All 18 later-query strata, no LN; F/U denote the donor policy. Reversals are retained.

Wind	Donor	Index	ΔE	$\Delta\rho$	Δ cost	Cost 95% interval
0.0	F	350	0.0511	0.0395	-0.3468	[-1.7566, 1.0631]
0.0	F	850	0.0586	0.0909	-0.5493	[-1.5287, 0.4301]
0.0	F	1350	0.0463	-0.0018	-0.3289	[-1.2455, 0.5878]
0.0	U	350	0.0383	-0.2128	0.5772	[-0.4611, 1.6156]
0.0	U	850	0.0279	-0.0456	0.4096	[-0.7858, 1.6049]
0.0	U	1350	0.0277	-0.2921	0.5199	[-0.7587, 1.7985]
6.1	F	350	0.0567	-0.0985	0.9225	[-2.6099, 4.4548]
6.1	F	850	0.0550	0.0659	1.5662	[-0.6078, 3.7403]
6.1	F	1350	0.0617	-0.0813	1.0351	[-0.1218, 2.1920]
6.1	U	350	0.1247	-0.3783	5.4560	[-1.3266, 12.2385]
6.1	U	850	0.0577	-0.5257	3.7300	[1.6902, 5.7699]
6.1	U	1350	0.0923	-0.3183	1.5277	[-0.9153, 3.9707]
8.5	F	350	0.0320	-0.1838	4.4562	[1.0361, 7.8764]
8.5	F	850	0.1349	-0.2125	6.2254	[-0.7621, 13.2130]
8.5	F	1350	0.0848	0.1597	-0.9120	[-3.0261, 1.2021]
8.5	U	350	0.1500	-0.3335	4.3890	[0.1815, 8.5966]
8.5	U	850	0.0721	-0.2651	6.0716	[0.3554, 11.7877]
8.5	U	1350	-0.0479	-0.1768	0.6039	[-2.1283, 3.3361]

Planning and plant sensitivity. Planning settings change only horizon/candidate count from the original H50/N64. The eight plant perturbations separately change mass (baseline 1 kg), both linear and quadratic drag coefficients, body-rate time constant (baseline 0.08 s), or gust correlation time (baseline 1 s). Controllers retain nominal design parameters; learned checkpoints, normalizers, seeds and all other plant quantities stay fixed. These are zero-shot physical perturbations, not retraining under new simulators. All 27,920 new episodes passed the original warmup, provenance and failure-scoring contracts. With the 2,560 original baseline episodes there are 30,480 unique episodes, including 1,680 unique references. Reused references do not create additional trials. No numerical failures occurred. Tables below retain every setting and wind; all four arms and registered LN/reference contrasts remain in the complete analysis artifacts. Settings are not independent random draws from a physics population.

Table 16: Planning/plant sensitivity (1/2). Four-arm RMSE means (m); F/U denote Frozen/Updated, I/L denote no LN/LN. Interval is paired U/I minus F/I.

Setting	Wind	F/I	F/L	U/I	U/L	Δ 95% interval
baseline	0.0	0.1093	0.1144	0.1398	0.1144	[-0.0445, 0.1054]
baseline	4.9	0.1570	0.1354	0.2292	0.2714	[0.0106, 0.1337]
baseline	6.1	0.2100	0.1879	0.3232	0.3682	[0.0250, 0.2012]
baseline	8.5	0.3682	0.3469	0.5712	0.6355	[0.0401, 0.3659]
h50_n32	0.0	0.1174	0.1220	0.1386	0.1076	[-0.0551, 0.0975]
h50_n32	4.9	0.1579	0.1410	0.2299	0.2656	[0.0133, 0.1309]
h50_n32	6.1	0.2131	0.1925	0.3236	0.3632	[0.0287, 0.1923]
h50_n32	8.5	0.3784	0.3516	0.5635	0.6140	[0.0531, 0.3171]
h50_n128	0.0	0.0980	0.1034	0.1362	0.1194	[-0.0295, 0.1058]
h50_n128	4.9	0.1520	0.1298	0.2286	0.2743	[-0.0052, 0.1583]
h50_n128	6.1	0.2084	0.1835	0.3261	0.3854	[-0.0044, 0.2398]
h50_n128	8.5	0.3631	0.3407	0.5760	0.6544	[0.0374, 0.3885]
h25_n64	0.0	0.0502	0.0595	0.0790	0.0598	[-0.0055, 0.0631]
h25_n64	4.9	0.2034	0.1585	0.2887	0.3758	[-0.1052, 0.2758]
h25_n64	6.1	0.3135	0.2479	0.4259	0.5321	[-0.1743, 0.3990]
h25_n64	8.5	0.6209	0.5381	0.7321	0.8467	[-0.2546, 0.4770]
mass_0p9	0.0	0.1315	0.1393	0.1599	0.1495	[-0.0269, 0.0836]
mass_0p9	4.9	0.1876	0.1654	0.2641	0.3044	[0.0030, 0.1498]
mass_0p9	6.1	0.2488	0.2265	0.3704	0.4118	[0.0204, 0.2228]
mass_0p9	8.5	0.4177	0.4009	0.6186	0.6837	[0.0497, 0.3521]
mass_1p1	0.0	0.1384	0.1374	0.1666	0.1467	[-0.0401, 0.0967]
mass_1p1	4.9	0.1631	0.1434	0.2264	0.2727	[0.0079, 0.1186]
mass_1p1	6.1	0.2037	0.1858	0.3115	0.3608	[0.0228, 0.1928]
mass_1p1	8.5	0.3446	0.3157	0.5366	0.6226	[0.0422, 0.3419]

Table 17: Planning/plant sensitivity (2/2). Four-arm RMSE means (m); F/U denote Frozen/Updated, I/L denote no LN/LN. Interval is paired U/I minus F/I.

Setting	Wind	F/I	F/L	U/I	U/L	Δ 95% interval
drag_0p8	0.0	0.1090	0.1151	0.1404	0.1218	[-0.0378, 0.1008]
drag_0p8	4.9	0.1291	0.1119	0.1805	0.2153	[0.0129, 0.0899]
drag_0p8	6.1	0.1638	0.1423	0.2459	0.2885	[0.0104, 0.1539]
drag_0p8	8.5	0.2765	0.2507	0.4268	0.4857	[0.0381, 0.2624]
drag_1p2	0.0	0.1095	0.1138	0.1391	0.1148	[-0.0425, 0.1016]
drag_1p2	4.9	0.1923	0.1688	0.2914	0.3303	[0.0144, 0.1837]
drag_1p2	6.1	0.2664	0.2391	0.4084	0.4680	[0.0235, 0.2606]
drag_1p2	8.5	0.4710	0.4498	0.7133	0.7964	[0.0564, 0.4282]
tau_0p06	0.0	0.1308	0.1275	0.1487	0.1281	[-0.0704, 0.1060]
tau_0p06	4.9	0.1612	0.1420	0.2304	0.2641	[0.0062, 0.1322]
tau_0p06	6.1	0.2218	0.1986	0.3235	0.3641	[0.0109, 0.1924]
tau_0p06	8.5	0.3894	0.3644	0.5641	0.6231	[0.0253, 0.3241]
tau_0p10	0.0	0.0946	0.1043	0.1299	0.1112	[-0.0272, 0.0979]
tau_0p10	4.9	0.1545	0.1319	0.2303	0.2788	[0.0015, 0.1501]
tau_0p10	6.1	0.2044	0.1849	0.3270	0.3898	[0.0263, 0.2188]
tau_0p10	8.5	0.3569	0.3377	0.5834	0.6606	[0.0803, 0.3729]
gust_tau_0p5	0.0	0.1105	0.1140	0.1387	0.1147	[-0.0476, 0.1039]
gust_tau_0p5	4.9	0.1547	0.1361	0.2313	0.2691	[0.0117, 0.1416]
gust_tau_0p5	6.1	0.2103	0.1848	0.3243	0.3710	[0.0193, 0.2086]
gust_tau_0p5	8.5	0.3658	0.3489	0.5714	0.6394	[0.0545, 0.3568]
gust_tau_2p0	0.0	0.1097	0.1148	0.1384	0.1157	[-0.0445, 0.1019]
gust_tau_2p0	4.9	0.1540	0.1352	0.2258	0.2692	[0.0145, 0.1291]
gust_tau_2p0	6.1	0.2105	0.1866	0.3259	0.3636	[0.0215, 0.2093]
gust_tau_2p0	8.5	0.3714	0.3471	0.5730	0.6378	[0.0362, 0.3670]

Ungated decoder replication. The history GRU encoder is unchanged. Replace the rollout cell only by $c_h = \tanh(W_x x_h + b_x + W_m N_r(m_h) + b_m)$; the same linear head reads raw c_h, and the same Frozen/Updated assignment carries m_0 or raw c_h. Non-affine LN uses $\epsilon = 10^{-5}$. Each arm has 28,489 parameters; equal width does not match the GRU family's 39,497 parameters. Five original datasets, four arms and three new seeds (80–82) yield 60 models and 180 prediction artifacts. FP32, ID-only checkpoint selection and the original training recipe are retained. All 2,400 baseline learned control episodes pass acceptance; the 160 original reference episodes are reused once. This tests a second rollout transition, not an entirely GRU-free architecture, new physical platform or optimized RNN benchmark.

Table 18: Elman offline aggregate error: all conditions and arms. The interval is no-LN Updated minus Frozen.

Condition	F/I	F/L	U/I	U/L	Δ 95% interval
Static 3.7	0.0297	0.0316	0.0245	0.0307	[-0.0074, -0.0029]
Static 6.1	0.0420	0.0447	0.0358	0.0424	[-0.0091, -0.0033]
Static 8.5	0.0737	0.0748	0.0662	0.0720	[-0.0108, -0.0041]
Slow dynamic	0.0354	0.0379	0.0294	0.0355	[-0.0082, -0.0037]
Fast dynamic	0.0434	0.0475	0.0342	0.0398	[-0.0113, -0.0071]

Table 19: Elman prediction LN effects and LN-by-memory interaction; descriptive dataset-level uncertainty for normalized prediction error E.

Condition	Contrast	Mean	95% interval	+
Static 3.7	LN(F)	0.0019	[0.0005, 0.0034]	5/5
Static 3.7	LN(U)	0.0062	[-0.0007, 0.0131]	5/5
Static 3.7	Interaction	0.0043	[-0.0026, 0.0112]	4/5
Static 6.1	LN(F)	0.0027	[0.0011, 0.0043]	5/5
Static 6.1	LN(U)	0.0066	[0.0014, 0.0119]	5/5
Static 6.1	Interaction	0.0039	[-0.0019, 0.0096]	4/5
Static 8.5	LN(F)	0.0011	[-0.0018, 0.0039]	3/5
Static 8.5	LN(U)	0.0058	[-0.0012, 0.0128]	5/5
Static 8.5	Interaction	0.0047	[-0.0030, 0.0124]	4/5
Slow dynamic	LN(F)	0.0026	[0.0013, 0.0038]	5/5
Slow dynamic	LN(U)	0.0061	[0.0001, 0.0121]	5/5
Slow dynamic	Interaction	0.0035	[-0.0025, 0.0095]	4/5
Fast dynamic	LN(F)	0.0041	[0.0031, 0.0051]	5/5
Fast dynamic	LN(U)	0.0055	[0.0001, 0.0110]	5/5
Fast dynamic	Interaction	0.0014	[-0.0039, 0.0068]	2/5

Table 20: Elman control LN effects and LN-by-memory interaction; descriptive dataset-level uncertainty. Control is RMSE in meters.

Condition	Contrast	Mean	95% interval	+
0.0	LN(F)	-0.0045	[-0.0405, 0.0314]	3/5
0.0	LN(U)	0.0213	[-0.0294, 0.0719]	3/5
0.0	Interaction	0.0258	[-0.0491, 0.1007]	3/5
4.9	LN(F)	0.0059	[-0.0390, 0.0509]	3/5
4.9	LN(U)	0.0789	[-0.0218, 0.1795]	5/5
4.9	Interaction	0.0729	[-0.0553, 0.2012]	4/5
6.1	LN(F)	-0.0032	[-0.0612, 0.0547]	2/5
6.1	LN(U)	0.0960	[-0.0134, 0.2054]	5/5
6.1	Interaction	0.0993	[-0.0380, 0.2365]	5/5
8.5	LN(F)	-0.0411	[-0.1558, 0.0736]	2/5
8.5	LN(U)	0.1096	[0.0248, 0.1945]	5/5
8.5	Interaction	0.1507	[0.0105, 0.2909]	5/5

Table 21: Elman rollout decoder: no-LN tracking RMSE (m). $\Delta = \mathrm{Updated} - \mathrm{Frozen}$. Descriptive paired intervals use five datasets; + counts positive dataset differences. The history encoder remains GRU.

Wind	Frozen	Updated	Δ	95% interval	+
0.0	0.1050	0.1236	0.0186	[-0.0306, 0.0678]	4/5
4.9	0.1525	0.2027	0.0502	[0.0084, 0.0920]	5/5
6.1	0.2245	0.2854	0.0609	[0.0051, 0.1167]	4/5
8.5	0.4402	0.4921	0.0520	[-0.1171, 0.2211]	4/5

J PAIRED REVERSAL AND QUERY BOUNDARIES

These post-review analyses use the existing accepted checkpoints and traces. They are descriptive and result-informed; no new significance tests are added. The independent units remain five training datasets, not individual windows, candidates, seeds or entries of a dataset-by-episode matrix.

Joint reversal and evaluation instances. At 6.1 and 8.5 m/s, respectively, 12/15 and 13/15 paired model-seed comparisons improve logged prediction while worsening control. Aggregating seeds first gives 4/5 datasets at each wind (Figure 2); dataset D1 is the exception in both. After averaging the fixed trained roster, all ten original evaluation episodes have positive control differences at both winds. These episode summaries are conditional on those trained models, not additional independent training replicates. Full seed pairs and dataset-by-episode matrices accompany the numerical supplement.

Horizon and physical-state components. We recompute query errors from the saved model predictions and the independent physical evaluation-noise stream, checking agreement with every original terminal metric before stratifying. Table 22 retains all four winds and five pre-existing horizons. The aggregate gap is already positive at step 1, primarily through angular rate. This is compatible with different trained weights; it cannot identify drift caused by repeated memory updates, since identical weights give identical first predictions. At both shared held-out winds, physical position differences remain negative at step 10 but turn positive by step 25. Thus the logged-action advantage fails early for one component, while position consequences develop later; no unique internal mechanism is established.

Table 22: Terminal query errors, Updated minus Frozen without LN. All are means across five datasets. Δv, ΔR, $\Delta\omega$ use the original normalized errors; Δp is physical position error in meters.

Wind	Step	ΔE	Δv	ΔR	$\Delta\omega$	Δp
0.0	1	0.03667	-0.00159	-0.00245	0.11405	-0.00000
0.0	5	0.04317	0.00271	0.04049	0.08631	-0.00018
0.0	10	0.04183	0.00878	0.02151	0.09521	0.00115
0.0	25	0.07393	0.01896	0.07392	0.12891	0.00687
0.0	50	0.07475	0.05678	0.05825	0.10922	0.04105
4.9	1	0.04170	-0.00215	-0.00314	0.13040	-0.00003
4.9	5	0.06172	0.00092	0.04721	0.13705	-0.00061
4.9	10	0.08253	0.00652	0.06599	0.17509	-0.00020
4.9	25	0.14057	0.03908	0.09353	0.28911	0.02207
4.9	50	0.19565	0.09483	0.19491	0.29721	0.11945
6.1	1	0.04239	-0.00245	-0.00312	0.13275	-0.00004
6.1	5	0.06504	-0.00151	0.04717	0.14946	-0.00102
6.1	10	0.10283	0.00421	0.08073	0.22354	-0.00162
6.1	25	0.16671	0.04810	0.12270	0.32934	0.02566
6.1	50	0.24300	0.11698	0.25284	0.35919	0.15983
8.5	1	0.04077	-0.00348	-0.00284	0.12862	-0.00014
8.5	5	0.06865	-0.00794	0.04404	0.16984	-0.00261
8.5	10	0.13679	-0.00391	0.10220	0.31207	-0.00674
8.5	25	0.20063	0.04438	0.20307	0.35445	0.01982
8.5	50	0.35974	0.15075	0.37475	0.55373	0.20650

Candidate amplitude. For perturbation $\epsilon \in \mathbb{R}^{50\times 3}$, define $a = \sqrt{\text{mean}[(\epsilon/\sigma)^2]}$ before action clipping, where $\sigma = (0.04, 0.08, 0.08)$. Fixed bins separate zero, $(0, 0.5)$, $[0.5, 1)$ and $[1, \infty)$. Per wind, the ten anchors contain 20 zero candidates, 496 in $(0, 0.5)$, 124 in $[0.5, 1)$ and none in $[1, \infty)$. Two candidates coincide at zero at each first decision (nominal and empty warm start); both are retained to match the actual planner bank. Each nonempty bin contains candidates at all ten anchors. We average candidates and evaluation-noise draws within anchor, then equally average anchors and seeds within dataset. Across every populated bin, the mean terminal aggregate gap is positive at all five registered horizons, including zero perturbation. This does not identify a large-action threshold; larger amplitudes are absent from this first-decision sample. Bins are observational subsets, not randomized interventions.

Table 23: Amplitude-stratified terminal errors at step 50, no LN. Intervals are paired dataset-level descriptive 95% intervals for ΔE; all empty bins are reported explicitly above. Full horizons/components remain in the supplement.

Wind	Amplitude	ΔE	Δp (m)	ΔE interval
0.0	zero	0.0382	0.0353	[0.0086, 0.0679]
0.0	(0,0.5)	0.0725	0.0391	[0.0309, 0.1140]
0.0	[0.5,1)	0.0906	0.0497	[0.0301, 0.1510]
4.9	zero	0.1946	0.1175	[-0.0141, 0.4032]
4.9	(0,0.5)	0.1940	0.1177	[0.0044, 0.3835]
4.9	[0.5,1)	0.2010	0.1260	[0.0050, 0.3971]
6.1	zero	0.2464	0.1559	[0.0066, 0.4862]
6.1	(0,0.5)	0.2403	0.1573	[0.0167, 0.4640]
6.1	[0.5,1)	0.2506	0.1698	[0.0228, 0.4783]
8.5	zero	0.3502	0.1973	[0.0556, 0.6449]
8.5	(0,0.5)	0.3547	0.2020	[0.0835, 0.6258]
8.5	[0.5,1)	0.3792	0.2244	[0.1091, 0.6492]

Finite-noise precision. The paired first-decision physical solution-cost differences at winds 0/4.9/6.1/8.5 m/s are 0.3610/1.2132/1.7228/2.6318; their Monte Carlo standard errors are 0.00379/0.01134/0.01311/0.02355. This calculation preserves shared noise across models and treats the ten independently seeded anchors as noise units. It quantifies evaluation-noise uncertainty conditional on the constructed solutions; it does not include oracle-construction uncertainty or replace the five-dataset intervals. Later-query donor/time strata remain distinct in Appendix I, including adverse and uncertain cases.

K NEW CLOSED-LOOP EVALUATION INSTANCES

After the initial results, we fixed a new evaluation panel before inspecting its outcomes. All 30 original no-LN GRU checkpoints were retained (five datasets, three training seeds and two memory designs), without retraining, model selection or controller changes. Each model ran 20 new episodes at each shared held-out wind, totaling 1,200 episodes. Episode $e = 0, \ldots, 19$ uses trajectory seed $430000 + e$, environment-noise seed $440000 + e$ and planning seed $450000 + e$, paired across models and winds. These seeds do not overlap the original panel. The plant, trajectory generator, H50/N64 planner, 100-step PD warmup, 1,500 scored steps and failure caps are unchanged. Four full-length pilot episodes used a separate namespace and never enter these summaries. All formal episodes passed provenance, action-trace, common-warmup and scoring checks; zero numerical failures occurred. Runs used RTX PRO 6000 GPUs with the existing CUDA evaluation environment.

Table 24: New-trajectory no-LN control results. RMSE and differences are in meters. Dataset intervals average three training seeds and 20 episodes within each of five datasets. Episode intervals average the fixed 15 trained models per arm and are conditional on that roster. Neither interval treats 100 dataset–episode cells as independent replicates.

Wind	Frozen	Updated	Δ	Dataset 95% interval	Episode 95% interval
6.1	0.20905	0.32105	0.11199	[0.02240, 0.20159]	[0.10693, 0.11706]
8.5	0.37040	0.57026	0.19986	[0.05031, 0.34941]	[0.18923, 0.21049]

Table 25: All five training-dataset differences on new trajectories: Updated minus Frozen RMSE (m). Dataset labels follow the original paired training-dataset order.

Wind	D1	D2	D3	D4	D5
6.1	-0.00421	0.18460	0.15432	0.10124	0.12403
8.5	0.00454	0.33229	0.24692	0.20775	0.20781

The Updated minus Frozen difference is positive for 4/5 datasets at 6.1 m/s and 5/5 at 8.5 m/s. The original exception D1 still slightly favors Updated at 6.1 m/s, and slightly favors Frozen at 8.5 m/s on the new panel. After averaging the fixed trained roster, all 20 individual episode contrasts are positive at each wind. This does not mean every model wins on every episode. The supplement retains the complete 5×20 dataset–episode matrices, three-seed summaries and all per-arm outcomes; old and new episodes are not pooled.

The fresh panel supports persistence of the mean control ordering across new trajectory and noise draws from this generator. It does not add independent training datasets, simulators, robot platforms or physical trials, and does not establish a universal memory-design principle. No additional hypothesis tests or outcome-dependent seed expansion were performed.

L TARGETED ACTION-SOURCE AND CONTEXT-ACCESS TESTS

These result-informed follow-ups address specific ambiguities in the original comparison. All use the five independently generated training datasets as the uncertainty unit, averaging three paired model seeds and flights or anchors within each dataset. Intervals are descriptive paired Student-t 95% intervals with four degrees of freedom, not simultaneous intervals. No new p-values, outcome-selected seeds, wind thresholds or pooled old/new evaluation samples are introduced. Formal manifests were fixed before their evaluations; development pilots are excluded. The complete machine-readable summaries retain individual dataset differences and component/horizon results.

L.1 OFFLINE COMPONENTS AND PHYSICAL POSITION

We audit the 30 original no-LN models and 30 new auxiliary-context models on all static test windows. Each model is evaluated in the original 85,500-window order with batches of 256, retaining the 57,000 windows at the two primary winds (2,850 per flight, ten flights per wind). Thus the position audit contains 3,420,000 model–window pairs, not independent experimental replicates. For $h = 1, \ldots, 50$,

$$\hat{p}_h = p_0 + 0.02 \sum_{j=1}^{h} \hat{v}_j, \qquad e_{p,h} = \|\hat{p}_h - p_h^{\text{recorded}}\|_2.$$

The prefix metric averages these physical-meter errors over all 50 steps; it is distinct from terminal error and from the previously normalized velocity-integral metric. A secondary calculation integrates recorded velocities to separate the recorded-position target from the numerical integration convention.

The original normalized velocity, rotation, rate, aggregate and displacement arrays are reproduced exactly for every model (maximum absolute difference zero). An original-model pilot on RTX hardware failed the locked orientation-reproduction gate; rerunning that model unchanged on its original L4 hardware passed exactly. The failed artifact is retained. Formal original models use L4 and new models use RTX, with paired arms on the same hardware within each family. Tolerances were not relaxed. This verifies the accepted predictions on their evaluation hardware, rather than demonstrating hardware-invariant numerics.

Table 26: Original-model physical-position errors in meters. Differences are Updated minus Frozen; both prefix and terminal comparisons favor Updated.

Wind	Outcome	Frozen	Updated	Difference	95% interval
6.1	Mean over 50 steps	0.06357	0.05746	-0.00611	$[-0.00812, -0.00410]$
6.1	Terminal step 50	0.17320	0.15903	-0.01417	$[-0.01914, -0.00920]$
8.5	Mean over 50 steps	0.16232	0.10566	-0.05666	$[-0.06678, -0.04653]$
8.5	Terminal step 50	0.44630	0.30013	-0.14618	$[-0.17230, -0.12006]$

L.2 CHANGING ONLY FUTURE ACTION SOURCE

At step 100 of each of ten static test flights in each dataset and primary wind, we retain the complete observed history and physical simulator state. Replay through step 150 is checked against the original CSV fields with a 5.1×10^{-9} absolute tolerance for their eight-decimal serialization. The action comparison uses one logged sequence versus a fresh 64-candidate MPPI bank, initialized from the original PD plan with empty warm start. This is a newly constructed first-decision bank, not a claim to recover a historical controller's bank. The physical branches share eight paired noise samples; both sources use float32 actions and 10 integration substeps per control step. All six original no-LN models from the corresponding dataset evaluate each anchor, totaling 100 anchors and 600 model records. No anchor is omitted.

Both the logged-action and candidate-action branches are evaluated under the same fresh physical noise. The actual realized logged future is an additional check, not substituted for the paired logged-action branch. Model errors use the original training-only scales and Equation 3.

Table 27: Original no-LN Updated minus Frozen normalized component errors. Numbered rows use the error at that step, not the prefix average. Early component exceptions are retained. Dataset values and intervals are in the accompanying numerical records.

Wind	Step	Velocity	Rotation	Angular rate	E
6.1	1	-0.00038	$+0.00030$	-0.00011	-0.00006
6.1	5	-0.00139	-0.00007	$+0.00039$	-0.00036
6.1	10	-0.00215	-0.00097	-0.00058	-0.00123
6.1	25	-0.00298	-0.00094	-0.00376	-0.00256
6.1	50	-0.00076	-0.00127	-0.00323	-0.00176
6.1	Mean 1–50	-0.00213	-0.00087	-0.00261	-0.00187
8.5	1	-0.00248	$+0.00069$	$+0.00017$	-0.00054
8.5	5	-0.01039	$+0.00018$	$+0.00226$	-0.00265
8.5	10	-0.01840	-0.00132	$+0.00120$	-0.00617
8.5	25	-0.03561	-0.00140	-0.00512	-0.01405
8.5	50	-0.04821	-0.00206	-0.00400	-0.01809
8.5	Mean 1–50	-0.03243	-0.00128	-0.00280	-0.01217

Table 28: Action-source experiment: Updated minus Frozen H50 prefix E. The final row at each wind is a difference of those paired differences.

Wind	Source / contrast	Difference	95% interval
6.1	Logged actions, paired noise	-0.00160	$[-0.00493, +0.00173]$
6.1	Candidate actions, paired noise	$+0.10456$	$[+0.04440, +0.16472]$
6.1	Actual logged future	-0.00154	$[-0.00585, +0.00278]$
6.1	Candidate minus logged contrast	$+0.10616$	$[+0.04764, +0.16468]$
8.5	Logged actions, paired noise	-0.01210	$[-0.02352, -0.00068]$
8.5	Candidate actions, paired noise	$+0.14237$	$[+0.05437, +0.23037]$
8.5	Actual logged future	-0.01246	$[-0.02340, -0.00153]$
8.5	Candidate minus logged contrast	$+0.15447$	$[+0.06406, +0.24487]$

This test removes history, initial-state and physical-noise changes as explanations of the action-source contrast. At 6.1 m/s the logged-action interval crosses zero; it is the relative deterioration on candidates that is consistent across all five datasets. The experiment does not separate the bank's nominal action sequence from its perturbations, identify an internal latent mechanism, or quantify how much of the tracking RMSE gap it explains.

L.3 Auxiliary context with dynamic rollout memory

Equation 4 adds a 64-dimensional auxiliary input to the original GRUCell. Both arms update the rollout state; only $q_h = m_0$ versus $q_h = m_h$ differs. The history encoder, width, output head, physical integrator, losses, data, ID checkpoint rule and maximum 120-epoch training recipe remain shared. All five datasets and seeds 70–72 are retained, with separate training for each arm. Both have 51,785 trainable parameters, 12,288 more than the original family. Paired initialization gives identical first-step output at identical weights; separately trained weights need not do so. All auxiliary-input weights are active in gradient computations. Parameter-count matching does not equate effective function classes, because the dynamic input duplicates hidden state in some gate transformations. Comparison with original Frozen/Updated models changes architecture and capacity, not only context access.

Query evaluation reuses the original 20 first-decision anchors at 6.1/8.5 m/s, with identical snapshots, candidate banks and construction/evaluation noise. All 600 new model–anchor records are accepted. Physical candidate costs agree with the original traces to 10^{-12}; the original saved no-LN metrics are regenerated before accepting new results. Both new arms run in CPU FP32 for this query contrast; cross-family numerical comparisons are not a hardware test. The actual softmin-weighted sequence is independently simulated with eight noise samples, as in Appendix I.

Closed-loop evaluation reuses the original ten episodes per wind and unchanged 50-step/64-candidate controller: 30 models $\times$ two winds $\times$ ten episodes gives 600 accepted episodes, with

zero numerical or threshold failures. The 100-step PD warmup, 1,500 scored steps and failure-aware scoring are unchanged. No original-model episodes are rerun or duplicated as independent observations.

Table 29: Auxiliary-context comparison. Differences are fixed minus dynamic auxiliary context; both arms evolve their rollout state. All metrics are errors or costs (lower is better), but their units differ.

Wind	Outcome	Fixed aux.	Dynamic aux.	Difference	95% interval
6.1	Logged E	0.04044	0.04272	-0.00228	$[-0.00363, -0.00093]$
6.1	Logged position (m)	0.05980	0.05743	$+0.00237$	$[+0.00119, +0.00355]$
6.1	Candidate E	0.34912	0.45055	-0.10143	$[-0.16811, -0.03476]$
6.1	Physical solution cost	27.65943	28.97478	-1.31535	$[-1.95957, -0.67113]$
6.1	Tracking RMSE (m)	0.22863	0.33278	-0.10415	$[-0.19658, -0.01171]$
8.5	Logged E	0.06987	0.06924	$+0.00063$	$[-0.00168, +0.00294]$
8.5	Logged position (m)	0.13677	0.11018	$+0.02659$	$[+0.02177, +0.03141]$
8.5	Candidate E	0.49453	0.62871	-0.13418	$[-0.22899, -0.03936]$
8.5	Physical solution cost	64.22108	66.08398	-1.86290	$[-3.41646, -0.30933]$
8.5	Tracking RMSE (m)	0.45733	0.56279	-0.10546	$[-0.23875, +0.02784]$

Persistent access improves common-query prediction and physical solution cost in all five datasets at both winds. Tracking RMSE improves in 5/5 datasets at 6.1 m/s and 4/5 at 8.5 m/s, where the interval crosses zero. It worsens logged physical-position prediction in every dataset at both winds. Original Frozen models still have lower mean tracking RMSE (0.210/0.368 m) than either new arm. These mixed results support evaluating context access under candidate actions, not claiming that the hybrid dominates or that improved offline position accuracy is necessary for improved tracking. The secondary 3.7 m/s logged E difference is -0.00197 ($[-0.00352, -0.00042]$); this interpolation condition was not added to the primary query/control tests after observing results.

M NORMALIZATION CONTROLS AND INDEPENDENT PREDICTION REPLICATION

Normalization and recurrence. LN was motivated in part by recurrent learning (Ba et al., 2016); RMSNorm removes its centering step (Zhang & Sennrich, 2019). Prior work studies normalization geometry (Gupta et al., 2024), architecture-dependent algorithmic generalization (Ouellette et al., 2023), and effective learning rates in reinforcement learning (Lyle et al., 2024). Our normalization result is a specific LN-by-memory interaction in a controlled dynamics family. The simulation checkpoints are evaluated on both prediction and closed-loop control, where rankings diverge.

M.1 THE NORMALIZATION EFFECT DEPENDS ON MEMORY UPDATING

Table 30: Static held-out prediction: absolute normalized error, averaged over ten paired seeds per study (lower is better). Every arm has 39,497 parameters. Updated can have lower absolute error even when its LN penalty is larger. Paired contrasts and intervals are reported in the text and Appendix C.

Study	Frozen (no LN)	Frozen (LN)	Updated (no LN)	Updated (LN)
Initial factorial	0.13407	0.13612	0.11030	0.12339
ID-selected rates	0.13499	0.13400	0.11111	0.12387

Table 30 gives the primary comparison. With Updated, LN increases error by 0.01309; the interaction is 0.01105. Both have Holm $p = 0.02539$ within the initial 13-test family, and both are positive in all ten seeds. The Frozen effect is smaller and inconclusive (0.00205, Holm $p = 0.96680$). A normalization-only comparison would miss the dependence on how the memory is carried. Conversely, the absolute means show that Updated with no LN predicts best. The result concerns the joint design, rather than a general advantage of freezing.

Equal-budget ID learning-rate selection retains the Updated penalty (0.01277, descriptive 95% interval $[0.00979, 0.01575]$) and interaction (0.01375, $[0.00892, 0.01859]$), both Holm $p = 0.00586$ on ten new paired seeds. An eight-arm placement study distinguishes memory-input LN from readout LN: the former retains the interaction, whereas the latter has an inconclusive interaction of -0.00041 (Holm $p = 0.92188$). Appendices C and D retain the complete search, absolute means, null results and direct placement contrasts. These checks bound an architecture-specific effect; they do not imply that LN generally harms recurrent dynamics models.

M.2 THE INTERACTION RECURS BEYOND A FIXED SET OF LOGS

Table 31: Independent simulated prediction. LN effects and their interaction are in normalized error. Interactions and Updated penalties are positive in all five datasets in every row. Intervals are descriptive paired 95% intervals over five datasets, not simultaneous; no p-values are used. Dynamic wind is $4.2 + 2.4\sin(\nu t)$ m/s, with ν in rad/s.

Condition	Δ_{Frozen}	Δ_{Updated}	$\mathcal{I}$	95% interval for $\mathcal{I}$
Static 3.7 m/s	$+0.00078$	$+0.00812$	$+0.00734$	$[+0.00376, +0.01092]$
Static 6.1 m/s	$+0.00072$	$+0.01081$	$+0.01008$	$[+0.00661, +0.01356]$
Static 8.5 m/s	-0.00169	$+0.01748$	$+0.01917$	$[+0.01338, +0.02496]$
Dynamic $\nu = 0.25$	$+0.00068$	$+0.00889$	$+0.00821$	$[+0.00488, +0.01153]$
Dynamic $\nu = 1$	$+0.00098$	$+0.00801$	$+0.00703$	$[+0.00383, +0.01024]$

To vary the data that train the models, we generate five independent datasets in a quadrotor simulator with aerodynamic drag, gusts and actuation noise. Each crosses disjoint whole-flight trajectory identities with wind conditions: 16 training flights, eight ID validation flights and ten test trajectories per condition. Training winds match the real-log range. Tests cover static winds 3.7, 6.1 and 8.5 m/s, plus $4.2 + 2.4\sin(\nu t)$ m/s at $\nu \in \{0.25, 1\}$ rad/s. The four primary arms use three model seeds within each dataset, totaling 60 trained models. Each dataset fits its own training-only normalizer and metric scales; simulator physics is shared across datasets.

The LN penalty for Updated and the interaction are positive in all five datasets in every condition (Table 31). LN effects for Frozen remain small and mixed, including a negative mean at static 8.5 m/s whose interval crosses zero. The pattern therefore survives independently sampled trajectories and changing wind in this simulator. It is not merely a repeated-seed result on a single trajectory collection. All four absolute means are in Appendix G; Updated with no LN has the lowest grand mean in all five conditions. This gives a clear predictor ranking to test in downstream control.

The evidence also bounds a simple extrapolation story. In the real-log learning-rate study, the interaction is larger at held-out 3.7 m/s than at 6.1 m/s (0.02195 versus 0.00555). In the separate figure-8 session, it reverses at static 12.1 m/s (-0.06617), while its other condition means stay positive. These descriptive decompositions (Appendix F) do not support a penalty that increases monotonically with wind novelty. That stress session changes trajectory and session along with wind; it does not locate a physical threshold. Replication establishes a recurring pattern within the tested conditions, not a universal law of latent memory.

N Persistent context with an evolving rollout state

The original comparison couples memory updating to direct access to m_0. We train two additional no-LN arms with the same dynamic state but different auxiliary context inputs:

$$c_h = \mathrm{GRUCell}([x_h, q_h], m_h), \qquad m_{h+1} = c_h, \qquad q_h = \begin{cases} m_0 & \text{fixed auxiliary context,} \\ m_h & \text{dynamic auxiliary context.} \end{cases} \tag{4}$$

Both start at m_0, use the same head and integrator, and have 51,785 parameters. Thirty models cover all five datasets and three paired seeds. Parameter counts match, but effective function classes need not: the dynamic arm duplicates its hidden input in some gate transformations. Neither arm is capacity-matched to the original 39,497-parameter models.

On the original 20 common-query anchors at 6.1/8.5 m/s, fixed auxiliary context reduces both candidate prediction error and physical solution cost in all five datasets. Cost differences are -1.315 ($[-1.960, -0.671]$) and -1.863 ($[-3.416, -0.309]$). In 600 paired closed-loop episodes, its RMSE advantage is 0.104 m ($[0.012, 0.197]$) at 6.1 m/s; at 8.5 m/s the mean advantage is 0.105 m but the interval crosses zero. It also *worsens* logged physical-position prediction at both winds, and its mean tracking remains worse than the original Frozen model. Persistent context access is therefore a useful design variable in this tested family, not a universal remedy or a newly dominant controller.

O Training action coverage and memory design

This result-informed follow-up was fixed before formal training and evaluated without tuning on planner-query or control outcomes. It crosses the original no-LN Frozen/Updated architectures with Re-simulated/Mixed training. All 60 models have 39,497 parameters: five training datasets, three paired seeds (70–72), and four arms. This is a controlled data intervention, not another architecture family or a literal replication of the original overlapping-window training recipe.

Training-only matched branches. For each dataset, the same 16 original training flights at $0, 1.3, 2.5, 4.9$ m/s supply 96 equally spaced valid centers per flight. We replay simulator state and gusts and verify recorded positions, velocities, rotations, angular rates and actions against the eight-decimal CSV representation. Histories and current model inputs use the unchanged original parser. No validation/test history, state, wind or trajectory supplies a training branch. At each center, a fresh 64-candidate, 50-step MPPI bank uses the unchanged PD nominal plan, empty warm start and perturbation scales $(0.04, 0.08, 0.08)$. Eight candidates are sampled uniformly without replacement, before observing their outcomes. Eight fresh physical-noise draws pair each logged-action branch with its candidate-action branch at identical full initial state and gust. Actions are float32 for both model and physical simulator; truth uses ten physical substeps. Proposal, selection and noise seed words are $[15, d, f, c, k]$, where d is dataset index, f original flight-row index, c center, and $k = 0, 1, 2$, respectively. Finite branches are never filtered by error.

Re-simulated training takes all eight logged-action branches. Mixed training takes logged-action branches 0–3 and candidate branches 4–7, giving 50% candidate exposure. Both recipes therefore

contain 12,288 examples per dataset with identical history multiplicities, exogenous noise, sample order and target counts. Original training normalizers, metric scales and splits remain unchanged. The treatment includes future states induced by different actions; it does not isolate action coordinates while holding their physical consequences fixed. The 80 generated flight files passed replay, shape, finite-array, rotation and increment checks; full-horizon vectorized physical integration was checked against the scalar engine.

Equal optimization budget and selection. All arms use the original losses, AdamW with learning rate 10^{-3} and weight decay 10^{-4}, gradient clipping at 1, OneCycle, FP32 and batch size 256. Each trains for exactly 120 epochs, 48 batches per epoch and 5,760 optimizer updates. Patience 121 prevents unequal early stopping. Selection uses only the original ID validation windows, stride 5 and minimum improvement 10^{-4}. Paired arms share initialization seeds and shuffled index streams. All formal arms use RTX PRO 6000 GPUs; compatibility/timing pilots also tested B200. Four full-length training pilots were excluded from the formal roster. All 60 checkpoints and 120 ID/static-test NPZ files passed provenance, budget, parameter-count, finite-array, summary, horizon and paired-window checks.

Fresh queries and closed-loop control. At winds 4.9, 6.1 and 8.5 m/s, episode $e = 0, \ldots, 9$ uses environment seed $610000 + e$, trajectory seed $620000 + e$ and planning seed $630000 + e$. Thirty first-decision anchors follow the same 100-step PD warmup. All 60 models receive identical history/state and 64 candidate sequences at an anchor. Eight construction and eight independent evaluation draws score the actual softmin-weighted solutions using physical simulation. Noise seeds are $[12012026, 1, 150000 + i, 100, p]$, with anchor index i and phase $p = 1, 2$. All 30 anchors (1,800 model records) passed complete trace reconstruction with no nonfinite model predictions. Closed-loop evaluation crosses the same 60 models with three winds and ten fresh episodes (1,800 episodes), retaining the unchanged 1,500-step score, planner budget and failure cap. All 1,800 episodes passed full acceptance with zero controller failures. Pilot evaluations are separate and excluded. Queries and closed-loop control share seed mappings but answer different questions; first-decision physical solution cost does not decompose an episode's closed-loop tracking error.

Estimation and scope. Let G_r denote Updated minus Frozen under training recipe r. The planned interaction is $G_{\text{Mixed}} - G_{\text{Re-simulated}}$; a negative value means a relative reduction in the updating disadvantage. We first average model seeds and evaluation anchors/episodes within each training dataset, then use the five paired dataset differences for descriptive Student-t 95% intervals (four degrees of freedom). Intervals are not simultaneous; no simulation p-values, outcome-dependent seed expansion or model selection is performed. Winds 6.1/8.5 m/s are primary and 4.9 m/s secondary. Original static logged-action evaluation separately uses its existing 3.7/6.1/8.5 m/s strata. Training branches cover neither policy-induced history shifts nor later warm-start distributions. The test is limited to this 50% mixture, proposal, data budget and physics family.

Table 32: Action-coverage factorial: all four absolute means. Three seeds and ten evaluation episodes/anchors are averaged within each of five training datasets; logged errors average original test windows. Lower is better.

Outcome	Wind	Re-simulated F	Re-simulated U	Mixed F	Mixed U
Query E	4.9	0.33929	0.37536	0.06503	0.06501
Query E	6.1	0.37311	0.41009	0.07532	0.07449
Query E	8.5	0.48085	0.51416	0.11519	0.11164
Physical solution cost	4.9	23.01036	24.42049	22.74231	22.66374
Physical solution cost	6.1	34.42120	36.10670	34.04508	34.07344
Physical solution cost	8.5	72.47213	74.79277	71.88362	72.00624
Query position (m)	4.9	0.07911	0.10875	0.06074	0.05297
Query position (m)	6.1	0.09867	0.12774	0.07967	0.06322
Query position (m)	8.5	0.18763	0.19722	0.18119	0.12796
Tracking RMSE (m)	4.9	0.20634	0.17371	0.14521	0.10981
Tracking RMSE (m)	6.1	0.24488	0.22136	0.19520	0.14974
Tracking RMSE (m)	8.5	0.39792	0.37757	0.33199	0.25588
Logged E	3.7	0.05809	0.05803	0.02472	0.02284
Logged E	6.1	0.07521	0.07498	0.04030	0.03658
Logged E	8.5	0.11688	0.10709	0.07884	0.06945

Table 33: All memory gaps and coverage interactions. Brackets are descriptive paired 95% intervals across five datasets, not simultaneous intervals. Mixed intervals crossing zero do not establish equivalence.

Outcome	Wind	Re-simulated U–F [CI]	Mixed U–F [CI]	Interaction [CI]
Query E	4.9	+0.0361 [+0.0078,+0.0643]	-0.0000 [-0.0137,+0.0137]	-0.0361 [-0.0682,-0.0040]
Query E	6.1	+0.0370 [+0.0045,+0.0695]	-0.0008 [-0.0167,+0.0150]	-0.0378 [-0.0755,-0.0002]
Query E	8.5	+0.0333 [-0.0055,+0.0721]	-0.0035 [-0.0262,+0.0191]	-0.0368 [-0.0877,+0.0140]
Physical solution cost	4.9	+1.4101 [+0.7385,+2.0817]	-0.0786 [-0.3784,+0.2212]	-1.4887 [-2.2449,-0.7325]
Physical solution cost	6.1	+1.6855 [+0.7917,+2.5793]	+0.0284 [-0.4242,+0.4809]	-1.6571 [-2.5228,-0.7915]
Physical solution cost	8.5	+2.3206 [+1.1230,+3.5183]	+0.1226 [-0.6570,+0.9022]	-2.1980 [-3.3624,-1.0337]
Query position (m)	4.9	+0.0296 [+0.0175,+0.0418]	-0.0078 [-0.0184,+0.0028]	-0.0374 [-0.0558,-0.0190]
Query position (m)	6.1	+0.0291 [+0.0215,+0.0366]	-0.0165 [-0.0319,-0.0010]	-0.0455 [-0.0628,-0.0283]
Query position (m)	8.5	+0.0096 [-0.0052,+0.0244]	-0.0532 [-0.0814,-0.0251]	-0.0628 [-0.0942,-0.0314]
Tracking RMSE (m)	4.9	-0.0326 [-0.0563,-0.0090]	-0.0354 [-0.0784,+0.0077]	-0.0028 [-0.0241,+0.0185]
Tracking RMSE (m)	6.1	-0.0235 [-0.0552,+0.0082]	-0.0455 [-0.0978,+0.0069]	-0.0220 [-0.0481,+0.0042]
Tracking RMSE (m)	8.5	-0.0204 [-0.0699,+0.0292]	-0.0761 [-0.1284,-0.0238]	-0.0558 [-0.0924,-0.0191]
Logged E	3.7	-0.0001 [-0.0054,+0.0053]	-0.0019 [-0.0033,-0.0005]	-0.0018 [-0.0084,+0.0048]
Logged E	6.1	-0.0002 [-0.0059,+0.0055]	-0.0037 [-0.0066,-0.0008]	-0.0035 [-0.0117,+0.0047]
Logged E	8.5	-0.0098 [-0.0160,-0.0036]	-0.0094 [-0.0140,-0.0048]	+0.0004 [-0.0094,+0.0102]

Table 34: Every training-dataset coverage interaction, in original dataset order. Negative values indicate that Mixed training reduces the Updated minus Frozen gap.

Outcome	Wind	D1	D2	D3	D4	D5
Query E	4.9	-0.0686	-0.0311	-0.0565	-0.0086	-0.0156
Query E	6.1	-0.0735	-0.0305	-0.0656	-0.0109	-0.0086
Query E	8.5	-0.0793	-0.0310	-0.0769	-0.0111	+0.0140
Physical solution cost	4.9	-2.3463	-1.7967	-1.2218	-0.7410	-1.3378
Physical solution cost	6.1	-2.6700	-1.9350	-1.5435	-0.8086	-1.3286
Physical solution cost	8.5	-3.1866	-2.7397	-2.6335	-1.4653	-0.9650
Query position (m)	4.9	-0.0417	-0.0581	-0.0385	-0.0175	-0.0314
Query position (m)	6.1	-0.0552	-0.0644	-0.0342	-0.0418	-0.0321
Query position (m)	8.5	-0.0876	-0.0826	-0.0328	-0.0721	-0.0390
Tracking RMSE (m)	4.9	-0.0057	-0.0252	+0.0016	-0.0069	+0.0223
Tracking RMSE (m)	6.1	-0.0418	-0.0424	-0.0127	-0.0204	+0.0075
Tracking RMSE (m)	8.5	-0.0989	-0.0478	-0.0171	-0.0634	-0.0517
Logged E	3.7	-0.0007	-0.0072	-0.0057	-0.0020	+0.0064
Logged E	6.1	-0.0020	-0.0097	-0.0082	-0.0046	+0.0070
Logged E	8.5	-0.0006	-0.0036	-0.0056	-0.0024	+0.0141

P TRAINING-RECIPE AND OPTIMIZATION-BUDGET BRIDGE

Question and fixed protocol. The first coverage experiment changes the original mean closed-loop ranking even in its Re-simulated arm. The follow-up asks which conditional changes in training support, re-simulation and optimization schedule affect that ranking. It was specified after observing the earlier study and before running the new models; it is a result-informed extension, not retrospective preregistration. The four recipes cross two budgets and Frozen/Updated, using the same five training datasets and seeds 70–72: 240 models, including 60 unchanged short-budget Re-simulated/Mixed checkpoints from Appendix O. The other 180 are newly trained. All have 39,497 parameters, no LN, the original loss, training-only normalizers, split manifests and ID checkpoint rule. Sixteen full training pilots use a separate namespace and are excluded from the formal roster.

Recipe definitions and contrasts. Every logical epoch consists of 48 full batches of 256. Dense draws 12,288 windows without replacement from the original 45,600 admissible training centers, reshuffling independently each epoch. Repeated uses the 1,536 training anchors of the first coverage experiment, with each original recorded window repeated eight times in the same branch ordering. Re-simulated uses the accepted eight noisy logged-action futures for those anchors; Mixed uses the unchanged 50% candidate-action mixture, including its induced future states. Thus Repeated–Dense changes support and repetition; Re-simulated–Repeated changes future-noise realization and numerical replay; Mixed–Re-simulated changes future actions and states. These are conditional comparisons, not a fully crossed support-by-noise decomposition. Code labels `recorded` and `log` correspond to Repeated and Re-simulated, respectively.

Short/long budgets use 120/445 logical epochs, or 5,760/21,360 optimizer steps. The latter equals 120 passes of 178 full batches in the original dataset, but does not reproduce the original early stopping or validation cadence. Each budget has its complete OneCycle schedule; early stopping is disabled with patience exceeding the epoch count. Validation occurs after each logical epoch, and the unchanged minimum improvement of 10^{-4} determines ID checkpoint selection. Consequently, budget includes learning-rate duration and selection opportunities, not a pure compute-only intervention. New models use RTX PRO 6000, matching the reused coverage models. All 180 models and 360 ID/static-test NPZ files passed complete provenance, budget, selection, finite-array, summary and paired-window acceptance before evaluation. Finite unfavorable models are retained; no outcome-dependent seed expansion or retraining is used.

Common fresh evaluation and reference models. All 240 models and the 30 original no-LN models receive identical evaluation identities: winds $\{0, 4.9, 6.1, 8.5\}$ m/s and episodes $e = 0, \ldots, 9$, with environment, trajectory and planner seeds $710000+e$, $720000+e$ and $730000+e$. The original models serve as references, not additional factorial treatments. Forty first-decision anchors follow 100 PD warmup steps and use the same 64 candidates and $H = 50$. Eight construction and eight independent evaluation noise draws score the actual weighted returned solution; noise words are $[12012026, 1, 160000 + i, 100, p]$, with anchor i and phase $p = 1, 2$. Query E and physical-position error average the first 50 rollout steps. All 10,800 query-model records are finite and pass full reconstruction. Closed-loop control uses the unchanged planner, 1,500-step scoring and capped failure-aware RMSE; all 10,800 episodes pass numerical acceptance with zero failures. Four query pilots and 72 full control pilots are excluded. Static logged tests use the original 3.7/6.1/8.5 m/s splits: 28,500 windows per wind and 720 model–wind records, with matching window identities across all 240 models.

Estimation and full reporting. The primary endpoint is tracking RMSE at 6.1/8.5 m/s; query metrics are secondary and zero/4.9 m/s are boundary strata. Episode and seed scores are averaged within each training dataset. The five paired dataset scores yield descriptive, nonsimultaneous Student-t 95% intervals with four degrees of freedom, without simulation p-values. Write $G_{r,b} = U_{r,b} - F_{r,b}$. Planned contrasts include each $G_{r,b}$; adjacent recipe changes in G at each budget; long–short changes in G for every recipe; and the budget change in the Mixed–Re-simulated interaction. Absolute within-memory recipe and budget changes are retained alongside these interactions. The following tables report all arm means and U–F intervals for every metric/wind, plus every planned primary tracking contrast. The compact supplement retains all five scores and all 117 offline, 480 query and 160 closed-loop control contrasts, including boundary and secondary con-

trasts. An independent verifier reconstructs means and intervals from the complete compact record rosters; raw numerical acceptance is a separate check on the full cluster-resident arrays and traces.

What the bridge adds and does not resolve. At both primary winds, long Dense/Repeated/Re-simulated recipes favor Frozen, whereas long Mixed training favors Updated. At the long budget, Mixed versus Re-simulated training improves Updated in all five datasets but worsens Frozen in all five; the latter's mean physical solution cost nevertheless improves. At the short budget, the new-roster query-cost interaction intervals cross zero at both primary winds, unlike the earlier coverage roster. Both panels remain reported; these are different evaluation anchors for the same 60 checkpoints. No averaging across panels is used to increase apparent precision. The new results therefore strengthen the training-dependent boundary while limiting any universal coverage claim. Support contrasts and several re-simulation contrasts remain uncertain. These results do not identify an internal memory mechanism, reproduce the original early-stop experiment, or establish a universal control advantage. Matched physical references were subsequently evaluated on this same roster; Appendix Q reports their separate, fully accepted comparison.

Table 35: Failure-aware tracking RMSE (m). All means and paired U–F contrasts from the complete training bridge. Intervals are descriptive and nonsimultaneous across five datasets. Original checkpoints are contextual anchors, not factorial cells.

Wind	Recipe / budget	Frozen	Updated	U–F	95% interval
0	Dense / S	0.1805	0.1454	-0.0351	$[-0.0634, -0.0067]$
0	Dense / L	0.1118	0.1557	$+0.0440$	$[-0.0195, 0.1075]$
0	Repeated / S	0.1999	0.1587	-0.0412	$[-0.0601, -0.0223]$
0	Repeated / L	0.1517	0.1489	-0.0028	$[-0.0377, 0.0321]$
0	Re-simulated / S	0.1995	0.1498	-0.0497	$[-0.0759, -0.0235]$
0	Re-simulated / L	0.1578	0.1172	-0.0406	$[-0.0684, -0.0127]$
0	Mixed / S	0.1223	0.0828	-0.0396	$[-0.0918, 0.0127]$
0	Mixed / L	0.1443	0.0864	-0.0579	$[-0.1330, 0.0172]$
0	Original	0.1133	0.1428	$+0.0295$	$[-0.0407, 0.0997]$
4.9	Dense / S	0.1530	0.1459	-0.0071	$[-0.0217, 0.0075]$
4.9	Dense / L	0.1502	0.2338	$+0.0836$	$[0.0472, 0.1200]$
4.9	Repeated / S	0.1685	0.1681	-0.0004	$[-0.0316, 0.0308]$
4.9	Repeated / L	0.1562	0.2230	$+0.0668$	$[0.0330, 0.1006]$
4.9	Re-simulated / S	0.1737	0.1457	-0.0280	$[-0.0494, -0.0066]$
4.9	Re-simulated / L	0.1224	0.1458	$+0.0234$	$[-0.0031, 0.0500]$
4.9	Mixed / S	0.1296	0.1003	-0.0293	$[-0.0624, 0.0038]$
4.9	Mixed / L	0.1914	0.0826	-0.1089	$[-0.2058, -0.0120]$
4.9	Original	0.1506	0.2242	$+0.0736$	$[0.0116, 0.1357]$
6.1	Dense / S	0.1951	0.2034	$+0.0083$	$[-0.0216, 0.0382]$
6.1	Dense / L	0.2064	0.3279	$+0.1215$	$[0.0741, 0.1689]$
6.1	Repeated / S	0.2131	0.2176	$+0.0046$	$[-0.0325, 0.0416]$
6.1	Repeated / L	0.2166	0.3078	$+0.0911$	$[0.0350, 0.1473]$
6.1	Re-simulated / S	0.2092	0.1847	-0.0245	$[-0.0593, 0.0103]$
6.1	Re-simulated / L	0.1606	0.1978	$+0.0373$	$[0.0038, 0.0707]$
6.1	Mixed / S	0.1782	0.1377	-0.0405	$[-0.0812, 0.0002]$
6.1	Mixed / L	0.2631	0.1119	-0.1512	$[-0.2711, -0.0312]$
6.1	Original	0.2123	0.3184	$+0.1060$	$[0.0111, 0.2010]$
8.5	Dense / S	0.3544	0.3862	$+0.0318$	$[-0.0283, 0.0919]$
8.5	Dense / L	0.3630	0.5628	$+0.1998$	$[0.1394, 0.2602]$
8.5	Repeated / S	0.3712	0.3898	$+0.0186$	$[-0.0113, 0.0484]$
8.5	Repeated / L	0.3952	0.5401	$+0.1449$	$[0.0600, 0.2298]$
8.5	Re-simulated / S	0.3505	0.3333	-0.0172	$[-0.0549, 0.0205]$
8.5	Re-simulated / L	0.2918	0.3727	$+0.0809$	$[0.0134, 0.1485]$
8.5	Mixed / S	0.3064	0.2413	-0.0651	$[-0.1102, -0.0199]$
8.5	Mixed / L	0.4354	0.1994	-0.2360	$[-0.3759, -0.0961]$
8.5	Original	0.3651	0.5761	$+0.2110$	$[0.0461, 0.3760]$

Table 36: Physical solution cost. All means and paired U–F contrasts from the complete training bridge. Intervals are descriptive and nonsimultaneous across five datasets. Original checkpoints are contextual anchors, not factorial cells.

Wind	Recipe / budget	Frozen	Updated	U–F	95% interval
0	Dense / S	3.972	4.127	$+0.156$	$[0.045, 0.267]$
0	Dense / L	4.059	4.207	$+0.148$	$[0.074, 0.222]$
0	Repeated / S	3.960	4.159	$+0.199$	$[0.083, 0.314]$
0	Repeated / L	4.038	4.229	$+0.191$	$[0.093, 0.290]$
0	Re-simulated / S	3.918	4.092	$+0.174$	$[0.094, 0.254]$
0	Re-simulated / L	3.978	4.145	$+0.166$	$[0.053, 0.279]$
0	Mixed / S	3.919	3.865	-0.053	$[-0.114, 0.007]$
0	Mixed / L	3.874	3.849	-0.025	$[-0.076, 0.025]$
0	Original	4.063	4.182	$+0.118$	$[0.009, 0.228]$
4.9	Dense / S	14.683	15.753	$+1.070$	$[0.653, 1.486]$
4.9	Dense / L	15.127	16.164	$+1.038$	$[0.551, 1.524]$
4.9	Repeated / S	14.664	15.496	$+0.832$	$[0.384, 1.279]$
4.9	Repeated / L	15.108	16.199	$+1.090$	$[0.643, 1.537]$
4.9	Re-simulated / S	14.687	15.195	$+0.508$	$[0.232, 0.785]$
4.9	Re-simulated / L	14.788	15.781	$+0.993$	$[0.484, 1.503]$
4.9	Mixed / S	14.650	14.824	$+0.174$	$[-0.044, 0.392]$
4.9	Mixed / L	14.530	14.512	-0.018	$[-0.126, 0.089]$
4.9	Original	15.122	16.055	$+0.933$	$[0.419, 1.447]$
6.1	Dense / S	23.998	25.468	$+1.469$	$[1.073, 1.866]$
6.1	Dense / L	24.514	26.066	$+1.552$	$[0.685, 2.419]$
6.1	Repeated / S	23.968	25.046	$+1.078$	$[0.593, 1.564]$
6.1	Repeated / L	24.563	26.136	$+1.573$	$[0.918, 2.227]$
6.1	Re-simulated / S	24.002	24.714	$+0.712$	$[0.315, 1.109]$
6.1	Re-simulated / L	24.149	25.403	$+1.254$	$[0.495, 2.013]$
6.1	Mixed / S	23.921	24.149	$+0.228$	$[-0.004, 0.461]$
6.1	Mixed / L	23.690	23.731	$+0.041$	$[-0.142, 0.223]$
6.1	Original	24.556	26.050	$+1.494$	$[0.599, 2.389]$
8.5	Dense / S	55.322	57.421	$+2.098$	$[1.672, 2.525]$
8.5	Dense / L	56.213	58.563	$+2.351$	$[0.800, 3.902]$
8.5	Repeated / S	55.176	56.942	$+1.766$	$[1.078, 2.455]$
8.5	Repeated / L	56.000	58.315	$+2.315$	$[1.099, 3.532]$
8.5	Re-simulated / S	55.173	56.551	$+1.378$	$[0.781, 1.975]$
8.5	Re-simulated / L	55.425	57.999	$+2.574$	$[1.614, 3.534]$
8.5	Mixed / S	54.998	55.402	$+0.404$	$[-0.125, 0.933]$
8.5	Mixed / L	54.478	54.623	$+0.145$	$[-0.265, 0.556]$
8.5	Original	56.355	58.290	$+1.935$	$[0.484, 3.387]$

Table 37: Prediction error E on the common candidate banks. All means and paired U–F contrasts from the complete training bridge. Intervals are descriptive and nonsimultaneous across five datasets. Original checkpoints are contextual anchors, not factorial cells.

Wind	Recipe / budget	Frozen	Updated	U–F	95% interval
0	Dense / S	0.1797	0.1920	$+0.0124$	$[0.0044, 0.0204]$
0	Dense / L	0.1102	0.1656	$+0.0554$	$[0.0110, 0.0998]$
0	Repeated / S	0.1849	0.2008	$+0.0159$	$[0.0045, 0.0272]$
0	Repeated / L	0.1132	0.1708	$+0.0576$	$[0.0172, 0.0980]$
0	Re-simulated / S	0.1838	0.1958	$+0.0120$	$[-0.0002, 0.0243]$
0	Re-simulated / L	0.1020	0.1514	$+0.0494$	$[0.0196, 0.0792]$
0	Mixed / S	0.0270	0.0227	-0.0043	$[-0.0068, -0.0019]$
0	Mixed / L	0.0147	0.0121	-0.0027	$[-0.0045, -0.0009]$
0	Original	0.1135	0.1713	$+0.0578$	$[0.0002, 0.1153]$
4.9	Dense / S	0.2194	0.2309	$+0.0115$	$[-0.0058, 0.0287]$
4.9	Dense / L	0.1270	0.1992	$+0.0722$	$[0.0208, 0.1236]$
4.9	Repeated / S	0.2263	0.2431	$+0.0168$	$[-0.0004, 0.0341]$
4.9	Repeated / L	0.1300	0.2075	$+0.0775$	$[0.0314, 0.1237]$
4.9	Re-simulated / S	0.2201	0.2337	$+0.0137$	$[-0.0002, 0.0276]$
4.9	Re-simulated / L	0.1117	0.1799	$+0.0682$	$[0.0343, 0.1021]$
4.9	Mixed / S	0.0376	0.0345	-0.0030	$[-0.0091, 0.0030]$
4.9	Mixed / L	0.0231	0.0215	-0.0016	$[-0.0028, -0.0003]$
4.9	Original	0.1298	0.2032	$+0.0734$	$[0.0115, 0.1352]$
6.1	Dense / S	0.2482	0.2598	$+0.0116$	$[-0.0078, 0.0310]$
6.1	Dense / L	0.1468	0.2329	$+0.0861$	$[0.0261, 0.1462]$
6.1	Repeated / S	0.2554	0.2744	$+0.0190$	$[-0.0007, 0.0387]$
6.1	Repeated / L	0.1491	0.2434	$+0.0943$	$[0.0421, 0.1465]$
6.1	Re-simulated / S	0.2471	0.2629	$+0.0158$	$[-0.0002, 0.0318]$
6.1	Re-simulated / L	0.1251	0.2068	$+0.0817$	$[0.0404, 0.1230]$
6.1	Mixed / S	0.0464	0.0418	-0.0046	$[-0.0115, 0.0023]$
6.1	Mixed / L	0.0285	0.0256	-0.0028	$[-0.0043, -0.0014]$
6.1	Original	0.1509	0.2365	$+0.0856$	$[0.0115, 0.1597]$
8.5	Dense / S	0.3355	0.3497	$+0.0142$	$[-0.0101, 0.0385]$
8.5	Dense / L	0.2245	0.3590	$+0.1345$	$[0.0430, 0.2260]$
8.5	Repeated / S	0.3445	0.3672	$+0.0226$	$[-0.0053, 0.0506]$
8.5	Repeated / L	0.2244	0.3707	$+0.1463$	$[0.0623, 0.2303]$
8.5	Re-simulated / S	0.3342	0.3538	$+0.0196$	$[-0.0051, 0.0443]$
8.5	Re-simulated / L	0.1892	0.3142	$+0.1249$	$[0.0591, 0.1908]$
8.5	Mixed / S	0.0802	0.0703	-0.0099	$[-0.0180, -0.0019]$
8.5	Mixed / L	0.0567	0.0458	-0.0109	$[-0.0168, -0.0049]$
8.5	Original	0.2270	0.3598	$+0.1328$	$[0.0241, 0.2415]$

Table 38: Physical position error (m) on the common candidate banks. All means and paired U–F contrasts from the complete training bridge. Intervals are descriptive and nonsimultaneous across five datasets. Original checkpoints are contextual anchors, not factorial cells.

Wind	Recipe / budget	Frozen	Updated	U–F	95% interval
0	Dense / S	0.0343	0.0388	$+0.0044$	$[-0.0050, 0.0139]$
0	Dense / L	0.0431	0.0521	$+0.0090$	$[0.0016, 0.0165]$
0	Repeated / S	0.0322	0.0409	$+0.0087$	$[0.0000, 0.0174]$
0	Repeated / L	0.0409	0.0523	$+0.0114$	$[0.0037, 0.0192]$
0	Re-simulated / S	0.0311	0.0357	$+0.0046$	$[-0.0066, 0.0157]$
0	Re-simulated / L	0.0328	0.0466	$+0.0138$	$[0.0026, 0.0251]$
0	Mixed / S	0.0314	0.0228	-0.0086	$[-0.0120, -0.0051]$
0	Mixed / L	0.0239	0.0183	-0.0057	$[-0.0100, -0.0014]$
0	Original	0.0440	0.0538	$+0.0098$	$[-0.0025, 0.0220]$
4.9	Dense / S	0.0646	0.0753	$+0.0107$	$[0.0050, 0.0164]$
4.9	Dense / L	0.0725	0.0837	$+0.0112$	$[-0.0010, 0.0234]$
4.9	Repeated / S	0.0575	0.0782	$+0.0207$	$[0.0168, 0.0246]$
4.9	Repeated / L	0.0666	0.0797	$+0.0132$	$[0.0054, 0.0209]$
4.9	Re-simulated / S	0.0477	0.0626	$+0.0149$	$[0.0065, 0.0234]$
4.9	Re-simulated / L	0.0452	0.0600	$+0.0148$	$[0.0069, 0.0228]$
4.9	Mixed / S	0.0478	0.0407	-0.0071	$[-0.0217, 0.0074]$
4.9	Mixed / L	0.0355	0.0345	-0.0010	$[-0.0045, 0.0026]$
4.9	Original	0.0760	0.0853	$+0.0093$	$[-0.0028, 0.0214]$
6.1	Dense / S	0.0918	0.1015	$+0.0097$	$[0.0018, 0.0177]$
6.1	Dense / L	0.1029	0.1142	$+0.0113$	$[-0.0004, 0.0231]$
6.1	Repeated / S	0.0819	0.1037	$+0.0218$	$[0.0174, 0.0262]$
6.1	Repeated / L	0.0920	0.1088	$+0.0168$	$[0.0063, 0.0274]$
6.1	Re-simulated / S	0.0656	0.0835	$+0.0179$	$[0.0079, 0.0280]$
6.1	Re-simulated / L	0.0549	0.0711	$+0.0162$	$[0.0062, 0.0261]$
6.1	Mixed / S	0.0672	0.0511	-0.0161	$[-0.0360, 0.0038]$
6.1	Mixed / L	0.0472	0.0413	-0.0059	$[-0.0101, -0.0016]$
6.1	Original	0.1068	0.1141	$+0.0074$	$[-0.0115, 0.0263]$
8.5	Dense / S	0.1954	0.1872	-0.0082	$[-0.0251, 0.0086]$
8.5	Dense / L	0.2217	0.2227	$+0.0010$	$[-0.0262, 0.0283]$
8.5	Repeated / S	0.1813	0.1850	$+0.0037$	$[-0.0059, 0.0134]$
8.5	Repeated / L	0.2020	0.2117	$+0.0096$	$[-0.0067, 0.0260]$
8.5	Re-simulated / S	0.1537	0.1569	$+0.0032$	$[-0.0125, 0.0190]$
8.5	Re-simulated / L	0.1304	0.1336	$+0.0032$	$[-0.0054, 0.0117]$
8.5	Mixed / S	0.1657	0.1089	-0.0568	$[-0.0826, -0.0311]$
8.5	Mixed / L	0.1305	0.0904	-0.0401	$[-0.0534, -0.0268]$
8.5	Original	0.2243	0.2202	-0.0041	$[-0.0406, 0.0325]$

Table 39: Prediction error E on original static logged-action tests. All means and paired U–F contrasts from the complete training bridge. Intervals are descriptive and nonsimultaneous across five datasets. Original checkpoints are contextual anchors, not factorial cells.

Wind	Recipe / budget	Frozen	Updated	U–F	95% interval
3.7	Dense / S	0.0561	0.0530	-0.0031	$[-0.0089, 0.0026]$
3.7	Dense / L	0.0289	0.0275	-0.0014	$[-0.0027, -0.0002]$
3.7	Repeated / S	0.0612	0.0630	$+0.0018$	$[-0.0036, 0.0071]$
3.7	Repeated / L	0.0299	0.0317	$+0.0018$	$[0.0006, 0.0030]$
3.7	Re-simulated / S	0.0581	0.0580	-0.0001	$[-0.0054, 0.0053]$
3.7	Re-simulated / L	0.0251	0.0225	-0.0026	$[-0.0043, -0.0009]$
3.7	Mixed / S	0.0247	0.0228	-0.0019	$[-0.0033, -0.0005]$
3.7	Mixed / L	0.0171	0.0161	-0.0009	$[-0.0014, -0.0004]$
6.1	Dense / S	0.0730	0.0680	-0.0050	$[-0.0114, 0.0014]$
6.1	Dense / L	0.0411	0.0386	-0.0026	$[-0.0039, -0.0012]$
6.1	Repeated / S	0.0780	0.0796	$+0.0016$	$[-0.0032, 0.0063]$
6.1	Repeated / L	0.0410	0.0424	$+0.0014$	$[-0.0001, 0.0029]$
6.1	Re-simulated / S	0.0752	0.0750	-0.0002	$[-0.0059, 0.0055]$
6.1	Re-simulated / L	0.0364	0.0319	-0.0045	$[-0.0062, -0.0028]$
6.1	Mixed / S	0.0403	0.0366	-0.0037	$[-0.0066, -0.0008]$
6.1	Mixed / L	0.0278	0.0259	-0.0019	$[-0.0031, -0.0007]$
8.5	Dense / S	0.1137	0.0972	-0.0165	$[-0.0229, -0.0102]$
8.5	Dense / L	0.0753	0.0627	-0.0126	$[-0.0151, -0.0102]$
8.5	Repeated / S	0.1185	0.1101	-0.0085	$[-0.0148, -0.0021]$
8.5	Repeated / L	0.0731	0.0657	-0.0074	$[-0.0123, -0.0026]$
8.5	Re-simulated / S	0.1169	0.1071	-0.0098	$[-0.0160, -0.0036]$
8.5	Re-simulated / L	0.0705	0.0545	-0.0160	$[-0.0206, -0.0113]$
8.5	Mixed / S	0.0788	0.0694	-0.0094	$[-0.0140, -0.0048]$
8.5	Mixed / L	0.0600	0.0520	-0.0079	$[-0.0131, -0.0027]$

Table 40: Complete planned training-recipe and budget contrasts in tracking RMSE (m) at 6.1 m/s. Each row uses all five paired dataset scores; negative values favor the first condition for the stated arm, or relatively favor Updated for a change in U–F.

Contrast	Mean	95% interval	Negative / 5
S: Repeated-Dense: Δ(U–F)	-0.0037	$[-0.0534, 0.0459]$	3/5
S: Repeated-Dense: Frozen	$+0.0180$	$[-0.0171, 0.0530]$	2/5
S: Repeated-Dense: Updated	$+0.0142$	$[-0.0037, 0.0322]$	1/5
S: Re-simulated-Repeated: Δ(U–F)	-0.0291	$[-0.0627, 0.0045]$	4/5
S: Re-simulated-Repeated: Frozen	-0.0038	$[-0.0289, 0.0212]$	4/5
S: Re-simulated-Repeated: Updated	-0.0329	$[-0.0486, -0.0172]$	5/5
S: Mixed-Re-simulated: Δ(U–F)	-0.0160	$[-0.0337, 0.0017]$	4/5
S: Mixed-Re-simulated: Frozen	-0.0310	$[-0.0557, -0.0063]$	5/5
S: Mixed-Re-simulated: Updated	-0.0470	$[-0.0845, -0.0095]$	5/5
L: Repeated-Dense: Δ(U–F)	-0.0304	$[-0.0784, 0.0177]$	4/5
L: Repeated-Dense: Frozen	$+0.0102$	$[-0.0407, 0.0612]$	3/5
L: Repeated-Dense: Updated	-0.0202	$[-0.0723, 0.0320]$	2/5
L: Re-simulated-Repeated: Δ(U–F)	-0.0539	$[-0.1076, -0.0002]$	5/5
L: Re-simulated-Repeated: Frozen	-0.0561	$[-0.0863, -0.0259]$	5/5
L: Re-simulated-Repeated: Updated	-0.1100	$[-0.1669, -0.0530]$	5/5
L: Mixed-Re-simulated: Δ(U–F)	-0.1884	$[-0.3019, -0.0749]$	5/5
L: Mixed-Re-simulated: Frozen	$+0.1025$	$[0.0041, 0.2009]$	0/5
L: Mixed-Re-simulated: Updated	-0.0859	$[-0.1279, -0.0440]$	5/5
Dense: L–S: Δ(U–F)	$+0.1132$	$[0.0378, 0.1887]$	0/5
Dense: L–S: Frozen	$+0.0113$	$[-0.0408, 0.0635]$	2/5
Dense: L–S: Updated	$+0.1246$	$[0.0773, 0.1718]$	0/5
Repeated: L–S: Δ(U–F)	$+0.0866$	$[0.0606, 0.1126]$	0/5
Repeated: L–S: Frozen	$+0.0036$	$[-0.0155, 0.0227]$	2/5
Repeated: L–S: Updated	$+0.0902$	$[0.0565, 0.1238]$	0/5
Re-simulated: L–S: Δ(U–F)	$+0.0617$	$[0.0079, 0.1156]$	0/5
Re-simulated: L–S: Frozen	-0.0486	$[-0.1026, 0.0053]$	4/5
Re-simulated: L–S: Updated	$+0.0131$	$[-0.0080, 0.0342]$	1/5
Mixed: L–S: Δ(U–F)	-0.1107	$[-0.2649, 0.0436]$	4/5
Mixed: L–S: Frozen	$+0.0848$	$[-0.0502, 0.2199]$	1/5
Mixed: L–S: Updated	-0.0258	$[-0.0567, 0.0051]$	4/5
Long–short change in Mixed–Re-simulated interaction	-0.1724	$[-0.2868, -0.0580]$	5/5

Table 41: Complete planned training-recipe and budget contrasts in tracking RMSE (m) at 8.5 m/s. Each row uses all five paired dataset scores; negative values favor the first condition for the stated arm, or relatively favor Updated for a change in U–F.

Contrast	Mean	95% interval	Negative / 5
S: Repeated-Dense: Δ(U–F)	-0.0132	$[-0.0910, 0.0646]$	3/5
S: Repeated-Dense: Frozen	$+0.0168$	$[-0.0563, 0.0898]$	2/5
S: Repeated-Dense: Updated	$+0.0035$	$[-0.0175, 0.0245]$	2/5
S: Re-simulated-Repeated: Δ(U–F)	-0.0358	$[-0.0831, 0.0115]$	4/5
S: Re-simulated-Repeated: Frozen	-0.0207	$[-0.0618, 0.0205]$	4/5
S: Re-simulated-Repeated: Updated	-0.0565	$[-0.0942, -0.0187]$	5/5
S: Mixed-Re-simulated: Δ(U–F)	-0.0479	$[-0.0656, -0.0301]$	5/5
S: Mixed-Re-simulated: Frozen	-0.0441	$[-0.1028, 0.0145]$	4/5
S: Mixed-Re-simulated: Updated	-0.0920	$[-0.1448, -0.0392]$	5/5
L: Repeated-Dense: Δ(U–F)	-0.0549	$[-0.1552, 0.0454]$	3/5
L: Repeated-Dense: Frozen	$+0.0323$	$[-0.0296, 0.0941]$	1/5
L: Repeated-Dense: Updated	-0.0226	$[-0.0977, 0.0525]$	3/5
L: Re-simulated-Repeated: Δ(U–F)	-0.0640	$[-0.1671, 0.0391]$	4/5
L: Re-simulated-Repeated: Frozen	-0.1034	$[-0.1479, -0.0590]$	5/5
L: Re-simulated-Repeated: Updated	-0.1674	$[-0.2601, -0.0747]$	5/5
L: Mixed-Re-simulated: Δ(U–F)	-0.3170	$[-0.4218, -0.2121]$	5/5
L: Mixed-Re-simulated: Frozen	$+0.1436$	$[0.0305, 0.2567]$	0/5
L: Mixed-Re-simulated: Updated	-0.1734	$[-0.2445, -0.1022]$	5/5
Dense: L–S: Δ(U–F)	$+0.1680$	$[0.0592, 0.2768]$	0/5
Dense: L–S: Frozen	$+0.0085$	$[-0.0876, 0.1047]$	3/5
Dense: L–S: Updated	$+0.1765$	$[0.1393, 0.2138]$	0/5
Repeated: L–S: Δ(U–F)	$+0.1264$	$[0.0654, 0.1873]$	0/5
Repeated: L–S: Frozen	$+0.0240$	$[0.0091, 0.0390]$	0/5
Repeated: L–S: Updated	$+0.1504$	$[0.0968, 0.2040]$	0/5
Re-simulated: L–S: Δ(U–F)	$+0.0982$	$[-0.0060, 0.2023]$	1/5
Re-simulated: L–S: Frozen	-0.0587	$[-0.1438, 0.0263]$	4/5
Re-simulated: L–S: Updated	$+0.0394$	$[0.0004, 0.0784]$	0/5
Mixed: L–S: Δ(U–F)	-0.1709	$[-0.3449, 0.0030]$	4/5
Mixed: L–S: Frozen	$+0.1290$	$[-0.0388, 0.2968]$	1/5
Mixed: L–S: Updated	-0.0420	$[-0.0799, -0.0040]$	4/5
Long–short change in Mixed–Re-simulated interaction	-0.2691	$[-0.3724, -0.1659]$	5/5

Q MATCHED PHYSICAL REFERENCES FOR THE TRAINING BRIDGE

Fixed comparison. After the training bridge, we evaluated Nominal-physics MPPI and True-mean-wind MPPI on its exact 40-episode roster: the same trajectories, physical noise, planning seeds, 100-step PD warmup, 64 candidates per decision, 50-step horizon, cost and 1,500-step failure-aware scoring. The references are the unchanged physical predictors defined in Appendix H. True-mean-wind MPPI receives the fixed mean wind, not future gusts or physical noise; it is an informed reference, not a guaranteed lower bound. Eight pilot episodes were excluded from the 80 formal episodes. All 80 passed complete numerical acceptance with zero failures; the two zero-wind controllers produced identical actions and errors. No learned checkpoint was retrained or selected.

Paired estimation. For each learned cell and training dataset, we average three seeds within each episode, subtract the matching physical-reference score and average the ten paired differences. Five dataset differences yield descriptive, nonsimultaneous t 95% intervals (four degrees of freedom), conditional on this common evaluation roster. The physical references are not five independently trained models. The compact supplement retains every reference episode and all 144 paired contrasts across 18 learned cells, four winds and two references. An independent implementation reconstructs all contrasts from the complete retained records.

Absolute benefit and remaining gap. Long-budget Mixed Updated has lower RMSE than Nominal-physics MPPI at 4.9/6.1/8.5 m/s, with all three paired intervals below zero. However, it remains worse than True-mean-wind MPPI at all winds and worse than both references at zero wind, with intervals above zero. These results establish a useful wind-dependent gain over the nominal predictor while preserving an informed physical reference's advantage. They neither establish state-of-the-art flight control nor change the conditional Frozen/Updated contrasts.

Table 42: Matched physical references and long-budget Mixed models. Failure-aware tracking RMSE (m), on identical episodes. Physical-reference means average ten episodes; learned means also average all five datasets and three model seeds.

Wind (m/s)	Nominal-physics MPPI	True-mean-wind MPPI	Mixed Frozen	Mixed Updated
0.0	0.0378	0.0378	0.1443	0.0864
4.9	0.2459	0.0542	0.1914	0.0826
6.1	0.3103	0.0671	0.2631	0.1119
8.5	0.4538	0.1318	0.4354	0.1994

Table 43: Long-budget Mixed minus each matched reference. Negative tracking RMSE differences (m) favor the learned model. All intervals use five training datasets, conditional on the shared episode roster. The complete 144-contrast roster is retained in the compact supplement.

Wind	Memory design	Reference	Difference	95% interval
0.0	Frozen	Nominal-physics MPPI	$+0.1065$	$[0.0343, 0.1786]$
0.0	Frozen	True-mean-wind MPPI	$+0.1065$	$[0.0343, 0.1786]$
0.0	Updated	Nominal-physics MPPI	$+0.0486$	$[0.0209, 0.0763]$
0.0	Updated	True-mean-wind MPPI	$+0.0486$	$[0.0209, 0.0763]$
4.9	Frozen	Nominal-physics MPPI	-0.0544	$[-0.1316, 0.0228]$
4.9	Frozen	True-mean-wind MPPI	$+0.1373$	$[0.0601, 0.2145]$
4.9	Updated	Nominal-physics MPPI	-0.1633	$[-0.1867, -0.1399]$
4.9	Updated	True-mean-wind MPPI	$+0.0284$	$[0.0050, 0.0518]$
6.1	Frozen	Nominal-physics MPPI	-0.0472	$[-0.1399, 0.0455]$
6.1	Frozen	True-mean-wind MPPI	$+0.1959$	$[0.1033, 0.2886]$
6.1	Updated	Nominal-physics MPPI	-0.1984	$[-0.2330, -0.1638]$
6.1	Updated	True-mean-wind MPPI	$+0.0448$	$[0.0102, 0.0794]$
8.5	Frozen	Nominal-physics MPPI	-0.0184	$[-0.1234, 0.0866]$
8.5	Frozen	True-mean-wind MPPI	$+0.3035$	$[0.1985, 0.4085]$
8.5	Updated	Nominal-physics MPPI	-0.2544	$[-0.2999, -0.2089]$
8.5	Updated	True-mean-wind MPPI	$+0.0675$	$[0.0220, 0.1130]$

R PROSPECTIVE TEST ON FIVE NEW TRAINING DATASETS

Question and independent data. After the training bridge in Appendix P, we fixed a separate protocol to test its recipe-by-memory-by-budget pattern on new data. The five new training datasets use the same physical parameters, whole-flight split design and training wind conditions as the original simulation study, with disjoint trajectory and noise seed blocks. Each dataset has 16 training flights, eight ID validation flights and ten test trajectories per condition. Its normalizer and metric scales use its own training data only. This is a prospective test of a result-selected finding within the same physics family. The original and confirmation cohorts are reported separately.

Fixed model matrix. Re-simulated and Mixed cross the short/long budgets, Frozen/Updated and three model seeds (70–72) across five new datasets, giving 120 formal models. All use the same no-LN architecture with 39,497 parameters, losses, optimizer and ID checkpoint rule as the corresponding training-bridge cells. Each training flight contributes 96 histories and eight paired noisy futures per history; Mixed replaces half with candidate actions and their induced future states. Both recipes therefore have 12,288 examples per logical epoch. The short/long budgets are 120/445 epochs, or 5,760/21,360 optimizer steps, with complete OneCycle schedules and early stopping disabled. Eight full training pilots cover the entire matrix at dataset 0, seed 70, in a separate namespace and are excluded from all formal comparisons. All formal training uses RTX PRO 6000.

Fresh paired evaluation. Every model receives the same 40 new control episodes: winds $\{0, 4.9, 6.1, 8.5\}$ m/s and $e = 0, \ldots, 9$, with environment, trajectory and planning seeds $1830000+e$, $1840000 + e$ and $1850000 + e$. These differ from the original and training-bridge rosters. The 100-step PD warmup, 64 candidates, 50-step horizon, controller costs, action constraints and 1,500-step failure-aware scoring are unchanged. This gives 4,800 formal control episodes. The 40 corresponding first-decision query banks also evaluate all 120 models, giving 4,800 model–anchor records. Eight physical-noise draws construct costs and eight independent draws evaluate the full returned solution. Static logged tests at 3.7/6.1/8.5 m/s retain 28,500 windows per wind, giving 360 model–wind records. Four query pilots and 32 control pilots are excluded.

Prespecified interpretation and uncertainty. The primary endpoint is failure-aware tracking RMSE at 6.1 and 8.5 m/s. The opposed-effects target is a negative long-budget Mixed–Re-simulated difference for Updated and a positive difference for Frozen at both winds. We distinguish recurrence of these four mean signs from interval precision. The protocol defines precise confirmation of the full pattern as all four descriptive paired 95% intervals excluding zero in their expected directions. If signs recur but an interval crosses zero, the conclusion is directional recurrence with unresolved precision. This is a descriptive interpretation rule, not a multiplicity-adjusted hypothesis test.

Seeds and episodes are averaged within each new training dataset before forming the five paired scores and Student-t intervals with four degrees of freedom. Intervals condition on the common new evaluation roster and are nonsimultaneous. Short-budget, zero-wind and 4.9 m/s results remain in the complete report. No dataset, seed or finite unfavorable outcome is removed, and the two cohorts are not pooled into a ten-dataset confirmation analysis. The compact supplement retains all arm scores and all 51 offline, 204 query and 68 control contrasts, with independent numerical reconstruction.

Accepted results and their limits. All 120 formal models, 240 offline evaluation files, 40 full query banks and 4,800 control episodes pass numerical acceptance, with no nonfinite query outputs or failed control episodes. The four primary long-budget recipe effects have the expected signs in every new dataset, and all four descriptive intervals exclude zero (Table 2). The memory ranking also switches in all five datasets: Re-simulated favors Frozen, whereas Mixed favors Updated at both primary winds. The long–short change in the Mixed–Re-simulated U–F interaction is -0.153 m ($[-0.271, -0.035]$) at 6.1 m/s and -0.226 m ($[-0.418, -0.033]$) at 8.5 m/s.

This recurrence has clear boundaries. At the short budget, both designs improve in mean with Mixed at both primary winds, but all four recipe-effect intervals cross zero. At 4.9 m/s, the long-budget opposite effects also recur in every dataset with intervals excluding zero. At zero wind, Mixed improves both designs' means; the Frozen recipe-effect interval crosses zero. These secondary outcomes do not extend the primary conclusion to every wind and budget.

The query/control distinction also recurs. At the long budget, Mixed improves both designs' physical solution costs at both primary winds, even as Frozen's tracking worsens. U–F physical solution-cost gaps shrink from $2.774/4.228$ under Re-simulated to $0.052/0.083$ under Mixed at 6.1/8.5 m/s; both Mixed intervals cross zero, without establishing equivalence. The logged and candidate-prediction tables retain all memory contrasts, including uncertain ones. In particular, the long-budget Mixed logged-prediction U–F interval at 8.5 m/s has a small positive upper endpoint and crosses zero.

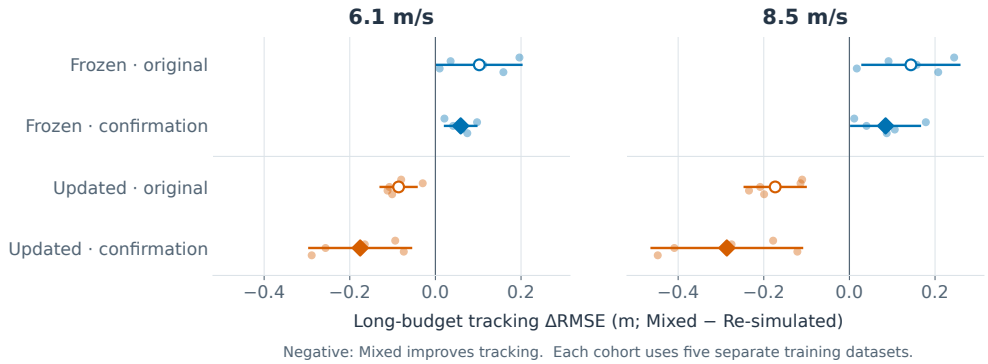

Figure 4: All primary long-budget Mixed–Re-simulated tracking contrasts. Small dots are the five training-dataset differences within each cohort; large markers and bars show cohort means and descriptive paired 95% intervals. Original and confirmation datasets have no between-cohort pairing and are analyzed separately. Negative differences favor Mixed training.

Table 44: Independent confirmation: Failure-aware tracking RMSE (m). All 120 models retained; S/L denotes the short/long budget. Means and paired U–F contrasts use five new training datasets. Intervals are descriptive and nonsimultaneous; extra digits preserve near-zero endpoint signs.

Wind	Recipe / budget	Frozen	Updated	U–F	95% interval
0	Re-simulated / S	0.1870	0.1539	−0.0331	[−0.0802, 0.0140]
0	Re-simulated / L	0.1542	0.1408	−0.0135	[−0.0631, 0.0362]
0	Mixed / S	0.1371	0.0736	−0.0634	[−0.1218, −0.0051]
0	Mixed / L	0.1447	0.0781	−0.0666	[−0.1201, −0.0130]
4.9	Re-simulated / S	0.1865	0.1947	+0.0082	[−0.0350, 0.0513]
4.9	Re-simulated / L	0.1436	0.2204	+0.0768	[0.0366, 0.1170]
4.9	Mixed / S	0.1640	0.1158	−0.0482	[−0.1327, 0.0363]
4.9	Mixed / L	0.1898	0.0958	−0.0941	[−0.1480, −0.0402]
6.1	Re-simulated / S	0.2375	0.2606	+0.0231	[−0.0636, 0.1098]
6.1	Re-simulated / L	0.1827	0.3037	+0.1210	[0.0548, 0.1872]
6.1	Mixed / S	0.2165	0.1576	−0.0589	[−0.1373, 0.0194]
6.1	Mixed / L	0.2418	0.1281	−0.1137	[−0.1823, −0.0451]
8.5	Re-simulated / S	0.3946	0.4365	+0.0419	[−0.0816, 0.1653]
8.5	Re-simulated / L	0.3182	0.5074	+0.1891	[0.0946, 0.2837]
8.5	Mixed / S	0.3704	0.2672	−0.1032	[−0.2231, 0.0167]
8.5	Mixed / L	0.4027	0.2211	−0.1817	[−0.2871, −0.0763]

Table 45: Independent confirmation: Physical solution cost. All 120 models retained; S/L denotes the short/long budget. Means and paired U–F contrasts use five new training datasets. Intervals are descriptive and nonsimultaneous; extra digits preserve near-zero endpoint signs.

Wind	Recipe / budget	Frozen	Updated	U–F	95% interval
0	Re-simulated / S	3.695	4.014	$+0.319$	$[0.201, 0.437]$
0	Re-simulated / L	3.893	4.302	$+0.409$	$[0.211, 0.607]$
0	Mixed / S	3.569	3.587	$+0.018$	$[-0.101, 0.137]$
0	Mixed / L	3.535	3.518	-0.017	$[-0.051, 0.016]$
4.9	Re-simulated / S	18.715	19.654	$+0.939$	$[0.338, 1.541]$
4.9	Re-simulated / L	18.971	21.120	$+2.150$	$[1.835, 2.465]$
4.9	Mixed / S	18.579	18.672	$+0.094$	$[-0.286, 0.473]$
4.9	Mixed / L	18.376	18.423	$+0.047$	$[-0.085, 0.179]$
6.1	Re-simulated / S	29.882	31.128	$+1.247$	$[0.621, 1.873]$
6.1	Re-simulated / L	30.371	33.146	$+2.774$	$[2.327, 3.222]$
6.1	Mixed / S	29.598	29.874	$+0.276$	$[-0.311, 0.864]$
6.1	Mixed / L	29.316	29.369	$+0.052$	$[-0.076, 0.181]$
8.5	Re-simulated / S	69.483	71.700	$+2.217$	$[1.279, 3.155]$
8.5	Re-simulated / L	70.936	75.164	$+4.228$	$[2.941, 5.515]$
8.5	Mixed / S	69.051	69.841	$+0.790$	$[-0.277, 1.858]$
8.5	Mixed / L	68.515	68.598	$+0.083$	$[-0.176, 0.342]$

Table 46: Independent confirmation: Prediction error E on common candidate banks. All 120 models retained; S/L denotes the short/long budget. Means and paired U–F contrasts use five new training datasets. Intervals are descriptive and nonsimultaneous; extra digits preserve near-zero endpoint signs.

Wind	Recipe / budget	Frozen	Updated	U–F	95% interval
0	Re-simulated / S	0.2295	0.2707	$+0.0412$	$[-0.0025, 0.0850]$
0	Re-simulated / L	0.1332	0.2373	$+0.1042$	$[0.0170, 0.1914]$
0	Mixed / S	0.0340	0.0281	-0.0059	$[-0.0087, -0.0031]$
0	Mixed / L	0.0184	0.0163	-0.0022	$[-0.0052, 0.0009]$
4.9	Re-simulated / S	0.3331	0.4124	$+0.0794$	$[-0.0139, 0.1726]$
4.9	Re-simulated / L	0.1912	0.4089	$+0.2176$	$[0.0018, 0.4335]$
4.9	Mixed / S	0.0498	0.0455	-0.0043	$[-0.0089, 0.0004]$
4.9	Mixed / L	0.0303	0.0285	-0.0018	$[-0.0046, 0.0009]$
6.1	Re-simulated / S	0.3899	0.4812	$+0.0913$	$[-0.0196, 0.2022]$
6.1	Re-simulated / L	0.2359	0.4965	$+0.2606$	$[0.0086, 0.5126]$
6.1	Mixed / S	0.0641	0.0597	-0.0044	$[-0.0117, 0.0029]$
6.1	Mixed / L	0.0401	0.0375	-0.0026	$[-0.0063, 0.0010]$
8.5	Re-simulated / S	0.5431	0.6515	$+0.1084$	$[-0.0277, 0.2445]$
8.5	Re-simulated / L	0.3871	0.7420	$+0.3549$	$[0.0501, 0.6597]$
8.5	Mixed / S	0.1160	0.1121	-0.0039	$[-0.0190, 0.0113]$
8.5	Mixed / L	0.0836	0.0718	-0.0118	$[-0.0239, 0.0004]$

Table 47: Independent confirmation: Physical position error (m) on common candidate banks. All 120 models retained; S/L denotes the short/long budget. Means and paired U–F contrasts use five new training datasets. Intervals are descriptive and nonsimultaneous; extra digits preserve near-zero endpoint signs.

Wind	Recipe / budget	Frozen	Updated	U–F	95% interval
0	Re-simulated / S	0.0422	0.0527	$+0.0105$	$[0.0062, 0.0148]$
0	Re-simulated / L	0.0477	0.0769	$+0.0292$	$[0.0165, 0.0419]$
0	Mixed / S	0.0314	0.0259	-0.0055	$[-0.0157, 0.0047]$
0	Mixed / L	0.0263	0.0225	-0.0038	$[-0.0110, 0.0034]$
4.9	Re-simulated / S	0.0698	0.0921	$+0.0223$	$[0.0132, 0.0313]$
4.9	Re-simulated / L	0.0767	0.1046	$+0.0278$	$[0.0144, 0.0413]$
4.9	Mixed / S	0.0446	0.0407	-0.0038	$[-0.0082, 0.0005]$
4.9	Mixed / L	0.0383	0.0372	-0.0011	$[-0.0049, 0.0027]$
6.1	Re-simulated / S	0.0972	0.1233	$+0.0260$	$[0.0144, 0.0376]$
6.1	Re-simulated / L	0.0943	0.1297	$+0.0354$	$[0.0208, 0.0501]$
6.1	Mixed / S	0.0662	0.0549	-0.0112	$[-0.0171, -0.0053]$
6.1	Mixed / L	0.0523	0.0479	-0.0044	$[-0.0092, 0.0003]$
8.5	Re-simulated / S	0.2065	0.2253	$+0.0188$	$[-0.0055, 0.0431]$
8.5	Re-simulated / L	0.1827	0.2251	$+0.0423$	$[0.0208, 0.0638]$
8.5	Mixed / S	0.1857	0.1411	-0.0446	$[-0.0588, -0.0304]$
8.5	Mixed / L	0.1628	0.1216	-0.0412	$[-0.0626, -0.0198]$

Table 48: Independent confirmation: Prediction error E on static logged-action tests. All 120 models retained; S/L denotes the short/long budget. Means and paired U–F contrasts use five new training datasets. Intervals are descriptive and nonsimultaneous; extra digits preserve near-zero endpoint signs.

Wind	Recipe / budget	Frozen	Updated	U–F	95% interval
3.7	Re-simulated / S	0.0792	0.0845	$+0.0054$	$[-0.0031, 0.0139]$
3.7	Re-simulated / L	0.0362	0.0386	$+0.0025$	$[-0.0073, 0.0123]$
3.7	Mixed / S	0.0312	0.0306	-0.0006	$[-0.0027, 0.0014]$
3.7	Mixed / L	0.0217	0.0212	-0.0005	$[-0.0010, -0.0001]$
6.1	Re-simulated / S	0.0991	0.1026	$+0.0035$	$[-0.0034, 0.0104]$
6.1	Re-simulated / L	0.0506	0.0504	-0.0002	$[-0.0068, 0.0064]$
6.1	Mixed / S	0.0508	0.0488	-0.0020	$[-0.0061, 0.0021]$
6.1	Mixed / L	0.0353	0.0340	-0.0013	$[-0.0039, 0.0013]$
8.5	Re-simulated / S	0.1545	0.1463	-0.0082	$[-0.0155, -0.0009]$
8.5	Re-simulated / L	0.0994	0.0835	-0.0159	$[-0.0241, -0.0077]$
8.5	Mixed / S	0.1010	0.0939	-0.0071	$[-0.0163, 0.0020]$
8.5	Mixed / L	0.0788	0.0694	-0.00945	$[-0.01891, 0.00002]$

Table 49: Independent confirmation: complete recipe and budget contrasts in tracking RMSE (m) at 0 m/s. Each contrast uses five new paired training datasets; intervals are descriptive and nonsimultaneous. U–F contrasts appear in the preceding mean table.

Contrast	Mean	95% interval	Negative / 5
S: Mixed–Re-simulated, Δ(U–F)	-0.0303	$[-0.0991, 0.0384]$	4/5
S: Mixed–Re-simulated, Frozen	-0.0499	$[-0.1047, 0.0048]$	4/5
S: Mixed–Re-simulated, Updated	-0.0803	$[-0.1268, -0.0337]$	5/5
L: Mixed–Re-simulated, Δ(U–F)	-0.0531	$[-0.1496, 0.0434]$	4/5
L: Mixed–Re-simulated, Frozen	-0.0096	$[-0.0617, 0.0426]$	3/5
L: Mixed–Re-simulated, Updated	-0.0627	$[-0.1186, -0.0067]$	4/5
Re-simulated: L–S, Δ(U–F)	$+0.0196$	$[-0.0229, 0.0621]$	2/5
Re-simulated: L–S, Frozen	-0.0328	$[-0.0687, 0.0031]$	4/5
Re-simulated: L–S, Updated	-0.0132	$[-0.0256, -0.0007]$	5/5
Mixed: L–S, Δ(U–F)	-0.0031	$[-0.0627, 0.0564]$	1/5
Mixed: L–S, Frozen	$+0.0076$	$[-0.0451, 0.0603]$	3/5
Mixed: L–S, Updated	$+0.0044$	$[-0.0264, 0.0353]$	3/5
L–S change in Mixed–Re-simulated Δ(U–F)	-0.0228	$[-0.1136, 0.0681]$	3/5

Table 50: Independent confirmation: complete recipe and budget contrasts in tracking RMSE (m) at 4.9 m/s. Each contrast uses five new paired training datasets; intervals are descriptive and nonsimultaneous. U–F contrasts appear in the preceding mean table.

Contrast	Mean	95% interval	Negative / 5
S: Mixed–Re-simulated, Δ(U–F)	-0.0564	$[-0.1758, 0.0631]$	4/5
S: Mixed–Re-simulated, Frozen	-0.0225	$[-0.0747, 0.0297]$	4/5
S: Mixed–Re-simulated, Updated	-0.0788	$[-0.1568, -0.0009]$	5/5
L: Mixed–Re-simulated, Δ(U–F)	-0.1709	$[-0.2622, -0.0795]$	5/5
L: Mixed–Re-simulated, Frozen	$+0.0462$	$[0.0253, 0.0671]$	0/5
L: Mixed–Re-simulated, Updated	-0.1246	$[-0.2114, -0.0379]$	5/5
Re-simulated: L–S, Δ(U–F)	$+0.0686$	$[0.0168, 0.1203]$	0/5
Re-simulated: L–S, Frozen	-0.0429	$[-0.0585, -0.0272]$	5/5
Re-simulated: L–S, Updated	$+0.0257$	$[-0.0136, 0.0650]$	1/5
Mixed: L–S, Δ(U–F)	-0.0459	$[-0.0991, 0.0074]$	4/5
Mixed: L–S, Frozen	$+0.0258$	$[-0.0160, 0.0676]$	1/5
Mixed: L–S, Updated	-0.0201	$[-0.0342, -0.0059]$	5/5
L–S change in Mixed–Re-simulated Δ(U–F)	-0.1145	$[-0.2034, -0.0256]$	5/5

Table 51: Independent confirmation: complete recipe and budget contrasts in tracking RMSE (m) at 6.1 m/s. Each contrast uses five new paired training datasets; intervals are descriptive and nonsimultaneous. U–F contrasts appear in the preceding mean table.

Contrast	Mean	95% interval	Negative / 5
S: Mixed–Re-simulated, Δ(U–F)	-0.0820	$[-0.2428, 0.0787]$	4/5
S: Mixed–Re-simulated, Frozen	-0.0210	$[-0.0808, 0.0389]$	3/5
S: Mixed–Re-simulated, Updated	-0.1030	$[-0.2277, 0.0216]$	5/5
L: Mixed–Re-simulated, Δ(U–F)	-0.2347	$[-0.3674, -0.1020]$	5/5
L: Mixed–Re-simulated, Frozen	$+0.0590$	$[0.0225, 0.0956]$	0/5
L: Mixed–Re-simulated, Updated	-0.1757	$[-0.2944, -0.0569]$	5/5
Re-simulated: L–S, Δ(U–F)	$+0.0979$	$[0.0234, 0.1724]$	0/5
Re-simulated: L–S, Frozen	-0.0548	$[-0.0806, -0.0289]$	5/5
Re-simulated: L–S, Updated	$+0.0431$	$[-0.0102, 0.0964]$	1/5
Mixed: L–S, Δ(U–F)	-0.0548	$[-0.1051, -0.0044]$	5/5
Mixed: L–S, Frozen	$+0.0252$	$[-0.0153, 0.0658]$	1/5
Mixed: L–S, Updated	-0.0295	$[-0.0408, -0.0183]$	5/5
L–S change in Mixed–Re-simulated Δ(U–F)	-0.1527	$[-0.2706, -0.0347]$	5/5

Table 52: Independent confirmation: complete recipe and budget contrasts in tracking RMSE (m) at 8.5 m/s. Each contrast uses five new paired training datasets; intervals are descriptive and nonsimultaneous. U–F contrasts appear in the preceding mean table.

Contrast	Mean	95% interval	Negative / 5
S: Mixed–Re-simulated, Δ(U–F)	-0.1451	$[-0.3878, 0.0976]$	4/5
S: Mixed–Re-simulated, Frozen	-0.0243	$[-0.1305, 0.0820]$	3/5
S: Mixed–Re-simulated, Updated	-0.1693	$[-0.3407, 0.0020]$	5/5
L: Mixed–Re-simulated, Δ(U–F)	-0.3708	$[-0.5605, -0.1812]$	5/5
L: Mixed–Re-simulated, Frozen	$+0.0845$	$[0.0044, 0.1646]$	0/5
L: Mixed–Re-simulated, Updated	-0.2863	$[-0.4619, -0.1107]$	5/5
Re-simulated: L–S, Δ(U–F)	$+0.1473$	$[0.0391, 0.2555]$	0/5
Re-simulated: L–S, Frozen	-0.0764	$[-0.1360, -0.0168]$	5/5
Re-simulated: L–S, Updated	$+0.0709$	$[0.0094, 0.1324]$	0/5
Mixed: L–S, Δ(U–F)	-0.0785	$[-0.1736, 0.0167]$	5/5
Mixed: L–S, Frozen	$+0.0324$	$[-0.0573, 0.1220]$	2/5
Mixed: L–S, Updated	-0.0461	$[-0.0658, -0.0264]$	5/5
L–S change in Mixed–Re-simulated Δ(U–F)	-0.2258	$[-0.4182, -0.0333]$	5/5