UDC 004.85:519.876
DOI https://doi.org/10.32782/2521-6643-2026-2-72.10

Zaika B. Yu., Postgraduate Student of Computer Science,
Junior Research Fellow at the Laboratory of Applied Informatics Problems V. M. Glushkov
Institute of Cybernetics of the NAS of Ukraine
ORCID: 0009-0001-9567-8361

Yershov S. V., Doctor of Physical and Mathematical Sciences, Senior Research Fellow, Head of the Department of Methods and Technological Means of Building Intelligent Software Systems
V. M. Glushkov Institute of Cybernetics of the NAS of Ukraine
ORCID: 0000-0002-9895-777X

STOCHASTIC OPERATORS OF ACTION IMPACT: FORMALISATION AND MULTI-STEP REGULARISED LEARNING

This paper proposes a formalisation of action impact in stochastic dynamical systems as a dedicated stochastic operator acting on system states. Accurate modelling of action impact is an important problem in sequential decision-making under uncertainty, since in many real-world systems actions are applied repeatedly and their consequences propagate through system dynamics over time. While modern machine learning approaches, including reinforcement learning and conditional density estimation, can approximate short-term transitions, the behaviour of learned models under recursive multi-step application remains insufficiently studied. In most existing frameworks, transition dynamics are embedded within policy optimisation or trajectory prediction objectives and are rarely treated as independent modelling entities. In the proposed approach, the action impact operator maps the current system state and applied action to a conditional distribution of future states and is defined with explicit compositional structure. This enables the analysis of recursive operator application across multiple time steps. A learning objective is introduced that combines one-step negative log-likelihood with a multi-step consistency term derived from operator composition. The central hypothesis of the study is that one-step maximum likelihood training does not guarantee stable long-horizon behaviour when the learned operator is recursively applied. To investigate this hypothesis, empirical evaluation is conducted in a fully observable stochastic dynamical system using a minimal realisable linear Gaussian model. The empirical results show that purely one-step training leads to long-horizon degradation, including accumulation of trajectory error and systematic underestimation of predictive uncertainty. Introducing explicit multi-step regularisation significantly improves long-horizon stability and uncertainty calibration, and the improvement persists beyond the training horizon. The proposed formulation establishes a basis for modelling action impact in stochastic dynamical systems and provides a machine-learning framework for robust modelling of recursively applied transitions. This provides a foundation for further research in partially observable environments, nonlinear architectures, and decision-support systems.

Key words: *stochastic dynamical systems; action impact operator; recursive composition; multi-step regularisation; uncertainty calibration; long-horizon stability; machine learning.*

Заіка Б. Ю., Єршов С. В. Стохастичні оператори впливу дій: формалізація та багатокрокове регуляризоване навчання

У статті запропоновано формалізацію впливу дій у стохастичних динамічних системах у вигляді окремого стохастичного оператора, що діє на стани системи. Точне моделювання впливу дій є важливою проблемою послідовного прийняття рішень за умов невизначеності, оскільки в багатьох реальних системах дії застосовуються повторно, а їхні наслідки поширюються через динаміку системи з плином часу. Хоча сучасні підходи машинного навчання, зокрема навчання з підкріпленням та оцінювання умовних щільностей розподілу, здатні апроксимувати короткострокові переходи, поведінка навчених моделей при рекурсивному багатокроковому застосуванні залишається недостатньо дослідженою. У більшості існуючих підходів динаміка переходів інтегрована в задачі оптимізації політики або прогнозування траєкторій і рідко розглядається як самостійний об'єкт моделювання. У запропонованому підході оператор впливу дії відображає поточний стан системи та застосовану дію в умовний розподіл майбутніх станів і визначається з явною композиційною структурою. Це дозволяє аналізувати рекурсивне застосування оператора протягом кількох кроків часу. Запропоновано цільову функцію навчання, яка поєднує однокрокову негативну логарифмічну правдоподібність із додатковим членом багатокрокової узгодженості, отриманим із композиції оператора. Центральна гіпотеза дослідження полягає в тому, що однокрокове навчання за принципом максимальної правдоподібності не гарантує стабільної довгострокової поведінки у випадку рекурсивного застосування навченого оператора. Для перевірки цієї гіпотези проведено емпіричне дослідження у повністю спостережуваній стохастичній динамічній

системі з використанням мінімальної реалізовної лінійної гаусівської моделі. Емпіричні результати показують, що однокрокове навчання призводить до суттєвої деградації багатокрокових прогнозів, зокрема до накопичення похибки траєкторії та систематичної недооцінки прогностичної невизначеності. Запровадження явної багатокрокової регуляризації суттєво покращує довгострокову стабільність і калібрування невизначеності, причому позитивний ефект зберігається за межами горизонту навчання. Запропонована формалізація встановлює основи моделювання впливу дій у стохастичних динамічних системах і пропонує машинно-навчальну основу для надійного моделювання рекурсивно застосовуваних переходів. Це створює підґрунтя для подальших досліджень у частково спостережуваних середовищах, нелінійних архітектурах та системах підтримки прийняття рішень.

Ключові слова: *стохастичні динамічні системи, оператор впливу дії, рекурсивна композиція, багатокрокова регуляризація, калібрування невизначеності, довгострокова стабільність, машинне навчання.*

Formulation of the problem. Complex dynamical systems operating under uncertainty arise in artificial intelligence [1], project management [2], epidemic modeling [3], student profiling [4], friction brake systems [5], inventory management [6] and other well-known fields. In such systems, actions are applied repeatedly over time, and their consequences propagate through stochastic dynamics. A central difficulty is not only the selection of locally effective decisions, but the understanding of how action effects accumulate under recursive application of the transition mechanism. Questions of uncertainty propagation, stability, and long-term behaviour have been studied in dynamic optimisation and decision-support contexts [1–6].

In modern data-driven settings, however, the transition mechanism is rarely known a priori and must be inferred from observational or experimental data. This shifts the problem from analytical specification of dynamics to statistical learning of action-dependent stochastic transitions. Within machine learning, such tasks are typically addressed through predictive modelling, reinforcement learning, or conditional density estimation. Yet even in these formulations, the learned transition mechanism is commonly embedded within broader objectives such as return maximisation or trajectory reconstruction [7–9].

This paper adopts a different perspective. Rather than treating action-conditioned dynamics as an internal component of a larger predictive or control framework, it considers action impact as a dedicated stochastic operator acting on system states. The central object of interest is therefore not a policy or value function, but a parametric operator that maps the current state and applied action to a conditional distribution over future states. This makes it possible to study not only immediate action effects, but also the behaviour of recursively composed operators over multiple time steps.

The purpose of the article is to formalise and investigate the action impact operator as a standalone stochastic mapping that governs the distributional evolution of system states under applied actions. Specifically, the goals of the paper are:

1. to introduce a general theoretical framework in which the effect of an action is represented as a stochastic operator with explicit compositional structure;
2. to formulate a learning objective that enforces both one-step distributional fidelity and multi-step consistency under recursive operator composition;
3. to evaluate, in a fully observable setting, whether explicit regularisation under recursive composition improves long-horizon stability and uncertainty calibration compared to purely one-step estimation.

By achieving these objectives, the study aims to establish methodological foundations for operator-level modelling of action impact in stochastic dynamical systems, providing a basis for subsequent extensions toward partially observable environments and broader decision-support applications.

Analysis of recent research and publications. Understanding how actions impact the evolution of a dynamical system is central across control theory, reinforcement learning, and causal modeling. However, most existing approaches embed action-dependent transitions into broader optimisation objectives, such as trajectory prediction or policy optimisation, rather than treating the action-induced stochastic transformation as an autonomous object of estimation. To clarify the conceptual position of the proposed approach, this section reviews existing frameworks that model action-dependent dynamics.

State-space modelling (SSM) remains one of the principal paradigms for controlled dynamical systems. In general form, such systems are represented as

$$x_{k+1} = f(x_k, u_k, w_k), \quad y_k = g(x_k, v_k), \tag{1}$$

where x_k denotes the latent state, u_k is the control input, y_k is the observed output and w_k, v_k represent stochastic factors. Recent developments extend this framework through neural and latent-variable parameterisations. Forgione and Piga [10] develop Bayesian neural SSMs for nonlinear system identification with posterior predictive uncertainty. Hu et al. [11] reinterpret SSM architectures as scalable neural operators suitable for long-horizon and chaotic dynamics. Volkmann et al. [12] focus on generative reconstruction of latent dynamical geometry from short time series. Despite methodological differences, these approaches share a predictive orientation: they aim to accurately reconstruct trajectories or system evolution under observed inputs, without isolating the stochastic operator of action impact as an independent modelling entity.

Model-based reinforcement learning (MBRL) integrates learned dynamics directly into policy optimisation. World-model approaches such as DreamerV3 [13] employ recurrent latent SSMs trained via reconstruction and reward objectives, enabling imagination-based planning entirely in latent space. Hybrid recurrent state-space formulations under exogenous noise [14] further combine representation learning with actor–critic optimisation, while uncertainty-aware Dyna-style algorithms such as MACURA [15] adapt rollout length based on epistemic uncertainty. Although these methods learn transition models, the dynamics component remains instrumental: its value is measured by improvement in expected return rather than by the fidelity of the action-induced stochastic transformation itself. Thus, even in uncertainty-aware or latent formulations, the learned model is not interpreted as a standalone operator mapping actions to distributions over future states.

For partially observable systems, belief-state approaches propagate uncertainty over latent states conditioned on observation–action history. Planning in belief space is rigorously analysed by Barenboim et al. [16], who demonstrate that common pruning strategies in Monte Carlo planning may bias value estimation in hybrid belief partially observable Markov decision processes. Arcieri et al. [17] address model inference and robust policy learning under parameter uncertainty, combining posterior sampling of transition and observation models with deep reinforcement learning and domain randomisation. Although these approaches explicitly maintain probabilistic beliefs and propagate uncertainty, belief states ultimately serve policy optimisation and value estimation, rather than being interpreted as representations of action-induced transformations of system dynamics.

A more structural perspective on interventions arises in dynamical extensions of causal modelling. Peters et al. [18] formulate causal kinetic models for continuous-time systems, separating intervention mechanisms from stochastic disturbance structure. Lozano-Durán and Arranz [19] develop an information-theoretic view of causality and control based on entropy and information flux, while causal reinforcement-learning formulations [20] incorporate causal structure into transition and decision models to improve robustness and generalisation. Although these approaches clarify intervention semantics, they do not usually formulate action impact itself as a separately learned stochastic operator over future-state distributions.

Complementary Bayesian approaches to sequential decision-making under limited data, such as the pessimistic framework for dynamic treatment regimes proposed by Zhou et al. [21], emphasise uncertainty quantification and conservative policy evaluation in offline reinforcement learning. By constructing posterior-based lower confidence bounds for value functions, such methods provide regret guarantees under incomplete coverage. This improves uncertainty-aware policy evaluation, but the modelling emphasis remains on value estimation and policy choice rather than on direct estimation of the stochastic transformation induced by actions.

Thus, across state-space modelling, model-based reinforcement learning, causal dynamical analysis, belief-based inference, and Bayesian offline decision-making, action-dependent transitions are usually embedded within prediction, planning, or evaluation pipelines. In contrast, the present work considers the action impact mechanism itself as a standalone stochastic operator and studies its learning under recursive composition.

Theoretical formalisation of the action impact operator. The analysis of existing approaches indicates that, despite substantial progress in modelling dynamical systems, reinforcement learning, belief-state inference, and causal reasoning, the effect of an action is rarely formalized as an autonomous stochastic operator acting on system evolution. To address this gap, we adopt a formulation in which the primary object of interest is not a policy or value function, but a stochastic operator of action impact governing the transformation of system states over a finite time interval.

In general, real dynamical systems may be only partially observable. However, in the present study, we restrict attention to the fully observable setting. This assumption is adopted deliberately in order to isolate and analyse the properties of the stochastic action impact operator itself, without introducing additional modelling complexity related to state inference. Treatment of partial observability, belief-state construction, and learned filtering mechanisms constitutes a distinct methodological direction and lies beyond the scope of the current work. These aspects will be investigated in subsequent studies as a natural continuation of the proposed framework, where the foundations established here will be extended to latent-state and partially observable environments. By focusing on the fully observable case, we ensure that the operator learning principles can be examined in a controlled and conceptually transparent setting before addressing the additional challenges of state estimation.

Let $x \in X$ denote the state of a dynamical system and $u(t) \in U$ is the applied action. The full system dynamics can be expressed in general form as

$$x(t + \Delta t) = f(x(t), u(t), \xi(t)), \tag{2}$$

where $\xi(t)$ represents exogenous stochastic factors. Instead of directly modelling the function f, we introduce the action impact operator $\mathcal{I}_{\Delta t}$, defined on a fixed time interval Δt, such that

$$x(t + \Delta t) \sim \mathcal{I}_{\Delta t}(\cdot \mid x(t), u(t)). \tag{3}$$

Here, $\mathcal{I}_{\Delta t}(\cdot \mid x, u)$ denotes the conditional distribution of the future state induced by action u applied in state x. In this formulation, the action is interpreted not merely as an input variable, but as a generator of a stochastic transformation of the state distribution.

The cumulative effect of actions over a horizon $H = m\Delta t$ is described through operator composition

$$\mathcal{I}_H = \mathcal{I}_{\Delta t}^{(m)} = \underbrace{\mathcal{I}_{\Ädt} \circ \ldots \circ \mathcal{I}_{\Ädt}}_{m \text{ operators}}, \tag{4}$$

which captures the trajectory-level impact of sequential interventions.

In practice, however, the action impact operator $\mathcal{I}_{\Delta t}$ is not directly accessible. Instead, we are given data $\mathcal{I}$ generated by the system under a sequence of applied actions:

$$\mathcal{I} = \{x(t), u(t)\}_{t=0}^{T}. \tag{5}$$

These data consist of time-indexed records of system evolution, from which the underlying transition mechanism must be inferred. The learning objective is therefore to construct a parametric approximation $\widehat{\mathcal{I}}_{\Delta t}^{\theta}$ of the action impact operator $\mathcal{I}_{\Delta t}$ such that, for each admissible pair $(x(t), u(t))$, it reproduces the conditional distribution of action-induced consequences over the interval Δt:

$$x(t + \Delta t) \sim \widehat{\mathcal{I}}_{\Delta t}^{\theta}(\cdot \mid x(t), u(t)). \tag{6}$$

Here, θ denotes the model parameters to be estimated from data. The resulting model defines a predictive distribution over future states and serves as a learned approximation of the unknown impact mechanism.

Beyond one-step transitions, it is essential that the learned operator preserves consistency under temporal composition. For a horizon $H = m\Delta t$, the model-induced multi-step operator is defined analogously to (4) as

$$x(t + H) \sim \widehat{\mathcal{I}}_{H}^{\theta}(\cdot \mid x(t), u(t : t + h)), \quad \widehat{\mathcal{I}}_{H}^{\theta} = \underbrace{\widehat{\mathcal{I}}_{\Ädt}^{\theta} \circ \ldots \circ \widehat{\mathcal{I}}_{\Ädt}^{\theta}}_{m \text{ operators}}, \tag{7}$$

and must yield coherent predictions of the system state distribution at $t + H$ time under a sequence of actions $u(t : t + h)$, where $h = H - \Delta t$. Ensuring such multi-step consistency constitutes a requirement of the proposed operator-learning framework.

The problem addressed in this work can be stated as follows: given data $\mathcal{I}$, construct a stochastic operator model $\widehat{\mathcal{I}}_{\Delta t}^{\theta}$ whose one-step and multi-step compositions accurately capture the distributional impact of actions on system evolution.

The above formulation establishes the object of learning at a conceptual level. To make this objective precise, the notion of accurate capture must be expressed in distributional terms.

Let $\mathbb{D}(\cdot\|\cdot)$ denote a divergence between probability measures. The one-step approximation quality is characterized by the divergence functional

$$\mathcal{D}_{\Delta t}(\theta) = \mathbb{E}\left[\mathbb{D}\left(\mathcal{I}_{\Delta t}(\cdot|x(t), u(t)) \,\|\, \widehat{\mathcal{I}}_{\Delta t}^{\theta}(\cdot|x(t), u(t))\right)\right]. \tag{8}$$

This functional quantifies how well the learned operator reproduces the conditional distribution of the next state induced by an action over a single time increment Δt.

In addition to one-step fidelity, the learned operator must remain consistent under temporal composition. Using the composed operators $\mathcal{I}_H$ and $\widehat{\mathcal{I}}_H^{\theta}$ defined in (4), (7), the cumulative discrepancy over a horizon $H = m\Delta t$ is described by the divergence functional

$$\mathcal{D}_H(\theta) = \mathbb{E}\left[\mathbb{D}\left(\mathcal{I}_H(\cdot|x(t), u(t : t + h)) \,\|\, \widehat{\mathcal{I}}_H^{\theta}(\cdot|x(t), u(t : t + h))\right)\right]. \tag{9}$$

Controlling both $\mathcal{D}_{\Delta t}(\theta)$ and $\mathcal{D}_H(\theta)$ ensures that the learned action impact operator captures not only local stochastic transitions but also the distributional consequences of sequential actions over extended horizons.

Training objective. As was mentioned before, the true operators $\mathcal{I}_{\Delta t}$ and $\mathcal{I}_H$ are unknown. Therefore, the divergence functionals defined in (8) and (9) cannot be evaluated directly. In practice, learning must rely on the dataset $\mathcal{I}$ introduced in (5).

For notational convenience and compactness, the discrete-time representation of dataset $\mathcal{I}$ is introduced:

$$\mathcal{I} = \{(x_k, u_k, x_{k+1})\}_{k=0}^{K-1}, \tag{10}$$

where $x_k := x(t_k)$, $u_k := u(t_k)$, $t_k = k\Delta t$, $k = 0, 1, \ldots, K$.

The theoretical divergence functional $\mathcal{D}_{\Delta t}(\theta)$ is then approximated empirically by replacing the expectation over the unknown data-generating distribution with an average over observed transitions. When the divergence $\mathbb{D}$ is instantiated as the Kullback–Leibler divergence between conditional distributions, the corresponding empirical objective reduces to the negative log-likelihood of the observed next states under the learned operator model. The corresponding empirical one-step objective $\widehat{\mathcal{L}}_{\Delta t}(\theta)$ is defined as

$$\widehat{\mathcal{L}}_{\Delta t}(\theta) = -\frac{1}{K}\sum_{k=0}^{K-1} \log \widehat{\mathcal{I}}^{\theta}_{\Delta t}(x_{k+1} \mid x_k, u_k). \tag{11}$$

Multi-step consistency is incorporated by evaluating how well the learned operator assigns probability to observed state outcomes over horizon $H = m\Delta t$. Concretely, for each starting index k, we consider the observed action subsequence $u_{k:k+m-1}$ and the corresponding observed future state x_{k+m} available in the trajectory. The learned one-step operator $\widehat{\mathcal{I}}^{\theta}_{\Delta t}$ is applied recursively (analogously to the composition in (10)) to obtain the model-implied conditional distribution of x_{k+m} given the initial state x_k and the action sequence $u_{k:k+m-1}$. The empirical multi-step objective is then defined through the negative log-likelihood of the observed horizon state under this composed operator model. The corresponding multi-step objective $\widehat{\mathcal{L}}_H(\theta)$ is given by

$$\widehat{\mathcal{L}}_H(\theta) = -\frac{1}{K-m}\sum_{k=0}^{K-m-1} \log \widehat{\mathcal{I}}^{\theta}_{H}(x_{k+m} | x_k, u_{k:k+m-1}). \tag{12}$$

The overall training objective $\widehat{\mathcal{L}}(\theta)$ is defined as a weighted combination of the one-step $\widehat{\mathcal{L}}_{\Delta t}(\theta)$ and multi-step $\widehat{\mathcal{L}}_H(\theta)$ objectives:

$$\widehat{\mathcal{L}}(\theta) = \widehat{\mathcal{L}}_{\Delta t}(\theta) + \lambda \widehat{\mathcal{L}}_H(\theta), \tag{13}$$

where $\lambda \geq 0$ balances one-step accuracy and multi-step stability.

The learning problem is formulated as:

$$\min_{\theta} \widehat{\mathcal{L}}(\theta). \tag{14}$$

Minimal realisable stochastic operator model. To study the learning behaviour of the action impact operator under controlled conditions, we consider a minimal realisable parametric model. The goal of this construction is not to build the most expressive model, but to isolate the structural properties of operator learning and recursive composition in a setting where the data-generating mechanism can be represented within the chosen model class.

In the experiments, the conditional distribution of the next state given the current state and action is instantiated as a Gaussian stochastic operator. Specifically, the learned operator defines a conditional distribution

$$\widehat{\mathcal{I}}^{\theta}_{\Delta t}(\cdot \mid x_k, u_k) = \mathcal{N}(\mu_\theta(x_k, u_k), \Sigma_\theta), \tag{15}$$

where $\mu_\theta(x_k, u_k) \in \mathbb{R}^d$ and $\Sigma_\theta \in \mathbb{R}^{d\times d}$ denote the mean and covariance of the distribution predicted by a parametric model with parameters θ, d is the dimension of the state space. Under this formulation, the learned operator defines a probabilistic transition mechanism that maps each state–action pair (x_k, u_k) to a distribution over future states. The empirical training objective introduced in (13)–(14) then corresponds to maximising the likelihood of observed transitions under this conditional distribution while enforcing multi-step consistency through recursive operator application.

The Gaussian formulation provides a simple and analytically tractable instantiation of the stochastic operator while remaining sufficiently expressive for the synthetic environment considered in this study. It is important to note that the proposed operator-learning framework is not limited to the specific choice of conditional distribution. More expressive density models could be incorporated within the same framework, but their investigation lies beyond the scope of the present work.

Multi-step rollout evaluation metrics. To ensure an unbiased assessment of compositional behaviour, training and evaluation are performed on disjoint datasets. The dataset $\mathcal{T}(10)$ is partitioned into two subsets, within which m-step segments are formed so that no segment crosses the split boundary:

$$\mathcal{T}_m^{train} = \left\{\left(x_k, u_{k:k+m-1}, x_{k+m}\right)\right\}_{k=0}^{K_{split}-m}, \tag{16}$$

$$\mathcal{T}_m^{valid} = \left\{\left(x_k, u_{k:k+m-1}, x_{k+m}\right)\right\}_{k=K_{split}+1}^{K-m}, \tag{17}$$

where $0 \leq K_{split} \leq K - m$, K_{split} is the threshold for dividing $\mathcal{T}$ dataset in training and validation datasets.

The training objective $\widehat{\mathcal{L}}(\theta)$ is optimised using only transitions from $\mathcal{T}_m^{train}$, while all horizon $H = m\Delta t$ performance metrics are computed exclusively on $\mathcal{T}_m^{valid}$. This separation prevents optimistic bias in the evaluation of multi-step degradation.

The horizon-wise negative log-likelihood is then defined as

$$NLL^{valid}(m) = -\frac{1}{K_m^{valid}}\sum_{k=K_{split}+1}^{K-m} \log \widehat{\mathcal{I}}^{\theta}_{H}(x_{k+m} \mid x_k, u_{k:k+m-1}), \tag{18}$$

where $K_m^{valid} = K - m - K_{split}$ is the number of samples for evaluation.

To further analyse deterministic drift accumulation, the horizon-wise mean rollout error on $\mathcal{T}_m^{valid}$ is computed:

$$RMSE^{valid}(m) = \sqrt{\frac{1}{K_m^{valid}} \sum_{k=K_{split}+1}^{K-m} \left\| \delta_{k,m} \right\|^2}, \tag{19}$$

where $\delta_{k,m} = x_{k+m} - \mu_{H,\theta}(x_k, u_{k:k+m-1})$, $\|\cdot\|$ is the Euclidean norm in $\mathbb{R}^d$.

To additionally assess calibration of predictive uncertainty, we calculate the Prediction Interval Coverage Probability (PICP) computed on $\mathcal{T}^{valid}$. For each validation segment, we consider the squared Mahalanobis distance

$$D_{k,m}^2 = \delta_{k,m}^T \Sigma_{m,\theta}^{-1} \delta_{k,m}, \tag{20}$$

and compare it to the $\mathcal{X}_{\nu,0.95}^2$, where $\mathcal{X}_{\nu,0.95}^2$ is the 95 % quantile of the chi-squared distribution with ν degrees of freedom. The horizon-wise PICP is then defined as

$$PICP_{0.95}^{valid}(m) = \frac{1}{K_m^{valid}} \sum_{k=K_{split}+1}^{K-m} 1\left\{ D_{k,m}^2 \le \mathcal{X}_{\nu,0.95}^2 \right\}, \tag{21}$$

where $1\{\cdot\}$ denotes the indicator function.

By evaluating $NLL^{valid}(m)$, $RMSE^{valid}(m)$ and $PICP_{0.95}^{valid}(m)$ for increasing m we obtain complementary perspectives on compositional behaviour: distributional accuracy, mean trajectory stability, and uncertainty calibration under recursive application. Since all metrics are computed strictly on $\mathcal{T}_m^{valid}$, these profiles reflect genuine generalisation behaviour rather than training-set memorisation effects. The subsequent section reports empirical results based on this validation protocol.

Empirical evaluation. The purpose of the empirical study is to evaluate the compositional behaviour of the learned stochastic action impact operator under recursive application. In particular, we examine horizon-wise distributional degradation, accumulation of trajectory error, and uncertainty calibration.

All experiments are conducted on a controlled synthetic environment given by a fully observable linear Gaussian dynamical system

$$x_{k+1} = Ax_k + Bu_k + \varepsilon_k,\ \varepsilon_k \sim \mathcal{N}(0,\Sigma), \quad \Sigma = \begin{pmatrix} 0.05 & 0.02 \\ 0.02 & 0.04 \end{pmatrix}, \tag{22}$$

where $x_k \in \mathbb{R}^2$; $u_k \in \mathbb{R}$; A is a randomly generated transition matrix rescaled to have spectral radius approximately equal to 0.995; B is the control matrix sampled from a zero-mean Gaussian distribution and scaled to moderate magnitude; Σ is the process noise covariance.

A single trajectory of length $K = 2000$ transitions is generated. Actions are piecewise constant over randomly sampled segments of length between 5 and 25 time steps. The dataset is split chronologically according to (16)-(17) with $K_{split} = 1500$.

The parameterisation of the synthetic environment (22) is chosen to make recursive forecasting non-trivial while preserving full analytical control. Rescaling the transition matrix A to spectral radius approximately 0.995 places the system near the stability boundary, where small one-step inaccuracies can accumulate over long rollouts. Piecewise-constant actions generate temporally coherent interventions rather than isolated random perturbations, making multi-step action impact structurally observable. Moderate cross-dimensional noise correlation further prevents the problem from becoming artificially axis-aligned and yields a more informative test of distributional propagation.

The learned operator $\widehat{\mathcal{I}}_{\Delta t}^{\theta}$ belongs to the minimal realisable linear Gaussian class introduced in (15), so the experiments isolate the effect of the training objective rather than model misspecification. The spectral norm of A_θ is constrained during optimisation in order to prevent trivial divergence.

Parameter estimation minimises the composite objective $\widehat{\mathcal{L}}(\theta)$ defined in (13), where the weight λ controls the contribution of multi-step consistency.

Evaluation is performed exclusively on the validation trajectory. For evaluation horizons $m_{valid} = 1, \ldots, 30$, we compute multi-step *NLL* (18), *RMSE* (19) and $PICP_{0.95}$ (21), capturing distributional degradation, trajectory drift, and uncertainty calibration, respectively.

The empirical analysis proceeds in two stages. First, a compact comparative table reports metric values at three evaluation horizons $m_{valid} \in \{1; 15; 30\}$, enabling simultaneous inspection of short-, medium- and long-horizon behaviour across combinations of $\lambda \in \{0; 0.5; 1; 2\}$ and $m_{train} \in \{5; 10; 20\}$ ($\lambda = 0$ serves as a baseline without compositional regularisation). Second, default λ and m_{train} values are selected based on this table, and full degradation profiles are visualised as metric-horizon line curves, where the vertical axis represents the metric value and the horizontal axis denotes the evaluation horizon. For fixed λ, curves corresponding to different m_{train} are compared, while for fixed m_{train}, varying λ isolates the influence of multi-step regularisation strength.

Table 1 shows validation metrics at three representative evaluation horizons $m_{valid} \in \{1; 15; 30\}$ for $\lambda \in \{0; 0.5; 1; 2\}$ and $m_{train} \in \{5; 10; 20\}$.

Table 1

Validation metrics at $m_{valid} \in \{1; 15; 30\}$ for one-step ($\lambda = 0$) and multi-step ($\lambda > 0$) training with different m_{train}

m_{valid}	1			15			30		
m_{train}	**5**	**10**	**20**	**5**	**10**	**20**	**5**	**10**	**20**
NLL									
$\lambda = 0$	0.906			4.393			4.395		
$\lambda = 0.5$	1.173	1.285	1.370	3.477	3.277	3.274	3.901	3.686	3.644
$\lambda = 1$	1.229	1.349	1.404	3.326	3.174	3.218	3.754	3.591	3.608
$\lambda = 2$	1.281	1.443	1.463	3.217	3.072	3.123	3.661	3.516	3.555
RMSE									
$\lambda = 0$	0.449			1.615			1.623		
$\lambda = 0.5$	0.431	0.451	0.463	1.381	1.265	1.206	1.597	1.548	1.512
$\lambda = 1$	0.444	0.474	0.483	1.306	1.154	1.113	1.570	1.474	1.438
$\lambda = 2$	0.460	0.494	0.501	1.224	1.071	1.048	1.527	1.393	1.366
$PICP_{0.95}$									
$\lambda = 0$	0.9873			0.6751			0.6581		
$\lambda = 0.5$	0.9915	0.9873	0.9915	0.8981	0.966	0.9894	0.8153	0.9278	0.9723
$\lambda = 1$	0.9872	0.9745	0.9787	0.9469	0.9873	1	0.8874	0.966	0.9872
$\lambda = 2$	0.9787	0.9214	0.9384	0.9681	0.9873	0.9957	0.9256	0.9787	0.9915

At the one-step level $m_{valid} = 1$, the baseline model ($\lambda = 0$) achieves the best local accuracy, with NLL equal to 0.906 and RMSE equal to 0.449. Adding multi-step regularisation results in slightly higher one-step NLL values, ranging from 1.173 to 1.463 across $\lambda > 0$ configurations. RMSE values remain comparable and in some cases are even marginally lower (e.g., RMSE = 0.431 for $\lambda = 0.5$, $m_{train} = 5$). These results indicate a modest trade-off in local one-step optimality when explicit multi-step consistency is enforced. Prediction interval coverage remains close to the nominal 0.95 level at this horizon for most configurations. However, for larger regularisation and training horizons ($\lambda = 2$, $m_{train} \in \{10; 20\}$) $PICP_{0.95}$ decreases from 0.9873 (baseline $\lambda = 0$ model) to 0.92–0.94, suggesting that excessive regularisation may influence short-horizon calibration.

At the intermediate horizon $m_{valid} = 15$, the compositional differences become pronounced. The purely one-step model exhibits NLL of 4.393 and RMSE of 1.615, while its calibration drops sharply to $PICP_{0.95} = 0.675$. In contrast, multi-step training reduces distributional degradation, with NLL values decreasing to about 3.07–3.48 and RMSE values dropping to 1.04–1.38 across $\lambda > 0$. Most importantly, $PICP_{0.95}$ substantially improves under multi-step training, reaching values between 0.898 and 1.00 depending on the configuration. Even under the worst observed multi-step configuration, coverage increases by more than 0.22 relative to the baseline. The dominant discrepancy at this horizon is therefore the severe under-dispersion exhibited by the one-step model.

At the long horizon $m_{valid} = 30$, trends seen in $m_{valid} = 15$ horizon amplify further. The baseline model exhibits NLL of 4.395, RMSE of 1.623, and $PICP_{0.95}$ of only 0.658, indicating persistent underestimation of predictive uncertainty under recursive application. Multi-step training mitigates these effects across all m_{train} and $\lambda > 0$, with the most notable difference again in $PICP_{0.95}$ values, where minimal improvement from baseline model equal 0.157.

Overall, Table 1 reveals a clear and systematic pattern. Optimising exclusively the one-step objective yields slightly superior short-horizon likelihood but leads to substantial degradation under recursive composition, particularly in uncertainty calibration, where coverage drops from approximately 0.95 to 0.66 by $m_{valid} = 30$. Introducing explicit multi-step regularisation modestly increases local loss yet substantially improves long-horizon distributional stability and restores probabilistic calibration. These results empirically demonstrate that one-step optimality does not imply multi-step compositional robustness.

Based on the analysis of Table 1, we select $\lambda = 0.5$ and $m_{train} = 5$ as default hyperparameter values for detailed degradation analysis. This configuration represents a minimally regularised regime that yields substantial improvements in long-horizon stability and uncertainty calibration, while adding only modest degradation in one-step accuracy. Moreover, the choice of a relatively short training horizon allows us to examine whether compositional improvements generalise beyond the horizon explicitly enforced during training. This provides a conservative and structurally informative setting for analysing recursive degradation profiles.

Figure 1 presents detailed degradation profiles of the validation metrics as functions of the evaluation horizon $m_{valid} \in \{1, \ldots, 30\}$. The figure contains six panels arranged in two columns and three rows, with rows corresponding

to NLL, RMSE, and $PICP_{0.95}$. The left column fixes $m_{train} = 5$ and compares different values λ, thereby isolating the effect of multi-step regularisation strength. The right column compares different training horizons m_{train} at fixed $\lambda = 0.5$.

When analysing NLL as a function of m_{valid} (the top left graph in Figure 1), the baseline model ($\lambda = 0$) exhibits rapid growth, reaching around 4.4 by $m_{valid} = 30$. Introducing multi-step regularisation substantially reduces this degradation. Increasing λ flattens the NLL curve, with $\lambda = 2$ producing the slowest growth across horizons and achieving similar values as $\lambda = 1$. This confirms that even short-horizon multi-step enforcement ($m_{train} = 5$) significantly improves compositional stability.

A similar pattern is observed for RMSE (the middle left graph in Figure 1). The baseline trajectory error grows quickly, whereas larger λ values systematically reduce long-horizon drift. The improvement is monotonic in λ, suggesting that stronger compositional regularisation directly mitigates accumulation of mean error under recursive application.

The most striking effect appears in the $PICP_{0.95}$ panel (the bottom left graph in Figure 1). For $\lambda = 0$, coverage drops sharply below the nominal 0.95 level after $m_{valid} = 3$ and stabilises near 0.66, indicating systematic underestimation of predictive uncertainty. In contrast, increasing λ dramatically improves calibration. While $\lambda = 0.5$ still shows gradual decline at longer horizons, $\lambda = 1$ and $\lambda = 2$ maintain coverage much closer to the target level across all horizons. This demonstrates that multi-step regularisation primarily corrects uncertainty miscalibration induced by recursive composition.

Overall, left column of Figure 1 shows that increasing λ improves long-horizon stability and calibration, while preserving reasonable short-horizon accuracy. Although training is enforced only up to $m_{train} = 5$, the most pronounced reductions in NLL and RMSE are observed in the range $m_{valid} \approx 10–15$, and for $PICP_{0.95}$ in the range $m_{valid} \approx 5–15$. This indicates that compositional improvements generalise beyond the explicitly enforced training horizon and remain beneficial at longer evaluation horizons.

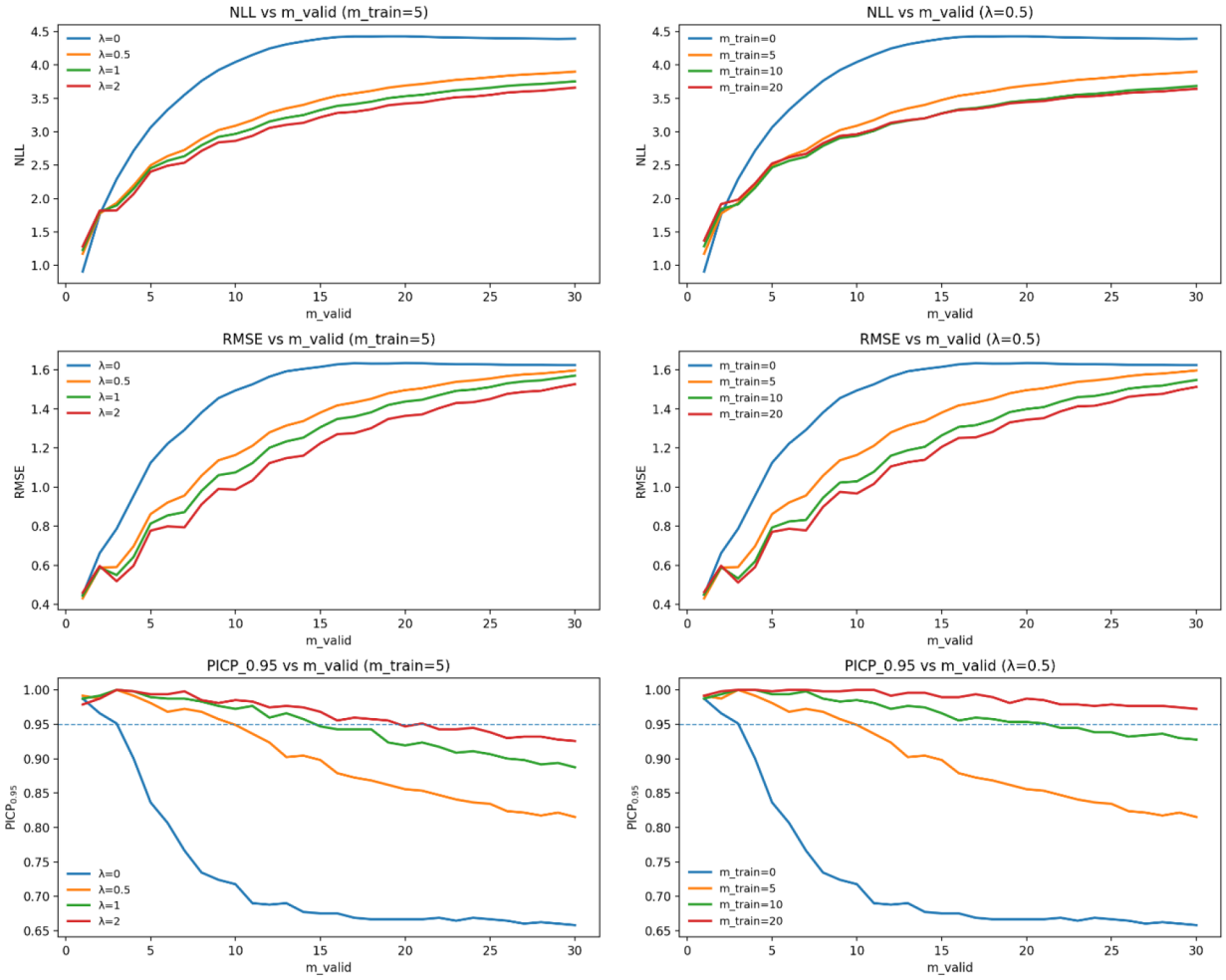

Fig. 1. Degradation profiles across validation horizons m_{valid} for varying regularisation strength λ at fixed $m_{train} = 5$ (left) and varying training horizons m_{train} at fixed $\lambda = 0.5$ (right)

The right column of Figure 1 analyses how the training horizon influences degradation when the regularisation strength is fixed at a moderate value ($\lambda = 0.5$). The curve labelled $m_{train} = 0$ corresponds to the baseline one-step model and provides a direct point of reference.

For NLL, increasing m_{train} consistently reduces long-horizon degradation. While the baseline model ($m_{train} = 0$) rapidly approaches values above 4.3, larger training horizons (10 and 20) produce noticeably flatter curves, even beyond the enforced training range. Importantly, improvements remain visible well beyond m_{train}, indicating partial generalisation of compositional consistency.

The RMSE curves exhibit analogous behaviour. Larger m_{train} values lead to reduced accumulation of trajectory drift, especially beyond $m_{valid} \approx 10$. The difference between $m_{train} = 5$ and $m_{train} = 10$ is already substantial, while further increase to $m_{train} = 20$ yields additional but diminishing improvements.

In terms of calibration, the baseline model shows severe under-coverage, stabilising near 0.66. With $\lambda = 0.5$, increasing m_{train} significantly mitigates this collapse. Although coverage gradually declines with horizon, larger training horizons remain closer to the nominal 0.95 level. This suggests that multi-step consistency over longer segments enhances uncertainty calibration under recursive rollout.

Thus, Figure 1 confirms the conclusions drawn from Table 1. The regularisation strength λ primarily controls the magnitude of compositional correction, whereas the training horizon m_{train} determines how far this correction generalises under recursive application. The baseline one-step model exhibits rapid metric degradation and severe uncertainty underestimation, while multi-step training improves long-horizon stability and probabilistic calibration.

Importantly, these improvements are not limited to horizons up to m_{train}. In many cases, the gap between the baseline and multi-step models continues to widen for $m_{valid} > m_{train}$. This suggests that multi-step regularisation induces structural adjustments to the learned operator that generalise beyond the explicitly optimised segment length. Overall, the results provide empirical evidence that one-step optimality does not imply compositional stability and that explicit training under operator composition yields more robust long-horizon behaviour.

Conclusions and directions for future work. This paper investigated the problem of learning stochastic action impact operators under recursive composition, with particular emphasis on the distinction between one-step optimality and multi-step compositional stability. We formalised the operator as a conditional stochastic mapping and introduced a training objective that augments the standard one-step negative log-likelihood with an explicit multi-step consistency term.

Experiments on a fully observable linear Gaussian system showed that one-step maximum-likelihood training alone does not ensure robust recursive behaviour. Although the baseline model achieves the best short-horizon likelihood, it undergoes rapid degradation and strong underestimation of predictive uncertainty under repeated application. In contrast, explicit multi-step regularisation improves long-horizon stability and calibration, and these gains persist beyond the enforced training horizon, indicating structural rather than merely local improvements in the learned operator.

Taken together, the theoretical formulation and empirical results support the central claim of the paper: one-step optimality is insufficient when the operator is intended for recursive use, and explicit training under composition is needed for stable and well-calibrated long-horizon behaviour.

Directions for future work include relaxing full observability through latent-state representations, extending the framework to nonlinear and high-dimensional settings, investigating adaptive weighting of the multi-step objective, and linking compositional operator to downstream sequential decision-making and decision-support tasks.

Bibliography:

1. Tsironis G. Artificial intelligence and complex dynamical systems. Cham: Springer, 2025. 296 p. (Understanding Complex Systems). https://doi.org/10.1007/978-3-031-81946-9

2. Симонов Д. І. Метод ентропії як інструмент оптимізації складних систем. Журнал обчислювальної та прикладної математики. 2024. № 1. С. 49–58. https://doi.org/10.17721/2706-9699.2024.1.04

3. Cheng C., Ichinose G., Small M., Moreno Y. Uncertainty quantification in complex dynamical systems. Physica D: Nonlinear Phenomena. 2025. Vol. 481. Art. 134838. https://doi.org/10.1016/j.physd.2025.134838

4. Poquet O., Jovanovic J., Pardo A. Student profiles of change in a university course: A complex dynamical systems perspective. In: Proceedings of the 13th International Learning Analytics and Knowledge Conference (LAK 2023). New York : ACM, 2023. P. 197–207. https://doi.org/10.1145/3576050.3576077

5. Geier C., Hamdi S., Chancelier T., Dufrénoy P., Hoffmann N., Stender M. Machine learning-based state maps for complex dynamical systems: Applications to friction-excited brake system vibrations. Nonlinear Dynamics. 2023. Vol. 111, No. 24. P. 22137–22151. https://doi.org/10.1007/s11071-023-08739-6

6. Симонов Д. І., Горбачук В. М. Метод пошуку рішень у динамічній моделі управління запасами за невизначеності. Вісник Київського національного університету імені Тараса Шевченка. Серія фізико-математичні науки. 2022. № 4. С. 31–39. https://doi.org/10.17721/1812-5409.2022/4.4

7. Li J., Guo S., Ma R., et al. Comparison of the effects of imputation methods for missing data in predictive modelling of cohort study datasets. BMC Medical Research Methodology. 2024. Vol. 24, No. 1. Art. 41. https://doi.org/10.1186/s12874-024-02173-x

8. Char I., Abbate J., Bardoczi L., et al. Offline model-based reinforcement learning for tokamak control. In: Proceedings of The 5th Annual Learning for Dynamics and Control Conference. Vol. 211. PMLR, 2023. P. 1357–1372.

9. Graffeuille O., Koh Y. S., Wicker J. S., Lehmann M. K. Semi-supervised conditional density estimation with Wasserstein Laplacian regularisation. Proceedings of the AAAI Conference on Artificial Intelligence. 2022. https://doi.org/10.1609/aaai.v36i6.20630

10. Forgione M., Piga D. Neural state-space models: Empirical evaluation of uncertainty quantification. IFAC-PapersOnLine. 2023. Vol. 56, No. 2. P. 4082–4087. https://doi.org/10.1016/j.ifacol.2023.10.1736

11. Hu Z., Ahmadi Daryakenari N., Shen Q., Kawaguchi K., Karniadakis G. E. State-space models are accurate and efficient neural operators for dynamical systems. Neural Networks. 2026. Vol. 197. Art. 108496. https://doi.org/10.1016/j.neunet.2025.108496

12. Volkmann E., Brändle A., Durstewitz D., Koppe G. A scalable generative model for dynamical system reconstruction from neuroimaging data. Advances in Neural Information Processing Systems. 2024. Vol. 37. P. 80328–80362.

13. Hafner D., Pasukonis J., Ba J., Lillicrap T. Mastering diverse control tasks through world models. Nature. 2025. Vol. 640, No. 8059. P. 647–653. https://doi.org/10.1038/s41586-025-08744-2

14. Sun R., Zang H., Li X., Islam R. Learning latent dynamic robust representations for world models. In: Proceedings of the 41st International Conference on Machine Learning (ICML 2024). PMLR, 2024.

15. Frauenknecht B., Eisele A., Devdutt S., Solowjow F., Trimpe S. Trust the model where it trusts itself: Model-based actor-critic with uncertainty-aware rollout adaption. In: Proceedings of the 41st International Conference on Machine Learning (ICML 2024). PMLR, 2024.

16. Barenboim M., Shienman M., Indelman V. Monte Carlo planning in hybrid belief POMDPs. IEEE Robotics and Automation Letters. 2023. Vol. 8, No. 8. P. 4410–4417. https://doi.org/10.1109/LRA.2023.3282773

17. Arcieri G., Hoelzl C., Schwery O., et al. POMDP inference and robust solution via deep reinforcement learning: An application to railway optimal maintenance. Machine Learning. 2024. Vol. 113, No. 10. P. 7967–7995. https://doi.org/10.1007/s10994-024-06559-2

18. Peters J., Bauer S., Pfister N. Causal models for dynamical systems. In: Probabilistic and Causal Inference: The Works of Judea Pearl. 2022. P. 671–690.

19. Lozano-Durán A., Arranz G. Information-theoretic formulation of dynamical systems: Causality, modeling, and control. Physical Review Research. 2022. Vol. 4, No. 2. Art. 023195. https://doi.org/10.1103/PhysRevResearch.4.023195

20. Zeng Y., Cai R., Sun F., Huang L., Hao Z. A survey on causal reinforcement learning. IEEE Transactions on Neural Networks and Learning Systems. 2024.

21. Zhou Y., Qi Z., Shi C., Li L. Optimizing pessimism in dynamic treatment regimes: A Bayesian learning approach. In: Proceedings of the 26th International Conference on Artificial Intelligence and Statistics (AISTATS 2023). PMLR, 2023. P. 6704–6721.

References:

1. Tsironis, G. (2025). Artificial intelligence and complex dynamical systems. Springer. https://doi.org/10.1007/978-3-031-81946-9

2. Symonov, D. I. (2024). Entropy method as a tool for optimization of complex systems. Journal of Computational and Applied Mathematics, (1), 49–58. https://doi.org/10.17721/2706-9699.2024.1.04

3. Cheng, C., Ichinose, G., Small, M., & Moreno, Y. (2025). Uncertainty quantification in complex dynamical systems. Physica D: Nonlinear Phenomena, 481, 134838. https://doi.org/10.1016/j.physd.2025.134838

4. Poquet, O., Jovanovic, J., & Pardo, A. (2023). Student profiles of change in a university course: A complex dynamical systems perspective. In Proceedings of the 13th International Learning Analytics and Knowledge Conference (pp. 197–207). ACM. https://doi.org/10.1145/3576050.3576077

5. Geier, C., Hamdi, S., Chancelier, T., Dufrénoy, P., Hoffmann, N., & Stender, M. (2023). Machine learning-based state maps for complex dynamical systems: Applications to friction-excited brake system vibrations. Nonlinear Dynamics, 111(24), 22137–22151. https://doi.org/10.1007/s11071-023-08739-6

6. Symonov, D. I., & Horbachuk, V. M. (2022). Method for finding solutions in a dynamic inventory management model under uncertainty. Bulletin of Taras Shevchenko National University of Kyiv. Series Physics and Mathematics, (4), 31–39. https://doi.org/10.17721/1812-5409.2022/4.4

7. Li, J., Guo, S., Ma, R., He, J., Zhang, X., Rui, D., Ding, Y., Li, Y., Jian, L., Cheng, J., & Guo, H. (2024). Comparison of the effects of imputation methods for missing data in predictive modelling of cohort study datasets. BMC Medical Research Methodology, 24(1), 41. https://doi.org/10.1186/s12874-024-02173-x

8. Char, I., Abbate, J., Bardoczi, L., Boyer, M., Chung, Y., Conlin, R., Erickson, K., Mehta, V., Richner, N., Kolemen, E., & Schneider, J. (2023). Offline model-based reinforcement learning for tokamak control. In Proceedings of the 5th Annual Learning for Dynamics and Control Conference (Vol. 211, pp. 1357–1372). PMLR.

9. Graffeuille, O., Koh, Y. S., Wicker, J. S., & Lehmann, M. K. (2022). Semi-supervised conditional density estimation with Wasserstein Laplacian regularisation. In Proceedings of the AAAI Conference on Artificial Intelligence. https://doi.org/10.1609/aaai.v36i6.20630

10. Forgione, M., & Piga, D. (2023). Neural state-space models: Empirical evaluation of uncertainty quantification. IFAC-PapersOnLine, 56(2), 4082–4087. https://doi.org/10.1016/j.ifacol.2023.10.1736

11. Hu, Z., Ahmadi Daryakenari, N., Shen, Q., Kawaguchi, K., & Karniadakis, G. E. (2026). State-space models are accurate and efficient neural operators for dynamical systems. Neural Networks, 197, 108496. https://doi.org/10.1016/j.neunet.2025.108496

12. Volkmann, E., Brändle, A., Durstewitz, D., & Koppe, G. (2024). A scalable generative model for dynamical system reconstruction from neuroimaging data. Advances in Neural Information Processing Systems, 37, 80328–80362.

13. Hafner, D., Pasukonis, J., Ba, J., & Lillicrap, T. (2025). Mastering diverse control tasks through world models. Nature, 640(8059), 647–653. https://doi.org/10.1038/s41586-025-08744-2

14. Sun, R., Zang, H., Li, X., & Islam, R. (2024). Learning latent dynamic robust representations for world models. In Proceedings of the 41st International Conference on Machine Learning. PMLR.

15. Frauenknecht, B., Eisele, A., Devdutt, S., Solowjow, F., & Trimpe, S. (2024). Trust the model where it trusts itself: Model-based actor-critic with uncertainty-aware rollout adaption. In Proceedings of the 41st International Conference on Machine Learning. PMLR.

16. Barenboim, M., Shienman, M., & Indelman, V. (2023). Monte Carlo planning in hybrid belief POMDPs. IEEE Robotics and Automation Letters, 8(8), 4410–4417. https://doi.org/10.1109/LRA.2023.3282773

17. Arcieri, G., Hoelzl, C., Schwery, O., Straub, D., Papakonstantinou, K. G., & Chatzi, E. (2024). POMDP inference and robust solution via deep reinforcement learning: An application to railway optimal maintenance. Machine Learning, 113(10), 7967–7995. https://doi.org/10.1007/s10994-024-06559-2

18. Peters, J., Bauer, S., & Pfister, N. (2022). Causal models for dynamical systems. In Probabilistic and causal inference: The works of Judea Pearl (pp. 671–690).

19. Lozano-Durán, A., & Arranz, G. (2022). Information-theoretic formulation of dynamical systems: Causality, modeling, and control. Physical Review Research, 4(2), 023195. https://doi.org/10.1103/PhysRevResearch.4.023195

20. Zeng, Y., Cai, R., Sun, F., Huang, L., & Hao, Z. (2024). A survey on causal reinforcement learning. IEEE Transactions on Neural Networks and Learning Systems.

21. Zhou, Y., Qi, Z., Shi, C., & Li, L. (2023). Optimizing pessimism in dynamic treatment regimes: A Bayesian learning approach. In Proceedings of the 26th International Conference on Artificial Intelligence and Statistics (Vol. 206, pp. 6704–6721). PMLR.

Дата першого надходження статті до видання: 10.03.2026
Дата прийняття статті до друку після рецензування: 30.03.2026
Дата публікації (оприлюднення) статті: 30.05.2026