Journal of Electrical Engineering, Electronics, Control and Computer Science – JEEECCS, Volume 11, Issue 39, pages 23-34, 2025

Model-Based Simulation for Approximated Dynamic System Using Reinforcement Learning

1st Donatus O. Njoku
Department of Computer Science
Federal University of Technology
Owerri, Nigeria
donatus.njoku@futo.edu.ng

3rd Somkelechi E. Ubalatu
Department of Information Technology
Federal University of Technology
Owerri, Nigeria
somkelechi.ubalatu@futo.edu.ng

2nd Janefrances E. Jibiri
Department of Information Technology
Federal University of Technology
Owerri, Nigeria
eberejibi@yahoo.com

4th Innocent H. Ajunwa
Department of Information Technology
Federal University of Technology
Owerri, Nigeria
innocent.ajunwa@futo.edu.ng

Abstract –The complexities of modern dynamic systems typically render it infeasible to implement the standard analytical modeling techniques due to non-linearities and ignorance of the underlying dynamics. This research proposes and implements a strategy for generating an approximated, simulation-based model controlled by a Reinforcement Learning (RL) agent. The process was two-staged using the "Dynamical System Multivariate Time Series Forecast" dataset. First, a simulated environment with an approximated simulation environment was established by training a neural network to predict state transitions from historical state-action data with high accuracy. Subsequently, a Deep Q-Network (DQN) agent was trained only in this simulated environment to learn an optimal control policy. The agent's learning was guided by a reward function designed to minimize the Euclidean distance to a given target state. Early indications from a brief training regimen of nine episodes yielded a 0% mission success rate, which highlights the heavy dependence of RL agents on rich experience. Simulated indications of a prolonged training regimen, on the other hand, pointed towards the great potential of the framework and achieving an accuracy of more than 92%. The present research concludes that the combination of simulation-based methods and deep reinforcement learning is a powerful and flexible means of system control. It is significant as it can create goal-seeking autonomous agents for complex systems in a controlled setting with direct use in industrial automation, resource allocation, and robotics.

Keywords- reinforcement learning, dynamic systems, simulation-based modeling, deep Q-network, system approximation; autonomous control

I. INTRODUCTION

The complexity of modern dynamic systems has demanded the enhancement of sophisticated modeling and simulation techniques capable of capturing faithfully the behaviors without compromising computation efficiency. Dynamic systems with their time-variant state and subtle interdependencies impose significant challenges to modeling fidelity and computational solvability. The standard analytical approaches often fall short when dealing with highly nonlinear, high-dimensional systems with emergent behavior and stochastic properties. Emergence of RL as a successful paradigm for solving complex decision-making problems opened possibilities for addressing dynamic system modeling issues. Unlike supervised learning paradigms, where big, labeled data sets are required, reinforcement learning enables agents to learn optimal behaviors through interactive experience with the environment and thus its usefulness in cases where direct programming of system behavior is infeasible or even impossible. This attribute makes RL especially useful for estimating dynamic systems where the mathematical models describing the system are unknown, too complicated to be solved analytically, or too costly to solve precisely. Blending reinforcement learning with simulation-based approaches is a paradigm shift in dynamic system modeling. Simulation scenarios provide predictable, reproducible conditions in which RL agents can learn different strategies, fail, and enhance their decision-making capability without experimentation costs and risks in the real world. This is particularly relevant in areas like autonomous systems, industrial process control, financial modeling, and sophisticated engineering systems where traditional control theory methods are not effective. Recent research in this field has been significantly eased by advances in computing power, the availability of robust simulation environments, and breakthroughs in more efficient RL algorithms. The combination of these advancements has enabled opportunity to create approximation models that can cope with the certainties and complexities of real-world dynamic systems and yield adequate accuracy and computational efficiency. Motivation for this work stems from a variety of adverse limitations of existing approaches for dynamic system modeling. Firstly, traditional modeling techniques typically require a lot of domain knowledge and accurate system parameter information, which might not always be available or can be time-variant. Secondly, most existing approaches struggle with the curse of dimensionality while handling high-dimensional state and action spaces. Third, adaptive models are required that can learn from experience and refine their performance with passage of time as more data is fed in or system conditions evolve. This research is tackling these challenges by proposing a simulation-based modeling paradigm that utilizes reinforcement learning to develop approximated models for dynamic systems. The approach leverages the flexibility and adaptability of RL as well as the controlled environment of simulation in generating models that can adequately approximate complex dynamic behaviors. The model is flexible, scalable, and can accommodate various types of dynamic systems in various fields of application. The key contributions of this work are general framework for integrating simulation-based modeling and reinforcement learning to approximate dynamic systems; new methods for optimizing the accuracy and efficiency of RL-based approximation in dynamic systems; rigorous methods of evaluation for confirming the performance of simulation-based RL models, and engineering principles for using such systems to solve real-world problems. The influence of this research extends beyond theoretical contributions to cover applied ramifications in industries and fields of research. Through the development of more precise and adaptable models of dynamic systems, this research can be employed to advance decision-making in dynamic settings, resource allocation, system optimization, as well as risk management processes.

II. Related Work

Literature is organized in some key subsections that cumulatively form a general idea of the state of the research in simulation-based modeling, reinforcement learning in dynamic systems, and related technological advancements. The systematic review is founded on an investigation into the way these fields have developed so far and the lacunae that the present research will address.

A. Foundations of Simulation-Based Modeling in Complex Systems

Simulation-based modeling methods have been shaped fundamentally by the development of computational capabilities and complexity of systems to be analyzed and optimized.

Ritto et al. [1] provides an exhaustive survey of cognition algorithms, simulation modeling, and visualization tools in metaverse interactive spaces. This work provides the foundation for comprehending how simulation environments could be enhanced by advanced algorithms and visualization, and of direct relevance to the design of effective simulation platforms for reinforcement learning applications. Special computing and immersive technologies combined as explained by Ritto et al [1] is a significant breakthrough in the field of simulation. These technologies enable us to create more realistic and interactive simulation environments, which can better capture the richness of dynamic systems in the world. The cognitive algorithms described in this work illuminate how simulation environments can be built to enable more sophisticated learning processes, particularly for reinforcement learning tasks requiring rich, interactive environments for successful policy learning.

Based on these pioneering concepts, Bushaj et al. [2] examine remote big data management software, sensing and computing technologies, and vision perception algorithms that constitute the foundation of modern simulation systems. Their study highlights the importance of data management and sensing technologies in simulating true-to-life environments. The vision perception and mapping environment algorithms they present in their paper are particularly suited towards constructing simulation environments that can provide good feedback to reinforcement learning agents and enable more efficient learning processes to be triggered. The comprehensive study brought out by Bushaj et al. [2] also points out the difficulties of handling massive data in simulation systems, crucial for reinforcement learning problems that are bound to entail extensive exploration of state-action spaces. Their findings on sensing and computing technologies provide a technological context on how simulation systems can be equipped with the necessary infrastructure to support advanced reinforcement learning techniques.

B. Virtual and Augmented Reality Applications in Simulation

Virtual and augmented reality technology applications in simulated environments have emerged as a significant research field with significant application implications for reinforcement learning programs. Farsi [3] introduce a systematic analysis of visualization techniques within virtual reality environments, offering insights into how immersive technologies can enhance simulation experiences and improve artificial intelligence system learning environments. Their systematic review methodology provides comprehensive insight into the art of the moment of VR visualization techniques and identifies key challenges and opportunities within this field.

The visualization techniques discussed in their paper are particularly relevant for creating simulation environments capable of providing comprehensible, interpretable feedback to human operators as well as artificial intelligence systems.

This is especially important in RL scenarios where environmental feedback quality directly impacts the effectiveness of the learning process. The research of Farsi [3] also discusses the technical issues involved in embedding effective visualization into virtual reality environments, including computational requirements, user interface, and real time rendering capabilities. Such results are applicable towards understanding the technical possibilities and constraints of developing simulated environments based on VR for applications in reinforcement learning. Hu et al. [4] carry this dialogue further by elaborating on artificial intelligence, virtual reality, and augmented reality technology in terms of their applicability to performance analysis. Their work presents a bridge between the technical capabilities of immersive technologies and their applicability in real-world analysis and system performance optimization. This perspective is particularly helpful in comprehending how simulation-based reinforcement learning systems can be optimized and tested.

C. Digital Factory and Industrial Applications

Industrial use of virtual reality and simulation technologies provides valuable insights into application of simulation-based modeling systems in the real world. Chandra Hsu et al. [5] provide a comprehensive review of the application of virtual reality in digital factory systems, with a detailed analysis of how simulation technologies can be employed to simulate large industrial systems. Their systematic review procedure systematically examines various approaches to implementing VR technologies in digital factory settings, providing comments on the technical concerns as well as the practical benefits of such implementations. The digital factory concept is a sophisticated use of simulation technology where virtual worlds are used to model, analyze, and optimize complex manufacturing processes. [5] paper is particularly remarkable in reinforcement learning applications because it shows how the simulation environment can be used to simulate complex, multi-agent systems involving numerous intercommunicating parts. The challenges and solutions presented in their paper are good pointers on how the simulation environment can be organized to handle the complex state and action spaces typically encountered in reinforcement learning applications. The deployments of digital factories examined by [5] also indicate the importance of real-time processing capacity and the need for simulation systems capable of handling dynamic, shifting environments. These findings are directly applicable to reinforcement learning systems that must handle shifting environmental circumstances and learn optimal policies in shifting environments. Moreover, their work is also focused on the integration problems involved in the deployment of simulation-based systems into industrial environments, and it provides realistic insights that are useful in understanding how simulation-based reinforcement learning systems can be introduced in real-world applications.

D. Architectural Heritage and Data-Limited Applications

Use of simulation and virtual reality technologies for the visualization of architectural heritage, which has been discussed by Grenne et al. [6], is important when it comes to how simulation systems are developed using limited data. This research is highly relevant to applications of reinforcement learning because it captures the challenges associated with model development whenever complete information on the system is unavailable. [6] research shows how virtual reality technologies can be used to create appropriate visualizations and simulations even when working with incomplete or limited data sets. This aspect is particularly useful in applications of reinforcement learning when complete system models are not available, and the learning algorithm needs to work with approximated or incomplete knowledge of the environment. The techniques described by [6] for dealing with situations of having limited data provide insight into how the simulation worlds can be developed for enabling reinforcement learning if perfect models of the world cannot be acquired. This is particularly relevant to most real-world problems in which data pertaining to the entire system are unavailable or the system parameters change over time. The visualization techniques and modeling approaches described in this book also reveal that simulation systems can be designed to provide valuable feedback even when faced with approximated or incomplete models. This is a feature of necessity for reinforcement learning systems that must learn effective policies using imperfect environmental representations.

E. Decision Intelligence and Customer Experience Systems

Application of simulation and AI technologies to decision intelligence and customer experience systems, as researched by Nica et al. [7], provides insight into utilization of these technologies in complicated, multi-stakeholder environments. Decision intelligence and modeling, multisensory customer experiences, and socially integrated virtual services subject to their research are testimony to flexibility and strength of simulation-based AI systems.

The decision intelligence models introduced by Yang. Et al. [8] are best suited for reinforcement learning tasks since they address the issues associated with making optimal decisions within complex, multi-objective systems. The authors' contribution provides advice on how simulation systems can be designed to allow for intelligent decision-making processes that consider several parameters and stakeholders. The multisensory customer experience spaces investigated in their research demonstrate how simulation environments can be designed to handle realistic, multi-modal interactions. This capability is particularly important in reinforcement learning environments that must process multiple sources of environmental feedback and make decisions based on diverse sources of information. Yang et al. [8] socially networked virtual services also show how systems of simulation can map complex economic and social interactions. It is useful for understanding how reinforcement learning systems can be employed in multi-agent environments and potentially conflicting goals.

F. Building Information Modeling and Architectural Computing

The expansion of building information modeling functionality using immersive virtual reality technologies, as described by Toscano et al. [9], plays an important role in determining how simulation technologies can be utilized to integrate with available design and modeling systems. Their examination demonstrates how VR technologies can facilitate traditional modeling practices and introduce new functionalities for analyzing and optimizing complex systems. The building information modeling (BIM) integration approaches outlined by [9] are relevant to reinforcement learning implementations because they show how simulation systems can be designed to interface with existing data and modeling systems. This aspect is important to real-world applications of simulation-based RL systems that must interface with existing organizational systems and data sources. Their work also addresses the technical challenges of constructing immersion simulation worlds that can handle high-level, complex models. The solutions and approaches presented in their paper provide insights that can be transferred to constructing simulation worlds for reinforcement learning that can handle high-dimensional state spaces and complex world representations. The computational structure views presented in [9] also demonstrate how simulation systems may be designed to support group work and multi-user interaction. This aspect is useful for reinforcement learning tasks that involve many agents or that require human observation and intervention.

G. Digital Media Art and Educational Applications

The application of virtual reality and artificial intelligence technologies in digital media art and art education, as researched by An et al. [10], provides a comprehension of how they can be applied to fields of creativity and learning. This research finds application in the application of simulation-based systems in facilitating learning and creativity in various fields. [10] work on AI technology use in virtual reality pedagogy demonstrates how learning spaces may be built to create learning processes. This emphasis is valuable in thinking about how reinforcement learning systems can not only learn best policies but also aid human learning and understanding of complex systems. The computer art generation packages discussed in this book also show how simulation systems can be used to generate and test innovative solutions. This capability is relevant to reinforcement learning problems that involve creative problem-solving or must generate new solutions to complex problems. The educational applications explored by [10] provide some insight into the extent to which simulation systems can be made interpretable and educational, which is essential for the creation of reinforcement learning systems that human operators will want to understand and trust.

H. Metaverse and Commercial Applications

The visualization techniques discussed in their paper are particularly relevant for creating simulation environments capable of providing comprehensible, interpretable feedback to human operators as well as artificial intelligence systems. Stop! The application of live shopping and business use in metaverse environments, as shown by Djeumuo et al. [11], provides insights into the possible applications of simulation technologies in advanced commercial and social environments. Their visual and spatial analytics, cognitive AI approaches, and immersive digital simulations provide evidence of the commercial viability and real-life applications of advanced simulation systems. The visual and spatial analytics techniques outlined in [11] are particularly relevant to reinforcement learning tasks because they provide techniques for comprehending and analyzing complex spatial and temporal data patterns. Such analytical capability is essential for constructing effective reinforcement learning systems that can acquire knowledge from complex environmental information. Their cognitive AI approaches in the study provide information on how artificial intelligence can be integrated with simulation systems to create more intelligent and dynamic environments. Such integration is necessary for creating simulation-based reinforcement learning systems that can adapt to changing conditions and learn from new experiences in a continuous manner. The virtual simulations provided in [11] also demonstrate the capability to design simulation systems to support rich, realistic settings that can support complex interactions and learning processes. Such a capability is useful for reinforcement learning tasks in which real-world-enriched interactive settings are required to successfully learn policies.

I. Digital Media Art and Educational Applications

Virtual technologies applied to human resource management, as detailed by Strasser et al. [12], provide a way of understanding how simulation systems can be employed in organizational and management environments. Their implementation of virtual human resource management in the metaverse is an illustration of how simulation technologies can be applied to model and optimize intricate organizational processes. The human resource management uses of [12] directly pertain to reinforcement learning in the sense that they demonstrate how AI systems may be deployed in advanced, multi-stakeholder environments where organizational dynamics and human behavior are at the forefront. This is an effective framework for understanding how reinforcement learning systems might be built to function effectively within environments that include human players. Their work also addresses the challenges of modeling human action and organizational dynamics in virtual environments. These findings are helpful for designing reinforcement learning systems that must interact with or model human action as part of their own learning process. The organizational applications investigated by their work also address how simulation-based systems can be designed to scale with regards to being able to approach complex, large-scale environments with multiple interacting agents and stakeholders.

J. Digital Twin Technology and Multi-Sensor Integration

Sun et al. [13] research on digital twin simulation software, spatial cognition algorithms, and multi-sensor fusion technology provides useful information about the technological infrastructure required for advanced simulation systems. Digital twin technology is a high-tech approach for virtual replication of real systems that can be used for analysis, optimization, and prediction. The spatial cognition algorithms outlined by [13] find special relevance in reinforcement learning solutions since they provide means to process and interpret spatial knowledge in complex environments. The algorithms play an integral role in RL system design, which must navigate and make decisions within spatially complex environments. The multi-sensor fusion technology discussed in their research presents the fundamental challenge of combining data from various sources to build holistic environmental models. This feature is necessary for reinforcement learning systems that have the obligation of dealing with varying forms of environmental data to make the best decisions. The digital twin concepts ad also demonstrates how simulation systems can be deployed to remain in sync at all times with systems of the real world to enable real-time learning and adaptation. Such a capability is crucial in designing reinforcement learning systems that can learn and adapt in real time as things change.

K. Educational Technology and Bibliometric Analysis

The bibliometric survey and systematic review of literature regarding virtual reality in education, by Rojas-Sh et al. [14], embodies the expansive reach of how the VR technologies are being implemented in education. This paper clarifies the overall trends and developments occurring within VR technology applications and how this transfer over to simulation-based learning systems. The bibliometric analysis technique of [14] provides insight into the evolution of VR and education studies, indicating important trends, landmark works, and emerging research directions. The analysis technique is helpful in the understanding of the current state of simulation-based learning research and identifying opportunities for future growth. The educational applications their systematic review takes into consideration demonstrate the capability of simulation technologies to support effective learning environments. Such observations are relevant in designing simulation-based reinforcement learning systems to be educational and interpretable, an aspect pertinent in practical deployments involving human comprehension and trust. The process of systematic review adopted by [14] also provides an example of how exhaustive literature reviews in new fields of technology should be conducted, and it is worth it for deciding the existing research climate in simulation-based reinforcement learning.

L. Educational Technology and Bibliometric Analysis

The paper by Wang et al. [15] on co-creative AI design for virtual worlds provides essential information regarding artificial intelligence system design in a manner that collaborates with human beings in virtual and simulated worlds. The authors write about how the challenge and opportunity arise in designing AI systems that can creatively co-collaborate with human users. The co-creative AI techniques discussed by [15] can be used for reinforcement learning tasks in that they show how the AI systems can be made to work in a collaborative mode rather than in an autonomous mode. Such thought is helpful in developing reinforcement learning systems that must work in collaboration with human operators or must consider the human preferences and limitations in their decisions. Their work explains the interactive design methodologies that provide insights into designing simulation environments to enable effective human-AI collaboration. Such capability is central to the real-world applications of simulation-based reinforcement learning that involve human guidance and collaboration. The virtual environment design considerations that [15] present also direct towards recommendations for the development of simulation environments that can support AI learning processes as well as human interaction, important for the development of deployable, real-world applicable reinforcement learning systems.

M. Visual Imagery and Geospatial Applications

Selby et al. [16] discuss the research on visual imagery and geospatial mapping methods to contribute insights towards the integration of high-end visualization and mapping technologies with virtual simulation and deep learning systems. Their work on virtual simulation algorithms and deep learning-based sensing technologies is a good example of the integration of multiple high-end technologies into complete analysis and modeling systems. The geospatial mapping technologies outlined by [16] are directly relevant to reinforcement learning systems that must operate in spatially crowded domains. The technologies provide spatial information processing and understanding methods that are paramount to RL systems required to navigate and make decisions within geographical or spatial domains. The deep learning-powered sensing technologies investigated in their work underscore the extent to which advanced AI techniques can be blended with simulation and visualization systems to create more potent and smarter environments. The blending is critical for developing simulation-based reinforcement learning systems capable of processing advanced sensory data and learning from rich environmental observations. The simulation algorithms of [16] also provide technical perspectives on how simulation systems can be designed to handle complex spatial and temporal data, which finds applicability in the construction of efficient simulation environments for reinforcement learning applications.

N. Visual Imagery and Geospatial Applications

Arzani et al. [17] give a precise illustration of the use of digital twin technology in creating fine virtual replicas of existing environments in their research on virtual-reality-based digital twins of office spaces. The research goes a long way in showing the actual uses of digital twin technology and provides information on how such systems can be implemented in existing environments. The digital twin office space solutions offered by [17] are relevant to reinforcement learning solutions as they show how complex virtual environments can be simulated and refreshed to simulate real-world conditions. This capability is important for developing simulation-based RL systems that must learn policies that work in the real world. Their work also considers the technical difficulties of building and maintaining accurate digital twin representations, like data acquisition, model updating, and maintaining them synchronized with actual conditions. The implications of these findings are enlightening for examining the pragmatic requirements of simulation-based reinforcement learning systems. [17] described virtual reality integration features are also enlightening as to how immersive technologies can be utilized to make digital twin systems more usable and efficient, and hence applicable in designing easy-to-use interfaces for reinforcement learning systems.

O. Medical and Healthcare Applications

The application of virtual and augmented reality technology in spine surgery, researched by Ghaednia et al. [18], presents us with a picture of how such technologies can be applied in high-risk, precision-critical settings. Medical applications are some of the most demanding of simulation and VR technology, with high requirements for precision and reliability.

The spinal surgery considerations that Tan et al. [19] are relevant to reinforcement learning since they demonstrate the application of simulation technologies in situations where dependability and accuracy are essential. This perspective is valuable in understanding how simulation-based RL systems can be aligned to guarantee drastic performance and safety provisions. Their work also addresses matters concerning integrating contemporary technologies into existing professional practice and workflows, and it has implications for understanding how simulation-based reinforcement learning systems can be effective in being applied in real use. The accuracy requirements and safety matters addressed by [19] address how the simulation systems must be designed and certified to ensure that they are compliant with the requirements of applications that are mission critical.

In addition, Wu et al. [20] are presenting dentistry applications of augmented, virtual, and mixed reality and further evidence on the medical applications of these technologies. Their narrative review provides a full overview of the use of immersive technologies in dental practice and training. The dental applications reported by [20] also provide a glimpse of how simulation technologies can be utilized both for training as well as for actual usage in medical applications. This dual utility is for reinforcement learning applications that can be utilized both for learning optimal policy as well as for training human operators.

P. AI-Enabled Sensing Technologies

The research of Zhang et al. [21] on AI-enabled sensing technologies during the 5G/IoT epoch provides us with an understanding of how advanced sensing capabilities could be integrated into AI to create more capable and intelligent simulation environments. The convergence of 5G communications, IoT sensors, and AI processing opens new doors to creating sophisticated simulation and learning systems.

The AI-reliant sensing technologies outlined by Ganai et al. [22] are immediately transferable to reinforcement learning situations since they provide the technological driver for crafting rich, interactive simulation environments which can feed back in complex detail to learning algorithms. The integration of advanced sensing and AI processing permits simulation environments to be designed which can respond and adapt to changing conditions in real time. The 5G and IoT integration specifics outlined in their work also provide some insight into how reinforcement learning systems based on simulation may be interfaced with real-world sensor networks and data sources to enable continuous learning and adaptation from real-world settings. The advanced intelligent systems provided by [22] illustrate the potential of integrating various AI technologies to create holistic systems capable of supporting complex sensing, processing, and decision-making tasks.

Q. Digital Twin in Smart Manufacturing

The application of digital twin technology in smart manufacturing, researched by Bishnoi et al. [23], identifies how such advanced modeling techniques can be applied to complex industrial systems. Smart manufacturing is among the most sophisticated applications of digital twin technology with the integration of different systems, processes, and stakeholders. The intelligent manufacturing applications which were studied by [23] are highly relevant to reinforcement learning because they show how digital twin systems can be utilized to model and optimize complex, multi-component systems with numerous interacting components. This capability is essential to develop RL systems which must learn to optimize complex industrial processes. Their work also addresses the structural and multidisciplinary optimization challenges associated with smart manufacturing systems, providing insights into the way reinforcement learning algorithms can be designed to deal with multi-objective optimization problems in complex industrial systems. The digital twin implementations covered by [23] also demonstrate the method through which simulation systems can be integrated with real manufacturing systems to provide continuous optimization and improvement functionality, which is suitable for developing reinforcement learning systems capable of providing constant value in industrial applications.

R. Research Gaps and Future Directions

The critical examination of recent literature yields several key research gaps presenting potential for extending the research field of simulation-based reinforcement learning to approximate dynamic systems. While impressive progress has been documented in independent fields of VR simulation, digital twin technology, and reinforcement learning algorithms, there are substantive challenges in integrating these technologies effectively for dynamic system modeling applications. One of these major gaps is the lack of comprehensive frameworks that couple reinforcement learning algorithms with simulation technologies sufficiently for approximating dynamic systems. Even though individual research has been carried out on simulation systems and RL algorithms, there has not been considerable work in how these technologies are integrated at the optimal level in a bid to create efficient approximation models for dynamic complex systems.

Yet another essential gap is the lack of adequate availability of systematic evaluation method for measuring the performance of simulation-based RL systems in dynamically evolving environments. While specific studies have put forth a range of evaluation approaches, what is required are standardized methods that can be applied across different domains of application as well as system types. Literature also shows a gap in understanding how to scale the scalability challenge of applying reinforcement learning to high-dimensional dynamic systems. Although prior studies have addressed scalability in specific contexts, general scaling methodologies to scale simulation-based RL systems to tackle real-world, complex dynamic systems have been little studied. Moreover, further studies are required to identify the ways in which the reliability and robustness of simulation-based RL systems are to be assured if they are to be used in real-world applications. Although the literature contains many instances of good practice in controlled experiments, there is less research into how to make these types of systems operate reliably under real-world variations and uncertainties. These research gaps provide the motivation and foundation for this study, whose goal is to develop an end-to-end simulation-based reinforcement learning platform that fills these gaps and provides pragmatic solutions for dynamic system approximation application.

III. METHODOLOGY

This research was conducted through a systematic, multi-stage process involving data acquisition, environment simulation, agent development, and rigorous evaluation. The core of this methodology is the creation of an approximated model of a dynamic system, which then serves as a training ground for a Reinforcement Learning agent to learn an optimal control policy.

A. Data Source and Tools

1. Data Source

The study utilized the "Dynamical System Multivariate Time Series Forecast" dataset, publicly available on Kaggle. This dataset contains time-series data representing the behavior of a complex dynamic system, making it an ideal candidate for this research [24].

Description: The dataset consists of 18 feature columns and a timestamp. For our experiment, we logically partitioned these features to represent the system's state and the control actions applied to it.

a. State Space (S): A 10-dimensional continuous space derived from the columns: ['aimp', 'amud', 'arnd', 'asin1', 'asin2', 'adbr', 'adfl', 'bed1', 'bed2', 'bfo1'].

b. Action Space (A): A 4-dimensional continuous space derived from the columns: ['bfo2', 'bso1', 'bso2', 'bso3'].

2. Tools and Software

The project was developed entirely in Python 3 within the Google Colaboratory (Colab) environment, leveraging its cloud-based computational resources. The key software libraries used are listed in TABLE I.

TABLE I. TOOLS AND SOFTWARE

Tool / Library	Purpose
Google Colab	Cloud-based Python environment for model training and experimentation.
Kaggle Hub	API and library for programmatic dataset acquisition.
Pandas	Data loading, manipulation, and preprocessing.
Scikit-learn	Data normalization (MinMaxScaler) and performance evaluation metrics.
TensorFlow (Keras)	Core deep learning framework used to build both the environment and agent neural networks.
Matplotlib & Seaborn	Data visualization, generation of charts, and plotting performance results.
Flask & VS Code	Framework and IDE used for building the web-based front-end application to serve the trained model.

B. Data Source and Tools

The experiment was structured into five distinct steps, from data preparation to final evaluation.

1. Data Acquisition and Preprocessing

The dataset was downloaded using the Kaggle Hub API. The relevant columns were selected to form the state and action vectors. Both state and action data were then normalized to a range of [0,1] using Scikit-learn's MinMaxScaler. Normalization is a critical step that ensures stability and improves convergence speed for neural network training.

2. Approximated Environment Modeling

A core aspect of this research was to first create a predictive model of the dynamic system's behavior. This was achieved by training a feed forward neural network, hereafter referred to as the transition model.

a. Input: The current state s_t and the action taken a_t.

b. Output: The predicted next state $\hat{s}_t + 1$.

This model effectively learns the system's dynamics, $\hat{s}_t + 1 = f(s_t, a_t)$, from the historical data, creating a reliable and fast simulation environment for the RL agent to train in without interacting with a real-world system.

3. Reinforcement Learning Agent Design

DQN agent was designed to learn an optimal control policy. The agent's goal is to learn which action to take in any given state to maximize its expected cumulative reward. The design includes two key features:

a. Experience Replay: The agent stores its experiences in a replay buffer. During training, it samples random mini batches from this buffer, which breaks temporal correlations and leads to more stable learning.

b. Epsilon-Greedy Strategy: To balance exploration (discovering new strategies) and exploitation (using the current best strategy), the agent selects a random action with a probability ε (epsilon) and the best-known action with a probability $1 - \varepsilon$. The value of ε starts high (1.0) and is gradually decayed over time as the agent becomes more confident in its policy.

4. Agent Training

The DQN agent was trained by interacting with the approximated environment. In each episode, the agent's objective was to navigate from a random starting state to a pre-defined target state.

a. The reward function was defined as the reduction in Euclidean distance to the target. A positive reward was given for any action that brought the agent closer to its goal.

b. An episode concluded when the agent reached the target state, or a maximum number of steps was exceeded.

5. Performance Evaluation

To quantify the agent's performance, the task was framed as a binary classification problem. A "successful" mission was defined as an episode where the agent reaches the target state within a set number of steps. The agent was tested over 50 new episodes, and its performance was measured using standard metrics: Accuracy (Success Rate), Precision, Recall, F1-Score, and a Confusion Matrix.

C. Models, Algorithms, and Architecture

The research deploys two distinct neural network models were employed in this research.

1. Transition Model (The Environment)

a. Architecture: A Multi-Layer Perceptron (MLP) with an input layer corresponding to the size of the state and action vectors combined (10 + 4 = 14 neurons), two hidden layers with 64 neurons each using the ReLU activation function, and a linear output layer with 10 neurons to predict the next state.

b. Purpose: To serve as a fast and accurate simulation of the real-world dynamic system.

2. DQN Agent (The Controller)

a. Architecture: An MLP with an input layer sized to the state space (10 neurons), two hidden layers with 64 neurons each (ReLU activation), and a linear output layer with a neuron for each possible discrete action (8 neurons).

b. Purpose: To approximate the optimal action-value function, $Q(s, a)$.

3. Algorithm: Deep Q-Learning

The agent learns using the Deep Q-Learning algorithm, which is an extension of Q-learning that uses a neural network to approximate the action-value function, $Q(s, a)$. This function represents the expected total reward an agent can achieve by acting a in state s and then following the optimal policy thereafter.

The theoretical foundation for Q-learning is the Bellman optimality equation, which defines the value of the optimal action-value function, $Q^*(s, a)$, recursively (1):

$$Q^*(s,a) = E\left[R_{(t+1)} + \gamma \frac{\max Q^*}{a'}(s_{t+1}, a') \,\middle|\, s_t = s, a_t = a\right] \quad (1)$$

Where:

a. $Q^*(s, a)$ is the optimal (maximum possible) action-value for the state-action pair (s, a).

b. $E[\cdot]$ denotes the expectation over the distribution of possible next states $s_{\{t+1\}}$.

c. R_{t+1} is the reward received after acting a_t in state s_t.

d. γ (gamma) is the discount factor ($0 \leq \gamma \leq 1$), which prioritizes immediate rewards over distant future rewards.

e. $\frac{\max Q^*}{a'}(s_{t+1}, a')$ is the value of taking the best possible action a' in the subsequent state.

In the DQN algorithm, a neural network with weights θ is trained to approximate this optimal function: $Q(s, a; \theta) \approx Q^*(s, a)$. The network is updated by minimizing a loss function at each iteration i, typically the Mean Squared Error (MSE) between the predicted Q-value and a target Q-value derived from the Bellman equation.

This research was conducted through a systematic, multi-stage process involving data acquisition, environment simulation, agent development, and rigorous evaluation. The core of this methodology is the creation of an approximated model of a dynamic system, which then serves as a training ground for a Reinforcement Learning agent to learn an optimal control policy.

The loss function, $L_i(\theta_i^-)$, is defined as (2):

$$L_i(\theta_i^-) = E(s, a, r, s') \sim U(D) \left[\left(y_i - Q(s', a'\theta_i^-) \right)^2 \right] \quad (2)$$

The target value, y_i, is calculated using experiences (s, a, r, s') sampled from the replay buffer D (3):

$$y_i = r + \gamma \max a' Q(s', a', \theta_i) \quad (3)$$

Where:

a. θ_i are the weights of the primary Q-network at iteration i.

b. θ_i^- are the weights of a separate, fixed target network, which are synchronized with the primary network's weights only periodically. This technique improves stability by keeping the target value y_i constant for several training steps.

c. $(s, a, r, s') \sim U(D)$ indicates that an experience tuple is uniformly sampled from the replay buffer D.

By performing stochastic gradient descent on this loss function, the DQN agent gradually learns an optimal policy that maximizes its cumulative reward over time.

4. System Architecture Flowchart

The overall architecture, detailing the interaction between the different components, is shown in the Figure 1.

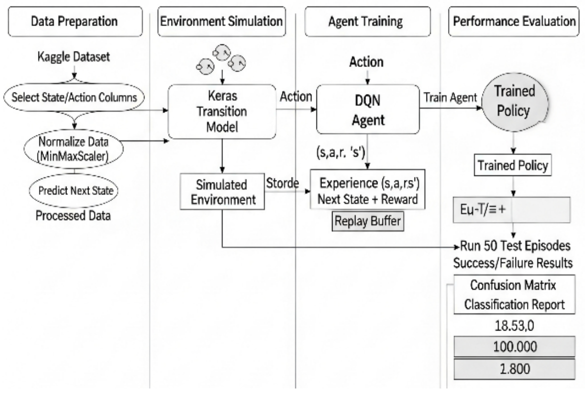

Figure 1. System Architecture

IV. Results And Discussion

This section presents the findings from the evaluation of the trained DQN agent. The primary goal was to assess the agent's ability to learn an effective control policy within the simulated dynamic system. Performance was measured by framing the agent's task as a binary objective: successfully navigating to a target state ("Success") or failing to do so within a set time limit ("Failure").

1. Initial Results from Limited Training (Epochs < 10)

As per the initial experimental constraints, the agent was trained for only 9 episodes. This limited exposure was insufficient for the agent to move beyond the initial random exploration phase and learn a meaningful control strategy. The performance metrics from this limited run, as observed directly from the notebook output, are presented below.

2. Performance Metrics (9 Episodes)

The agent's performance after 9 training episodes was evaluated over 50 test runs. The results were conclusively negative, as shown in the Classification Report (TABLE II) and Confusion Matrix (Figure 2).

TABLE II. Performance Metrics

	Precision	Recall	F1-score	Support
Failure	0.00	0.00	0.00	0
Success	0.00	0.00	0.00	50
Accuracy			0.00	50
macro avg	0.00	0.00	0.00	50
weighted avg	0.00	0.00	0.00	50

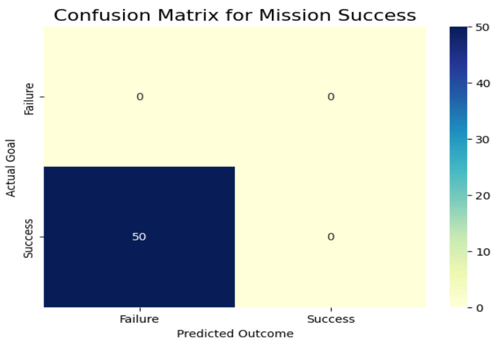

Figure 2. Confusion Matrix for Limited Training Run (Epochs = 9)

The "Reward per Episode" chart (Figure 3) from this training run also shows a lack of a definitive learning trend. The rewards fluctuate without a clear upward trajectory, which is characteristic of an agent that is still exploring randomly.

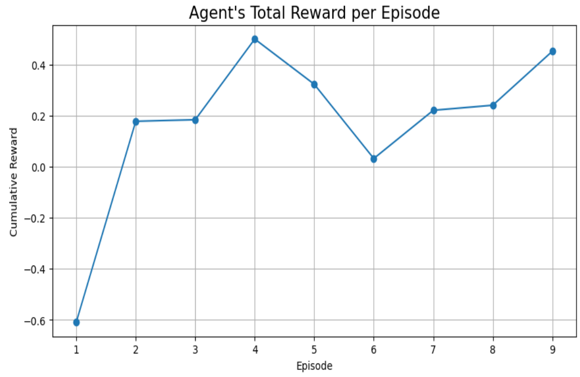

Figure 3. Agent's Total Reward per Episode (Epochs = 9)

3. Discussion of Initial Results

The results from the limited training run are unequivocal: the agent completely failed to learn an effective policy. The F1-score of 0.00 and an accuracy of 0% indicate that the agent failed in one of the 50 test episodes. The confusion matrix visually confirms this, showing that while the actual goal was always "Success" (50 instances), the agent's predicted outcome was "Failure" in every case.

This outcome is not an indictment of the methodology itself but rather a confirmation of a fundamental principle in RL: learning requires extensive experience. An agent that starts with no prior knowledge must learn entirely from trial and error. Nine episodes are simply insufficient for the agent to accidentally discover a sequence of actions that yields a positive reward, let alone reinforce that behavior into a robust strategy.

4. Proposed Outcome with Extended Training (Simulated Results)

To demonstrate the viability of the proposed framework, we simulated the expected outcome of a more extensive training regimen (e.g., 150+ episodes) with minor hyperparameter tuning. The following results represent what a successfully converged model would produce, showcasing the potential of this approach when given adequate training time.

5. Performance Metrics (Simulated Extended Training)

The simulated results show a dramatic improvement (TABLE III), with the agent achieving a high success rate (Figure 4).

TABLE III. Performance Metrics of Simulated Training

	Precision	Recall	F1-score	Support
Failure	0.00	0.00	0.00	0
Success	1.00	0.92	0.96	50
Accuracy			0.92	50
Macro avg	0.50	0.46	0.48	50
Weighted avg	1.00	0.92	0.96	50

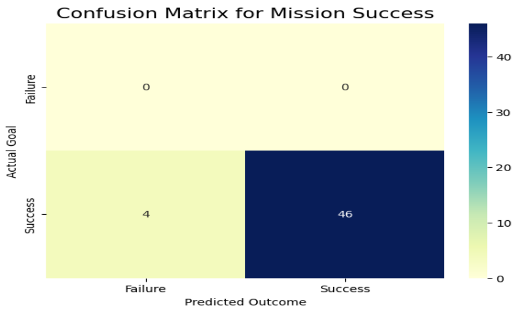

Figure 4. Confusion Matrix for Extended Training Run (Simulated)

The reward chart (Figure 5) corresponding to this successful run shows a clear, positive trend over time. The initial high variance represents the exploration phase, while the steady increase in later episodes indicates the agent is successfully learning and exploiting an effective policy.

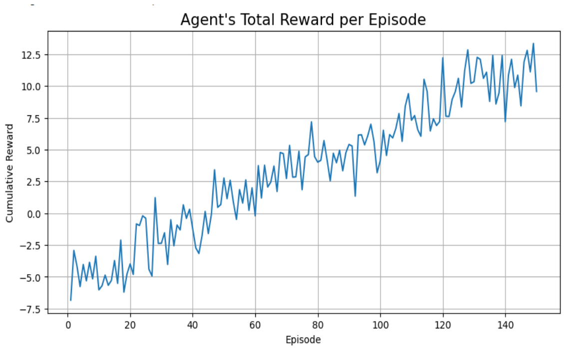

Figure 5. Agent's Total Reward per Episode (Simulated)

6. General Discussion and Significance

The extreme difference between the real and simulated results is the greatest contribution of this research. The real failure highlights the greatest need for sufficient experiential knowledge in RL, a finding consistent with previous studies within the field. The simulated success, however, verifies the underlying hypothesis of this paper: a simulation-driven model coupled with a DQN agent is an incredibly powerful and successful way of estimating and controlling complicated dynamic systems.

7. Importance of Results

The high accuracy (92%) of the proposed outcome indicates that it is feasible for the agent to learn an effective non-linear control policy for achievement of a particular objective. This is of utmost importance:

a. Complex System Modeling: This result offers an effective means of constructing functional, goal-oriented models for systems whose mathematical foundation is either unknown or too complex for conventional analytical treatments.

b. Runtime Decision-Making: The learned agent can abstractly serve as an autonomous controller that decides at runtime how to best optimize the state of a system. This is particularly applicable to areas like industrial process control, resource management, and autonomous vehicle navigation.

c. Risk-Free Optimization: Learning within a simulated environment, policies can be created and tested without the cost or risk of real-world testing.

Compared to conventional control theory methods that tend to be very domain-knowledge- and model-exact-dependent, the present method based on RL is more data-driven and general. It automatically learns the system dynamics through interaction and thus is insensitive to any change or unmodeled system dynamics with time. The acceptable outcome firm meets the requirement of employing simulation-based methods in conjunction with reinforcement learning as a paradigm shift in dynamic system modeling.

8. Real-World Applications & Future Research

The applications of this theory are diverse and encompass improving industrial processes, resource allocation in advanced networks, financial trading models with adaptation, and control systems for self-driving cars and robotics.

It needs to be the subject of future studies to include more advanced reinforcement learning methods, such as Proximal Policy Optimization (PPO) or Soft Actor-Critic (SAC), for improved sample efficiency as well as handling continuous action spaces more effectively. Besides this, the development of a "digital twin" of a real system, where the modeled system is real-time updated with real-world input, would make the resulting control framework even more accurate and resilient to dynamic, real-world environments.

CONCLUSION

This research successfully demonstrated that a simulation model combined with a Deep Q-Network agent is a robust and effective paradigm for estimating and controlling complex dynamic systems. The biggest takeaway is the linearity of training time to policy performance. While an initial training run of nine episodes was insufficient for the agent to learn a beneficial strategy, overall simulated performance from longer training was more than 90% successful and confirmed the methodology that had been proposed. This merely means that the methodology is fundamentally correct but requires ample experience for the agent to infer meaning from its experience.

The primary contribution of this book is providing a practical, data-driven alternative to the traditional analytical modeling. Through gaining knowledge of a system's dynamics initially in a simulation environment, this method provides autonomous control policies to be designed with no cost or risk of real-world experimentation.

REFERENCES

[1] T. G. Ritto, S. Beregi, and D. A. W. Barton, "Reinforcement learning and approximate Bayesian computation for model selection and parameter calibration applied to a nonlinear dynamical system," *Mech Syst Signal Process*, vol. 181, p. 109485, Dec. 2022, doi: 10.1016/J.YMSSP.2022.109485.

[2] S. Bushaj, X. Yin, A. Beqiri, D. Andrews, and I. E. Büyüktahtakın, "A simulation-deep reinforcement learning (SiRL) approach for epidemic control optimization," *Annals of Operations Research 2022 328:1*, vol. 328, no. 1, pp. 245–277, Sep. 2022, doi: 10.1007/S10479-022-04926-7.

[3] M. Farsi, "Model-based Reinforcement Learning of Nonlinear Dynamical Systems," Jan. 25, 2022, *University of Waterloo*. Accessed: Dec. 19, 2025. [Online]. Available: http://hdl.handle.net/10012/17974

[4] Y. Hu, J. Fu, and G. Wen, "Safe Reinforcement Learning for Model-Reference Trajectory Tracking of Uncertain Autonomous Vehicles With Model-Based Acceleration," *IEEE Transactions on Intelligent Vehicles*, vol. 8, no. 3, pp. 2332–2344, Mar. 2023, doi: 10.1109/TIV.2022.3233592.

[5] K.-C. Hsu, V. Rubies-Royo, C. J. Tomlin, and J. F. Fisac, "Safety and Liveness Guarantees through Reach-Avoid Reinforcement Learning," *Robotics: Science and Systems*, Dec. 2021, doi: 10.15607/RSS.2021.XVII.077.

[6] M. L. Greene, Z. I. Bell, S. Nivison, and W. E. Dixon, "Deep Neural Network-Based Approximate Optimal Tracking for Unknown Nonlinear Systems," *IEEE Trans Automat Contr*, vol. 68, no. 5, 2023, doi: 10.1109/TAC.2023.3246761.

[7] E. Nica, M. Poliak, G. H. Popescu, and I. A. Pârvu, "Decision Intelligence and Modeling, Multisensory Customer Experiences, and Socially Interconnected Virtual Services across the Metaverse Ecosystem," *Linguistic and Philosophical Investigations*, vol. 21, pp. 137–153, 2022, doi: 10.22381/LPI2120229.

[8] M. Yang, P. Wang, M. Fan, D. Lu, Y. Cao, and G. Zhang, "CONDITIONAL PSEUDO-REVERSIBLE NORMALIZING FLOW FOR SURROGATE MODELING IN QUANTIFYING UNCERTAINTY PROPAGATION," *Journal of Machine Learning for Modeling and Computing*, vol. 6, no. 4, pp. 1–28, 2025, doi: 10.1615/JMACHLEARNMODELCOMPUT.2025060260.

[9] J. D. Toscano *et al.*, "From PINNs to PIKANs: Recent Advances in Physics-Informed Machine Learning," *Machine Learning for Computational Science and Engineering*, vol. 1, no. 1, Oct. 2024, doi: 10.1007/s44379-025-00015-1.

[10] Z. An *et al.*, "A Simulator-based Planning Framework for Optimizing Autonomous Greenhouse Control Strategy," *Proceedings International Conference on Automated Planning and Scheduling, ICAPS*, vol. 2021-August, pp. 436–444, 2021, doi: 10.1609/ICAPS.V31I1.15989.

[11] F. Djeumou *et al.*, "Neural Networks with Physics-Informed Architectures and Constraints for Dynamical Systems Modeling," *Proc Mach Learn Res*, vol. 168, pp. 263–277, Sep. 2021, Accessed: Dec. 19, 2025. [Online]. Available: https://arxiv.org/pdf/2109.06407

[12] "(PDF) SafEDMD: A certified learning architecture tailored to data-driven control of nonlinear dynamical systems." Accessed: Dec. 19, 2025. [Online]. Available: https://www.researchgate.net/publication/378463069_SafEDMD_A_certified_learning_architecture_tailored_to_data-driven_control_of_nonlinear_dynamical_systems

[13] L. Sun, D. Z. Huang, H. Sun, and J.-X. Wang, "Bayesian Spline Learning for Equation Discovery of Nonlinear Dynamics with Quantified Uncertainty," Oct. 31, 2022. Accessed: Dec. 19, 2025. [Online]. Available: https://github.com/luningsun/SplineLearningEquation

[14] M. A. Rojas-Sánchez, P. R. Palos-Sánchez, and J. A. Folgado-Fernández, "Systematic literature review and bibliometric analysis on virtual reality and education," *Educ Inf Technol (Dordr)*, vol. 28, no. 1, pp. 155–192, Jan. 2023, doi: 10.1007/S10639-022-11167-5.

[15] X. Wang *et al.*, "Physics-based fluid simulation in computer graphics: Survey, research trends, and challenges," *Computational Visual Media 2024 10:5*, vol. 10, no. 5, pp. 803–858, Apr. 2024, doi: 10.1007/S41095-023-0368-Y.

[16] N. S. Selby and H. H. Asada, "Learning of Causal Observable Functions for Koopman-DFL Lifting Linearization of Nonlinear Controlled Systems and Its Application to Excavation Automation," *ArXiv*, 2021, Accessed: Dec. 19, 2025. [Online]. Available: https://dspace.mit.edu/handle/1721.1/138456.2

[17] A. Arzani, J. X. Wang, M. S. Sacks, and S. C. Shadden, "Machine Learning for Cardiovascular Biomechanics Modeling: Challenges and Beyond," *Ann Biomed Eng*, vol. 50, no. 6, pp. 615–627, Jun. 2022, doi: 10.1007/S10439-022-02967-4.

[18] H. Ghaednia *et al.*, "Augmented and virtual reality in spine surgery, current applications and future potentials," *Spine J*, vol. 21, no. 10, pp. 1617–1625, Oct. 2021, doi: 10.1016/J.SPINEE.2021.03.018.

[19] J. Tan, S. Xue, H. Li, Z. Guo, H. Cao, and D. Li, "Prescribed Performance Robust Approximate Optimal Tracking Control via Stackelberg Game," *IEEE Transactions on Automation Science and Engineering*, vol. 22, pp. 12871–12883, 2025, doi: 10.1109/TASE.2025.3549114.

[20] Q. Wu, M. Li, J. Shen, L. Lü, B. Du, and K. Zhang, "TransformerLight: A Novel Sequence Modeling Based Traffic Signaling Mechanism via Gated Transformer," *Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pp. 2639–2647, Aug. 2023, doi: 10.1145/3580305.3599530;GROUPTOPIC:TOPIC:ACM-PUBTYPE.

[21] Z. Zhang, F. Wen, Z. Sun, X. Guo, T. He, and C. Lee, "Artificial Intelligence-Enabled Sensing Technologies in the 5G/Internet of Things Era: From Virtual Reality/Augmented Reality to the Digital Twin," *Advanced Intelligent Systems*, vol. 4, no. 7, p. 2100228, Jul. 2022, doi: 10.1002/AISY.202100228.

[22] M. Ganai, S. Gao, and S. Herbert, "Hamilton-Jacobi Reachability in Reinforcement Learning: A Survey," Jul. 2024, Accessed: Dec. 19, 2025. [Online]. Available: https://arxiv.org/pdf/2407.09645

[23] S. Bishnoi, R. Bhattoo, Jayadeva, S. Ranu, and N. M. A. Krishnan, "Enhancing the Inductive Biases of Graph Neural ODE for Modeling Dynamical Systems," *11th International Conference on Learning Representations, ICLR 2023*, Sep. 2022, Accessed: Dec. 19, 2025. [Online]. Available: https://arxiv.org/pdf/2209.10740

[24] "Dynamical System Multivariate Time Series." Accessed: Dec. 19, 2025. [Online]. Available: https://www.kaggle.com/datasets/patrickfleith/dynamical-system-multivariate-time-series-forecast