Transactions on Computational and Scientific Methods | Vo. 6, No. 03, 2026
ISSN: 2998-8780
https://pspress.org/index.php/tcsm
Pinnacle Science Press

Data-Driven Modeling of Soft Robots via Koopman Operator Theory

Leif Corvinus
University of North Texas, Denton, Texas, USA
Leif Corvinus@gmail.com

Abstract: To address the problem of accurate modeling of soft robots, a data-driven modeling method based on the Koopman operator is proposed. The method aims to achieve global linearization of highly nonlinear dynamical systems in an infinite-dimensional lifted state space. By employing the Extended Dynamic Mode Decomposition (EDMD) algorithm, the infinite-dimensional linear Koopman operator is approximated, and both linear and nonlinear models of the soft robot dynamical system are established. The results demonstrate that the two Koopman-based models achieve higher accuracy than conventional state-space models, and the Koopman nonlinear model exhibits the best modeling and prediction performance.

Keywords: Soft robots, Koopman operator, data-driven modeling, nonlinear systems

1. Introduction

Soft robots have emerged as a new research field in recent years. They are mainly constructed from low-modulus soft materials or driven by electroactive materials. Compared with traditional rigid robots, these characteristics enable soft robots to exhibit natural compliance, low impedance, and strong adaptability. Owing to their unique properties and advantages, soft robots have found extensive applications in manipulation, medical treatment, wearable devices, and aerospace engineering [1-2]. However, the inherent nonlinearity and infinite degrees of freedom of soft robots pose significant challenges to modeling, which in turn limits their functionality. In addition, due to the continuous deformation and elastic behavior of soft robots, their dynamics are characterized by complex nonlinear properties. As a result, mature modeling techniques developed for traditional rigid robots cannot be directly applied to soft robots. Reliable and accurate models are therefore crucial for achieving precise control of soft robots, making accurate modeling an essential prerequisite.

Modeling approaches for soft robots can generally be divided into two categories: model-based methods and data-driven model-free methods. Model-based approaches rely on accurate or simplified approximate models of soft robots, but they suffer from limitations such as restricted applicability, high computational complexity, and relatively low modeling accuracy. Data-driven modeling methods mainly include machine learning and deep learning techniques [3]. These methods exhibit a certain degree of generality and can leverage their “black-box” characteristics to construct models by mapping inputs to outputs. Moreover, such approaches have been demonstrated to effectively predict the behavior of soft robots under various configurations. However, for subsequent controller design, the nonlinear models established by these methods are often not suitable for linear control strategies that are simple and technically mature.

Koopman operator theory provides a data-driven modeling framework that can effectively avoid physical simplification assumptions while enabling the derivation of simpler and more explicit linear models for subsequent control design. Bruder et al. [4] investigated a Koopman-operator-based approach for constructing explicit dynamic models of rigid-link robots and achieved satisfactory modeling performance. Mamakoukas et al. [5] applied this method to the modeling of tail-actuated robotic fish and obtained promising results. In addition, this theory has been applied to modeling problems in other fields, such as high-speed trains [6], wind farms [7], and micro aerial vehicles [8].

In this work, a data-driven modeling method based on the Koopman operator is applied to a corrugated-tube soft robot, and both linear and nonlinear models of its nonlinear dynamical system are established. By comparing the modeling accuracy and prediction performance with those of state-space modeling methods, the feasibility of the proposed approach is validated.

2. Related Work

Soft robotics has attracted increasing attention due to its inherent compliance, adaptability, and safety in unstructured environments. Compared with traditional rigid robots, soft robotic systems exhibit highly nonlinear and infinite-dimensional dynamics, posing significant challenges for accurate modeling and control. Early studies have primarily focused on the design, fabrication, and control of soft robots, establishing foundational insights into their physical characteristics and application potential [9-11]. However, these works also reveal the limitations of conventional physics-based modeling methods when dealing with complex deformation and strong nonlinearities. To address these challenges, data-driven approaches have been widely explored. Machine learning and deep learning techniques provide flexible modeling paradigms without requiring explicit physical assumptions and have demonstrated effectiveness in various domains, including industrial fault diagnosis under data imbalance [12], financial anomaly detection and fraud detection [13-14], enterprise risk prediction [15-16], and healthcare data modeling using electronic health records [17]. In addition, recent advances in cross-modal representation learning [18], uncertainty-aware decision-making [19], intelligent agent-based optimization [20], and causal reasoning frameworks [21-22] further enhance modeling capability in complex systems. Among these approaches, Koopman operator theory has emerged as a powerful framework for nonlinear dynamical system modeling. By lifting nonlinear dynamics into a higher-dimensional observable space, the Koopman operator enables linear evolution of inherently nonlinear systems. Foundational works established Koopman invariant subspaces and linear representations for nonlinear systems [23], while subsequent studies developed data-driven approximation methods such as dynamic mode decomposition and its variants [24-25]. Further integration with control frameworks, including model predictive control, enables practical applications in system regulation and control design [26], while extensions based on ergodic theory and spectral analysis deepen the theoretical understanding of Koopman operators in complex dynamical systems [27].

In robotics and control applications, Koopman-based approaches have demonstrated strong potential for data-driven modeling and control of nonlinear systems. Prior work shows that Koopman operator theory enables accurate modeling of robotic dynamics and supports control design without relying on explicit physical models [28], while soft robotic systems continue to benefit from advances in material design and control strategies. Compared with purely black-box learning methods, Koopman-based approaches provide a better balance between interpretability and predictive accuracy. Meanwhile, advances in physics-informed learning and neural dynamics modeling further improve generalization and stability in complex systems [29-30]. In parallel, recent developments in time-series modeling and probabilistic learning have improved the ability to handle non-stationary and uncertain environments. Residual-regulated learning methods have been proposed for non-stationary time series forecasting [31], while Wasserstein-based generative modeling enables robust optimization under distributional uncertainty [32]. Graph-based and spatiotemporal learning approaches further enhance the modeling of structured dependencies in complex systems [33]. Beyond traditional modeling paradigms, recent studies increasingly integrate causal reasoning, large language models (LLMs), and retrieval-augmented learning into intelligent systems. Causal representation learning improves interpretability and robustness in financial modeling [34-35], while attention alignment and logical constraint mechanisms enhance reasoning reliability [36]. Retrieval-augmented generation and context-aware ranking further improve knowledge utilization and decision consistency [37], and hybrid frameworks combining LLMs and causal inference enable automated reasoning and diagnostic capabilities in enterprise scenarios [38]. In addition, reinforcement learning and multi-agent optimization methods have been explored for system-level decision-making and resource scheduling, including hierarchical reinforcement learning for adaptive scheduling [39] and multi-objective optimization for large-scale decision systems [40]. Despite these advances, existing approaches often face trade-offs between accuracy, interpretability, and control compatibility. Therefore, a unified data-driven modeling framework that achieves global linearization while preserving nonlinear dynamics remains highly desirable, motivating the Koopman-based approach proposed in this work.

3. Koopman-Operator-Based Modeling Method

3.1 Koopman-Based Modeling of Soft Robots

Consider a discrete-time nonlinear dynamical system with external inputs, described as

$$x_{k+1} = F(x_k, u_k)$$

where $x_k \in X \subset \mathbb{R}^n$ and $u_k \in U \subset \mathbb{R}^m$ denote the system state and input at time step k, respectively, and $F(\cdot)$ represents the nonlinear state transition mapping.

Instead of directly modeling the nonlinear dynamics in the original state space, the Koopman operator framework describes the system evolution through observable functions defined in a lifted space. Let $\psi(x,u)$ denote a vector of observable functions. The evolution of these observables is governed by the Koopman operator K_d, such that

$$\psi(x_{k+1}, u_{k+1}) = K_d\,\psi(x_k, u_k)$$

Although the Koopman operator is infinite-dimensional in theory, it enables the representation of nonlinear dynamics using linear evolution in the lifted space.

3.2 Data-Driven Approximation via EDMD

In practice, a finite-dimensional approximation of the Koopman operator is obtained using the Extended Dynamic Mode Decomposition (EDMD) algorithm. By collecting input-output data pairs from system trajectories and selecting appropriate basis functions, the Koopman operator is approximated through a least-squares regression problem. This data-driven procedure yields a finite-dimensional Koopman matrix that best captures the evolution of the lifted observables in a linear sense.

By appropriately separating the observable functions into state-related and input-related components, the lifted linear dynamics can be equivalently expressed as a linear state-space model of the form

$$z_{k+1} = Az_k + Bu_k, \quad x_k = Cz_k$$

where z_k denotes the lifted state vector, and A, B, and C are system matrices identified from data. This formulation enables the application of mature linear control and analysis techniques while preserving the nonlinear characteristics of the original system through the lifting process.

3.3 Nonlinear Model Reconstruction

In addition to the linear Koopman representation, a nonlinear discrete-time model can be reconstructed by projecting the lifted dynamics back into the original state space. By selecting observable functions that explicitly include the original state variables, the nonlinear state transition mapping can be recovered from the finite-dimensional Koopman operator. This approach provides a nonlinear predictive model that maintains consistency with the data-driven Koopman formulation while offering higher modeling fidelity for strongly nonlinear soft robotic systems.

3.4 Remarks on Model Characteristics

The Koopman-based modeling framework establishes a unified representation for both linear and nonlinear models of soft robots. The linear lifted model facilitates controller design, whereas the reconstructed nonlinear model improves prediction accuracy. By relying purely on measured data, the proposed approach avoids strong physical assumptions and is well suited for soft robots characterized by high nonlinearity and infinite degrees of freedom.

4. Implementation of Koopman-Operator-Based Modeling

4.1 Experimental Platform

The soft pneumatic robot was fabricated using a commercial fused deposition modeling (FDM) 3D printer (Flash Forge Creator Pro, Flash Forge Corporation). The overall dimensions of the printed two-dimensional soft robot are 108 mm × 14 mm × 25 mm, as shown in Figure 1.

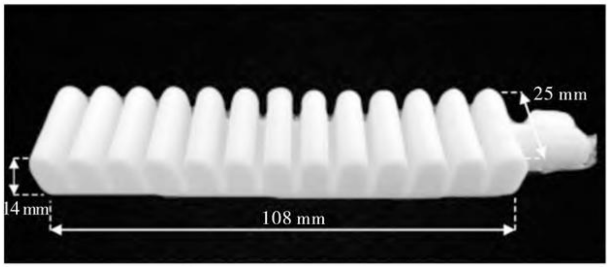


Figure 1. Dimensions of the 3D-printed soft robot

The experimental platform was constructed to acquire input-output data for evaluating and comparing the performance of Koopman-operator-based models. The system architecture, illustrated in Figure 2, was designed based on an open-source soft robotics control framework. The platform mainly consists of a solenoid valve, a bending sensor, a pressure sensor, an Arduino control board, and an air pump. The solenoid valve (SMC VQ110U-5M) regulates the airflow supplied to the pneumatic actuator, thereby adjusting the driving pressure. The valve switching signal is generated by the Arduino board using a 60 Hz pulse-width modulation (PWM) signal. By adjusting the duty cycle of the PWM signal, the output pressure can be effectively regulated to the desired level, providing a simple and cost-effective method for pressure control of the soft actuator.

A pressure sensor (XGZP6847) is employed to measure the internal pressure of the soft robot, while a bending sensor (FS-L-0095-103-ST) is used to measure the resulting bending angle, which serves as feedback for control and modeling. The Arduino Mega 2560 board interfaces with both the bending sensor and the pressure sensor to acquire feedback signals.

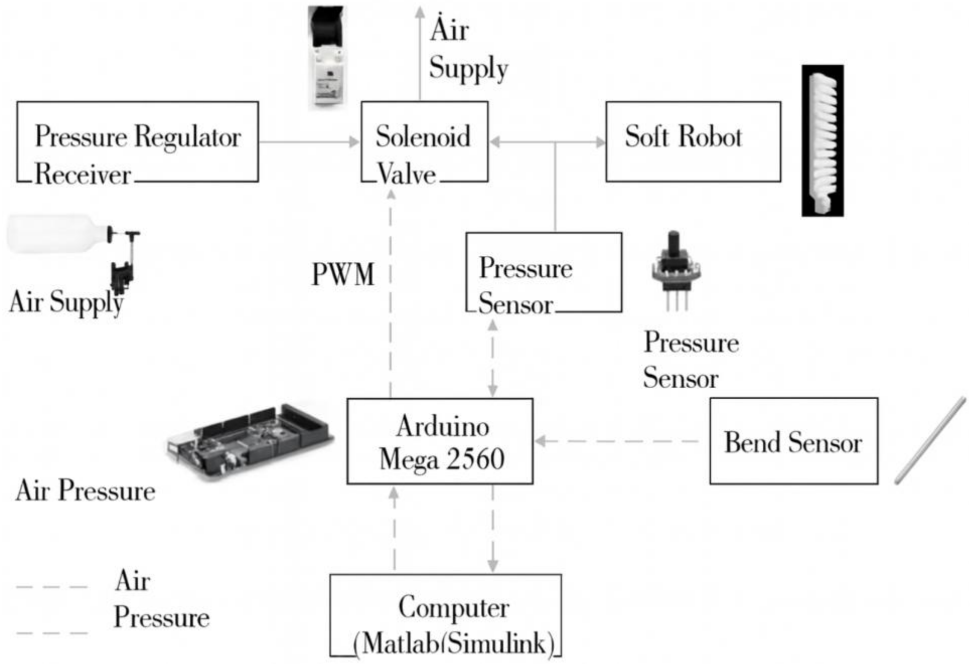


Figure 2. Schematic diagram of the pneumatic system

4.2 Data Acquisition and Processing

The soft robot used in this study contains a single pneumatic chamber and can therefore be modeled as a single-input single-output (SISO) system. A total of 11,460 input-output data pairs were collected through experiments, among which 10,000 samples were used for training and 1,460 samples were reserved for testing. These data were used to identify both linear and nonlinear Koopman models.

To improve the applicability of the data to the Koopman modeling framework, preprocessing was performed to normalize the input and output signals. Data normalization facilitates more accurate capture of the relationships between independent and dependent variables in different models, thereby improving prediction accuracy. A linear min-max normalization method was applied to scale both input and output data to the interval [−1, 1].

4.3 Model Comparison

The Koopman linear model, Koopman nonlinear model, and conventional state-space linear model were all constructed using the same experimentally obtained training dataset. The modeling accuracy of each approach was evaluated by comparing the predicted trajectories with the measured trajectories of the soft robot on both training and testing datasets.

The normalized root mean square error (NRMSE) was used to quantify the trajectory fitting performance, which serves as a measure of model accuracy, and is defined as

$$\mathrm{NRMSE} = \frac{\sqrt{\frac{1}{m}\sum_{i=1}^{m}\left(s_{\mathrm{real}} - s_{\mathrm{sim}}\right)^2}}{\max(s_{\mathrm{real}}) - \min(s_{\mathrm{real}})}$$

where s_{real} and s_{sim} denote the measured and simulated state values, respectively, and m represents the number of measurement points.

5. Results and Analysis

5.1 Modeling Results of the Soft Robot

The modeling results of the soft robot obtained using the state-space linear model are shown in Figure 3.

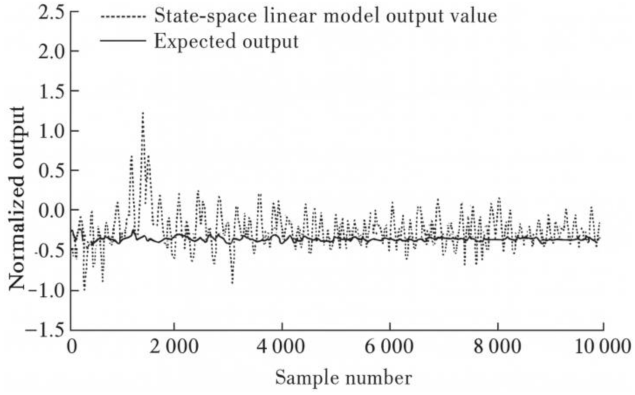


Figure 3. Results of establishing the state-space linear model for the soft robot

The modeling results of the Koopman linear model and the Koopman nonlinear model are illustrated in Figure 4 and Figure 5, respectively. In these models, the maximum degree of the basis functions and the number of time delays are both set to 2.

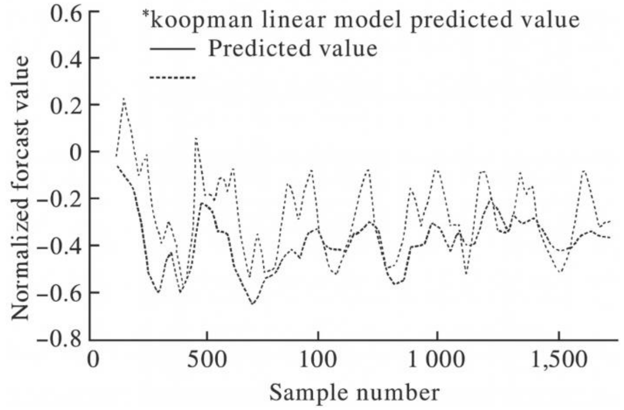


Figure 4. Results of establishing the linear model for the soft robot via Koopman operator

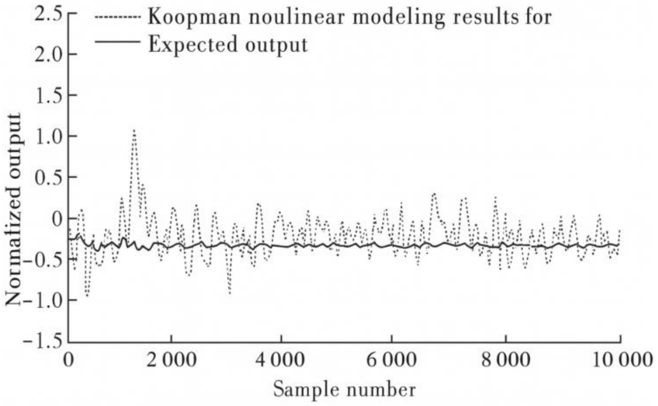


Figure 5. Results of establishing the nonlinear model for the soft robot via Koopman operator

A quantitative comparison of the modeling performance among the Koopman linear model, the Koopman nonlinear model, and the conventional state-space model is summarized in Table I. The comparison includes the maximum degree of the basis functions, the number of delays, and the improvement ratio with respect to the state-space model.

Table 1: Comparison of Modeling Accuracy Among Koopman Linear Model, Koopman Nonlinear Model, and State-Space Model

Model Type	Maximum Degree	Number of Delays	NRMSE	Improvement Ratio (%)
State-space model	-	-	0.1188	-
Koopman linear model	2	1	0.0796	32.98
Koopman nonlinear model	2	1	0.0791	33.42

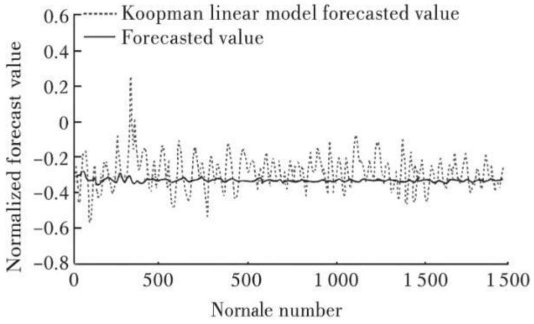
5.2 Model Prediction Results

The prediction results of the established soft robot models can be further categorized into those obtained from the Koopman linear model, the Koopman nonlinear model, and the state-space linear model, as shown in Figure 6.

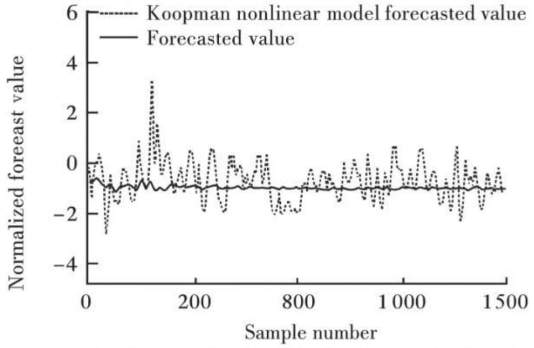
A quantitative comparison of the prediction performance among the Koopman linear model, the Koopman nonlinear model, and the state-space linear model is presented in Table 2. The comparison includes the maximum degree of the basis functions, the number of delays, and the improvement ratio relative to the state-space linear model.

Table 2: Quantitative Comparison of Prediction Performance Among Different Models

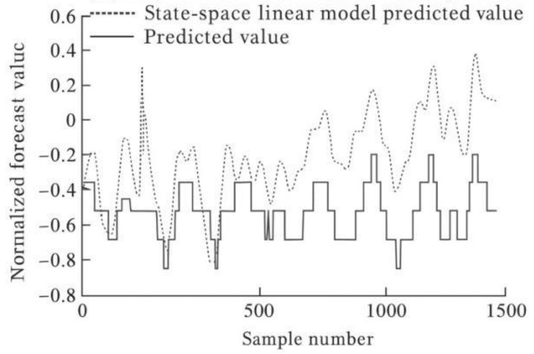
Model Type	Maximum Degree	Number of Delays	NRMSE	Improvement Ratio (%)
State-space model	-	-	0.3674	-
Koopman linear model	2	1	0.1119	69.54
Koopman nonlinear model	2	1	0.0966	73.71



(a) Koopman Linear Model Forecasting Results



(b) Koopman Nonlinear Model Forecasting Results



(c) State-Space Linear Model Forecasting Results

Figure 6. Comparison of prediction results among three models

5.3 Discussion

According to the results presented in Table 1 and Table 2, the Koopman-operator-based global linearization modeling strategy for soft robots achieves higher modeling accuracy than the local linearization approach based on the conventional state-space model. Based on the modeling results, the modeling accuracy of the Koopman linear model is improved by 32.98%, while the Koopman nonlinear model achieves an improvement of 33.42% compared with the state-space linear modeling method.

In terms of prediction accuracy, the Koopman linear model improves prediction performance by 69.54% relative to the state-space linear model, whereas the Koopman nonlinear model achieves a higher improvement of 73.71%. These results demonstrate that the Koopman-based modeling approach is more effective in capturing the nonlinear dynamics of soft robots and provides superior performance in both modeling and prediction tasks.

6. Conclusion

In this study, a data-driven modeling approach for soft robots based on the Koopman operator was proposed. Both linear and nonlinear models of the soft robot dynamical system were established. The proposed method enables accurate modeling of soft robots and facilitates data-driven control in the absence of explicit physical models by exploiting the black-box characteristics of measured data.

The results demonstrate that both the Koopman linear and Koopman nonlinear models achieve higher accuracy than the conventional state-space linear model, with the Koopman nonlinear model providing the highest modeling accuracy, followed by the Koopman linear model. Although the nonlinear Koopman model outperforms the linear model in terms of modeling accuracy, the increased complexity of controller design and solution procedures must be considered. In contrast, linear models allow the use of mature and computationally efficient linear control strategies. Therefore, the Koopman linear model represents a practical and effective choice for soft robot modeling, balancing modeling accuracy and control implementation complexity.

References

[1] R. Rus and M. T. Tolley, “Design, fabrication and control of soft robots,” Nature, vol. 521, no. 7553, pp. 467-475, 2015.

[2] C. Laschi, B. Mazzolai, and M. Cianchetti, “Soft robotics: Technologies and systems pushing the boundaries of robot abilities,” Science Robotics, vol. 1, no. 1, eaat3690, 2016.

[3] Bhattacharya D, Cheng L K, Xu W. Sparse machine learning discovery of dynamic differential equation of an esophageal swallowing robot[J]. IEEE Transactions on Industrial Electronics, 2019, 67(6): 4711-4720.

[4] Bruder D, Fu X, Gillespie R B, et al. Data-driven control of soft robots using Koopman operator theory[J]. IEEE Transactions on Robotics, 2020, 37(3): 948-961.

[5] Mamakoukas G, Castano M L, Tan X, et al. Derivative-based Koopman operators for real-time control of robotic systems[J]. IEEE Transactions on Robotics, 2021, 37(6): 2173-2192.

[6] Chen B, Huang Z, Zhang R, et al. Data-driven Koopman model predictive control for optimal operation of high-speed trains[J]. IEEE Access, 2021, 9: 82233-82248.

[7] Cassamo N, Wingerden J V. On the potential of reduced order models for wind farm control: A Koopman dynamic mode decomposition approach[J]. Energies, 2020, 13(24): 6513-6533.

[8] Shi L, Teng H, Kan X, et al. A data-driven hierarchical control structure for systems with uncertainty[C]// IEEE Conference on Control Technology and Applications (CCTA). Montreal: IEEE, 2020: 57-63.

[9] D. Rus and M. T. Tolley, "Design, fabrication and control of soft robots," Nature, vol. 521, no. 7553, pp. 467-475, 2015.

[10] C. Laschi, B. Mazzolai and M. Cianchetti, "Soft robotics: Technologies and systems pushing the boundaries of robot abilities," Science Robotics, vol. 1, no. 1, eaah3690, 2016.

[11] D. Trivedi, C. D. Rahn, W. M. Kier and I. D. Walker, "Soft robotics: Biological inspiration, state of the art, and future research," Applied Bionics and Biomechanics, vol. 5, no. 3, pp. 99-117, 2008.

[12] J. Huang, J. Zhan, Q. Wang, J. Jia and B. Zhang, "Stable fault diagnosis under data imbalance via self-supervised learning in industrial IoT," 2026.

[13] N. Chen, Y. Zhang, W. Wang, Z. Pan, Y. Wang and Y. Lu, "CoReAD: Context-aware retrieval-augmented deep anomaly detection for evolving business tabular data."

[14] N. Chen, S. Sun, Y. Wang, Z. Li, A. Zhu and Y. Lu, "Few-shot financial fraud detection using meta-learning and large language models," Proceedings of the 6th International Conference on Computer Science and Management Technology, pp. 822-826, 2025.

[15] R. Fang, "A machine learning framework for enterprise risk prediction: Unified feature embedding and lightweight attention," 2026.

[16] A. Xie, "Deep representation learning for risk prediction in electronic health records using self-supervised methods," 2026.

[17] X. Zhang, Q. Wang and X. Wang, "Joint cross-modal representation learning of ECG waveforms and clinical reports for diagnostic classification," Transactions on Computational and Scientific Methods, vol. 6, no. 2, 2026.

[18] Q. Liu, Y. Zhang, S. Chen, Z. Liu, Y. Xu and H. Cui, "Uncertainty-aware marketing attribution inference and budget decision-making with intelligent agents," 2026.

[19] X. Yang, S. Sun, Y. Li, Y. Xing, M. Wang and Y. Wang, "CaliCausalRank: Calibrated multi-objective ad ranking with robust counterfactual utility optimization," arXiv preprint arXiv:2602.18786, 2026.

[20] H. Chen, Y. Lu, Y. Wei, J. Lyu, R. Wu and C. Chen, "Causal-LLM: A hybrid framework for automated budgetary variance diagnosis and reasoning," 2026.

[21] J. Li, Q. Gan, R. Wu, C. Chen, R. Fang and J. Lai, "Causal representation learning for robust and interpretable audit risk identification in financial systems," 2025.

[22] C. Chen, R. Fang and J. Lai, "Causal representation learning for robust and interpretable audit risk identification in financial systems," Proceedings of ICEMME 2025, p. 454, 2026.

[23] S. L. Brunton, B. W. Brunton, J. L. Proctor and J. N. Kutz, "Koopman invariant subspaces and finite linear representations of nonlinear dynamical systems for control," PLOS ONE, vol. 11, no. 2, e0150171, 2016.

[24] M. O. Williams, I. G. Kevrekidis and C. W. Rowley, "A data-driven approximation of the Koopman operator: Extending dynamic mode decomposition," Journal of Nonlinear Science, vol. 25, no. 6, pp. 1307-1346, 2015.

[25] J. L. Proctor, S. L. Brunton and J. N. Kutz, "Dynamic mode decomposition with control," SIAM Journal on Applied Dynamical Systems, vol. 15, no. 1, pp. 142-161, 2016.

[26] M. Korda and I. Mezić, "Linear predictors for nonlinear dynamical systems: Koopman operator meets model predictive control," Automatica, vol. 93, pp. 149-160, 2018.

[27] H. Arbabi and I. Mezić, "Ergodic theory, dynamic mode decomposition, and computation of spectral properties of the Koopman operator," SIAM Journal on Applied Dynamical Systems, vol. 16, no. 4, pp. 2096-2126, 2017.

[28] D. Bruder, X. Fu, R. B. Gillespie, C. D. Remy and R. Vasudevan, "Data-driven control of soft robots using Koopman operator theory," IEEE Transactions on Robotics, vol. 37, no. 3, pp. 948-961, 2020.

[29] A. Nagabandi, G. Kahn, R. S. Fearing and S. Levine, "Neural network dynamics for model-based deep reinforcement learning with model-free fine-tuning," IEEE International Conference on Robotics and Automation (ICRA), pp. 7559-7566, 2018.

[30] M. Lutter, C. Ritter and J. Peters, "Deep Lagrangian networks: Using physics as model prior for deep learning," arXiv preprint arXiv:1907.04490, 2019.

[31] Y. Ou, S. Huang, R. Yan, K. Zhou, Y. Shu and Y. Huang, "A residual-regulated machine learning method for non-stationary time series forecasting using second-order differencing," 2025.

[32] S. Huang, Y. Shu, K. Zhou, S. Sun, Y. Ou and R. Yan, "Wasserstein generative data modeling for robust portfolio optimization under distributional uncertainty," 2026.

[33] X. Liang, Y. Zhao, M. Chang, R. Zhou, K. Cao and Y. Zheng, "Spatiotemporal risk representation learning using transformers and graph structure," 2026.

[34] C. Chiang, "Collaborative machine learning for risk ranking under concurrent class imbalance and distribution shift," 2026.

[35] Y. Wang, "AI-augmented spatiotemporal learning for identifying risk propagation in pharmaceutical supply chains," 2026.

[36] J. Lai, "Attention alignment under logical constraints for reliable financial statement reasoning," Transactions on Computational and Scientific Methods, vol. 4, no. 12, 2024.

[37] T. Guan, S. Sun and B. Chen, "Faithfulness-aware multi-objective context ranking for retrieval-augmented generation," 2025.

[38] C. Nie, "Adaptive ETL task scheduling via hierarchical reinforcement learning with joint rewards for latency and load balancing," 2026.

[39] A. Sanchez-Gonzalez, J. Godwin, T. Pfaff, R. Ying, J. Leskovec and P. Battaglia, "Learning to simulate complex physics with graph networks," International Conference on Machine Learning, pp. 8459-8468, 2020.

[40] J. Li, Q. Gan, Z. Liu, C. Chiang, R. Ying and C. Chen, "An improved attention-based LSTM neural network for intelligent anomaly detection in financial statements," 2025.