# WHEN TOMORROW BECOMES TODAY: SELF-EVOLVING POLICIES FOR AGENTIC TIME-SERIES FORECASTING

**Anonymous authors**
Paper under double-blind review

## ABSTRACT

Agentic time series forecasting concerns systems whose underlying mechanisms evolve, making the relative effectiveness of numerical models, reasoning strategies, and intervention rules inherently time-varying. Consequently, a time series agent must adapt the forecasts it produces and the orchestration policy that determines which components to trust and how to coordinate them. The deployment process naturally provides supervision for this adaptation as forecast horizons elapse and realized targets reveal the effectiveness of earlier decisions. Committing all numerical expert forecasts and candidate agent paths before target observation allows each realized outcome to evaluate the entire alternative set, providing delayed feedback without additional annotation. However, existing time series agents primarily incorporate prior experience through forecast refinement, reflection, or retrieval, without systematically converting realized outcomes into persistent updates to the joint orchestration policy governing later origins. To exploit this delayed feedback systematically, we introduce TIMEVOLVE, a frozen-backbone time series agent that converts each realized outcome into persistent joint updates of expert trust, agent path selection, and intervention strength. A temporally ordered predict, reveal, and update protocol applies this feedback to subsequent forecasts. Experiments across eight Time-MMD domains show that TIMEVOLVE achieves the best average MSE and MAE ranks among fifteen methods and the lowest errors on both metrics in seven domains. These results demonstrate the value of learning forecasting policies from the futures encountered during deployment. The code is available in the supplementary material.

## 1 INTRODUCTION

Agentic time series forecasting addresses systems whose underlying mechanisms evolve as trends drift, regimes shift, and external events alter system dynamics (Chang et al., 2026; Zhao et al., 2025). Consequently, the relative effectiveness of numerical models, reasoning strategies, and intervention rules is inherently time-varying, and a component that supports accurate forecasting under one set of temporal conditions can become unreliable under another (Cheng et al., 2026; Kim et al., 2025). Therefore, a time series agent must adapt not only the forecasts it produces, but also the orchestration policy that determines which components to trust and how to coordinate them (Hu et al., 2026).

The deployment process naturally provides supervision for this adaptation. As forecast horizons elapse, realized targets reveal the errors of earlier forecasts, supplying delayed feedback for online learning (Joulani et al., 2013; Zhang et al., 2023). In sequential decision-making, evaluating alternative policies from logged interaction trajectories is an off-policy estimation problem (Jiang & Li, 2016). Non-performative forecasting instead permits direct comparison of precommitted forecasts against a shared outcome, following the full-information setting of prediction with expert advice (Cesa-Bianchi & Lugosi, 2006). We extend this feedback structure to numerical experts and candidate agent paths fixed before reveal, where each *agent path* is an evidence-grounded plan specifying expert selection and optional corrections to produce a candidate forecast. Each reveal provides task-native supervision for both the selected and unselected alternatives, enabling prior forecasting decisions to be evaluated directly from observed outcomes. When tomorrow becomes today, the environment evaluates both what the agent predicted and how that prediction was constructed.



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.

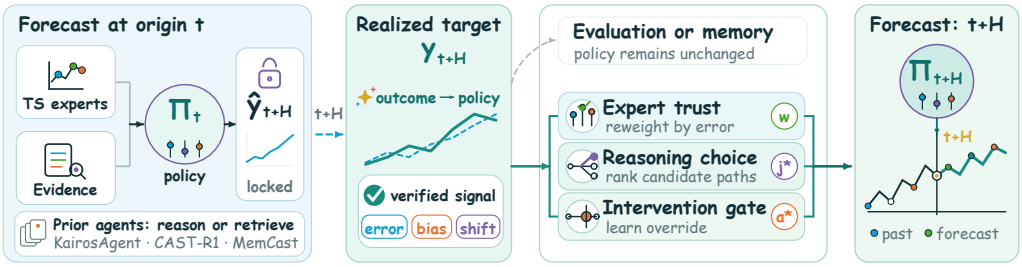


Figure 1: When outcomes become supervision, forecast evaluation gives way to policy evolution. Before observing the target, a time series agent commits its forecast and decision trace. Evaluation or passive memory leaves the orchestration policy unchanged, whereas TIMEVOLVE uses the realized outcome to update expert trust, agent path selection, and intervention strength for subsequent forecasts.

Recent work has expanded time series analysis and forecasting into agentic workflows (Cheng et al., 2026) involving tool use (Wu et al., 2026; Tao et al., 2026c), learned sequential decision policies (Tao et al., 2026b), semantic reasoning (Zhou et al., 2026; Guan et al., 2025), specialized numerical forecasters (Feng et al., 2026), iterative forecast refinement (Liao et al., 2026), reflection (Wang et al., 2024), and memory (Tao et al., 2026a; Lyu et al., 2026). As shown in Fig. 1, these systems improve current forecast construction and incorporate prior experience through refinement, reflection, or retrieval. However, they do not systematically convert realized deployment outcomes into persistent updates to an explicit joint orchestration policy governing later forecasting decisions. *This creates a mismatch between the time-varying effectiveness of individual components and the comparatively static rule used to coordinate models, evidence, reasoning, and intervention.* Revealed errors can thus inform retrieval or local forecast refinement without directly recalibrating the persistent decision rules used in subsequent forecasts.

Addressing this orchestration gap requires adapting three coupled decisions. ❶ The agent must determine which numerical experts to trust, since their relative reliability changes as temporal conditions evolve. ❷ It must select an evidence-grounded agent path, because alternative interpretations of shared numerical and contextual evidence can favor different expert subsets and corrections. ❸ It must determine how strongly the selected path should modify the numerical prior, balancing useful corrections against errors introduced by weak or misinterpreted evidence. These decisions are *interdependent* because expert trust influences both the numerical prior and the candidate forecasts, path selection determines the proposed correction relative to the prior, while intervention strength controls the extent to which this correction is included in the final forecast. Revising expert trust can therefore change which path is most useful and how strongly it should be adopted, while selecting a different path changes the required calibration of the correction. These dependencies necessitate a joint orchestration policy that uses the same realized outcome to inform coordinated updates to expert trust, agent path selection, and intervention strength, allowing subsequent decisions to adapt to changes in the reliability of both components and their interactions.

To realize this joint update, we introduce TIMEVOLVE, a frozen-backbone time series agent with three complementary modules that convert delayed outcome feedback into persistent updates to the corresponding decision rules. EvolveTrust updates the contribution of numerical experts to the time series prior, EvolveReason updates the preference over evidence-grounded agent paths, and EvolveIntervene updates how strongly the selected path modifies the numerical prior. Before observing the target, TIMEVOLVE commits all numerical expert forecasts, candidate paths, selected decisions, and the final forecast. Once the forecast horizon elapses, the realized target provides expert errors, candidate forecast errors, and a target for calibrating the selected path's influence on the numerical prior. The predict, reveal, and update protocol applies these updates to subsequent eligible forecasts. The language model, numerical forecasters, analytical tools, and prompts remain frozen while the orchestration policy evolves. Across eight Time-MMD domains, TIMEVOLVE achieves the best average MSE and MAE ranks against agentic systems, foundation models, and full-shot forecasting models. Our contributions are as follows.

- **Delayed-Feedback Learning.** We formulate deployment forecasting as policy learning from delayed outcome feedback and develop a temporally valid predict, reveal, and update protocol.

Each realized target scores the precommitted alternatives, and the resulting policy updates govern subsequent forecasts.

- **Policy Evolution.** We introduce TIMEVOLVE, which converts realized outcome feedback into persistent, coordinated updates to expert trust, agent path selection, and intervention strength while keeping the language model, numerical forecasters, analytical tools, and prompts frozen.
- **Empirical Validation.** Extensive experiments on eight Time-MMD domains show that TIMEVOLVE achieves the best average rank under both MSE and MAE and improves over its fixed-policy counterpart on both metrics in every domain.

## 2 Related Work

### 2.1 Agentic time series forecasting

Agentic time series forecasting extends model-centric prediction into a workflow that coordinates numerical forecasters, analytical tools, contextual evidence, semantic reasoning, iterative refinement, reflection, and memory (Cheng et al., 2026). Cast-R1 learns tool-augmented sequential decision policies through supervised and reinforcement learning (Tao et al., 2026b). KairosAgent integrates tool-grounded semantic reasoning with a time series foundation model and optimizes its reasoner using forecasting-oriented objectives (Feng et al., 2026). Last-mile agents use weakly structured contextual evidence to revise numerical forecasts and retain post-hoc reflections for later use (Liao et al., 2026). These methods broaden time series forecasting from a numerical mapping into an agentic process that selects tools, interprets evidence, invokes specialized forecasters, and refines candidate predictions. These approaches develop forecast construction and experience use. TIMEVOLVE makes the persistent orchestration policy the unit of adaptation, using realized outcomes to update expert trust, agent path selection, and intervention strength across deployment.

### 2.2 Outcome-driven adaptation

Learning from revealed outcomes is central to online prediction with expert advice, where observed losses update the relative influence of competing experts (Freund & Schapire, 1997; Cesa-Bianchi & Lugosi, 2006). OneNet addresses concept drift in time series forecasting through online ensembling (Zhang et al., 2023). MoE-F develops online gating for mixtures of language model experts in time series prediction (Saqur et al., 2025). These methods use outcome feedback to adapt expert allocation. A complementary line of work incorporates feedback through agent memory. Reflexion converts environmental feedback into textual experience that guides subsequent attempts (Shinn et al., 2023). In time series forecasting, MemCast constructs hierarchical experience memory and adapts memory-entry confidence during inference, allowing retrieved experience to influence trajectory selection and reflection (Tao et al., 2026a). TIMEVOLVE combines these adaptation targets in a joint orchestration policy. Each realized outcome scores the precommitted numerical expert forecasts and candidate agent paths and provides an intervention target, supervising expert trust, agent path selection, and intervention strength together.

## 3 Preliminary: Delayed-Feedback Policy Learning

We consider a scalar target series $\{x_1, \ldots, x_T\}$ and a fixed forecast horizon $H$. At forecasting origin $t$, the history $x_{1:t}$ and time-aligned context $c_t$ are available, while the target window is

$$\boldsymbol{y}_t = [x_{t+1}, \ldots, x_{t+H}]^\top \in \mathbb{R}^H. \tag{1}$$

Both $\hat{\boldsymbol{y}}_t$ and $\boldsymbol{y}_t$ use the forecast issue time $t$ as their index. The information set $\mathcal{I}_t$ contains the available history and context, fixed predeployment calibration artifacts, and observed outcomes of windows completed by $t$. A forecast issued at $s$ supplies feedback when $s + H \leq t$. Overlapping forecasts remain outstanding until their corresponding target windows have elapsed.

At reveal, let $\Omega_s \subseteq \{1, \ldots, H\}$ contain the observed coordinates of $\boldsymbol{y}_s$, and let $m_s = |\Omega_s|$. For $m_s > 0$, the masked inner product and MSE are

$$\langle \boldsymbol{u}, \boldsymbol{v} \rangle_{\Omega_s} = \sum_{h \in \Omega_s} u_h v_h, \qquad \mathrm{MSE}_{\Omega_s}(\boldsymbol{u}, \boldsymbol{v}) = \frac{\|\boldsymbol{u} - \boldsymbol{v}\|^2_{\Omega_s}}{m_s}, \tag{2}$$

where $\|\boldsymbol{u}\|^2_{\Omega_s} = \langle \boldsymbol{u}, \boldsymbol{u} \rangle_{\Omega_s}$. These observed coordinates define all target-dependent losses; windows with $m_s = 0$ leave the state unchanged. Pre-reveal features are computed from $\mathcal{I}_t$.

Let $\mathcal{K}_t \subseteq \{1, \ldots, K\}$ index the active numerical experts and $\mathcal{J}_t \subseteq \{1, \ldots, J\}$ the valid agent paths. They produce indexed forecast collections $\mathcal{F}_t = \{\hat{\boldsymbol{y}}_t^{(k)} : k \in \mathcal{K}_t\}$ and $A_t = \{\boldsymbol{a}_t^{(j)} : j \in \mathcal{J}_t\}$, with every curve in $\mathbb{R}^H$. A path is a structured decision whose execution produces $\boldsymbol{a}_t^{(j)}$; its index $j$ is local to that origin. The learner state is

$$\boldsymbol{\Pi}_t = \left(\boldsymbol{\Pi}_t^{\text{trust}}, \boldsymbol{\Pi}_t^{\text{reason}}, \boldsymbol{\Pi}_t^{\text{intervene}}\right), \qquad \hat{\boldsymbol{y}}_t = G(\mathcal{I}_t, \mathcal{F}_t, A_t; \boldsymbol{\Pi}_t). \tag{3}$$

Each component stores decision parameters and update statistics. In particular, $\boldsymbol{\Pi}_t^{\text{reason}}$ contains the selector coefficients $\boldsymbol{\theta}_t$, and $\boldsymbol{\Pi}_t^{\text{intervene}}$ contains the gate coefficients $\boldsymbol{\omega}_t$ shown in Fig. 2. The fixed map $G$ constructs and combines forecasts under this state.

Before observing $\boldsymbol{y}_s$, the agent stores its candidates, features, and decisions in a record $\mathcal{R}_s$. At $s + H$, its feedback updates the latest active state,

$$\boldsymbol{\Pi}_{s+H} = U\left(\boldsymbol{\Pi}_{s+H}^-, \mathcal{R}_s, \boldsymbol{y}_s, \Omega_s\right). \tag{4}$$

The state $\boldsymbol{\Pi}_{s+H}^-$ includes updates from other completed windows and immediately precedes this reveal. The learner processes feedback before issuing the forecast at $s + H$, with earlier forecasts already committed. Targets are unaffected by forecast selection, so all stored alternatives share the same observed outcome.

# 4 METHOD

## 4.1 OVERVIEW

TIMEVOLVE implements the loop in Fig. 2 through three linked decisions. EvolveTrust forms expert weights $\boldsymbol{p}_t$ and the numerical prior $\boldsymbol{b}_t$. A frozen language model proposes structured paths, whose execution under these weights produces $A_t$. EvolveReason selects $j_t$ and hence $\boldsymbol{a}_t = \boldsymbol{a}_t^{(j_t)}$, and EvolveIntervene chooses its influence $\alpha_t$ on the prior. Expert trust therefore affects both the prior and the candidates, while the selected path determines the correction calibrated by the gate. Revealed outcomes update these linked decision rules while the forecasting components remain frozen.

## 4.2 EVIDENCE-GROUNDED FORECAST CONSTRUCTION

**Forecasts and evidence.** The FORECAST-PACK supports frozen time series foundation models and statistical filters for level, trend, seasonality, and changing dynamics. The main evaluation uses a foundation-model-augmented numerical tool pool, with the same frozen forecasters shared across TIMEVOLVE and its policy ablations (App. C). Active experts produce aligned $H$-step forecasts from the available history. For each expert, $\boldsymbol{e}_{t,k}$ summarizes past reliability, current forecast plausibility, and agreement with other sources. Fixed analytical tools extract temporal structure from the history, while $c_t$ supplies reports and search evidence available by $t$. Numerical diagnostics characterize each forecast relative to the observed temporal pattern, while contextual evidence informs expert selection and local corrections. App. D details these diagnostics.

**Structured paths.** Each proposed path specifies an evidence interpretation, an expert subset $S_t^{(j)} \subseteq \mathcal{K}_t$, and an optional correction $\boldsymbol{\Delta}_t^{(j)} \in \mathbb{R}^H$. Current trust weights define the initial mixture

$$u_{t,k}^{(j)} = \frac{p_{t,k} \mathbb{1}[k \in S_t^{(j)}]}{\sum_{q \in S_t^{(j)}} p_{t,q}}. \tag{5}$$

Valid paths select nonempty subsets with positive trust mass. The configured source-budget projection yields normalized weights $\tilde{u}_{t,k}^{(j)}$ supported on the selected subset, limiting the influence of correlated expert families. The runtime then constructs

$$\boldsymbol{a}_t^{(j)} = \sum_{k \in \mathcal{K}_t} \tilde{u}_{t,k}^{(j)} \hat{\boldsymbol{y}}_t^{(k)} + \boldsymbol{\Delta}_t^{(j)}. \tag{6}$$

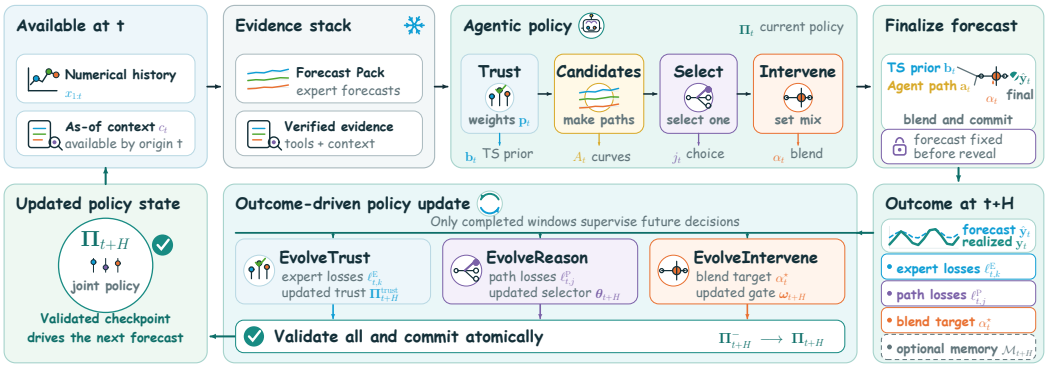


Figure 2: **TIMEVOLVE predict, reveal, and update pipeline.** History $x_{1:t}$ and context $c_t$ produce expert forecasts and the candidate set $A_t$. Expert weights $\boldsymbol{p}_t$ form the prior $\boldsymbol{b}_t$, selection $j_t$ yields $\boldsymbol{a}_t = \boldsymbol{a}_t^{(j_t)}$, and the gate commits $\hat{\boldsymbol{y}}_t = (1-\alpha_t)\boldsymbol{b}_t + \alpha_t \boldsymbol{a}_t$. At $t+H$, observed coordinates of $\boldsymbol{y}_t$ yield expert losses $\ell^{\mathrm{E}}_{t,k}$, path losses $\ell^{\mathrm{P}}_{t,j}$, and the blend target $\alpha^\star_t$. They update the latest state $\boldsymbol{\Pi}^-_{t+H}$ to $\boldsymbol{\Pi}_{t+H}$, including expert trust, selector coefficients $\boldsymbol{\theta}_{t+H}$, and gate coefficients $\boldsymbol{\omega}_{t+H}$. Episodic memory $\mathcal{M}_{t+H}$ is maintained as auxiliary context.

Paths with $\boldsymbol{\Delta}_t^{(j)} = 0$ can still differ from the prior through their selected expert mixture. Nonzero corrections add evidence-linked adjustments to that mixture. Valid paths form the nonempty set $\mathcal{J}_t$. The language model interprets the evidence, and the runtime executes and stores the resulting candidate forecasts together with their structured decisions.

### 4.3 EVOLVETRUST FOR EXPERT TRUST

EvolveTrust combines four views of expert reliability. The warm-up view uses training and validation forecasts, the online view uses completed deployment windows, the current-evidence view assesses forecast plausibility, and the pattern view summarizes performance under previously observed evidence patterns. Together, these views distinguish aggregate expert performance from reliability under different observed evidence patterns and current forecast plausibility. Each $\boldsymbol{p}_t^v$ is a probability vector over active experts, with zero mass outside $\mathcal{K}_t$. Their mixture forms the prior,

$$\boldsymbol{p}_t = \sum_v \beta_{t,v} \boldsymbol{p}_t^v, \quad \beta_{t,v} \geq 0, \quad \sum_v \beta_{t,v} = 1, \qquad \boldsymbol{b}_t = \sum_{k \in \mathcal{K}_t} p_{t,k} \hat{\boldsymbol{y}}_t^{(k)}. \tag{7}$$

The view maps and mixture schedule are fixed calibration choices. The online and pattern contributions enter through $\rho_t = N_t/(N_t+\tau)$, where $N_t$ counts completed deployment windows with observed targets and $\tau > 0$ controls cold start. The ramp increases their contribution as labeled windows accumulate, while warm-up and current evidence support prediction during cold start.

Let $\mathcal{C}_{t,k}$ contain deployment origins $s$ with $s + H \leq t$, $m_s > 0$, and a stored expert-$k$ forecast. With per-window losses $\ell^{\mathrm{E}}_{s,k} = \mathrm{MSE}_{\Omega_s}(\hat{\boldsymbol{y}}_s^{(k)}, \boldsymbol{y}_s)$, the online view uses inverse-error weighting,

$$\mathrm{RMSE}_{t,k} = \left( \frac{1}{|\mathcal{C}_{t,k}|} \sum_{s \in \mathcal{C}_{t,k}} \ell^{\mathrm{E}}_{s,k} \right)^{1/2}, \qquad p^{\mathrm{online}}_{t,k} \propto \frac{1}{\mathrm{RMSE}_{t,k} + \epsilon}. \tag{8}$$

Here $\epsilon > 0$, and the error estimate applies to experts with eligible observations. Each reveal scores all stored experts, including those outside the selected path, and updates reliability and pattern statistics. Updated trust changes both the numerical prior and future path mixtures.

### 4.4 EVOLVEREASON FOR AGENT PATH SELECTION

EvolveReason ranks candidates using five features available before reveal,

$$\boldsymbol{\psi}_t^{(j)} = [1,\ \text{continuity},\ \text{in-range},\ \text{smoothness},\ \text{structural score}]^\top \in \mathbb{R}^5, \tag{9}$$

computed from $\boldsymbol{a}_t^{(j)}$ and $x_{1:t}$. This representation relates candidate shape and boundary behavior to the observed history, allowing one selector to compare paths built from different expert subsets. A shared rolling ridge model predicts log-transformed MSE and selects the lowest-scoring candidate,

$$\hat{z}_t^{(j)} = {\boldsymbol{\psi}_t^{(j)}}^\top \boldsymbol{\theta}_t, \qquad j_t = \arg\min_{j \in \mathcal{J}_t} \hat{z}_t^{(j)}. \tag{10}$$

During cold start, the fixed structural score selects the path. Subsequently, every observed target supplies $z_s^{(j)} = \log(1+\ell_{s,j}^{\mathrm{P}})$, where $\ell_{s,j}^{\mathrm{P}} = \mathrm{MSE}_{\Omega_s}(\boldsymbol{a}_s^{(j)}, \boldsymbol{y}_s)$. The monotone log transform preserves observed loss ordering while compressing large errors in the regression targets. The shared rolling fit in App. D.2 pairs saved features from every valid candidate, including unselected ones, with the outcome-derived target for that candidate forecast.

### 4.5 EvolveIntervene for Intervention Strength

**Pre-reveal decision.** Let $\boldsymbol{\delta}_t = \boldsymbol{a}_t - \boldsymbol{b}_t$ be the selected path's deviation from the prior. This includes both expert reselection and any local correction. The final forecast is

$$\hat{\boldsymbol{y}}_t = \boldsymbol{b}_t + \alpha_t \boldsymbol{\delta}_t, \qquad \alpha_t = \mathrm{clip}(\boldsymbol{g}_t^\top \boldsymbol{\omega}_t, 0, 1). \tag{11}$$

Here $\mathrm{clip}(z, l, u) = \min\{u, \max\{l, z\}\}$. The pre-reveal features $\boldsymbol{g}_t \in \mathbb{R}^{d_g}$ contain an intercept, shift diagnostics, expert disagreement, source-weight concentration, and text availability. The endpoints recover the prior and selected candidate, while intermediate values interpolate between their forecasts.

**Outcome-derived target.** After reveal, set $\boldsymbol{r}_t = \boldsymbol{y}_t - \boldsymbol{b}_t$ and $D_t = \|\boldsymbol{\delta}_t\|_{\Omega_t}^2$. For the stored prior and path, the scalar blend minimizes

$$L_t(\alpha) = \mathrm{MSE}_{\Omega_t}(\boldsymbol{b}_t + \alpha\boldsymbol{\delta}_t, \boldsymbol{y}_t) = \frac{\|\boldsymbol{r}_t - \alpha\boldsymbol{\delta}_t\|_{\Omega_t}^2}{m_t}. \tag{12}$$

For $D_t > 0$, differentiating this quadratic gives the unconstrained optimum $\tilde{\alpha}_t$, and projection onto the admissible interval gives the supervision target,

$$\tilde{\alpha}_t = \frac{\langle \boldsymbol{\delta}_t, \boldsymbol{r}_t \rangle_{\Omega_t}}{D_t}, \qquad \alpha_t^\star = \arg\min_{\alpha \in [0,1]} L_t(\alpha) = \mathrm{clip}(\tilde{\alpha}_t, 0, 1). \tag{13}$$

The numerator measures alignment between the proposed correction and the prior's realized residual, while $D_t$ normalizes correction magnitude. Nonpositive alignment yields $\alpha_t^\star = 0$, and sufficiently strong positive alignment yields full adoption. The target measures how strongly the selected path should have modified its committed prior and supplies a regression label for subsequent gate updates.

**Loss geometry and weighted learning.** Completing the square yields

$$L_t(\alpha) = C_t + w_t(\alpha - \tilde{\alpha}_t)^2, \qquad w_t = \frac{D_t}{m_t}, \quad C_t = \frac{\|\boldsymbol{r}_t\|_{\Omega_t}^2}{m_t} - w_t\tilde{\alpha}_t^2. \tag{14}$$

The quadratic weight $w_t$ quantifies the effect of coefficient error on forecast MSE. For $\tilde{\alpha}_t \in [0, 1]$, the excess loss is $w_t(\alpha - \alpha_t^\star)^2$. App. D.3 derives the general excess-loss expression.

EvolveIntervene fits $\boldsymbol{g}_s$ to $\alpha_s^\star$ by rolling weighted ridge regression with $\bar{w}_s = \mathrm{clip}(w_s, 10^{-4}, 10^4)$. This energy-weighted supervised objective emphasizes identifiable differences between the prior and selected path. The fit retains windows with $D_s$ above numerical tolerance. The resulting coefficients $\boldsymbol{\omega}_t$ predict $\alpha_t$ from current pre-reveal features.

### 4.6 Outcome feedback and coordinated updates

The record $\mathcal{R}_s$ stores expert forecasts, executed candidates, evidence, pre-reveal features, and the committed decisions $(\boldsymbol{p}_s, \boldsymbol{b}_s, j_s, \boldsymbol{a}_s, \alpha_s, \hat{\boldsymbol{y}}_s)$. At $s + H$, the observed target supplies expert losses $\ell_{s,k}^{\mathrm{E}}$, path losses $\ell_{s,j}^{\mathrm{P}}$, and the blend target $\alpha_s^\star$. This produces a labeled example for each recorded expert and candidate, together with a blend label when the selected correction is identifiable. These targets jointly supervise expert reliability, path ranking, and intervention.

The learner applies these signals to the latest state in Eq. (4) and publishes the three validated updates as one checkpoint. Its updated selector and gate coefficients are $\boldsymbol{\theta}_{s+H}$ and $\boldsymbol{\omega}_{s+H}$, respectively. Optional episodic summaries are stored separately as $\mathcal{M}_{s+H}$. The updated trust, selector, and gate propagate feedback through expert aggregation, candidate execution, and subsequent forecasts.

Table 1: Results on the eight Time-MMD domains. Each domain cell reports MSE/MAE metrics, averaged equally over the four forecasting horizons. Average rank reports the corresponding ranks averaged across the eight domains among fifteen methods. The best and second-best results are in **bold** and underline respectively.

| Model | Agriculture | Climate | Economy | Energy | Environment | Security | Social Good | Traffic | Avg. rank |
|---|---|---|---|---|---|---|---|---|---|
| *Agentic systems* | | | | | | | | | |
| TIMEVOLVE | **0.129/0.227** | **0.314/0.402** | **0.181/0.329** | **0.181/0.303** | **0.274/0.372** | **58.937/3.285** | 0.808/0.381 | **0.087/0.176** | **1.250/1.375** |
| KairosAgent | 0.194/0.282 | 0.863/0.739 | 0.186/0.335 | 0.217/0.330 | 0.378/0.435 | 76.658/4.340 | **0.769/0.376** | 0.151/0.231 | 5.188/5.125 |
| MemCast | 0.218/0.313 | 0.897/0.752 | 0.196/0.339 | 0.207/0.318 | 0.289/0.390 | 69.481/3.876 | 0.797/0.378 | 0.132/0.200 | 4.375/5.250 |
| CastFlow | 0.220/0.303 | 0.890/0.748 | 0.193/0.338 | 0.204/0.316 | 0.286/0.388 | 68.900/3.860 | 0.810/0.383 | 0.130/0.198 | 4.000/4.125 |
| *Zero-shot foundation models* | | | | | | | | | |
| Aurora | 0.282/0.356 | 0.863/0.747 | 0.275/0.412 | 0.251/0.370 | 0.276/0.379 | 72.763/4.085 | 0.828/0.506 | 0.162/0.289 | 7.688/8.938 |
| TimesFM-3 | 0.195/0.270 | 0.555/0.516 | 0.428/0.482 | 0.334/0.384 | 0.326/0.391 | 63.468/3.822 | 0.834/0.457 | 0.414/0.469 | 8.125/8.375 |
| Chronos-2 | 0.133/0.233 | 0.321/0.417 | 0.264/0.408 | 0.209/0.316 | 0.340/0.397 | 64.695/3.818 | 1.124/0.497 | 0.315/0.431 | 7.375/6.688 |
| Toto-2.0 (313M) | 0.330/0.346 | 0.393/0.450 | 0.565/0.556 | 0.196/0.314 | 0.329/0.394 | 244.928/5.691 | 0.818/0.469 | 0.649/0.640 | 9.375/9.750 |
| Sundial | 0.327/0.366 | 0.920/0.765 | 0.216/0.348 | 0.234/0.337 | 0.379/0.443 | 83.403/4.836 | 0.819/0.377 | 0.228/0.292 | 9.375/8.375 |
| Moirai-Large | 0.239/0.306 | 0.982/0.792 | 0.198/0.345 | 0.261/0.347 | 0.412/0.446 | 74.249/4.129 | 0.868/0.391 | 0.186/0.263 | 8.750/7.750 |
| *Full-shot multimodal models* | | | | | | | | | |
| T3Time | 0.229/0.303 | 1.206/0.894 | 0.239/0.384 | 0.266/0.378 | 0.489/0.507 | 72.113/4.070 | 0.998/0.432 | 0.289/0.368 | 10.750/9.813 |
| TimeCMA | 0.318/0.360 | 1.282/0.926 | 0.262/0.412 | 0.351/0.447 | 0.536/0.533 | 72.011/4.113 | 1.092/0.578 | 0.297/0.412 | 12.250/13.188 |
| CALF | 0.241/0.311 | 1.199/0.895 | 0.223/0.370 | 0.258/0.373 | 0.537/0.509 | 73.267/4.040 | 0.890/0.416 | 0.227/0.305 | 10.438/9.438 |
| *Full-shot unimodal models* | | | | | | | | | |
| PatchTST | 0.248/0.308 | 1.176/0.891 | 0.223/0.380 | 0.243/0.353 | 0.496/0.513 | 76.105/4.445 | 0.959/0.475 | 0.209/0.316 | 10.188/10.750 |
| DLinear | 0.377/0.396 | 1.036/0.807 | 0.218/0.370 | 0.233/0.346 | 0.591/0.627 | 82.521/4.891 | 0.891/0.448 | 0.219/0.315 | 10.875/11.063 |

# 5 EXPERIMENTS

## 5.1 EXPERIMENTAL SETUP

We evaluate TIMEVOLVE on the eight Time-MMD (Liu et al., 2024) domains using chronological $70/10/20$ training, validation, and test splits and each domain's benchmark lookback and four forecast horizons. We report MSE and MAE in standardized target space, excluding missing target coordinates. Each horizon is evaluated separately and contributes equally to its domain-level result. We then rank all fifteen methods within each domain and average the ranks across domains. Test forecasts use stride one, and policy updates use only forecasts whose complete horizons have elapsed. App. B provides the domain-specific configurations and details.

We compare TIMEVOLVE with fourteen baselines spanning four groups. ① Agentic baselines: KairosAgent (Feng et al., 2026), MemCast (Tao et al., 2026a), and CastFlow (Pan et al., 2026). ② Zero-shot foundation models: Aurora (Wu et al., 2025), TimesFM-3 (Jain & Sen, 2026), Chronos-2 (Ansari et al., 2025), Toto-2.0 (313M) (Khwaja et al., 2026), Sundial (Liu et al., 2025c) and Moirai-Large (Woo et al., 2024). ③ Full-shot multimodal models: T3Time (Chowdhury et al., 2026), TimeCMA (Liu et al., 2025a), and CALF (Liu et al., 2025b). ④ Full-shot unimodal models include PatchTST (Nie et al., 2023) and DLinear (Zeng et al., 2023).

## 5.2 MAIN RESULTS

Tab. 1 compares TIMEVOLVE with agentic systems, foundation models, and full-shot forecasting models across eight Time-MMD domains. We highlight three findings. ① **Consistent cross-domain performance.** TIMEVOLVE achieves the best average MSE/MAE ranks of $1.250/1.375$ among fifteen methods, leading both metrics in seven of eight domains. Its aggregate advantage therefore reflects broad domain-level gains under both error criteria, rather than isolated improvements in a few favorable settings. ② **Gains over agentic baselines.** TIMEVOLVE outperforms CastFlow in all sixteen domain–metric comparisons and both MemCast and KairosAgent in fourteen each. These results establish that a frozen-backbone agent can outperform existing agentic systems with adaptation concentrated in the orchestration policy rather than the underlying forecasting components. ③ **Effective orchestration across heterogeneous domains.** The strongest standalone foundation model varies across domains, with Chronos-2 leading this group in Agriculture and Climate, Toto-2.0 in Energy, and Aurora in Environment. TIMEVOLVE surpasses these domain-specific competitors with a shared orchestration mechanism. This contrast supports adaptive coordination as an alternative to relying on a single forecaster whose relative strength varies across domains.

Table 2: Effect of policy evolution. Each ablation freezes one policy component at its initial checkpoint while preserving its forward computation; TIMEVOLVE-STACK freezes all three. Cells report standardized-space MSE/MAE averaged equally over four horizons. Lower is better; bold marks the best values, including ties at the reported precision.

| Model | Agriculture | Climate | Economy | Energy | Environment | Security | Social Good | Traffic |
|---|---|---|---|---|---|---|---|---|
| w/o EvolveTrust | 0.142/0.236 | 0.334/0.424 | 0.205/0.349 | 0.183/0.307 | 0.281/0.385 | 59.712/3.391 | 0.812/0.448 | 0.088/0.178 |
| w/o EvolveReason | 0.138/0.234 | 0.328/0.418 | 0.196/0.342 | 0.182/0.305 | 0.278/0.380 | 59.402/3.351 | 0.810/0.421 | 0.088/0.177 |
| w/o EvolveIntervene | 0.135/0.232 | 0.324/0.414 | 0.190/0.337 | 0.182/0.304 | 0.276/0.378 | 59.248/3.329 | 0.809/0.408 | 0.088/0.177 |
| TIMEVOLVE-STACK | 0.148/0.240 | 0.342/0.432 | 0.220/0.360 | 0.185/0.309 | 0.285/0.396 | 60.429/3.479 | 0.814/0.508 | 0.089/0.178 |
| TIMEVOLVE | **0.129/0.227** | **0.314/0.402** | **0.181/0.329** | **0.181/0.303** | **0.274/0.372** | **58.937/3.285** | **0.808/0.381** | **0.087/0.176** |

Table 3: Effect of feedback coverage. Rows vary whether expert and path losses supervise all precommitted alternatives or only those used by the selected path. Cells report standardized-space MSE/MAE averaged equally over four horizons. All updating variants retain intervention learning. Lower is better; bold marks the best values.

| Feedback regime | Agriculture | Climate | Economy | Energy | Environment | Security | Social Good | Traffic |
|---|---|---|---|---|---|---|---|---|
| Selected-path feedback | 0.134/0.231 | 0.321/0.409 | 0.187/0.334 | 0.182/0.305 | 0.277/0.379 | 59.374/3.348 | 0.810/0.414 | 0.088/0.177 |
| Selected-expert feedback | 0.140/0.236 | 0.331/0.419 | 0.198/0.344 | 0.184/0.308 | 0.282/0.386 | 59.756/3.401 | 0.812/0.451 | 0.088/0.178 |
| Selected-only feedback | 0.145/0.239 | 0.336/0.424 | 0.208/0.352 | 0.184/0.308 | 0.283/0.392 | 60.158/3.448 | 0.813/0.485 | 0.089/0.178 |
| Frozen policy | 0.148/0.240 | 0.342/0.432 | 0.220/0.360 | 0.185/0.309 | 0.285/0.396 | 60.429/3.479 | 0.814/0.508 | 0.089/0.178 |
| All-alternative feedback | **0.129/0.227** | **0.314/0.402** | **0.181/0.329** | **0.181/0.303** | **0.274/0.372** | **58.937/3.285** | **0.808/0.381** | **0.087/0.176** |

### 5.3 EFFECT OF POLICY EVOLUTION

We assess both the overall benefit of policy evolution and the contribution of each update in Tab. 2. All variants share the numerical experts, evidence and context pipeline, candidate-generation configuration, initialization, forecast origins, and reveal schedule. TIMEVOLVE-STACK holds expert trust, agent path selection, and intervention strength at their initial checkpoints throughout testing. Each *w/o* variant instead freezes only the named policy component while the other two continue to update. The forward computation is retained in every variant, isolating the contribution of outcome-driven updates. App. C.2 distinguishes these update interventions from restrictions on feedback coverage.

Compared with TIMEVOLVE-STACK, TIMEVOLVE lowers both errors in every domain. Averaging domain-wise relative error reductions gives a $6.3\%$ reduction in MSE and a $7.6\%$ reduction in MAE. The largest MSE reduction is $17.7\%$ in Economy, from $0.220$ to $0.181$, while the largest MAE reduction is $25.0\%$ in Social Good, from $0.508$ to $0.381$. Since both variants use the same frozen forecasters and forecast-construction pipeline, these improvements isolate the benefit of adapting orchestration rather than expanding the numerical tool pool.

The complete policy achieves lower error in all sixteen domain–metric comparisons against each single-component freezing variant. The average degradation is largest when EvolveTrust is frozen, followed by EvolveReason and EvolveIntervene. In Economy, freezing trust, path selection, and intervention increases MSE by $13.3\%$, $8.3\%$, and $5.0\%$, respectively. Expert reliability adaptation thus has the largest observed effect, while learned path selection and intervention each improve the use of the resulting numerical evidence. Each contrast measures the contribution of one update with the other two active, rather than an additive decomposition of the total gain.

### 5.4 VALUE OF FEEDBACK COVERAGE

We isolate the value of supervising unselected alternatives by varying expert and path feedback coverage. All variants share the forecast-construction configuration and reveal schedule. *All-alternative feedback* updates expert trust from every committed expert forecast and trains path selection on every valid candidate. *Selected-path feedback* restricts only path supervision to the chosen path, whereas *selected-expert feedback* restricts only expert supervision to the experts used by that path. *Selected-only feedback* applies both restrictions. Every updating variant retains the intervention-learning rule and computes its blend target from its stored prior and selected path. The frozen-policy reference applies no outcome-driven updates.

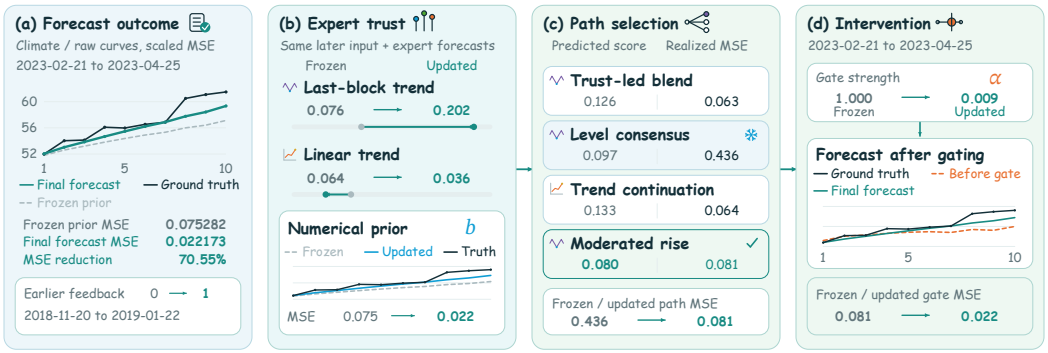


Figure 3: **Realized feedback reshapes future forecasting decisions.** A two-window Climate replay from Time-MMD using statistical forecasting experts. (a) Final forecast, ground truth, and Trust-frozen numerical prior. (b) Reweighting strengthens the prior's upward trend. (c) Path selection favors moderated rise over level consensus. (d) The learned gate limits premature flattening by the selected path. Panel (a) uses original units; all MSE values use training-split standardization.

Tab. 3 shows that supervising all precommitted alternatives improves all sixteen comparisons over selected-only feedback, with average domain-wise relative reductions of $5.0\%$ in MSE and $6.3\%$ in MAE. Selected-only feedback improves fourteen comparisons over the frozen policy and ties both Traffic metrics at the reported precision. Restricting expert supervision incurs at least as much error as restricting path supervision in every cell. These results support using revealed outcomes to update the policy and retaining loss feedback for unselected alternatives, with broader expert coverage providing the larger observed benefit.

### 5.5 Case study of policy evolution

To complement the aggregate comparisons, we inspect a post hoc selected two-window Climate replay using the statistical-expert configuration of TimEvolve on Time-MMD (Liu et al., 2024). The target continues rising and steepens toward the end of the horizon, whereas the Trust-frozen prior responds too weakly to this movement in Fig. 3(a). The case examines how earlier feedback revises later decisions. App. C.3 gives the sample dates, replay protocol, and comparison settings.

Fig. 3(b-d) reveals three complementary responses. EvolveTrust shifts reliance toward the recent-block trend and away from the full-window linear trend, producing a prior with stronger upward movement. EvolveReason changes the preferred interpretation from level consensus to a moderated rise, reducing the tendency to flatten the forecast. The selected path nevertheless levels off earlier than the revised prior. The intervention rule learned from earlier feedback assigns it limited influence, preserving the prior's stronger continuation in the final forecast.

The mechanism therefore separates path preference from intervention strength. Expert adaptation improves the numerical basis, path adaptation changes the preferred interpretation, and the gate controls its effect on the output. Their interaction illustrates how policy evolution retains useful numerical structure while revising subsequent decisions.

## 6 Conclusion

TimEvolve turns realized futures into supervision for expert trust, agent path selection, and intervention strength through a temporally ordered predict, reveal, and update loop. Across eight Time-MMD domains, it achieves the best average MSE and MAE ranks among fifteen methods and improves both metrics over its fixed-policy counterpart in every domain. The case analysis illustrates how these updates change the interpretation and use of numerical forecasts. Together, the results support adapting the orchestration policy while keeping forecasting components frozen. The agent does more than retain past outcomes. It uses them to revise how it forecasts what comes next.

## AI USE STATEMENT

We use generative AI tools to assist with manuscript drafting and polishing, literature retrieval and discovery, and figure preparation. We take responsibility for verifying all AI-assisted text, references, and figures and for the accuracy and integrity of the final content.

## ETHICS STATEMENT

As our work only focuses on the time series forecasting problem, there are no potential ethical risks.

## REPRODUCIBILITY STATEMENT

The main text defines the forecast construction and policy-update equations. App. B specifies dataset configurations, chronological splits, reveal eligibility, metrics, and aggregation. App. C describes the numerical tool pool, policy ablations, feedback restrictions, and case replay. Average ranks use all fifteen methods and mean ranks for ties at the reported precision.

## REFERENCES

Abdul Fatir Ansari, Lorenzo Stella, Caner Turkmen, Xiyuan Zhang, Pedro Mercado, Huibin Shen, Oleksandr Shchur, Syama Sundar Rangapuram, Sebastian Pineda Arango, Shubham Kapoor, Jasper Zschiegner, Danielle C. Maddix, Hao Wang, Michael W. Mahoney, Kari Torkkola, Andrew Gordon Wilson, Michael Bohlke-Schneider, and Yuyang Wang. Chronos: Learning the language of time series. *Transactions on Machine Learning Research*, 2024. ISSN 2835-8856. URL https://openreview.net/forum?id=gerNCVqqtR.

Abdul Fatir Ansari, Oleksandr Shchur, Jaris Küken, Andreas Auer, Boran Han, Pedro Mercado, Syama Sundar Rangapuram, Huibin Shen, Lorenzo Stella, Xiyuan Zhang, Mononito Goswami, Shubham Kapoor, Danielle C. Maddix, Pablo Guerron, Tony Hu, Junming Yin, Nick Erickson, Prateek Mutalik Desai, Hao Wang, Huzefa Rangwala, George Karypis, Yuyang Wang, and Michael Bohlke-Schneider. Chronos-2: From univariate to universal forecasting. *arXiv preprint arXiv:2510.15821*, 2025. URL https://arxiv.org/abs/2510.15821.

Nicolò Cesa-Bianchi and Gábor Lugosi. *Prediction, Learning, and Games*. Cambridge University Press, 2006. doi: 10.1017/CBO9780511546921. URL https://doi.org/10.1017/CBO9780511546921.

Ching Chang, Yidan Shi, Defu Cao, Wei Yang, Jeehyun Hwang, Haixin Wang, Jiacheng Pang, Wei Wang, Yan Liu, Wen-Chih Peng, and Tien-Fu Chen. A survey of reasoning and agentic systems in time series with large language models. *Transactions on Machine Learning Research*, 2026. ISSN 2835-8856. URL https://openreview.net/forum?id=l3QW42g6u3.

Mingyue Cheng, Xiaoyu Tao, Qi Liu, Ze Guo, and Enhong Chen. Position: Beyond model-centric prediction—agentic time series forecasting. *arXiv preprint arXiv:2602.01776*, 2026. URL https://arxiv.org/abs/2602.01776.

Abdul Monaf Chowdhury, Rabeya Akter, and Safaeid Hossain Arib. T3Time: Tri-modal time series forecasting via adaptive multi-head alignment and residual fusion. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 40, pp. 20597–20605, 2026. doi: 10.1609/aaai.v40i25.39196. URL https://ojs.aaai.org/index.php/AAAI/article/view/39196.

Abhimanyu Das, Weihao Kong, Rajat Sen, and Yichen Zhou. A decoder-only foundation model for time-series forecasting. In *Proceedings of the 41st International Conference on Machine Learning*, volume 235 of *Proceedings of Machine Learning Research*, pp. 10148–10167. PMLR, 2024. URL https://proceedings.mlr.press/v235/das24c.html.

Kun Feng, Ziwei Shan, Yuchen Fang, Yiyang Tan, Sihan Lu, Shuqi Gu, Lintao Ma, Xingyu Lu, and Kan Ren. KairosAgent: Agentic time series forecasting with fused semantic reasoning. *arXiv preprint arXiv:2605.30002v1*, 2026. URL https://arxiv.org/abs/2605.30002v1.

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. doi: 10.1006/jcss.1997.1504. URL https://www.sciencedirect.com/science/article/pii/S002200009791504X.

Tong Guan, Zijie Meng, Dianqi Li, Shiyu Wang, Chao-Han Huck Yang, Qingsong Wen, Zuozhu Liu, Sabato Marco Siniscalchi, Ming Jin, and Shirui Pan. TimeOmni-1: Incentivizing complex reasoning with time series in large language models. *arXiv preprint arXiv:2509.24803*, 2025. URL https://arxiv.org/abs/2509.24803.

Yifan Hu, Jie Yang, Xilin Dai, Wanxu Cai, Kuiye Ding, Yuante Li, Qinghua Liu, Enze Ma, Zhiyuan Qu, Yixin Wang, Binyan Xu, Kexin Zhang, Peiyuan Liu, Zhijian Xu, Guibin Zhang, Yujin Tang, Yanwei Yue, Kening Zheng, Chengze Li, Hanrong Zhang, Haoyan Xu, Naiqi Li, Tao Dai, Dawei Cheng, John Paparrizos, Kaize Ding, Tian Zhou, Qiang Xu, Shu-tao Xia, Shirui Pan, and Philip S. Yu. The landscape of agentic time series systems: Architectures, reliability, and frontiers. Preprint, 2026. URL https://github.com/TROUBADOUR000/Awesome-Agentic-Time-Series/blob/main/The%20Landscape%20of%20Agentic%20Time%20Series%20Systems.pdf.

Ayush Jain and Rajat Sen. TimesFM-3: A zero-shot foundation model for multivariate forecasting. Google Research Blog, August 2026. URL https://research.google/blog/timesfm-3-a-zero-shot-foundation-model-for-multivariate-forecasting/. Accessed: 2026-09-16.

Nan Jiang and Lihong Li. Doubly robust off-policy value evaluation for reinforcement learning. In *Proceedings of the 33rd International Conference on Machine Learning*, volume 48 of *Proceedings of Machine Learning Research*, pp. 652–661. PMLR, 2016. URL https://proceedings.mlr.press/v48/jiang16.html.

Pooria Joulani, Andras Gyorgy, and Csaba Szepesvari. Online learning under delayed feedback. In *Proceedings of the 30th International Conference on Machine Learning*, volume 28 of *Proceedings of Machine Learning Research*, pp. 1453–1461. PMLR, 2013. URL https://proceedings.mlr.press/v28/joulani13.html.

Emaad Khwaja, Chris Lettieri, Gerald Woo, Eden Belouadah, Marc Cenac, Guillaume Jarry, Enguerrand Paquin, Xunyi Zhao, Viktoriya Zhukov, Othmane Abou-Amal, Chenghao Liu, Ameet Talwalkar, and David Asker. Toto 2.0: Time series forecasting enters the scaling era. *arXiv preprint arXiv:2605.20119*, 2026. URL https://arxiv.org/abs/2605.20119.

Jongseon Kim, Hyungjoon Kim, HyunGi Kim, Dongjun Lee, and Sungroh Yoon. A comprehensive survey of deep learning for time series forecasting: Architectural diversity and open challenges. *Artificial Intelligence Review*, 58(7):216, 2025. doi: 10.1007/s10462-025-11223-9. URL https://link.springer.com/article/10.1007/s10462-025-11223-9.

Yuhua Liao, Zetian Wang, Qiangqiang Nie, and Zhenhua Zhang. Bridging the last mile of time series forecasting with LLM agents. *arXiv preprint arXiv:2606.02497*, 2026. URL https://arxiv.org/abs/2606.02497.

Chenxi Liu, Qianxiong Xu, Hao Miao, Sun Yang, Lingzheng Zhang, Cheng Long, Ziyue Li, and Rui Zhao. TimeCMA: Towards LLM-empowered multivariate time series forecasting via cross-modality alignment. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 39, pp. 18780–18788, 2025a. doi: 10.1609/aaai.v39i18.34067. URL https://ojs.aaai.org/index.php/AAAI/article/view/34067.

Haoxin Liu, Shangqing Xu, Zhiyuan Zhao, Lingkai Kong, Harshavardhan Kamarthi, Aditya B. Sasanur, Megha Sharma, Jiaming Cui, Qingsong Wen, Chao Zhang, and B. Aditya Prakash. Time-MMD: Multi-domain multimodal dataset for time series analysis. In *Advances in Neural Information Processing Systems*, volume 37, 2024. URL https://proceedings.neurips.cc/paper_files/paper/2024/hash/8e7768122f3eeec6d77cd2b424b72413-Abstract-Datasets_and_Benchmarks_Track.html. Datasets and Benchmarks Track.

Peiyuan Liu, Hang Guo, Tao Dai, Naiqi Li, Jigang Bao, Xudong Ren, Yong Jiang, and Shu-Tao Xia. CALF: Aligning LLMs for time series forecasting via cross-modal fine-tuning. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 39, pp. 18915–18923, 2025b. doi: 10.1609/aaai.v39i18.34082. URL https://ojs.aaai.org/index.php/AAAI/article/view/34082.

Yiding Liu, Yifan Hu, Hongjie Xia, Peiyuan Liu, Hongzhou Chen, Xilin Dai, Zewei Dong, and Jiang-Ming Yang. Falcon-X: A time series foundation model for heterogeneous multivariate modeling. *arXiv preprint arXiv:2605.27286*, 2026. URL https://arxiv.org/abs/2605.27286.

Yong Liu, Guo Qin, Zhiyuan Shi, Zhi Chen, Caiyin Yang, Xiangdong Huang, Jianmin Wang, and Mingsheng Long. Sundial: A family of highly capable time series foundation models. In *Proceedings of the 42nd International Conference on Machine Learning*, volume 267 of *Proceedings of Machine Learning Research*, pp. 39295–39317. PMLR, 2025c. URL https://proceedings.mlr.press/v267/liu25be.html.

Sisuo Lyu, Siru Zhong, Tiegang Chen, Weilin Ruan, Qingxiang Liu, Taiqiang Lv, Qingsong Wen, Raymond Chi-Wing Wong, and Yuxuan Liang. TS-Memory: Plug-and-play memory for time series foundation models. *arXiv preprint arXiv:2602.11550*, 2026. URL https://arxiv.org/abs/2602.11550.

Yuqi Nie, Nam H. Nguyen, Phanwadee Sinthong, and Jayant Kalagnanam. A time series is worth 64 words: Long-term forecasting with transformers. In *International Conference on Learning Representations*, 2023. URL https://openreview.net/forum?id=Jbdc0vTOcol.

Bokai Pan, Mingyue Cheng, Zhiding Liu, Shuo Yu, Xiaoyu Tao, Yuchong Wu, Qi Liu, Defu Lian, and Enhong Chen. CastFlow: Learning role-specialized agentic workflows for time series forecasting. *arXiv preprint arXiv:2604.27840*, 2026. URL https://arxiv.org/abs/2604.27840.

Raeid Saqur, Anastasis Kratsios, Florian Krach, Yannick Limmer, Jacob-Junqi Tian, John Willes, Blanka Horvath, and Frank Rudzicz. Filtered not mixed: Filtering-based online gating for mixture of large language models. In *International Conference on Learning Representations*, 2025. URL https://proceedings.iclr.cc/paper_files/paper/2025/hash/d4c2f25bf0c33065b7d4fb9be2a9add1-Abstract-Conference.html.

Noah Shinn, Federico Cassano, Ashwin Gopinath, Karthik Narasimhan, and Shunyu Yao. Reflexion: Language agents with verbal reinforcement learning. In *Advances in Neural Information Processing Systems*, volume 36, pp. 8634–8652, 2023. URL https://proceedings.neurips.cc/paper_files/paper/2023/hash/1b44b878bb782e6954cd888628510e90-Abstract-Conference.html.

Xiaoyu Tao, Mingyue Cheng, Ze Guo, Shuo Yu, Yaguo Liu, Qi Liu, and Shijin Wang. MemCast: Memory-driven time series forecasting with experience-conditioned reasoning. *arXiv preprint arXiv:2602.03164*, 2026a. URL https://arxiv.org/abs/2602.03164.

Xiaoyu Tao, Mingyue Cheng, Chuang Jiang, Tian Gao, Huanjian Zhang, and Yaguo Liu. Cast-R1: Learning tool-augmented sequential decision policies for time series forecasting. *arXiv preprint arXiv:2602.13802*, 2026b. URL https://arxiv.org/abs/2602.13802.

Xiaoyu Tao, Yuchong Wu, Mingyue Cheng, Ze Guo, and Tian Gao. AnomaMind: Agentic time series anomaly detection with tool-augmented reasoning. *arXiv preprint arXiv:2602.13807*, 2026c. URL https://arxiv.org/abs/2602.13807.

Xinlei Wang, Maike Feng, Jing Qiu, Jinjin Gu, and Junhua Zhao. From news to forecast: Integrating event analysis in LLM-based time series forecasting with reflection. In *Advances in Neural Information Processing Systems*, volume 37, pp. 58118–58153, 2024. URL https://proceedings.neurips.cc/paper_files/paper/2024/hash/6aef8bffb372096ee73d98da30119f89-Abstract-Conference.html.

Gerald Woo, Chenghao Liu, Akshat Kumar, Caiming Xiong, Silvio Savarese, and Doyen Sahoo. Unified training of universal time series forecasting transformers. In *Proceedings of the 41st International Conference on Machine Learning*, volume 235 of *Proceedings of Machine Learning Research*, pp. 53140–53164. PMLR, 2024. URL https://proceedings.mlr.press/v235/woo24a.html.

Xingjian Wu, Jianxin Jin, Wanghui Qiu, Peng Chen, Yang Shu, Bin Yang, and Chenjuan Guo. Aurora: Towards universal generative multimodal time series forecasting. *arXiv preprint arXiv:2509.22295*, 2025. URL https://arxiv.org/abs/2509.22295.

Xingjian Wu, Junkai Lu, Zhengyu Li, Xiangfei Qiu, Jilin Hu, Chenjuan Guo, Christian S. Jensen, and Bin Yang. TimeART: Towards agentic time series reasoning via tool-augmentation. *arXiv preprint arXiv:2601.13653*, 2026. URL https://arxiv.org/abs/2601.13653.

Ailing Zeng, Muxi Chen, Lei Zhang, and Qiang Xu. Are transformers effective for time series forecasting? In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 37, pp. 11121–11128, 2023. doi: 10.1609/aaai.v37i9.26317. URL https://ojs.aaai.org/index.php/AAAI/article/view/26317.

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. URL https://proceedings.neurips.cc/paper_files/paper/2023/hash/dd6a47bc0aad6f34aa5e77706d90cdc4-Abstract-Conference.html.

Haokun Zhao, Xiang Zhang, Jiaqi Wei, Yiwei Xu, Yuting He, Siqi Sun, and Chenyu You. TimeSeries-Scientist: A general-purpose AI agent for time series analysis. *arXiv preprint arXiv:2510.01538*, 2025. URL https://arxiv.org/abs/2510.01538.

Jiahui Zhou, Dan Li, Boxin Li, Xiao Zhang, Erli Meng, Lin Li, Zhuomin Chen, Jian Lou, and See-Kiong Ng. Time series reasoning via process-verifiable thinking data synthesis and scheduling for tailored LLM reasoning. *arXiv preprint arXiv:2602.07830*, 2026. URL https://arxiv.org/abs/2602.07830.

## A LIMITATIONS

This study focuses on forecasting components and outcomes that are unaffected by forecast selection. Therefore, its predictive scope depends partly on the available base models. Richer expert sets and longer deployment studies are natural extensions of this approach. The predict, reveal and update protocol provides a basis for examining these settings while maintaining the temporal separation between prediction and feedback.

## B TIME-MMD EVALUATION PROTOCOL

Time-MMD aligns numerical targets with temporally indexed reports and web search text (Liu et al., 2024). The main comparison uses Agriculture, Climate, Economy, Energy, Environment, Security, Social Good, and Traffic, which are the eight domains shared with the KairosAgent evaluation. The release also includes separate U.S. and Africa Health streams, but Health remains outside the shared comparison because the KairosAgent paper identifies overlap with its training corpus (Feng et al., 2026).

The text modality separates report facts, report predictions, search facts, and search predictions. An item is available only when its recorded availability interval ends no later than the forecasting origin, and items overlapping the target window are excluded. Tab. 4 lists the domain-specific input lengths and forecast horizons used by the evaluation configuration. All lengths count observations, not calendar units. Climate retains the benchmark configuration of eight input observations and horizons $\{6, 8, 10, 12\}$ even though its released numerical timestamps are weekly. This distinction also applies to the case study in App. C.3. Energy and Environment use their separate weekly and daily configurations.

Each series is placed in chronological order, including Economy, before the $70/10/20$ training, validation, and test split. The scaler uses only observed training values. Filled numerical inputs preserve continuity, while an immutable validity mask excludes filled targets from every loss and policy update. Text availability is evaluated at each forecasting origin independently of the numerical sampling frequency.

We score every complete test origin with stride one. A forecast issued at origin $s$ becomes eligible for an update at origin $t$ only when $s + H \leq t$. For each domain, we evaluate the four configured horizons independently and average them with equal weight,

$$\overline{m}_d = \frac{1}{|\mathcal{H}_d|} \sum_{h \in \mathcal{H}_d} m_{d,h}, \tag{15}$$

where $m_{d,h}$ denotes MSE or MAE for one domain and horizon. We rank the fifteen methods within each domain using the reported three-decimal values, assign ties their mean rank, and average the resulting ranks across the eight domains.

Table 4: Domain-specific forecast configurations. Cadence describes the numerical timestamps in the Time-MMD release; lookback and horizons count observations. Climate retains its benchmark shape configuration despite its weekly cadence.

| Domain | Cadence | Lookback | Forecast horizons |
|---|---|---|---|
| Agriculture | Monthly | 8 | 6, 8, 10, 12 |
| Climate | Weekly | 8 | 6, 8, 10, 12 |
| Economy | Monthly | 8 | 6, 8, 10, 12 |
| Energy | Weekly | 36 | 12, 24, 36, 48 |
| Environment | Daily | 96 | 48, 96, 192, 336 |
| Security | Monthly | 8 | 6, 8, 10, 12 |
| Social Good | Monthly | 8 | 6, 8, 10, 12 |
| Traffic | Monthly | 8 | 6, 8, 10, 12 |

# C Implementation and Evaluation Details

## C.1 Evaluated configuration

The numerical tool pool augments statistical experts with Chronos-2 (Ansari et al., 2025), Toto-2.0 (313M) (Khwaja et al., 2026), Falcon-X (Liu et al., 2026), TimesFM-2.5 (Das et al., 2024), and Chronos-Bolt (Ansari et al., 2024). All numerical forecasters remain frozen.

TimEvolve and TimEvolve-Stack share the same foundation-model-augmented ForecastPack forecasts, evidence and context pipeline, candidate-generation configuration, initialization, chronological origins, and aggregation rules. The complete method updates expert trust, agent path selection, and intervention strength after each eligible reveal. TimEvolve-Stack holds all three policy components at their initial checkpoint throughout testing. This pairing isolates the effect of outcome-driven policy evolution while preserving the forecast and evidence substrate.

## C.2 Policy ablations and feedback coverage

The policy ablations in Section 5.3 change which learned states may update. Freezing a component retains its inference-time computation and initial checkpoint; it does not remove numerical experts, candidate generation, or the final blending operation. Its outputs still depend on the current history and evidence. The feedback-coverage comparisons in Section 5.4 instead change which stored losses enter the update. A selected-path restriction still trains the selector, and a selected-expert restriction still updates trust. They therefore differ from freezing either learner.

Table 5: Update rules for the two families of comparisons. All and selected describe which committed alternatives supply loss feedback. Fixed means that the learned state remains at its initial checkpoint.

| Configuration | Expert feedback | Path feedback | Intervention |
|---|---|---|---|
| TimEvolve | All | All | Updated |
| TimEvolve-Stack | Fixed | Fixed | Fixed |
| w/o EvolveTrust | Fixed | All | Updated |
| w/o EvolveReason | All | Fixed | Updated |
| w/o EvolveIntervene | All | All | Fixed |
| Selected-path feedback | All | Selected path | Updated |
| Selected-expert feedback | Selected support | All | Updated |
| Selected-only feedback | Selected support | Selected path | Updated |

For a forecast issued at $s$, selected-expert feedback uses only $k \in S_s^{(j_s)}$, while selected-path feedback uses only $j_s$. These eligibility sets come from the pre-reveal record and remain fixed when the target arrives. In Eq. (8), an expert's eligible-origin set is restricted accordingly; in Eq. (16), the candidate sum is restricted to the recorded selected path. Intervention learning retains the blend target formed from the committed prior and selected candidate in every updating coverage regime. All-alternative feedback and the frozen-policy reference correspond to TimEvolve and TimEvolve-Stack, respectively, so their repeated rows denote shared configurations.

## C.3 Climate case replay

Section 5.5 uses the weekly Climate target column in the Time-MMD numerical release, whose source is historical drought information from Drought.gov (Liu et al., 2024). The later numerical forecast context consists of eight observations from December 27, 2022 to February 14, 2023. Its ten targets span February 21 to April 25, 2023. The earlier window uses numerical context from September 25 to November 13, 2018 and targets from November 20, 2018 to January 22, 2019. Past-reliability diagnostics use only observations available before each origin. The replay reveals the earlier target window, updates the policy once, and applies the resulting checkpoint to the later window. No forecasts or policy updates occur in the intervening period. The case is selected after outcome evaluation to inspect the behavior of all three update components.

The replay retains its recorded configuration of thirteen statistical filters and four structured candidate paths, with additive corrections disabled. This configuration provides an interpretable illustration of

the update mechanism, separate from the foundation-model-augmented pool used in the aggregate experiments. Comparisons reuse the recorded expert-support choices and recompute expert weights, candidate execution, selection, and blending under the relevant policy state. Expert forecasts, executed candidates, and comparison outputs are stored before each target reveal. Targets and predictions use the same training-split scaler for scoring; Fig. 3(a) converts the curves back to original units.

The panels in Fig. 3 distinguish component outputs from the final forecast. Panel (a) compares the final forecast with a numerical prior whose Trust state remains at its initial checkpoint. Panel (b) isolates prior construction under frozen and updated trust using the same expert forecasts. Panel (c) compares path selection under a fixed updated trust state. Its predicted scores are log-loss estimates available before reveal, while realized MSE is computed afterward. Panel (d) holds the updated prior and selected path fixed and changes only the intervention state. The final forecast includes the gated path contribution and is therefore close to, but distinct from, the updated prior. These comparisons explain the forecast's construction at a selected origin, separately from the aggregate evaluation over test origins.

Fig. 4 places these decisions in their temporal context. The observed history rises steadily, while the realized forecast window contains short plateaus followed by renewed growth and a sharper increase near its end. The selected candidate before gating follows the early level but becomes too flat to capture the subsequent rise. The final forecast stays close to the updated numerical prior, retaining its rising trajectory while attenuating the flatter candidate. It still underestimates the late increase. This behavior illustrates intervention as selective control of a candidate's influence, rather than an unconditional override of the numerical forecast.

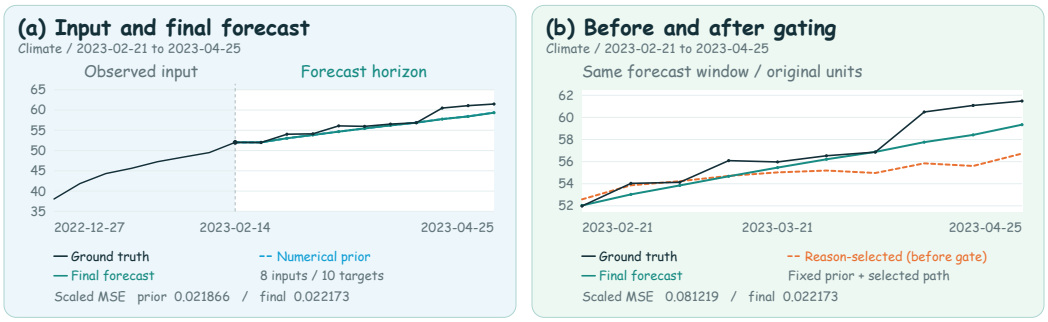


Figure 4: **Numerical input and forecast outputs for the Climate case.** (a) Eight observed inputs and the ten-target forecast window, together with the updated-trust numerical prior and the final forecast. (b) The selected candidate before gating and the final forecast over the same horizon. The numerical prior in panel (a) uses updated trust, unlike the frozen-trust reference in Fig. 3(a). Curves use original target units with panel-specific vertical ranges; the displayed MSE values use training-split standardization.

# D Method Details and Intervention Analysis

## D.1 Evidence and Candidate Execution

The Forecast-Pack interfaces with frozen time series foundation models and statistical experts for level, moving average, robust trend, seasonality, spectral structure, state space, and shifts. The main evaluation uses the foundation-model-augmented configuration described in App. C, while the Climate replay retains its recorded statistical-filter configuration (App. C.3). Each active expert receives the available numerical history and returns an aligned forecast, and unavailable members are excluded from the active set and its normalization.

Expert evidence includes historical RMSE/MAE, bias, exponentially weighted error, trend and seasonal residuals, boundary continuity, forecast range and volatility, source-relative consensus, and disagreement. A fixed tool bundle measures level, range, recent and global trend, volatility, mean absolute change, periodicity, autocorrelation, level shifts, and extremes. Context retains the distinction between report facts, report predictions, search facts, and search predictions. Items enter the context

when their recorded availability intervals end by the forecasting origin, and target-overlapping items are excluded.

Typed validation precedes entry into the language-model context. Path validation rejects empty expert selections, zero trust mass, and invalid corrections. The source-budget projection retains a probability distribution on the selected subset while constraining source concentration. Its configured budgets and the admissible correction bounds determine the execution map; the language model supplies the structured choices, not unconstrained mixture coefficients. A local correction $\boldsymbol{\Delta}_t^{(j)}$ can be zero even when the overall deviation $\boldsymbol{\delta}_t = \boldsymbol{a}_t - \boldsymbol{b}_t$ is nonzero because the path selects a different expert mixture.

### D.2 Rolling fits and stored state

Let $\mathcal{W}_t^{\mathrm{R}}$ and $\mathcal{W}_t^{\mathrm{I}}$ be the rolling sets of completed origins retained for path selection and intervention. Each retained origin satisfies $s + H \le t$ and $m_s > 0$, and intervention additionally requires identifiable correction energy. The ridge objectives corresponding to the two learners are

$$\boldsymbol{\theta}_t = \arg\min_{\boldsymbol{\theta}} \sum_{s \in \mathcal{W}_t^{\mathrm{R}}} \sum_{j \in \mathcal{J}_s} \left( {\boldsymbol{\psi}_s^{(j)}}^{\top} \boldsymbol{\theta} - z_s^{(j)} \right)^2 + \boldsymbol{\theta}^{\top} \Lambda_{\mathrm{R}} \boldsymbol{\theta}, \tag{16}$$

$$\boldsymbol{\omega}_t = \arg\min_{\boldsymbol{\omega}} \sum_{s \in \mathcal{W}_t^{\mathrm{I}}} \bar{w}_s \left( \boldsymbol{g}_s^{\top} \boldsymbol{\omega} - \alpha_s^{\star} \right)^2 + \boldsymbol{\omega}^{\top} \Lambda_{\mathrm{I}} \boldsymbol{\omega}. \tag{17}$$

The positive-semidefinite matrices $\Lambda_{\mathrm{R}}$ and $\Lambda_{\mathrm{I}}$ encode the configured ridge penalties, including intercept treatment. The feature maps, rolling-window settings, cold-start conditions, and penalty choices are calibration settings. Each component state contains its coefficients and the buffers or sufficient statistics needed to update them. The active checkpoint is replaced once validation of all three component updates completes; if validation fails, the current checkpoint is retained. Retaining original features pairs each outcome with the decisions available before it was observed.

### D.3 Geometry of the intervention target

Fix a committed prior and selected path, with $m_t > 0$ and $D_t > 0$. Expanding Eq. (12) gives

$$L_t(\alpha) = \frac{\|\boldsymbol{r}_t\|_{\Omega_t}^2 - 2\alpha \langle \boldsymbol{\delta}_t, \boldsymbol{r}_t \rangle_{\Omega_t} + \alpha^2 D_t}{m_t}. \tag{18}$$

Its derivative vanishes at $\tilde{\alpha}_t = \langle \boldsymbol{\delta}_t, \boldsymbol{r}_t \rangle_{\Omega_t} / D_t$, and its second derivative is $2w_t > 0$. This establishes both the unique unconstrained optimum and the clipped optimum in Eq. (13). Completing the square gives Eq. (14). For any admissible $\alpha \in [0, 1]$, the exact excess loss is

$$\begin{aligned} L_t(\alpha) - L_t(\alpha_t^{\star}) &= w_t(\alpha - \alpha_t^{\star})^2 + 2w_t(\alpha - \alpha_t^{\star})(\alpha_t^{\star} - \tilde{\alpha}_t) \\ &\ge w_t(\alpha - \alpha_t^{\star})^2. \end{aligned} \tag{19}$$

The second term vanishes for an unconstrained optimum inside $[0, 1]$. If $\tilde{\alpha}_t < 0$, both factors in that term are nonnegative; if $\tilde{\alpha}_t > 1$, both are nonpositive. This explains the boundary correction. With $D_t = 0$, the loss is constant in $\alpha$ on the observed coordinates, so that window contains no intervention-identification signal.

The regression in Eq. (17) predicts the clipped target using bounded energy weights. For any raw prediction $q$, projection satisfies $|\operatorname{clip}(q, 0, 1) - \alpha_t^{\star}| \le |q - \alpha_t^{\star}|$. The ridge loss therefore provides a stable target-fitting objective for the bounded gate. It is a supervised surrogate rather than an unconditional identity with forecast MSE, because boundary projection, bounded weights, and parameter regularization change the optimization objective.