CAUTIOUS ADAPTATION UNDER PERSISTENT–TRANSIENT AMBIGUITY FOR ONLINE FORECASTING

Anonymous authors
Paper under double-blind review

ABSTRACT

Online forecasting methods increasingly emphasize rapid adaptation to concept drift and equate model adaptation with immediate deployment, yet newly observed behavior may indicate either the beginning of a persistent shift or a transient episode caused by temporary events, sensor failure, or malicious poisoning. At the moment the data is received, these alternatives may be indistinguishable, although they imply opposite decisions for future deployment. We therefore formulate online adaptation as an operational deployment decision problem under persistent–transient ambiguity. A newly updated model is promoted only if its out-of-sample forecasts demonstrate statistically credible improvements over the previous model. We prove that any model promotion rule using only update-time information incurs a nonzero worst-case excess risk across indistinguishable persistent and transient continuations. The resulting framework shifts the emphasis from unconditional rapid adaptation toward cautious updating. Our method hedges the frozen models while out-of-sample evidence accumulates during the subsequent interval, then uses a Diebold–Mariano test to determine promotion. We further establish a pathwise within-cycle sublinear regret bound relative to the better frozen model. Experiments on five forecasting benchmarks show that our method is competitive with strong recent online forecasting baselines, while operational diagnostics characterize its deployment behavior beyond predictive MSE.

1 INTRODUCTION

Forecasting models used in many applications, including energy, logistics, healthcare, and finance, are rarely trained once and left unchanged, as the statistical patterns governing the data may evolve over time. This phenomenon is commonly referred to as *concept drift* (Tsymbal, 2004). In online time-series forecasting, such evolution can take different forms: patterns may recur after having been observed previously, entirely new patterns may emerge, or changes may occur over different temporal scales (Hong et al., 2026). Recent online forecasting methods therefore focus on the speed, contextual specificity, and structure of adaptation (Pham et al., 2023; Zhang et al., 2023; Chen et al., 2024; Lau et al., 2025; Zhao & Shen, 2025; Huang et al., 2026; Hong et al., 2026), aiming to make model updates faster or more accurate. However, they primarily address *how* to adapt, implicitly assuming that newly observed data are sufficiently reliable to update the deployed model and informative for future predictions.

Following the recommendation that practical stream learning should prioritize safety, trustworthiness, and reliable monitoring (Read & Zliobaite, 2025), we instead take a cautious view of model updates during deployment. Faster adaptation is not necessarily preferable: a forecaster that reacts quickly may track a genuine change in the data-generating process, but it can also absorb sensor faults, temporary events, or maliciously poisoned observations. Indeed, distribution shift and adversarial corruption can even be observationally indistinguishable (Sankararaman & Narayanaswamy, 2023). Unlike robust approaches studying robustness and vulnerability of forecasters to corrupted data (Cheng et al., 2024; Lin et al., 2024), our concern is cause-neutral: regardless of why newly observed behavior occurs, we ask whether it should influence future deployment. Conversely, a conservative forecaster limits exposure to transient behavior but may become obsolete when a change persists.

Rather than targeting a particular mechanism or conventional type of concept drift, we adopt a complementary viewpoint based on the *persistence* of newly observed behavior. The same observations may mark the beginning of a persistent change or instead correspond to a transient episode that soon

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.

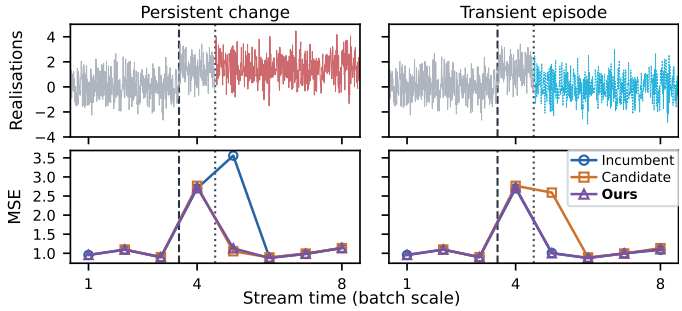

Figure 1: Proof-of-concept comparison on a scalar mean-shift process. The dashed vertical line marks the event batch $\mathcal{B}_4$, and the dotted vertical line marks batch $\mathcal{B}_5$. The candidate is preferable when the change persists, whereas the incumbent becomes preferable when the episode ends.

disappears. These alternatives can imply opposite deployment decisions: adapting may improve subsequent forecasts in one continuation, whereas retaining the previous model may be preferable in another. Crucially, the two continuations may be indistinguishable at update time. This distinction is orthogonal to taxonomies such as recurring versus emergent drift: our question is not *what kind of change occurred?*, but *should the observed behavior influence future predictions?*

A motivating paired example. Consider a scalar mean-shift process $z_t = \mu_t + \varepsilon_t$, with $\varepsilon_t \sim \mathcal{N}(0, 1)$. At batch $b = 4$, the mean shifts from $\mu_0 = 0$ to $\mu_1 = 1.5$. We construct two continuations sharing the same history and update batch $\mathcal{B}_4$: in the persistent continuation the shift remains, whereas in the transient continuation the original regime immediately returns. We call the currently retained model the *incumbent* and the model adapted on $\mathcal{B}_4$ the *candidate*. The candidate is preferable in the former, while the incumbent is preferable in the latter. Since both continuations are identical through $\mathcal{B}_4$, no update-time decision can distinguish them; as illustrated in Figure 1, the subsequent batch provides the first out-of-sample evidence for deciding which model should be retained.

This example suggests that adaptation should not imply immediate permanent promotion: newly observed behavior may be learned while its usefulness for future deployment is evaluated from subsequent out-of-sample evidence.

Contributions and outline. Our contributions are threefold:

- In Section 3, we formulate online forecasting under *persistent–transient ambiguity* as an operational deployment decision problem. We show that when indistinguishable persistent and transient continuations favor opposite models, any rule forced to commit at update time incurs a strictly positive worst-case excess risk, providing a decision-theoretic reason to delay permanent promotion.
- In Section 4, we introduce a two-model deployment-promotion cycle sketched in Figure 2. A candidate is adapted from the incumbent while the latter is preserved; during the subsequent interval, both models remain frozen and a delayed-feedback Hedge rule combines their predictions as out-of-sample evidence accumulates. Once the comparison batch is resolved, a one-sided Diebold–Mariano test determines whether the candidate is permanently promoted. We further establish a pathwise within-cycle sublinear regret guarantee relative to the better frozen model.
- In Section 5, we evaluate the resulting *Hedge and Promote* (HAP) framework on paired persistent–transient semi-synthetic experiments and standard online forecasting benchmarks. The former characterize the trade-off between exposure to transient episodes and delayed adaptation to persistent changes, while the latter show that cautious deployment remains competitive with strong recent adaptation methods.

2 Related Work

Our work connects three lines of research that address complementary aspects of reliable online forecasting: adaptation to distribution shifts, robustness to corrupted observations, and the ambiguity between persistent and transient changes.

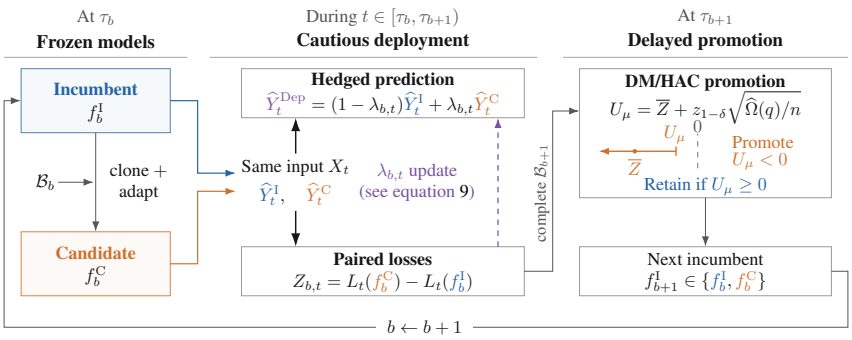

Figure 2: HAP deployment cycle. A candidate is trained on $\mathcal{B}_b$ while the incumbent is preserved. During the next interval, the frozen pair supplies both the hedged predictions and the paired losses used for delayed weight updates. At τ_{b+1}, DM/HAC selects the next incumbent, which initializes a new candidate trained on $\mathcal{B}_{b+1}$.

Adaptation strategies for online forecasting. Recent online forecasting methods explore fast adaptation, memory for recurring patterns, context-aware calibration, dynamic ensembling, and mechanisms for handling delayed feedback. FSNet (Pham et al., 2023) and OneNet (Zhang et al., 2023) modify the forecasting architecture itself, through fast adapters and memory for FSNet and a dynamic ensemble of complementary predictors for OneNet. More recently, ADAPT-Z (Huang et al., 2026) adjusts latent representations using current features and historical gradients, shifting adaptation from model parameters to feature space. More closely aligned with our setting, are the following methods whose focus is the online adaptation of the forecaster to newly observed data:

- SOLID (Chen et al., 2024) constructs a contextualized set for each test input by selecting historical samples according to temporal proximity, periodic alignment, and input similarity, and uses this set to adapt the prediction layer.
- Proceed (Zhao & Shen, 2025) proactively adjusts a model by estimating the corresponding concept drift between the latest fully observed training sample and the current input. Then, it uses a learned adaptation generator to produce input-specific parameter corrections.
- DSOF (Lau et al., 2025) addresses delayed feedback through a teacher–student predictor. Its slow stream updates both models using complete targets and replay, while its fast stream updates the student through temporal-difference learning using newly observed data. The student corrects the teacher's forecasts rather than competing with it for promotion.
- DynaME (Hong et al., 2026) dynamically identifies relevant periods and fits period-specific experts for recurring patterns. A learned gate combines these experts with a general expert. In response to prediction error spikes, it increases reliance on the general expert. In particular, this expert is updated online together with the shared backbone rather than remaining a frozen model.

Our focus is complementary to these adaptation strategies: rather than changing how the forecaster adapts, we separate candidate adaptation from permanent model retention and require subsequent out-of-sample evidence before promotion.

Robust forecasting and poisoning attacks. Another line of work considers robustness to corrupted observations. RobustTSF (Cheng et al., 2024) learns forecasting models from anomaly-contaminated training data. BackTime (Lin et al., 2024) demonstrates that a small number of structured poisoned timestamps can create effective backdoors in multivariate forecasters. Our base problem is broader and cause-neutral, since a transient episode need not be malicious.

Identifiability under drift and contamination. More fundamentally, adjacent work studies the identifiability of distributional changes from observations that may be contaminated. In adversarially robust change-point detection, Li & Yu (2021) characterize contamination-dependent regimes in which no estimator can consistently localize the underlying change points. In online non-stationary estimation, Sankararaman & Narayanaswamy (2023) give a simple exact construction in which a shift-only process and a corruption-only process generate the same observed sequence. Our work builds on this shared identifiability principle but shifts the focus to supervised forecasting with delayed outcomes: rather than inferring the cause of an observed change, we ask whether a candidate yields lower subsequent out-of-sample forecasting risk than the incumbent.

3 Problem Formulation and Persistent–Transient Ambiguity

We now formalize the batch-incremental online forecasting setting underlying our approach, illustrated in Figure 3. We first describe the prediction and model-update protocol, in which forecasts are issued continuously while model updates occur only at larger times. We then formulate the incumbent-candidate deployment decision and show how it gives rise to persistent–transient ambiguity and an unavoidable cost of immediate model promotion.

3.1 Online forecasting and update protocol

Let $\{z_t\}_{t\geq 1}$, with $z_t \in \mathbb{R}^d$, be a multivariate process. At each time t, we define

$$(L\text{-context}) \quad X_t = z_{t-L+1:t} \in \mathbb{R}^{L\times d},$$
$$(H\text{-prediction}) \quad Y_t = z_{t+1:t+H} \in \mathbb{R}^{H\times d}.$$

Pairs accumulate as their targets become observable. At time τ_{b+1}, the completed batch is

$$\mathcal{B}_{b+1} = \big((X_t, Y_t)\big)_{t\in I_{b+1}}, \tag{1}$$

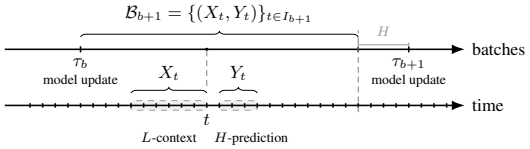

Figure 3: Batch and time views.

where $I_{b+1} = \{\tau_b, \ldots, \tau_{b+1} - H\}$ and $n_{b+1} = |I_{b+1}|$. We consider a *batch-incremental deployment* setting in which model updates occur at times $0 < \tau_0 < \tau_1 < \cdots < \tau_B$, with $\tau_{b+1} - \tau_b > H$.

At update time τ_b, the learner holds an *incumbent model* f_b^{I} and constructs a *candidate model* f_b^{C} through an adaptation procedure $\mathcal{A}$,

$$f_b^{\mathrm{C}} = \mathcal{A}\big(f_b^{\mathrm{I}}, \mathcal{B}_b; \xi_b\big)\,, \tag{2}$$

where ξ_b denotes optimization randomness. At the next update time τ_{b+1}, the learner decides whether to promote the candidate,

$$f_{b+1}^{\mathrm{I}} = \begin{cases} f_b^{\mathrm{C}}, & \text{(promote candidate)},\\ f_b^{\mathrm{I}}, & \text{(retain incumbent)}. \end{cases} \tag{3}$$

The losing model is discarded, and the selected incumbent is used to train the candidate for the next cycle, $f_{b+1}^{\mathrm{C}} = \mathcal{A}(f_{b+1}^{\mathrm{I}}, \mathcal{B}_{b+1}; \xi_{b+1})$. Thus, the forecaster maintains exactly two model instances: one incumbent and one candidate. The promotion rule is specified in Section 4.2.

Between update times τ_b and τ_{b+1}, the forecaster must continue issuing predictions at every time step, even though the next model update occurs only at τ_{b+1}. During this interval, both the incumbent f_b^{I} and the candidate f_b^{C} are available and remain fixed. The deployment rule must therefore specify how these two models are used while new data arrive. We address this in Section 4.1 through hedged deployment while evidence accumulates.

In both cases, the central question is whether the incumbent or the candidate should be favored, together with the consequences of making that choice, as formalized in the next subsection.

Remark 1 (Causality and information leakage)**.** *Online forecasting methods may suffer from information leakage when updates rely on targets that would not yet be observable at deployment time (Zhao & Shen, 2025). Throughout this work, all methods use only information available at the corresponding time. For HAP, every training update, Hedge update, validation statistic, and model-selection decision uses only forecast–target pairs whose targets are fully observable.*

3.2 Persistent–transient ambiguity

At update time τ_b, the future usefulness of the incumbent and candidate is not yet known. The behavior observed in $\mathcal{B}_b$ may either persist into subsequent intervals, defining a *persistent-change continuation* P, or disappear, defining a *transient-behavior continuation* T. These two continuations may favor different models.

To formalize this, let $\mathcal{O}_b = (\mathcal{H}_b, \mathcal{B}_b, \xi_b)$ denote the information available at time τ_b, where $\mathcal{H}_b$ contains the observed stream history and forecaster state. Conditional on $\mathcal{O}_b$, both f_b^{I} and f_b^{C} are

fixed. For any possible continuation $w \in \{P, T\}$ of the stream, we measure performance through the conditional average next-batch risk

$$R^w_{b+1}(f) = \frac{1}{n_{b+1}} \sum_{t \in I_{b+1}} \mathbb{E}[L_t(f) \mid \mathcal{O}_b, w], \tag{4}$$

where $L_t(f) = \ell(f(X_t), Y_t)$ for a forecasting loss $\ell : \mathbb{R}^{H\times d} \times \mathbb{R}^{H\times d} \to \mathbb{R}_+$. Hereafter, we omit the conditioning on $\mathcal{O}_b$ when no ambiguity arises.

To compare the two models, we define the paired loss difference $Z_{b,t} = L_t(f^{\mathrm{C}}_b) - L_t(f^{\mathrm{I}}_b)$ for $t \in I_{b+1}$, and the corresponding conditional average excess risk

$$\Delta^w_{b+1} = R^w_{b+1}(f^{\mathrm{C}}_b) - R^w_{b+1}(f^{\mathrm{I}}_b) = \frac{1}{n_{b+1}} \sum_{t \in I_{b+1}} \mathbb{E}[Z_{b,t} \mid \mathcal{O}_b, w]. \tag{5}$$

The sign of Δ^w_{b+1} indicates which of the two models is preferable in that continuation. In particular, $\Delta^w_{b+1} < 0$ favors the candidate and $\Delta^w_{b+1} > 0$ favors the incumbent. We can now formalize the ambiguity described in the introduction.

Definition 1 (Persistent–transient ambiguity)**.** *Consider two possible continuations P and T sharing the same information $\mathcal{O}_b$. For a given incumbent-candidate pair, persistent–transient ambiguity arises when these two continuations favor opposite models, i.e.,*

$$\Delta^P_{b+1} < 0, \qquad \Delta^T_{b+1} > 0. \tag{6}$$

Thus, the information available at update time is identical in the two worlds, yet the model minimizing future forecasting risk is different.

Remark 2 (Malicious corruption as a special transient case)**.** *Persistent–transient ambiguity is cause-agnostic and therefore also covers maliciously constructed update batches. Temporally localized attacks such as BackTime can alter future forecasts (Lin et al., 2024), but our protocol does not attempt to identify their origin: the candidate is evaluated on subsequent out-of-sample observations before it can replace the incumbent.*

The ambiguity has an unavoidable operational cost if the learner is forced to choose immediately.

Proposition 1 (Cost of immediate commitment)**.** *Let $g_{\mathsf{P}} = -\Delta^{\mathsf{P}}_{b+1} > 0$ and $g_{\mathsf{T}} = \Delta^{\mathsf{T}}_{b+1} > 0$ denote the excess risks of making the wrong decision in the persistent and transient continuations, respectively. Consider any possibly randomized decision rule $d(\mathcal{O}_b) = \mathbb{P}(\text{choose candidate}|\mathcal{O}_b)$ that must choose between the incumbent and candidate at time τ_b using only $\mathcal{O}_b$. Let $\mathcal{R}_w(d)$ denote its expected next-batch excess risk relative to the lower-risk model in world w. Then*

$$\inf_{d(\mathcal{O}_b)} \max_{w \in \{\mathsf{P},\mathsf{T}\}} \mathcal{R}_w(d) = \frac{g_{\mathsf{P}} g_{\mathsf{T}}}{g_{\mathsf{P}} + g_{\mathsf{T}}} > 0. \tag{7}$$

For deterministic rules, the minimax excess risk is $\min\{g_{\mathsf{P}}, g_{\mathsf{T}}\}$.

Proposition 1 shows that persistent–transient ambiguity makes immediate commitment intrinsically risky: no rule based only on $\mathcal{O}_b$ can avoid positive worst-case excess risk. Randomization can reduce this cost by balancing the two possible errors, but cannot eliminate it; a deterministic rule must incur at least $\min\{g_{\mathsf{P}}, g_{\mathsf{T}}\}$.

This also motivates our batch-incremental protocol, in which promotion decisions and model updates occur only at the "coarser" times $\{\tau_b\}_b$, rather than reacting to every newly arriving observation, which can be undesirable or even unsafe (Read & Zliobaite, 2025). Predictions must nevertheless be issued continuously between $\{\tau_b\}_b$, so delaying promotion decision does not remove the need for an intermediate deployment strategy. Our approach therefore combines cautious, hedged deployment while evidence accumulates with a delayed promotion decision once sufficient out-of-sample evidence is available.

4 Cautious Deployment and Evidence-Gated Promotion

Having motivated the two-stage deployment-promotion principle in Section 3, we now specify how to deploy the two models while evidence accumulates and how to decide permanent promotion once the comparison batch is resolved.

4.1 HEDGED DEPLOYMENT DURING EVIDENCE ACCUMULATION

During the deployment interval $[\tau_b, \tau_{b+1})$, two models are retained: the incumbent f_b^{I} and the candidate $f_b^{\mathrm{C}} = \mathcal{A}\big(f_b^{\mathrm{I}}, \mathcal{B}_b; \xi_b\big)$ which has been adapted on $\mathcal{B}_b$.
As new data arrive, we use a *Hedge rule* to combine f_b^{I} and f_b^{C} for prediction. At each forecast origin $t \in [\tau_b, \tau_{b+1})$, we output the deployed predictor

$$f_{b,t}^{\mathrm{dep}} \coloneqq (1 - \lambda_{b,t}) f_b^{\mathrm{I}} + \lambda_{b,t} f_b^{\mathrm{C}}, \tag{8}$$

where the weight $\lambda_{b,t} \in (0, 1)$ controls the relative influence of the candidate. The weight is updated from the paired losses whose targets are observable at time t. We define

$$\lambda_{b,t} = \sigma(a_{b,t}) = \frac{1}{1 + e^{-a_{b,t}}} \quad \text{with} \quad a_{b,t} = \eta \sum_{u=\tau_b}^{t-H} \left[L_u(f_b^{\mathrm{I}}) - L_u(f_b^{\mathrm{C}}) \right], \tag{9}$$

where $\eta > 0$ controls how rapidly the weights react to accumulated evidence, and an empty sum is defined as zero. Consequently, before any forecast from the current interval has been fully resolved, $a_{b,t} = 0$ and $\lambda_{b,t} = 1/2$, so the two models initially receive equal weight.
Whenever a new target becomes observable, its paired loss difference updates the score. If the candidate performs better, $L_u(f_b^{\mathrm{C}}) < L_u(f_b^{\mathrm{I}})$, then $a_{b,t}$ increases and the deployed prediction shifts toward the candidate; if the incumbent performs better, the score decreases and the prediction shifts toward the incumbent. With equal initial weights, this is exactly the two-expert exponential-weights (Hedge) update (Freund & Schapire, 1997), written in log-odds form.
The candidate can therefore influence deployment before being permanently promoted, while its influence remains continuously adjustable as new outcomes are observed.

We now characterize the cost of hedged deployment relative to the better of the two frozen models in hindsight. For this result, we instantiate the generic forecasting loss L_t from Section 3.2 as the normalized multivariate, multi-horizon squared loss, $L_t(f) = (Hd)^{-1}\|f(X_t) - Y_t\|_F^2$. This specialization is only required for the theoretical result below; the Hedge rule itself applies to any observable paired loss.

Theorem 1 (Delayed-Hedge pair-oracle regret)**.** *Let $t_1 < \cdots < t_n$ enumerate I_{b+1}, with $n = n_{b+1}$. The deployed predictor satisfies, pathwise,*

$$\underbrace{\sum_{i=1}^{n} L_{t_i}(f_{b,t_i}^{\mathrm{dep}}) - \min_{g \in \{\mathrm{I},\mathrm{C}\}} \sum_{i=1}^{n} L_{t_i}(f_b^g)}_{R_{b+1}^{\mathrm{pair}}} \le \frac{\log 2}{\eta} + \frac{\eta(2H-1)}{8} V_n - \sum_{i=1}^{n} \lambda_i (1 - \lambda_i) \Gamma_i. \tag{10}$$

Here, $V_n = \sum_{i=1}^{n} Z_{b,t_i}^2$ is the cumulative squared paired loss difference, $\Gamma_i = (Hd)^{-1}\|f_b^{\mathrm{C}}(X_{t_i}) - f_b^{\mathrm{I}}(X_{t_i})\|_F^2$ measures prediction disagreement, $\eta > 0$ is the Hedge rate, and $\lambda_i = \lambda_{b,t_i}$ is the candidate weight at forecast origin t_i.

Theorem 1 requires neither independence nor stationarity. The first term is the usual two-expert Hedge cost, while the second accounts for the realized loss differences and the H-step feedback delay. The final, nonpositive term is specific to squared-loss prediction mixing: disagreement between the two models can make their convex combination outperform the corresponding weighted average of their individual losses.

In particular, delayed feedback does not prevent sublinear within-cycle regret when the cumulative squared loss difference grows at most linearly with the cycle length.

Corollary 1 (Sublinear within-cycle regret)**.** *Under the conditions of Theorem 1, suppose $V_n \le v^2 n$ for some constant $v \ge 0$ independent of n. Setting $\eta = \eta_0/\sqrt{(2H-1)n}$, with $\eta_0 > 0$, gives, for fixed forecast horizon H,*

$$R_{b+1}^{\mathrm{pair}} \le \sqrt{(2H-1)n} \left(\frac{\log 2}{\eta_0} + \frac{\eta_0 v^2}{8} \right) = O(\sqrt{n}). \tag{11}$$

4.2 HAC-ADJUSTED CANDIDATE PROMOTION

At time τ_{b+1}, all forecast–target pairs in $\mathcal{B}_{b+1}$ are fully observable, and the temporary deployment phase ends. The accumulated out-of-sample evidence can therefore be used to decide whether the candidate should permanently replace the incumbent, as in equation 3.

This comparison is performed on $\mathcal{B}_{b+1}$ rather than on the training batch $\mathcal{B}_b$: the latter only measures how well the candidate fits the behavior on which it was adapted, whereas $\mathcal{B}_{b+1}$ provides the first complete block of subsequent forecasting evidence. We summarize this evidence through the average paired loss difference $\overline{Z}_{b+1} = n_{b+1}^{-1} \sum_{t \in I_{b+1}} Z_{b,t}$, where negative values favor the candidate. Because the paired loss differences are temporally dependent, both through the underlying series and through overlapping H-step targets, we assess this improvement using a one-sided Diebold–Mariano test (Diebold & Mariano, 1995) with a Bartlett–Newey–West estimate of the long-run variance (Newey & West, 1987).

Assumption 1 (Local comparison regularity)**.** *Let $t_1 < \cdots < t_n$ enumerate I_{b+1}, with $n = n_{b+1}$, and set $Z_i = Z_{b,t_i}$. Conditional on $\mathcal{O}_b$, the incumbent and candidate are fixed, and there exists a finite, nondegenerate long-run variance $\Omega_b \in (0, \infty)$ such that $\sqrt{n}(\overline{Z}_{b+1} - \mu_{b,n}) \xrightarrow{d} \mathcal{N}(0, \Omega_b)$ and $\widehat{\Omega}(\widehat{q}) \xrightarrow{p} \Omega_b$, where $\mu_{b,n} = \mathbb{E}[\overline{Z}_{b+1} \mid \mathcal{O}_b]$.*

Under Assumption 1, the candidate is promoted only when its average loss advantage is sufficiently large relative to its estimated long-run uncertainty. We consider the one-sided hypotheses $H_0 : \mu_{b,n} \geq 0$ against $H_1 : \mu_{b,n} < 0$ and define the corresponding upper confidence bound

$$U_b = \overline{Z}_{b+1} + z_{1-\delta} \sqrt{\frac{\widehat{\Omega}(\widehat{q})}{n}}, \tag{12}$$

where δ is the one-sided significance level. The promotion decision at τ_{b+1} is $\mathbf{1}\{U_b < 0\}$. Thus, the candidate is promoted only when its observed average improvement remains negative after accounting for estimation uncertainty. Failure to promote does not establish that the incumbent is superior; it means only that the candidate has not demonstrated sufficient evidence of improvement, in which case the incumbent is retained.

Remark 3. *We estimate the long-run variance with the Bartlett–Newey–West estimator $\widehat{\Omega}(q) = \widehat{\gamma}_0 + 2\sum_{k=1}^{q}(1 - k/(q+1))\widehat{\gamma}_k$, where $\widehat{\gamma}_k$ is the empirical lag-k autocovariance of the ordered paired differences $Z_1, \ldots, Z_n$. We use the automatic Newey–West bandwidth rule (Newey & West, 1994), with the additional constraint $q \geq H - 1$ to account for the dependence induced by overlapping targets; the complete specification is given in Appendix A.*

Remark 4 (Local meaning of promotion)**.** *The comparison on $\mathcal{B}_{b+1}$ provides evidence over the completed comparison interval, not a guarantee that the selected model will remain preferable indefinitely. Subsequent changes are handled by repeating the cycle as new data arrive.*

At τ_{b+1}, the promotion decision therefore converts the out-of-sample evidence accumulated during hedged deployment into a permanent model update. The selected model becomes the next incumbent f_{b+1}^{I}, the other model is discarded, and $\mathcal{B}_{b+1}$ is then used to train the next candidate f_{b+1}^{C}. The same deployment–promotion cycle can therefore be repeated as new data arrive, without requiring a separate validation stream.

5 Experiments

We compare the proposed *Hedge and Promote* (HAP) method against the information-leakage free baselines of Section 2. All methods use PatchTST (Nie et al., 2023) as the forecasting backbone; experimental details are provided in Appendix C. We first study a controlled setting exhibiting persistent–transient ambiguity, and then evaluate HAP on standard online forecasting benchmarks.

5.1 ETTm1-PT: paired daily-profile intervention

To isolate persistent–transient ambiguity in a more challenging setting than the example of Section 1, we construct ETTm1-PT, a semi-synthetic benchmark[1] derived from ETTm1 (Zhou et al., 2021).

ETTm1-PT dataset. Starting from standardized ETTm1, we apply a smooth $\sin^2$ daily-profile perturbation to its six load channels, increasing morning loads and decreasing evening loads, while leaving oil temperature unchanged. We consider three perturbation severities $\Delta \in \{3, 6, 9\}$ and construct matched versions of the same underlying time series:

[1]The dataset will be made publicly available.

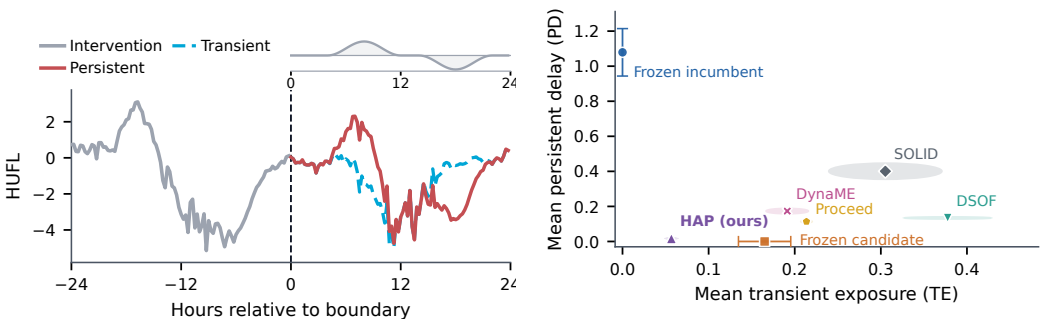

Figure 4: **ETTm1-PT benchmark.** (Left) Daily-profile intervention with $\Delta = 3$. (Right) Persistent-delay–transient-exposure trade-off averaged over $\Delta \in \{3, 6, 9\}$. Shaded ellipses and error bars indicate variability across seeds.

- TRANSIENT: the modified daily profile is present during the shared history and then disappears;
- PERSISTENT: the same modification is present during the shared history and continues afterward.

The persistent and transient versions are therefore identical up to the branching time and differ only in what follows. The exact perturbation and severity levels are specified in Appendix C.3. Figure 4 (Left) illustrates the intervention on the HUFL load channel for $\Delta = 3$.

Results. Averaged over the three perturbation magnitudes $\Delta \in \{3, 6, 9\}$, Table 1 shows that HAP achieves the lowest mean MSE in both continuations and improves the PT average by 5.33% over Proceed. To characterize how this gain relates to persistent–transient ambiguity, we additionally measure two operational costs using HAP's frozen incumbent and candidate as common references. Let $\ell^m_{w,t}$ denote the MSE of method m at forecast origin t under continuation w. We define the *transient exposure* TE^m and *persistent delay* PD^m as follows

$$\mathrm{TE}^m = \frac{1}{N}\sum_t \left[\ell^m_{\mathrm{T},t} - \ell^{\mathrm{I}}_{\mathrm{T},t}\right]_+, \qquad \mathrm{PD}^m = \frac{1}{N}\sum_t \left[\ell^m_{\mathrm{P},t} - \ell^{\mathrm{C}}_{\mathrm{P},t}\right]_+, \tag{13}$$

where N is the number of evaluated forecast origins, and $[x]_+ = \max(x, 0)$. TE therefore quantifies degradation relative to retaining the incumbent when the intervention disappears, whereas PD quantifies degradation relative to deploying the candidate when it persists. Figure 4 (right) shows that HAP Pareto-dominates the adaptive baselines in the mean TE–PD plane, indicating a better balance between limiting transient exposure and avoiding delayed adaptation. Additional operational results reported in Appendix D show that HAP maintains the lowest transient exposure and persistent delay across all perturbation magnitudes Δ.

Table 1: **ETTm1-PT forecasting performance.** Post-intervention MSE averaged over $\Delta \in \{3, 6, 9\}$ (mean $\pm$ standard deviation over three seeds). PT Average is the mean across the transient and persistent continuations; per-magnitude results are given in Appendix D.

Method	Transient	Persistent	PT Average
HAP (ours)	$\mathbf{0.8549} \pm 0.0159$	$\mathbf{0.6398} \pm 0.0148$	$\mathbf{0.7474} \pm 0.0151$
SOLID	1.0488 ± 0.0676	1.0320 ± 0.0477	1.0404 ± 0.0109
Proceed	0.8889 ± 0.0128	0.6900 ± 0.0033	0.7895 ± 0.0081
DSOF	1.0770 ± 0.0470	0.6852 ± 0.0118	0.8811 ± 0.0289
DynaME	0.8776 ± 0.0213	0.7795 ± 0.0198	0.8285 ± 0.0144

5.2 FORECASTING BENCHMARK

We next conduct extensive experiments on well-established forecasting datasets to assess HAP's performance across diverse domains and forecast horizons.

Datasets and protocol. We evaluate HAP on five multivariate forecasting benchmarks: ETTm1 and ETTh2 (Zhou et al., 2021), Electricity (ECL) (Trindade, 2015; Lai et al., 2018), Weather (Wu et al.,

Table 2: **Test MSE on forecasting benchmarks.** Results are reported as mean ± standard deviation over three seeds. Bold indicates the lowest mean MSE for each dataset–horizon pair. The final row reports the average rank across the 15 tasks.

Dataset	Horizon	HAP	SOLID	Proceed	DSOF	DynaME
ETTm1	24	0.4775 ± 0.0651	0.5136 ± 0.0097	$\mathbf{0.4723 \pm 0.0032}$	0.6349 ± 0.0050	0.4799 ± 0.0020
	48	$\mathbf{0.6051 \pm 0.0037}$	0.6986 ± 0.0113	0.6456 ± 0.0033	0.8289 ± 0.0072	0.6404 ± 0.0020
	96	0.7876 ± 0.0588	0.7860 ± 0.0150	0.7839 ± 0.0034	1.0597 ± 0.0211	$\mathbf{0.7298 \pm 0.0025}$
ETTh2	24	2.0621 ± 0.0191	1.9308 ± 0.0226	2.1681 ± 0.0451	$\mathbf{1.8927 \pm 0.0076}$	2.0608 ± 0.2475
	48	3.6680 ± 0.3006	3.5001 ± 0.0197	4.1970 ± 0.2534	3.5739 ± 0.0449	$\mathbf{3.2832 \pm 0.0188}$
	96	6.1267 ± 0.0534	6.0482 ± 0.0846	6.2635 ± 0.1065	6.4125 ± 0.0927	$\mathbf{6.0452 \pm 0.1143}$
Electricity	24	$\mathbf{4.1054 \pm 0.0116}$	4.1669 ± 0.0115	4.3131 ± 0.0194	5.3522 ± 0.0340	4.1268 ± 0.0106
	48	4.8296 ± 0.0163	4.8622 ± 0.0239	5.1412 ± 0.0177	6.5112 ± 0.1414	$\mathbf{4.8059 \pm 0.0147}$
	96	$\mathbf{6.0311 \pm 0.0168}$	6.0374 ± 0.0166	6.0317 ± 0.0131	7.4330 ± 0.1269	6.1163 ± 0.0147
Weather	24	$\mathbf{0.7379 \pm 0.0049}$	0.7385 ± 0.0020	0.7383 ± 0.0026	0.9010 ± 0.0045	0.7476 ± 0.0042
	48	1.0121 ± 0.0058	$\mathbf{1.0099 \pm 0.0029}$	1.0142 ± 0.0018	1.3840 ± 0.0289	1.0179 ± 0.0059
	96	$\mathbf{1.3518 \pm 0.0102}$	1.3592 ± 0.0109	1.3646 ± 0.0020	1.9621 ± 0.0071	1.3563 ± 0.0044
Traffic	24	$\mathbf{0.3664 \pm 0.0001}$	0.3666 ± 0.0004	0.3725 ± 0.0004	0.3975 ± 0.0003	0.3711 ± 0.0002
	48	0.3846 ± 0.0005	$\mathbf{0.3845 \pm 0.0003}$	0.3931 ± 0.0004	0.4298 ± 0.0001	0.3916 ± 0.0002
	96	0.3981 ± 0.0001	$\mathbf{0.3978 \pm 0.0002}$	0.3983 ± 0.0003	0.4516 ± 0.0001	0.4081 ± 0.0003
Average rank ↓		**2.07**	2.47	3.33	4.60	2.53

2021), and Traffic (Lai et al., 2018). Each dataset is evaluated at three forecast horizons. Dataset statistics, lookback lengths, update intervals, chronological splits and hyperparameter details are reported in Appendix C. For every dataset, horizon, and seed, the PatchTST component of each method is initialized from the same pretrained backbone checkpoint. Method-specific auxiliary components, such as the DSOF student, are initialized according to their official implementations. We report global test MSE in coordinates standardized using fixed pretraining statistics. Results are reported as mean ± standard deviation over seeds 2021–2023.

Results. Table 2 shows that HAP achieves the lowest MSE on 6 of the 15 dataset–horizon tasks and the best overall average rank of 2.07, ahead of SOLID (2.47) and DynaME (2.53). HAP is particularly competitive on Electricity, Weather, and Traffic, where it attains or closely matches the best performance across horizons, while results on ETTm1 and ETTh2 are more mixed. Overall, these results indicate that the additional caution introduced by hedged deployment and delayed promotion does not come at the cost of systematically degraded forecasting accuracy.

6 Conclusion

We formulated online forecasting under persistent–transient ambiguity as a deployment decision problem. Our method *Hedge and Promote* (HAP) implements this distinction through a two-model batch-incremental cycle. It combines frozen incumbent and candidate predictions through delayed-feedback Hedge, then uses a one-sided DM/HAC comparison to decide promotion from subsequent forecasting evidence. Our analysis establishes the unavoidable worst-case cost of immediate commitment under indistinguishable continuations and a pathwise within-cycle regret bound for hedged deployment. Across five popular forecasting datasets, HAP achieves the lowest mean MSE on more tasks. The paired ETTm1 case study further shows lower mean transient exposure and persistent delay than baselines, supporting the value of cautious deployment beyond predication MSE.

Limitations and perspective. The statistical interpretation of promotion relies on local comparison regularity, and no finite amount of past evidence can guarantee superiority after an unforeseen regime reversal. This is an intrinsic limitation of forecasting under concept drift rather than a peculiarity of HAP: future regimes may invalidate any previously supported deployment decision. HAP addresses this uncertainty operationally by treating promotion as local and revisable, and by repeating the deployment–promotion cycle as new evidence arrives. In this sense, following Read & Zliobaite (2025), our goal is not to eliminate non-stationarity but to make adaptation more cautious, reliable, and evidence-driven by separating the opportunity to adapt from the decision to retain an adaptation. Our semi-synthetic analysis is currently limited to one real-world dataset and one intervention family; broader evaluation across datasets and perturbation mechanisms would further assess robustness under persistent–transient ambiguity.

AI USE STATEMENT

In this work, we used generative AI tools for Generate semi-synthetic data sets, provide critical ingredients for proving mathematical claims, assist in the writing of proofs, propose or refine hypotheses and implement methods. We have not used generative AI tools for help develop theoretical models or conceptual frameworks, formulate mathematical claims, design or provide feedback on research methodology or experiments, clean and reformat dataset, interpret results, and assist with translation, support qualitative and thematic data analysis are not applicable to this work. Additionally, we used generative AI tools for create or modify scientific figures or images, summarize or analyse existing literature, identify relevant literature, format references. We have reviewed all AI-assisted work. LLM-generated code was verified and tested for correctness. We take responsibility for the final content of this work, including text, claims or artifacts produced with the aid of generative AI.

REPRODUCIBILITY STATEMENT

We provide an anonymous supplementary code archive for reproducibility. Appendix C documents the experimental details.

REFERENCES

Mouxiang Chen, Lefei Shen, Han Fu, Zhuo Li, Jianling Sun, and Chenghao Liu. Calibration of time-series forecasting: Detecting and adapting context-driven distribution shift. In *Proceedings of the 30th ACM SIGKDD Conference on Knowledge Discovery and Data Mining*, pp. 341–352, 2024. doi: 10.1145/3637528.3671926.

Hao Cheng, Qingsong Wen, Yang Liu, and Liang Sun. Robusttsf: Towards theory and design of robust time series forecasting with anomalies. In *International Conference on Learning Representations*, 2024.

Francis X. Diebold and Roberto S. Mariano. Comparing predictive accuracy. *Journal of Business & Economic Statistics*, 13(3):253–263, 1995. doi: 10.1080/07350015.1995.10524599.

Yoav Freund and Robert E Schapire. A decision-theoretic generalization of on-line learning and an application to boosting. *Journal of computer and system sciences*, 55(1):119–139, 1997.

Seungha Hong, Sukang Chae, Suyeon Kim, Sanghwan Jang, and Hwanjo Yu. Dynamic multi-period experts for online time series forecasting. In *Proceedings of the ACM Web Conference 2026*, 2026. doi: 10.1145/3774904.3792731.

Xiannan Huang, Shuhan Qiu, Jiayuan Du, and Chao Yang. Online time series prediction using feature adjustment. In *International Conference on Learning Representations*, 2026.

Guokun Lai, Wei-Cheng Chang, Yiming Yang, and Hanxiao Liu. Modeling long- and short-term temporal patterns with deep neural networks. In *Proceedings of the 41st International ACM SIGIR Conference on Research and Development in Information Retrieval*, pp. 95–104, 2018.

Ying-yee Ava Lau, Zhiwen Shao, and Dit-Yan Yeung. Fast and slow streams for online time series forecasting without information leakage. In *International Conference on Learning Representations*, 2025.

Mengchu Li and Yi Yu. Adversarially robust change point detection. In *Advances in Neural Information Processing Systems*, volume 34, 2021.

Xiao Lin, Zhining Liu, Dongqi Fu, Ruizhong Qiu, and Hanghang Tong. Backtime: Backdoor attacks on multivariate time series forecasting. In *Advances in Neural Information Processing Systems*, volume 37, 2024.

Whitney K. Newey and Kenneth D. West. A simple, positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix. *Econometrica*, 55(3):703–708, 1987.

Whitney K. Newey and Kenneth D. West. Automatic lag selection in covariance matrix estimation. *The Review of Economic Studies*, 61(4):631–653, 1994. doi: 10.2307/2297912.

Yuqi Nie, Nam H Nguyen, Phanwadee Sinthong, and Jayant Kalagnanam. A time series is worth 64 words: Long-term forecasting with transformers. In *The Eleventh International Conference on Learning Representations*, 2023.

Quang Pham, Chenghao Liu, Doyen Sahoo, and Steven C. H. Hoi. Learning fast and slow for online time series forecasting. In *International Conference on Learning Representations*, 2023.

Jesse Read and Indre Zliobaite. Supervised learning from data streams: An overview and update. *ACM Computing Surveys*, 57(12), 2025. ISSN 1557-7341.

Abishek Sankararaman and Balakrishnan Narayanaswamy. Online robust non-stationary estimation. In *Advances in Neural Information Processing Systems*, volume 36, 2023.

Artur Trindade. ElectricityLoadDiagrams20112014, 2015. URL `https://doi.org/10.24432/C58C86`. Dataset.

Alexey Tsymbal. The problem of concept drift: Definitions and related work. Technical Report TCD-CS-2004-15, Department of Computer Science, Trinity College Dublin, 2004.

Haixu Wu, Jiehui Xu, Jianmin Wang, and Mingsheng Long. Autoformer: Decomposition transformers with auto-correlation for long-term series forecasting. In *Advances in Neural Information Processing Systems*, volume 34, pp. 22419–22430, 2021.

Yifan Zhang, Qingsong Wen, Xue Wang, Weiqi Chen, Liang Sun, Zhang Zhang, Liang Wang, Rong Jin, and Tieniu Tan. Onenet: Enhancing time series forecasting models under concept drift by online ensembling. In *Advances in Neural Information Processing Systems*, volume 36, pp. 69949–69980, 2023.

Lifan Zhao and Yanyan Shen. Proactive model adaptation against concept drift for online time series forecasting. In *Proceedings of the 31st ACM SIGKDD Conference on Knowledge Discovery and Data Mining V.1*, pp. 2020–2031, 2025. doi: 10.1145/3690624.3709210.

Haoyi Zhou, Shanghang Zhang, Jieqi Peng, Shuai Zhang, Jianxin Li, Hui Xiong, and Wancai Zhang. Informer: Beyond efficient transformer for long sequence time-series forecasting. In *Proceedings of the Thirty-Fifth AAAI Conference on Artificial Intelligence (AAAI-21)*, pp. 11106–11115, 2021.

A HAC VARIANCE ESTIMATION AND BANDWIDTH SELECTION

We provide here the complete specification of the HAC estimator and bandwidth selection used in Section 4.2. For the ordered paired loss differences $Z_1, \ldots, Z_n$, let

$$\overline{Z} = \frac{1}{n}\sum_{i=1}^{n} Z_i, \qquad \widehat{\gamma}_k = \frac{1}{n}\sum_{i=k+1}^{n} (Z_i - \overline{Z})(Z_{i-k} - \overline{Z}), \tag{14}$$

the corresponding long-run variance is

$$\Omega = \gamma_0 + 2\sum_{k\geq 1} \gamma_k.$$

We estimate it using the Bartlett/Newey–West estimator

$$\widehat{\Omega}(q) = \widehat{\gamma}_0 + 2\sum_{k=1}^{q} \left(1 - \frac{k}{q+1}\right) \widehat{\gamma}_k. \tag{15}$$

The Bartlett weights yield a positive semi-definite spectral estimate and downweight increasingly distant sample autocovariances (Newey & West, 1987).

Automatic bandwidth selection. Rather than fixing the truncation lag q, we use the plug-in procedure of Newey & West (1994). The pilot lag is

$$m = \min\left\{n-1, \left\lfloor 4\left(\frac{n}{100}\right)^{2/9}\right\rfloor\right\}. \tag{16}$$

Using the corresponding sample autocovariances, define

$$\widehat{s}_0 = \widehat{\gamma}_0 + 2\sum_{k=1}^{m}\widehat{\gamma}_k, \qquad \widehat{s}_1 = 2\sum_{k=1}^{m} k\widehat{\gamma}_k. \tag{17}$$

When $\widehat{s}_0 > 0$, the plug-in bandwidth is

$$q_{\mathrm{NW}} = \left\lfloor 1.1447\left[n\left(\frac{\widehat{s}_1}{\widehat{s}_0}\right)^2\right]^{1/3}\right\rfloor. \tag{18}$$

If $\widehat{s}_0 \le 0$, we use the fallback $q_{\mathrm{NW}} = m$.

Because consecutive H-step forecasting targets overlap, dependence up to lag $H-1$ is structurally induced even in the absence of additional serial dependence. We therefore use

$$q = \min\{n-1, \max(H-1, q_{\mathrm{NW}})\}. \tag{19}$$

The lower bound ensures that the known target-overlap dependence is not ignored, while the upper bound restricts the truncation lag to the available comparison batch.

Finally, the statistic and confidence bound used for promotion are

$$T_{\mathrm{DM}} = \frac{\sqrt{n}\,\overline{Z}}{\sqrt{\widehat{\Omega}(q)}}, \qquad U_\mu = \overline{Z} + z_{1-\delta}\sqrt{\frac{\widehat{\Omega}(q)}{n}}, \tag{20}$$

and the candidate is promoted whenever $U_\mu < 0$.

B Proofs of the theoretical results

B.1 Proof of Proposition 1

Because the two continuations share the same $\mathcal{O}_b$, any randomized rule based only on $\mathcal{O}_b$ promotes with the same probability $p \in [0,1]$ in both worlds. In world P, retaining the incumbent incurs excess risk g_{P} and promoting incurs zero excess risk, hence $\mathcal{R}_{\mathsf{P}}(p) = (1-p)g_{\mathsf{P}}$. In world T, promotion incurs excess risk g_{T} and retention incurs zero excess risk, hence $\mathcal{R}_{\mathsf{T}}(p) = pg_{\mathsf{T}}$. The minimax randomized rule equalizes the two quantities,

$$(1-p)g_{\mathsf{P}} = pg_{\mathsf{T}},$$

which gives $p = g_{\mathsf{P}}/(g_{\mathsf{P}} + g_{\mathsf{T}})$ and the value in equation 7. Restricting p to $\{0,1\}$ gives $\min\{g_{\mathsf{P}}, g_{\mathsf{T}}\}$.

B.2 Proof of Theorem 1

Proof. Fix a realized comparison cycle. Write $S_0 = 0$ and $S_i = \sum_{s=1}^{i} Z_s$, and introduce the hypothetical immediate-feedback weights

$$\widetilde{\lambda}_i \coloneqq \sigma(-\eta S_{i-1}).$$

These weights use every preceding loss, whereas the deployed weights use only losses whose forecast horizons have elapsed.

1. Separate the mixing gain from the weight updates. Under normalized squared loss, prediction mixing gives the exact identity

$$\ell_i^{\mathrm{dep}} = \ell_i^{\mathrm{I}} + \lambda_i Z_i - \lambda_i(1-\lambda_i)\Gamma_i. \tag{21}$$

Let $G_n \coloneqq \sum_{i=1}^n \lambda_i(1-\lambda_i)\Gamma_i \geq 0$. The better fixed expert has cumulative loss $\sum_{i=1}^n \ell_i^{\mathrm{I}} + \min\{0, S_n\}$. Consequently,

$$\begin{aligned} R_{b+1}^{\mathrm{pair}} &= \underbrace{\sum_{i=1}^n \widetilde{\lambda}_i Z_i - \min\{0, S_n\}}_{\text{immediate-feedback regret}} \\ &\quad + \underbrace{\sum_{i=1}^n (\lambda_i - \widetilde{\lambda}_i) Z_i}_{\text{delay effect}} - \underbrace{G_n}_{\text{mixing gain}}. \end{aligned} \tag{22}$$

It remains to bound the first two terms.

2. Bound the immediate-feedback regret. Consider the relative exponential-weights potential

$$W_i \coloneqq \frac{1 + e^{-\eta S_{i-1}}}{2}, \qquad \frac{W_{i+1}}{W_i} = (1 - \widetilde{\lambda}_i) + \widetilde{\lambda}_i e^{-\eta Z_i}.$$

Hoeffding's lemma applied to the two-point losses 0 and Z_i, whose range is $|Z_i|$, yields

$$\widetilde{\lambda}_i Z_i \leq -\frac{1}{\eta} \log \frac{W_{i+1}}{W_i} + \frac{\eta}{8} Z_i^2.$$

Summing over i, using $W_1 = 1$ and $W_{n+1} \geq \frac{1}{2} e^{-\eta \min\{0, S_n\}}$, gives

$$\sum_{i=1}^n \widetilde{\lambda}_i Z_i - \min\{0, S_n\} \leq \frac{\log 2}{\eta} + \frac{\eta}{8} V_n. \tag{23}$$

This argument uses the realized range $|Z_i|$ at each origin, not a uniform bound on the expert losses.

3. Bound the effect of unresolved feedback. At origin i, the losses available to immediate-feedback Hedge but not to delayed Hedge are indexed by

$$\mathcal{D}_i \coloneqq \{s : 1 \leq s < i,\ i - s \leq H - 1\}.$$

The two logits therefore differ by $\eta \sum_{s \in \mathcal{D}_i} Z_s$. Since the logistic function is $1/4$-Lipschitz,

$$\begin{aligned} (\lambda_i - \widetilde{\lambda}_i) Z_i &\leq |Z_i|\, |\lambda_i - \widetilde{\lambda}_i| \\ &\leq \frac{\eta}{4} |Z_i| \sum_{s \in \mathcal{D}_i} |Z_s|. \end{aligned} \tag{24}$$

Combining this with equation 22 and equation 23 already establishes the sharper bound

$$\begin{aligned} R_{b+1}^{\mathrm{pair}} &\leq \frac{\log 2}{\eta} + \frac{\eta}{8} V_n \\ &\quad + \frac{\eta}{4} \sum_{i=1}^n |Z_i| \sum_{s \in \mathcal{D}_i} |Z_s| - G_n. \end{aligned} \tag{25}$$

To obtain the stated bound, use $|Z_i Z_s| \leq (Z_i^2 + Z_s^2)/2$:

$$\begin{aligned} \sum_{i=1}^n \sum_{s \in \mathcal{D}_i} |Z_i Z_s| &\leq \frac{1}{2} \sum_{i=1}^n \sum_{s \in \mathcal{D}_i} (Z_i^2 + Z_s^2) \\ &\leq (H-1) V_n. \end{aligned} \tag{26}$$

The last inequality holds because each index participates in at most $2(H-1)$ such pairs. Substituting into equation 25 proves equation 10. □

Proof of Corollary 1. Substitute $V_n \leq v^2 n$ and $\eta = \eta_0 / \sqrt{(2H-1)n}$ into equation 10. Dropping the nonpositive mixing term and taking the positive part gives equation 11. The rate $O(\sqrt{n})$ follows directly. □

C Experimental Details

C.1 Implementation and hyperparameter selection

We use the OnlineTSF implementations of SOLID and Proceed,[2] and the authors' implementations of DSOF[3] and DynaME.[4] Hyperparameters are selected independently for each dataset, horizon, and seed using only the validation stream. For HAP, we evaluate

$$E \in \{3, 5, 10, 20\}, \qquad \rho_{\mathrm{lr}} \in \{1, 3, 10\}, \tag{27}$$

where E is the number of candidate-training epochs, ρ_{lr} multiplies the dataset-specific base learning rate in OnlineTSF. Across all experiments, the one-sided DM/HAC significance level is fixed to $\delta_{\mathrm{DM}} = 0.20$ and is not included in the validation search. For each competitor, we retain the official method-specific configuration and tune only an online learning-rate multiplier over $0.3, 1, 3$

C.2 Datasets and forecasting configurations

All datasets are divided chronologically without shuffling, and use the first 20% of timestamps for pretraining, the following 5% for validation, and the remaining 75% for online testing. Table 3 summarizes the forecasting configuration. Here, C is the number of target channels, L is the lookback length, H is the forecast horizon, and U is the HAP update interval. The final column gives its corresponding physical duration. The ETTm1-PT intervention experiment uses $(L, H) = (336, 48)$.

Table 3: Dataset and forecasting configurations.

Dataset	Sampling	C	L	H	U	Physical duration
ETTm1	15 min	7	336	$\{24, 48, 96\}$	1,440	15 days
ETTm1-PT	15 min	7	336	48	1,440	15 days
ETTh2	1 hour	7	336	$\{24, 48, 96\}$	720	30 days
Electricity	1 hour	321	336	$\{24, 48, 96\}$	672	28 days
Weather	10 min	21	336	$\{24, 48, 96\}$	1,008	7 days
Traffic	1 hour	862	336	$\{24, 48, 96\}$	672	28 days

C.3 ETTm1-PT dataset

We modify ETTm1's six daily load profile in standardised coordinates, leaving oil temperature unchanged:

$$z^{w}_{t,j} = z^{\mathrm{ETT}}_{t,j} + \Delta\, e^{w}_{t}\, g(h_t), \qquad j \in \mathcal{S}_{\mathrm{load}}, \qquad \Delta \in \{3, 6, 9\}. \tag{28}$$

where z^{ETT}_t denotes the standardized ETTm1 background and $z_t{}^{w}$ is the resulting process under continuation w.

Here, $h_t \in [0, 24)$ is the hour of day, $e^{w}_{t} \in \{0, 1\}$ indicates whether the intervention is active and Δ controls the amplitude. The fixed daily pattern

$$g(h) = \begin{cases} \sin^2\!\left(\dfrac{\pi(h-4)}{8}\right), & 4 \le h < 12, \\ -\sin^2\!\left(\dfrac{\pi(h-14)}{8}\right), & 14 \le h < 22, \\ 0, & \text{otherwise} \end{cases} \tag{29}$$

increases the morning profile and decreases the evening profile, with equal and opposite daily contributions. See Figure 4 (Left) for illustration using HUFL load. At $\Delta = 3$, the intervention represents a domain-motivated reshaping of ETTm1's daily load profile, smoothly increasing morning loads and reducing evening loads while preserving the daily mean and background fluctuations.

[2] https://github.com/SJTU-DMTai/OnlineTSF

[3] https://github.com/yyalau/iclr2025_dsof

[4] https://github.com/shhong97/DynaME

$\Delta = 9$ corresponds to a severe anomalous perturbation, but not an adversarial attack designed to manipulate forecasters. After a shared 15-day modified history (11 June 2017 – 26 June 2017), the intervention then ends ($w = \mathrm{T}$) or persists ($w = \mathrm{P}$). Evaluation covers 30 days (26 June 2017 – 25 July 2017) after shared history.

D Additional results

D.1 Per-magnitude PT Average

See Table 4.

Table 4: PT Average for each intervention magnitude. Values are mean ± standard deviation over three seeds.

Method	$\Delta = 3$	$\Delta = 6$	$\Delta = 9$
HAP (ours)	0.6283 ± 0.0213	$\mathbf{0.7254 \pm 0.0156}$	$\mathbf{0.8884 \pm 0.0092}$
SOLID	0.7268 ± 0.0097	0.9975 ± 0.0120	1.3970 ± 0.0145
Proceed	$\mathbf{0.6112 \pm 0.0065}$	0.7614 ± 0.0091	0.9957 ± 0.0088
DSOF	0.7493 ± 0.0130	0.8718 ± 0.0371	1.0223 ± 0.0567
DynaME	0.6301 ± 0.0040	0.7824 ± 0.0116	1.0731 ± 0.0277

D.2 Operational analysis

Figure 5 shows that HAP achieves the lowest mean transient exposure and persistent delay at every tested magnitude. Its exposure and delay grow more slowly than that of the baselines as the intervention becomes stronger. In this case study, HAP limits the costs of persistence-transient ambiguity in both continuations.

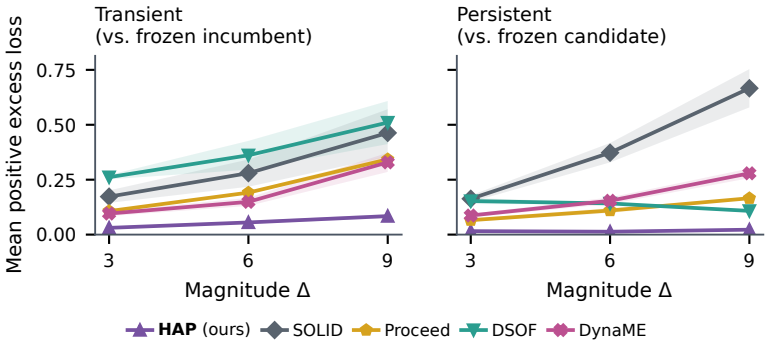

Figure 5: Mean positive excess loss in the transient and persistent continuations with standard deviation across seeds.

Figure 6 confirms the opposite model preferences underlying this case study: over the 30-day evaluation window, the frozen incumbent has lower MSE in all nine transient seed–magnitude scenarios, whereas the frozen candidate is preferable in all nine persistent scenarios.

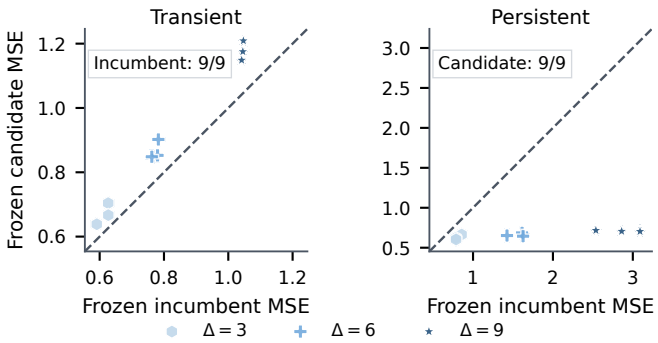

Figure 6: Frozen-reference MSE ordering. The diagonal denotes equal incumbent and candidate error.