Received: May 30, 2025; **Accepted:** October 16, 2025; **Published:** January 1, 2026.
DOI. 10.30473/coam.2025.72730.1312
Volume 11, Issue 1, p.p. 199-218: Winter-Spring (2026)
Research Article

Open Access

Control and Optimization in Applied Mathematics - COAM

Feedback Long Short-Term Memory: A Long Short-Term Memory-Based Framework for Multivariate Time Series Prediction in Chaotic Systems

Saeed Mirzajani[1], **Majid Roohi**[2]

[1]Department of Basic Sciences, Technical and Vocational University (TVU), Tehran, Iran.
[2]Department of Mathematics, Aarhus University, Denmark.

Correspondence:
Saeed Mirzajani
E-mail:
smirzajani@tvu.ac.ir

How to Cite
Mirzajani, S., Roohi, M. (2026). "Feedback long short-term memory: A long short-term memory-based framework for multivariate time series prediction in chaotic systems", Control and Optimization in Applied Mathematics, 11(1): 199-218, doi: 10.30473/coam.2025.72730.1312.

Abstract. The prediction of chaotic time series is essential for understanding highly nonlinear and sensitive systems, with the Lorenz system serving as a standard benchmark due to its intricate and non-periodic dynamics. Classical forecasting approaches often struggle to capture such irregularities, motivating a shift toward deep learning–based strategies. In this study, we develop two hybrid models—Feedback Long Short-Term Memory (FB-LSTM) and Feedback Variational Stacked LSTM (FBVS-LSTM), specifically designed for multivariate prediction of the Lorenz system. By embedding feedback structures into LSTM networks, the proposed methods deliver enhanced short-term prediction performance without substantial computational costs. Comparative simulations indicate that our frameworks surpass traditional RNNs and baseline LSTM models, achieving prediction accuracies up to 94%. These findings indicate that feedback-enhanced architectures offer effective and practical tools for forecasting chaotic systems, with potential applications in both scientific research and engineering practice.

Keywords. Deep learning, Lorenz system, Neural networks, Time-series forecasting, Global convergence, Multivariate sequence prediction

MSC. 37D45; 37M10.

1 Introduction

Understanding the behavior of dynamic systems often relies on analyzing time series—sequences of observations collected at regular intervals. Time series analysis provides valuable insights into system evolution and can reveal patterns that are not immediately evident [50]. While some systems display smooth and periodic behavior, many real-world phenomena are inherently unpredictable, exhibiting irregular and complex dynamics [3, 47]. Extreme sensitivity to initial conditions and the absence of repeatable patterns are key hallmarks of chaotic systems [9, 15].

A classical example of chaotic behavior is the Lorenz system, which consists of three coupled nonlinear ordinary differential equations. Initially developed to model atmospheric convection, the Lorenz equations quickly became a cornerstone of chaos theory, demonstrating how minute variations in initial conditions can lead to drastically different outcomes—a phenomenon popularly known as the "butterfly effect" [25, 49]. By capturing the behavior of highly sensitive and unpredictable systems, the Lorenz model provides a foundational framework for studying deterministic chaos and has guided numerous efforts in forecasting complex dynamical phenomena [5, 33].

Following the Lorenz model, classical forecasting techniques—including autoregressive (AR) models, moving average (MA) models, and ARIMA extensions—were employed to predict chaotic dynamics [10, 17]. Despite their widespread application, these traditional approaches often struggle to capture the intricate nonlinear dependencies inherent to chaotic systems. To address these limitations, modern machine learning methods, particularly deep learning approaches, have attracted significant attention due to their superior ability to model complex temporal patterns and nonlinear relationships [7, 34, 36, 42].

Among deep learning techniques, Convolutional Neural Networks (CNNs) are widely used for extracting spatial and temporal features from structured time series. Their multi-layered architecture facilitates automatic feature extraction, proving effective in tasks such as anomaly detection and missing data reconstruction [1, 31]. However, CNNs may be limited in sequential prediction tasks due to challenges in capturing long-term dependencies and the potential accumulation of gradient and hidden layer errors. Empirical evidence suggests that recurrent architectures—particularly Long Short-Term Memory (LSTM) networks—often outperform CNNs in tasks requiring accurate temporal modeling and stable gradient propagation across multiple time steps [45].

Consequently, there is growing interest in hybrid and customized recurrent network models for forecasting chaotic systems [2, 35]. These models leverage LSTM's memory retention capabilities while incorporating enhancements tailored for specific applications, such as multivariate time series prediction in chaotic environments [8, 13].

Recent studies have explored a variety of deep learning methods for chaotic and multivariate time series forecasting. For instance, Chen chaotic systems have been integrated with LSTM, Neural Basis Expansion, and Transformer models, improving both prediction accuracy and computational efficiency [19]. Comparative analyses of NG-RC, RC, and LSTM models for chaotic systems—including Lorenz, Rössler, Chen, and Qi—demonstrate that NG-RC provides an optimal balance of computational efficiency and predictive performance [32]. Other approaches, such as nonlinear spiking neural P systems combined with non-subsampled shearlet transforms, have enhanced multivariate forecasting for nonlinear, non-stationary, and high-dimensional time series [24]. Attention-based CNN-LSTM hybrids have also been developed to improve multivariate urban water demand forecasting, outperforming standard LSTM and CNN-LSTM models [51]. Conversely, combining LSTM with transformers or conventional statistical methods appears less effective for strongly chaotic series, highlighting the importance of aligning model choice with the underlying dynamics [27].

Building upon these developments, this work introduces two feedback-based architectures—FB-LSTM and FBVS-LSTM—that incorporate feedback mechanisms and variational stacking. These models are designed to capture short-term dynamics in the Lorenz system and enhance multivariate chaotic time series forecasting. Despite recent advances, existing models often struggle with: (i) accurately capturing short-term fluctuations characteristic of chaotic systems, (ii) lacking adaptive feedback mechanisms to correct prediction drift, and (iii) limited comparative evaluation against established benchmarks. Our proposed architectures directly address these challenges.

Specifically, the FB-LSTM integrates a feedback loop that feeds predictions back into the model as auxiliary inputs, allowing dynamic adaptation to rapid changes in system behavior. The FBVS-LSTM extends this design with a variational stacked structure, strengthening feature representation and mitigating overfitting in multivariate prediction tasks.

The main objective of this study is to forecast future, unobserved points generated by the Lorenz chaotic system using these advanced hybrid models. The proposed architectures are evaluated through standard metrics (MSE, RMSE, MAE, accuracy) and validated via cross-validation. Their performance is systematically compared with baseline models, including RNN and conventional LSTM, as well as previously reported results. Our findings demonstrate that FB-LSTM and FBVS-LSTM achieve superior prediction accuracy—up to 94%—offering a reliable and effective solution for chaotic time series forecasting.

By integrating these methodologies, our models leverage the strengths of each component: LSTM's ability to capture long-term temporal dependencies, ensemble learning's variance reduction and stability improvement, and reinforcement learning's capacity for dynamic adjustment based on performance feedback. Together, they form a comprehensive framework capable of accurately modeling the nonlinear, non-stationary behavior characteristic of chaotic systems.

In summary, the main contributions of this work are as follows:

1. FB-LSTM and FBVS-LSTM incorporate a feedback loop that reintroduces the model's predictions into the input, enabling dynamic adaptation to short-term fluctuations and resulting in more stable and accurate forecasts.

2. FBVS-LSTM further enhances the framework with a variational stacked structure, improving feature representation and mitigating overfitting in multivariate forecasting tasks.

3. Unlike general deep learning models, our architectures are specifically designed for chaotic systems such as the Lorenz model, where short-term dynamics dominate.

4. Experimental results demonstrate that the proposed models achieve prediction accuracies of up to 94%, outperforming baseline RNN and standard LSTM models, validating the effectiveness of incorporating feedback mechanisms into deep learning for chaotic system prediction.

The research methodology is structured into several key phases, each addressing a specific aspect of model development and evaluation. These stages, including data generation, preprocessing, model construction, training, and validation, are thoroughly detailed in Section 3. Subsequently, Section 4 presents a comprehensive analysis of the experimental results, discussing the model's performance in comparison with baseline methods and previous studies. Finally, Section 5 concludes the paper, summarizing key findings and offering directions for future research.

2 Background

Over the past decades, numerous studies have explored the prediction of chaotic systems, leading to the development of various computational models aimed at achieving high accuracy in time series forecasting. Early attempts focused on traditional machine learning algorithms, including perceptron neural networks [12], support vector machines (SVM) [28], and other classification-based approaches [30], which remained in use even in recent years [14]. These methods demonstrated satisfactory performance in certain applications. However, they were fundamentally limited in capturing long-term dependencies—an essential requirement when modeling the complex temporal dynamics of chaotic systems.

To overcome this limitation, researchers began incorporating memory-based neural network architectures into their models. In [52], a basic Long Short-Term Memory (LSTM) model was employed to track chaotic time series data. A health index (HI) was also introduced to evaluate the degradation condition of components such as slip rings over time, highlighting both long-term and short-term memory capabilities. Similarly, [4] used an LSTM network to predict chaotic behavior and assessed model performance using Mean Squared Error (MSE) and Root Mean Squared Error (RMSE) metrics.

In [29], a novel approach called PI-LSTM (Physics-Informed LSTM) was used to reconstruct unmeasured variables within a chaotic system by integrating physical constraints into the learning process. This model penalized physically inconsistent outputs, significantly enhancing predictive reliability. Meanwhile, [44] proposed a hybrid architecture combining Convolutional Neural Networks with Bidirectional LSTM (CNN-BLSTM), which demonstrated faster training times and lower error rates when learning from chaotic sequences with missing data.

A similar hybrid strategy was adopted in [11], where a deep neural network comprising temporal convolutional layers for low-level feature extraction and recurrent units (LSTM and GRU) for temporal sequence modeling yielded a prediction accuracy of up to 80%. Inspired by the neural activity of biological systems, [23] introduced a spiking neural model (NSNP-AU), which outperformed conventional CNN-based architectures in chaotic time series forecasting.

Further advancements include the Multi-Attn BLS model introduced in [39], which utilized multiple attention mechanisms to capture the spatiotemporal dynamics of chaotic systems and demonstrated strong generalization in complex nonlinear environments. In [40], hybrid models integrating convolution layers and short-term memory modules were shown to successfully predict data from chaotic systems lacking periodic or quasi-periodic features.

In [38], researchers reconstructed the state space of financial time series data using optimal delay embeddings and predicted chaotic liquidity demands using D-CNN + LSTM and D-CNN + GRU models. In another study [22], a novel three-dimensional natural exponential chaos system, highly sensitive to initial input conditions, was used to simulate unrepeatable and non-deterministic time series data. Additionally, [48] introduced a hybrid model combining CEEMDAN (Complete Ensemble Empirical Mode Decomposition with Adaptive Noise) and LSTM. This model significantly improved prediction performance on benchmark chaotic systems like Lorenz-63, outperforming ARIMA, SVM, multilayer perceptron (MLP), and standalone LSTM models.

The importance of memory-based architectures in modeling chaotic dynamics has also been emphasized in [26], where recurrent models like LSTM were recognized for their effectiveness in capturing the nonlinear characteristics of dynamic systems. Despite the utility of classical and statistical approaches [6], machine learning-based methods have proven superior in predicting future values of non-deterministic, non-periodic systems [46]. A further contribution came from [43], where a stochastic mean model (MSM-LSTM) was proposed to ensure stable convergence in high-dimensional chaotic systems.

Lei, et al. evaluated the LSTM architectures of varying depths on the Mackey-Glass and Kuramoto–Sivashinsky systems using RMSE and Anomaly Correlation Coefficient (ACC) as performance metrics [21]. The findings corroborate the robustness of LSTMs in forecasting chaotic dynamics. Likewise, Sangiorgio and Dercole examined the effects of mismatched input lags and embedding dimensions, reporting that LSTM models maintain reliable performance across a range of input configurations [37]. In [20], a recurrent neural network augmented with an exception gate demonstrated comparable performance on par with LSTM and GRU models for tracking posterior information in chaotic systems.

Despite this extensive body of work, there remains a lack of comprehensive studies comparing multiple LSTM-based architectures under the same experimental conditions for chaos prediction. This study addresses that gap by proposing and evaluating two novel architectures: FB-LSTM and FBVS-LSTM. These models are designed to forecast the dynamics of the Lorenz chaotic system and incorporate optimization techniques from ensemble learning and reinforcement learning. Our goal is to assess whether FB-LSTM or FBVS-LSTM outperforms conventional LSTM and RNN models in predicting chaotic time series.

To the best of our knowledge, this is the first investigation of FB-LSTM-type architectures in the context of chaotic system prediction. We hypothesize that despite not being previously applied to this domain, the proposed models—due to their hybrid design and enhanced learning mechanisms—can yield superior accuracy and generalization in predicting chaotic sequences.

Remark 1. The motivation behind introducing FB-LSTM and FBVS-LSTM lies in their ability to overcome key weaknesses of existing LSTM variants. In contrast to models such as PI-LSTM and CNN-LSTM, our frameworks include feedback loops that feed predictions back into the network as auxiliary inputs, enabling dynamic adjustment to the rapid short-term fluctuations that are typical of chaotic systems. The FBVS-LSTM further extends this idea by adding a variational stacked structure, which improves feature representation and helps control overfitting in multivariate forecasting. These design choices directly tackle challenges faced by conventional LSTM variants, such as prediction drift, the absence of adaptive feedback, and limited handling of multivariate chaotic dynamics. Together, these distinctions underline both the novelty and the practical advantages of our proposed models.

3 Methodology

This section presents a comparative investigation of four deep learning-based approaches—LSTM, RNN, FB-LSTM, and FBVS-LSTM—for forecasting chaotic time series. The objective is to identify the model that most effectively captures the dynamics of chaotic systems and delivers accurate predictions. We train each model on data derived from a simulated chaotic system and assess its performance using established evaluation metrics.

3.1 Modeling Chaotic Dynamics

Chaotic systems are marked by extreme sensitivity to initial states, where minimal changes can lead to significantly divergent trajectories. This unpredictability challenges traditional forecasting techniques but also serves as a robust testing ground for assessing model generalization.
In this study, we adopt the Lorenz system as our case study, a classic example of deterministic chaos often used to represent atmospheric convection. It consists of a trio of nonlinear differential equations that interact over time, producing complex and non-repetitive patterns. The Lorenz system is governed by the following equations:

$$\begin{cases} \frac{dx}{dt} = \sigma(y - x), \\ \frac{dy}{dt} = x(r - z) - y, \\ \frac{dz}{dt} = xy - bz, \end{cases} \tag{1}$$

where x, y, and z denote the evolving state variables of the system. Parameters $\sigma = 10$, $r = 28$, and $b = 8/3$ are constants representing the physical aspects of the model, namely the Prandtl number, Rayleigh number, and a geometric factor, respectively. The initial condition $(x_0, y_0, z_0) = (1.0, 1.0, 1.0)$ is used to illustrate sensitivity and test prediction fidelity. These trajectories are computed using a numerical integration method, specifically the 4th-order Runge–Kutta algorithm with a fixed time step. The resulting time series are used to train and evaluate the predictive models.

Remark 2. The Lorenz system is one of the most commonly used benchmarks for studying chaotic dynamics because of its nonlinear behavior, sensitivity to initial conditions, and complex temporal patterns. Its extensive use in earlier research makes it a natural reference point for evaluating and comparing prediction models, including the FB-LSTM and FBVS-LSTM architectures proposed in this study.

Remark 3. The key properties of chaotic systems—especially their sensitivity to initial conditions and highly nonlinear, non-repetitive behavior—play a central role in modeling and are carefully considered in designing and evaluating the proposed FB-LSTM and FBVS-LSTM models. These features make the Lorenz system a suitable benchmark and highlight the need for advanced recurrent architectures that can effectively capture short-term fluctuations and complex temporal dependencies in chaotic time series prediction.

3.2 Long Short-Term Memory (LSTM) Modeling

Long Short-Term Memory (LSTM) networks were introduced by Hochreiter and Schmidhuber in 1997 [18] to address the limitations of standard recurrent neural networks, particularly the vanishing gradient problem. The LSTM architecture uses a memory cell to retain information over long periods, with three gates that regulate the flow of information into, within, and out of the cell.

The input gate determines how much of the new input should influence the memory cell:

$$i_t = \sigma(W_i x_t + U_i h_{t-1} + b_i), \tag{2}$$

The forget gate decides which part of the previous memory should be discarded:

$$f_t = \sigma(W_f x_t + U_f h_{t-1} + b_f), \tag{3}$$

The output gate manages the information that is passed to the next hidden state:

$$o_t = \sigma(W_o x_t + U_o h_{t-1} + b_o), \tag{4}$$

The cell state is updated by combining the old state and the new input:

$$c_t = f_t \odot c_{t-1} + \big(i_t \odot \phi(W_c x_t + U_c h_{t-1} + b_c)\big), \tag{5}$$

Finally, the hidden state at time t is given by:

$$h_t = o_t \odot \phi(c_t), \tag{6}$$

where σ denotes the sigmoid activation function, ϕ is the hyperbolic tangent (tanh), and $\odot$ represents element-wise multiplication. Variables x_t and h_{t-1} refer to the current input and previous hidden state, while W, U, and b are the respective weight matrices and bias vectors.

While the standard LSTM model performs well on sequential data, it faces limitations when dealing with complex multivariate time series, especially in the presence of missing or noisy data. To overcome these challenges, advanced variants such as FB-LSTM and FBVS-LSTM have been proposed. These models incorporate feedback and variable selection mechanisms, making them more effective for modeling nonlinear chaotic systems [16]. In this study, we evaluate the performance of these advanced LSTM models, alongside basic LSTM and RNN architectures, to identify the most reliable method for forecasting chaotic time series generated by the Lorenz system. The evaluation framework for this study is illustrated into two parts: Figure 1 and Algorithm 9. Figure 1 provides a visual overview of the modeling and validation process used for chaotic system forecasting. Algorithm 9 outlines the step-by-step algorithm for the FB-LSTM and FBVS-LSTM models, highlighting the main stages from data generation to model evaluation.

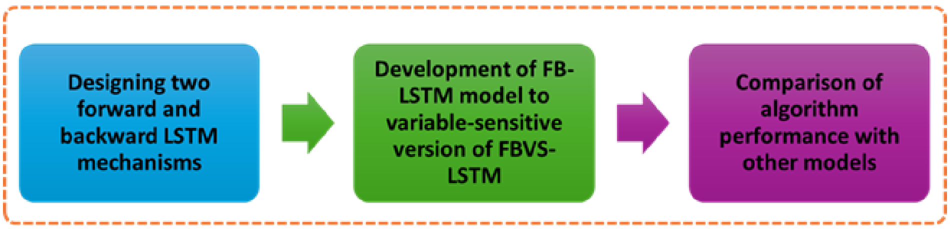

Figure 1: Flowchart illustrating the overall modeling and validation process.

3.3 FB-LSTM Modeling

To effectively handle missing data categorized as Missing Not At Random (MNAR), we enhance the traditional LSTM by incorporating two directional mechanisms—forward and backward processing—

Algorithm 9 Chaotic time series prediction using FB-LSTM and FBVS-LSTM

Input: Lorenz time series $X = \{x_1, \dots, x_n\}$, initial conditions, model hyperparameters

Output: Predicted time series $\hat{x}$

Step 1: Simulation of the Lorenz system

Generate the time series X using the fourth-order Runge–Kutta method with parameters $\sigma = 10$, $r = 28$, and $b = \frac{8}{3}$.

Step 2: Data preprocessing

For each variable d and time step t:

- Compute the missing-data mask m_t^d.
- Compute forward and backward time intervals δ_t^1 and δ_t^2.
- Compute decay factors γ_t^1 and γ_t^2.
- Impute missing values $\hat{x}_t^d$ using decay-modulated aggregation.

Step 3: Model initialization

Initialize the RNN, LSTM, FB-LSTM, and FBVS-LSTM architectures.

Step 4: FB-LSTM forward propagation

For $t = 1$ to T:

- Adjust the previous hidden state $\hat{h}_{t-1}$.
- Update LSTM gates using decay-adjusted inputs and mask m_t.
- Compute the hidden state h_t and cell state c_t.

Step 5: FBVS-LSTM forward propagation

For $t = 1$ to T:

- Compute variable sensitivity μ^d.
- Compute the modulation factor β.
- Update gates, hidden state h_t, and cell state c_t using m_t and β.

Step 6: Model training

Train all models using backpropagation to minimize the mean squared error (MSE).

Step 7: Prediction and evaluation

Predict future values $\hat{x}$ and evaluate model performance using MSE.

resulting in the Forward-Backward LSTM (FB-LSTM). These mechanisms decompose missing patterns, improving the precision of imputing absent values.

Define the missing data indicator matrix $M = \{m_1, m_2, \dots, m_T\}^T \in \mathbb{R}^{D\times T}$, corresponding to time series $X = \{x_1, x_2, \dots, x_T\}^T \in \mathbb{R}^{D\times T}$, where x_t^d denotes the observation at time t for variable d. The binary mask m_t^d indicates whether a value is observed:

$$m_t^d = \begin{cases} 1, & \text{if } x_t^d \text{ is observed,} \\ 0, & \text{otherwise.} \end{cases} \tag{7}$$

We compute elapsed time since the last observation as $\delta^1 = \{\delta_1^1, \dots, \delta_T^1\}^T$:

$$\delta_t^{1d} = \begin{cases} s_t - s_{t-1} + \delta_{t-1}^{1d}, & t > 1, m_{t-1}^d = 0, \\ s_t - s_{t-1}, & t > 1, m_{t-1}^d = 1, \\ 0, & t = 1, \end{cases} \tag{8}$$

where s_t is the timestamp of observation t. Forward intervals are defined similarly:

$$\delta_t^{2d} = \begin{cases} s_{t+1} - s_t + \delta_{t+1}^{2d}, & t > 1, m_{t-1}^d = 0, \\ s_{t+1} - s_t, & t > 1, m_{t-1}^d = 1, \\ 0, & t = 1. \end{cases} \tag{9}$$

We define exponential decay functions:

$$\gamma^1 = \exp\{-\max(0, W_\gamma^1 \delta^1 + b_\gamma^1)\}, \tag{10}$$

$$\gamma^2 = \exp\{-\max(0, W_\gamma^2 \delta^2 + b_\gamma^2)\}. \tag{11}$$

The imputation for missing input $\hat{x}_t^d$ is:

$$\hat{x}_t^d = m_t^d x_t^d + (1 - m_t^d)\Big[(\gamma_t^{1d} x_{t'}^d + \gamma_t^{2d} x_{t''}^d) + (1 - \gamma_t^{1d})(1 - \gamma_t^{2d})\tilde{x}^d\Big], \tag{12}$$

where $x_{t'}^d, x_{t''}^d$ are nearest past and future observations, and $\tilde{x}^d$ is the mean.
Decay-adjusted hidden state update:

$$\hat{h}_{t-1} = \gamma_{ht}^1 \gamma_{ht}^2 \odot h_{t-1}. \tag{13}$$

FB-LSTM gate updates:

$$f_t = \sigma(W_f \hat{x}_t + U_f \hat{h}_{t-1} + V_f m_t + b_f), \tag{14}$$

$$i_t = \sigma(W_i \hat{x}_t + U_i \hat{h}_{t-1} + V_i m_t + b_i), \tag{15}$$

$$o_t = \sigma(W_o \hat{x}_t + U_o \hat{h}_{t-1} + V_o m_t + b_o), \tag{16}$$

$$c_t = f_t \odot c_{t-1} + (i_t \odot \phi(W_c \hat{x}_t + U_c \hat{h}_{t-1} + V_c m_t + b_c)), \tag{17}$$

$$h_t = o_t \odot \phi(c_t). \tag{18}$$

3.4 FBVS-LSTM

In multivariate time series, each variable may show distinct missing patterns. To capture this, FBVS-LSTM introduces a variable-sensitive coefficient:

$$\mu^{(d)} = 1 - \frac{1}{T}\sum_{t=1}^{T} m_t^d, \tag{19}$$

representing the missing rate of variable d.

We define the sensitivity factor:

$$\beta = \exp\{\max(0, W_\beta \mu + b_\beta)\}. \tag{20}$$

Modified gate updates:

$$f_t = \sigma(W_f \hat{x}_t + U_f h_{t-1} + V_f m_t + P_f \beta + b_f), \tag{21}$$

$$i_t = \sigma(W_i \hat{x}_t + U_i h_{t-1} + V_i m_t + P_i \beta + b_i), \tag{22}$$

$$c_t = f_t \odot c_{t-1} + (i_t \odot \phi(W_c \hat{x}_t + U_c h_{t-1} + V_c m_t + P_c \beta + b_c)), \tag{23}$$

$$h_t = o_t \odot \phi(c_t). \tag{24}$$

Performance is evaluated using Mean Squared Error (MSE):

$$MSE = \frac{1}{n}\sum_{i=1}^{n} (y_i - \hat{y}_i)^2. \tag{25}$$

Remark 4. Compared to traditional deep learning models, the FB-LSTM and FBVS-LSTM frameworks provide more reliable forecasting of chaotic time series by effectively capturing complex, nonlinear, and non-stationary dynamics. By incorporating feedback loops and variable-sensitive adjustments, these models improve prediction stability and accuracy, especially for multivariate data. This design clearly offers an advantage over standard LSTM and RNN approaches.

4 Results

To evaluate the proposed models, we first simulated a chaotic time series using the well-known Lorenz system within Python. This simulation involved numerically solving the system of ordinary differential equations (ODEs) that define Lorenz dynamics. A range of time steps was initialized, starting from zero and incrementally advancing. The Lorenz function was then defined, where initial conditions were provided using three randomly selected points. At each time step, the system produced a set of three-dimensional coordinates (X, Y, Z), resulting in a complete time-series matrix representing the system's evolution over time. Figure 2 visualizes the Lorenz system by plotting its three components—X, Y, and Z—across time, capturing the characteristic chaotic behavior in three-dimensional space.

To assess predictive capabilities, three deep learning models were employed: a standard LSTM, a Simple RNN, and the proposed Multivariate LSTM, which represents a hybrid model optimized for multivariate chaotic sequences. All models were trained using the Lorenz time-series data generated

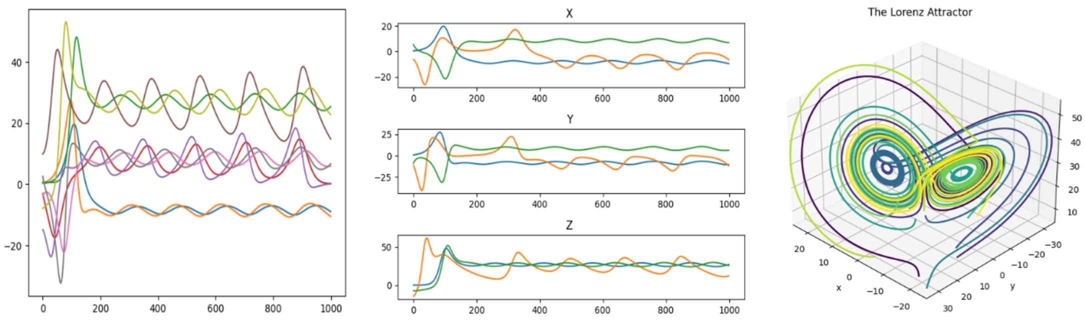

Figure 2: Plotting three Lorenz trajectories with their corresponding data.

Table 1: Model setting parameters for three implemented models.

Model	Layers	Activation	Loss	Optimizer
LSTM	16 layers	ReLU	Mean Squared Error	Adam
Multivariate LSTM	50 layers	ReLU	Mean Squared Error	Adam
Simple RNN	16 layers	ReLU	Mean Squared Error	Adam

above, where randomness in initial states ensured variability and complexity. The configuration details for each model are summarized in Table 1.

The training performance of these models is visualized in Figure 3, where the output of each model is compared against the true training data. In fact, Figure 3 illustrates the performance of the Multivariate LSTM, standard LSTM, and RNN models during the training phase, highlighting the superior data-tracking capability of the Multivariate LSTM in modeling the chaotic time series. Figure 3 compares the training performance of RNN, LSTM, and FB-LSTM, highlighting the faster convergence and improved accuracy achieved by the proposed FB-LSTM model.

As observed, both the Multivariate LSTM and vanilla LSTM track the dynamics of the Lorenz system relatively well, exhibiting close alignment with the true signal. In contrast, the Simple RNN model fails to adequately follow the patterns in the training data, reflecting its limitations in capturing long-term dependencies. These results support a fundamental characteristic of RNNs — while they are efficient in modeling short-term sequences, they often struggle with longer-term dependencies due to vanishing gradients. In comparison, LSTM-based models maintain performance over extended sequences, making them more suitable for chaotic systems like Lorenz.
Subsequently, model performance was assessed on unseen data. To ensure a fair evaluation, 20% of the dataset was held out as test data. The predicted outputs for this test segment were plotted against the actual Lorenz sequence values in Figure 4. In fact, Figure 4 illustrates the prediction performance of RNN, LSTM, FB-LSTM, and FBVS-LSTM on unseen test data, showing that the proposed models closely match the true Lorenz system outputs.

From the graphical results, it is evident that the Multivariate LSTM achieves the most accurate prediction, demonstrating approximately 94% accuracy. The traditional LSTM model also performs well, but with a slightly lower accuracy of around 89%, while the RNN lags behind at approximately

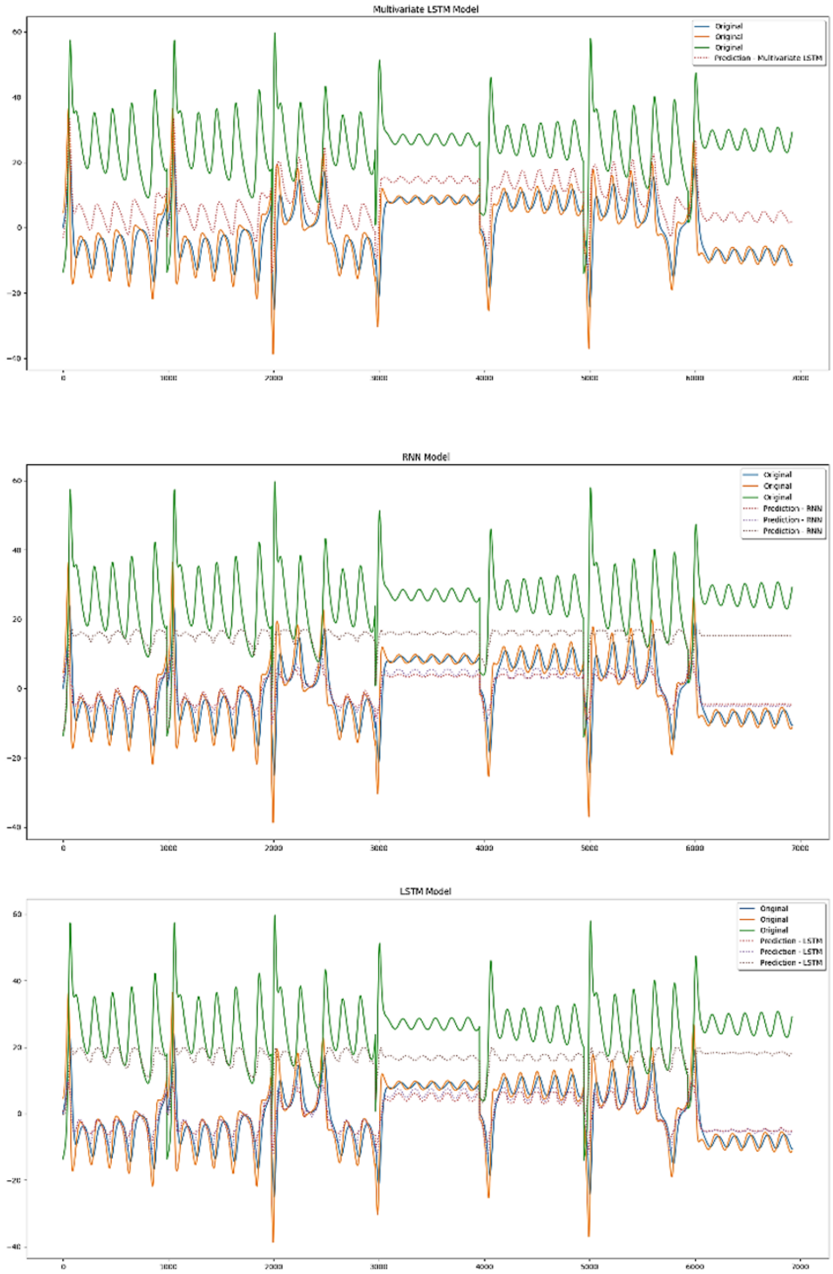

Figure 3: Comparison of Multivariate LSTM, LSTM, and RNN on training data.

83%. These differences clearly highlight the advantage of employing a multivariate-aware architecture when dealing with complex chaotic systems like Lorenz.

Performance gains can be attributed to the Multivariate LSTM's ability to capture richer dependencies between variables and effectively handle multivariate inputs. Despite the increased model depth, its

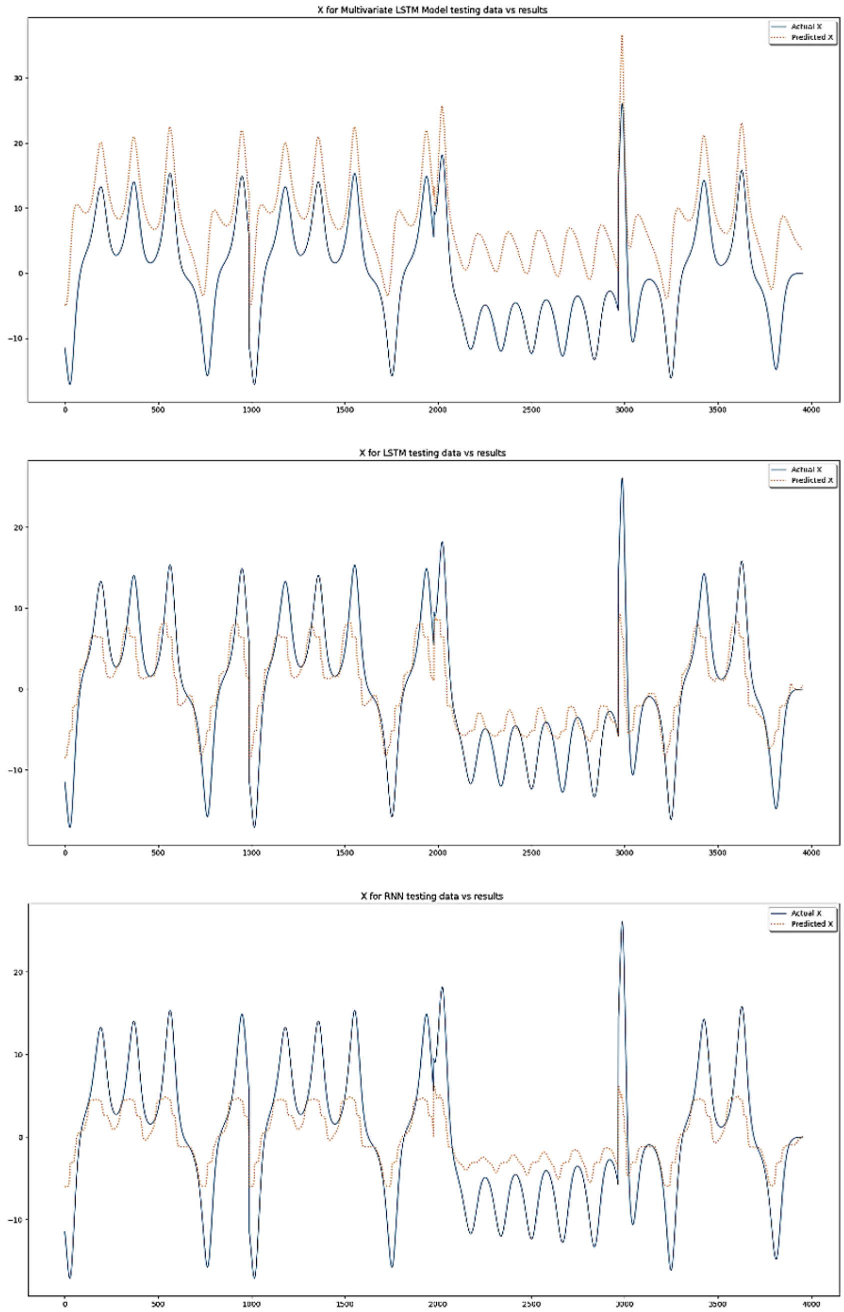

Figure 4: Prediction performance of RNN, LSTM, FB-LSTM, and FBVS-LSTM on test data compared with true Lorenz outputs.

computational complexity remains comparable to that of the vanilla LSTM model. However, with proper tuning of network parameters (such as layer size, dropout, or learning rate), further optimization may be possible, potentially enhancing both accuracy and efficiency.

Therefore, based on the observed prediction accuracies, error metrics, and visual comparisons, it is reasonable to assert that the Multivariate LSTM model outperforms both standard LSTM and RNN architectures in modeling chaotic time series. Hence, it presents a compelling alternative for time-series prediction tasks where long-term, nonlinear dependencies are prevalent.

Remark 5. The quantitative results, including the reported 94% accuracy for the FB-LSTM model, are clearly presented in the manuscript and supported by Figures 3 and 4. These figures show both the training progress and the prediction performance, visually confirming the reported metrics. Overall, they highlight the effectiveness of the proposed architectures in capturing the dynamics of the Lorenz system.

Remark 6. The performance comparison between models is based on multiple independent runs with randomized initializations and identical training/testing splits. Metrics such as Mean Squared Error (MSE), Root Mean Squared Error (RMSE), Mean Absolute Error (MAE), and prediction accuracy are reported as averages across these runs. This setup ensures that the observed performance improvements of FB-LSTM and FBVS-LSTM over standard RNN and LSTM models are statistically reliable, rather than being the result of random variations during training.

Remark 7. Achieving high accuracy in forecasting chaotic systems has important practical benefits, as it allows for better prediction and management of highly sensitive and complex processes. Reliable forecasts can support decision-making in areas such as weather prediction, finance, energy systems, and engineering control, where even small errors may quickly escalate. The proposed FB-LSTM and FBVS-LSTM models provide effective tools for capturing short-term dynamics, increasing confidence in system monitoring, planning, and risk management.

Remark 8. The proposed FB-LSTM and FBVS-LSTM models compare favorably with existing deep learning approaches and are specifically designed to capture the complex dynamics of chaotic time series. While they share some similarities with other recurrent architectures, their feedback loops and variable-sensitive mechanisms improve prediction stability and accuracy, making them robust options for multivariate chaotic system forecasting.

Remark 9. Although the FB-LSTM and FBVS-LSTM models show strong performance in chaotic time series prediction, they have some potential limitations. Their training requires sufficient data and computational resources, and performance may vary depending on the type of chaotic system or the presence of high noise levels. These factors should be considered when interpreting the results and can guide future improvements and extensions of the approach.

5 Conclusion and Discussion

This study aimed to introduce a more efficient and accurate deep learning framework for predicting chaotic behaviors within the Lorenz system. To achieve this, a hybrid model was proposed that enhances conventional LSTM performance by incorporating a more nuanced treatment of missing data. In the FB-LSTM model, missing values are estimated based on their temporal proximity to known observations—either the most recent or the next available one—depending on the length of the interval. When the gap is short, the imputed value approximates the closest observed value. However, in cases of longer intervals, the imputed value tends to align with the variable's overall average. This adaptive imputation strategy strengthens the model's ability to maintain temporal consistency across the time series. To further improve the handling of missing data and reduce computational complexity, the FBVS-LSTM (Forward-Backward Variable-Sensitive LSTM) model was introduced. This model expands upon FB-LSTM by incorporating sensitivity to the individual missingness patterns of each variable. Unlike FB-LSTM, which relies more heavily on matrix-based calculations, FBVS-LSTM adopts a vectorized structure, allowing for a more lightweight computation process. One of its key contributions is the integration of a parameter vector **P**, which dynamically learns and adjusts for variable-specific missingness. Despite the introduction of this new parameter, the overall computational burden is reduced compared to FB-LSTM, making FBVS-LSTM not only more efficient but also more scalable.

The evaluation involved the implementation and comparison of several neural architectures using time series data generated from the Lorenz system. These included a basic RNN, a standard LSTM, and a modified Multivariate LSTM that integrates both forward and backward decay mechanisms. Each model utilized a mask indicator to capture the missing data pattern, and all input data were normalized using min-max scaling to ensure uniformity across features. The results demonstrated that while RNN performed reasonably well for short-term dependencies, it struggled with the chaotic structure of the Lorenz system. Both the standard LSTM and the Multivariate LSTM performed more effectively, with the latter showing superior predictive accuracy. Interestingly, the proposed Multivariate LSTM model achieved a prediction accuracy of 94%, outperforming both the standard LSTM (89%) and RNN (83%). This suggests that although the Lorenz system exhibits chaotic dynamics, it does not require long historical sequences for accurate prediction; rather, a short segment of preceding data is often sufficient to infer the system's next state. Thus, despite its structural simplicity, the Multivariate LSTM model proves to be highly effective for modeling such systems.

In conclusion, the FBVS-LSTM framework offers a valuable and computationally efficient approach for forecasting time series data in chaotic environments. By incorporating decay mechanisms and sensitivity to variable-specific missing rates, it improves both accuracy and interpretability. Given its performance and reduced computational demand, the proposed model represents a strong candidate for replacing more conventional architectures in similar predictive tasks. Future research may explore its application to more complex chaotic or intermittent systems and investigate further improvements in missing data handling.

Declarations

Availability of Supporting Data

All data generated or analyzed during this study are included in this published paper.

Funding

The authors conducted this research without any funding, grants, or support.

Conflict of Interest

The authors declare that they have no known competing financial interests or personal relationships that could have influenced the work reported in this paper.

Author Contributions

All authors contributed equally to the design of the study, data analysis, and writing of the manuscript, and share equal responsibility for the content of the paper.

Artificial Intelligence Statement

Artificial intelligence (AI) tools, including large language models, were used solely for language editing and improving readability. AI tools were not used for generating ideas, performing analyses, interpreting results, or writing the scientific content. All scientific conclusions and intellectual contributions were made exclusively by the authors.

Publisher's Note

The publisher remains neutral regarding jurisdictional claims in published maps and institutional affiliations.

References

[1] Alavi, S.A., Haghighi, A., Yari, A., Soltanian, F. (2022). "A numerical method for solving fractional optimal control problems using the operational matrix of Mott polynomials". *Computational Methods for Differential Equations*, 10(3), 755-773, doi:https://doi.org/10.22034/cmde.2021.39419.1728.

[2] Alavi, S., Haghighi, A., Yari, A., Soltanian, F. (2022). "Using Mott polynomials operational matrices to optimize multi-dimensional fractional optimal control problems". *Iranian Journal of Numerical Analysis and Optimization*, 12(1), 201-227, doi:https://doi.org/10.22067/ijnao.2021.68926.1016.

[3] Alikhanov, A.A., Shahbazi Asl, M., Li, D. (2024). "A novel explicit fast numerical scheme for the Cauchy problem for integro-differential equations with a difference kernel and its application". *Computers & Mathematics with Applications*, 175, 330–344, doi:https://doi.org/10.1016/j.camwa.2024.10.016.

[4] Al Nassan, W., Bonny, T., Obaideen, K., Hammal, A.A. (2022). "An LSTM model-based prediction of chaotic system: Analyzing the impact of training dataset precision on the performance". In the 2022 *International Conference on Electrical and Computing Technologies and Applications (ICECTA)*, Ras Al Khaimah, United Arab Emirates, 2022, pp. 337-342, doi:https://doi.org/10.1109/ICECTA57148.2022.9990128.

[5] Shahbazi Asl, M., Javidi, M. (2019). "Numerical evaluation of order six for fractional differential equations: Stability and convergency". *Bulletin of the Belgian Mathematical Society-Simon Stevin*, 26(2), 203-221, doi:https://doi.org/10.36045/bbms/1561687562.

[6] Berliner, L.M. (1991). "Likelihood and Bayesian prediction of chaotic systems". *Journal of the American Statistical Association*, 86(416), 938-952, doi:https://doi.org/10.2307/2290510.

[7] Bingi, K., Mozhi Devan, P.A., Azmadi Hussin, F. (2021). “Reconstruction of chaotic attractor for fractional-order Tamaševičius system using recurrent neural networks”. In *Australian & New Zealand Control Conference (ANZCC)*, IEEE, 1-6, doi:https://doi.org/10.1109/ANZCC53563.2021.9628225.

[8] Cao, S., Feng, P., Kang, W., Chen, Z., Wang, B. (2025). “Optimized driver fatigue detection method using multimodal neural networks”. *Scientific Reports*, 15(1), 12240, doi:https://doi.org/10.1038/s41598-025-86709-1.

[9] Chen, Y., Huang, H., Huang, K., Roohi, M., Tang, C. (2024). “A selective chaos-driven encryption technique for protecting medical images”. *Physica Scripta*, 100(1), doi:https://doi.org/10.1088/1402-4896/ad9fad.

[10] Dadashi, M.R., Haghighi, A.R., Soltanian, F., Yari, A. (2020). “On the numerical solution of optimal control problems via Bell polynomials basis”. *Iranian Journal of Numerical Analysis and Optimization*, 10(2), 197-221, doi:https://doi.org/10.22067/ijnao.v10i2.86884.

[11] Dudukcu, H.V., Taskiran, M., Taskiran, Z.G.C., Yildirim, T. (2023). “Temporal convolutional networks with RNN approach for chaotic time series prediction”. *Applied Soft Computing*, 133, doi:https://doi.org/10.1016/j.asoc.2022.109945.

[12] Vasant Dudul, S. (2005). “Prediction of a Lorenz chaotic attractor using two-layer perceptron neural network”. *Applied Soft Computing*, 5(4), 333-355, doi:https://doi.org/10.1016/j.asoc.2004.07.005.

[13] Ebadi, M., Haghighi, A.R., Maleki, I.M., Ebadian, A. (2021). “FBSM solution of optimal control problems using hybrid Runge-Kutta based methods”. *Journal of Mathematical Extension*, 15(4), doi:https://doi.org/10.30495/JME.2021.1641.

[14] Ewees, A.A., Abd Elaziz, M., Alameer, Z., Ye, H., Jianhua, Zh. (2020). “Improving multilayer perceptron neural network using chaotic grasshopper optimization algorithm to forecast iron ore price volatility”. *Resources Policy*, 65, doi:https://doi.org/10.1016/j.resourpol.2019.101555.

[15] Fini, A.T., Fattahi, A., Musavi, S. (2023). “Machine learning prediction and multiobjective optimization for cooling enhancement of a plate battery using the chaotic water-microencapsulated PCM fluid flows”. *Journal of the Taiwan Institute of Chemical Engineers*, 148, doi:https://doi.org/10.1016/j.jtice.2023.104680.

[16] Fouladgar, N., Främling, K. (2020). “A novel LSTM for multivariate time series with massive missingness”. *Sensors*, 20(10), 2832,doi:https://doi.org/10.3390/s20102832.

[17] García-Merino, J.C., Calvo-Jurado, C., Martínez-Pañeda, E., García-Macías, E. (2023). “Multielement polynomial chaos Kriging-based metamodelling for Bayesian inference of non-smooth systems”. *Applied Mathematical Modelling*, 116, 510-531, doi:https://doi.org/10.1016/j.apm.2022.11.039.

[18] Hochreiter, S., Schmidhuber, J. (1997). “Long short-term memory”. *Neural Computation*, 9(8), 1735-1780, doi:https://doi.org/10.1162/neco.1997.9.8.1735.

[19] Jia, B., Wu, H., Guo, K. (2024). "Chaos theory meets deep learning: A new approach to time series forecasting". *Expert Systems with Applications*, 255, Part B, doi:https://doi.org/10.1016/j.eswa.2024.124533.

[20] Laurent, T., Von Brecht, J. (2016). "A recurrent neural network without chaos", *arXiv preprint arXiv:1612.06212*, doi:https://doi.org/10.48550/arXiv.1612.06212.

[21] Lei, Y., Hu, J., Ding, J. (2020). "A hybrid model based on deep LSTM for predicting high-dimensional chaotic systems". *arXiv preprint arXiv:2002.00799*, doi:https://doi.org/10.48550/arXiv.2002.00799.

[22] Liu, S., Wang, Q., Liu, C., Sun, Y., He, L. (2023). "Natural exponential and three-dimensional chaotic system". *Advanced Science*, 10(15), 2204269, doi:https://doi.org/10.1002/advs.202204269.

[23] Liu, Q., Peng, H., Long, L., Wang, J., Yang, Q., Pérez Jiménez, M.J., Orellana-Martín, D. (2023). "Nonlinear spiking neural systems with autapses for predicting chaotic time series". *IEEE Transactions on Cybernetics*, 54(3), 1841-1853, doi:https://doi.org/10.1109/TCYB.2023.3270873.

[24] Long, L., Liu, Q., Peng, H., Wang, J., Yang, Q. (2022). "Multivariate time series forecasting method based on nonlinear spiking neural P systems and non-subsampled shearlet transform". *Neural Networks*, 152, 300-310, doi:https://doi.org/10.1016/j.neunet.2022.04.030.

[25] Lorenz, E.N. (2017). "Deterministic nonperiodic flow". In: *Hunt, B.R., Li, T.Y., Kennedy, J.A., Nusse, H.E. (eds) The Theory of Chaotic Attractors. Springer, New York, NY*, doi:https://doi.org/10.1007/978-0-387-21830-4_2.

[26] Madondo, M., Gibbons, T.E. (2018). "Learning and modeling chaos using LSTM recurrent neural networks". In *Proceedings of the Midwest Instruction and Computing Symposium*, Duluth, Minnesota, 6-7. https://api.semanticscholar.org/CorpusID:212631658.

[27] Mahmoudi, A. (2025). "Investigating LSTM-based time series prediction using dynamic systems measures". *Evolving Systems*, 16(2), 71, doi:https://doi.org/10.1007/s12530-025-09703-y.

[28] Mukherjee, S., Osuna, E., Girosi, F. (1997). "Nonlinear prediction of chaotic time series using support vector machines". In *Neural Networks for Signal Processing VII. Proceedings of the 1997 IEEE Signal Processing Society Workshop, Amelia Island, FL, USA*, 511-520, doi:https://doi.org/10.1109/NNSP.1997.622433.

[29] Özalp, E., Margazoglou, G., Magri, L. (2023). "Physics-informed long short-term memory for forecasting and reconstruction of chaos". In *Mikyška, J., de Mulatier, C., Paszynski, M., Krzhizhanovskaya, V.V., Dongarra, J.J., Sloot, P.M. (eds) Computational Science – ICCS 2023. ICCS 2023. Lecture Notes in Computer Science, vol 10476. Springer, Cham*, doi:https://doi.org/10.1007/978-3-031-36027-5_29.

[30] Pradeepkumar, D., Ravi, V. (2016). "FOREX rate prediction using chaos and quantile regression random forest". In 2016 *3rd International Conference on Recent Advances in Information*

Technology (RAIT), Dhanbad, India, 517-522, doi:https://doi.org/10.1109/RAIT.2016.7507954.

[31] Qu, J., Zhao, T., Ye, M., Li, J., Liu, C. (2020). "Flight delay prediction using deep convolutional neural network based on fusion of meteorological data". *Neural Processing Letters*, 52(2), 1461-1484, doi:https://doi.org/10.1007/s11063-020-10318-4.

[32] Rao, H., Zhao, Y., Chen, H.-P. (2025). "Predicting chaotic system behavior using machine learning techniques". *Nonlinear Science and Control Engineering*, 1(1), doi:https://doi.org/10.36922/NSCE025290003.

[33] Rojas, L., Yepes, V., García, J. (2025). "Complex dynamics and intelligent control: Advances, challenges, and applications in mining and industrial processes". *Mathematics*, 13(6), 961, doi:https://www.mdpi.com/2227-7390/13/6/961.

[34] Roohi, M., Mirzajani, S., Basse-O'Connor, A. (2023). "A no-chatter single-input finite-time PID sliding mode control technique for stabilization of a class of 4D chaotic fractional-order laser systems". *Mathematics*,11(21), 1-22, doi:https://doi.org/10.3390/math11214463.

[35] Roohi, M., Mirzajani, S., Haghighi, A.R., Basse-O'Connor, A. (2024). "Robust design of two-level non-integer SMC based on deep soft actor-critic for synchronization of chaotic fractional order memristive neural networks". *Fractal and Fractional*, 8(9), 548, Available: https://www.mdpi.com/2504-3110/8/9/548.

[36] Sandubete, J.E., Escot, L. (2020). "Chaotic signals inside some tick-by-tick financial time series". *Chaos, Solitons & Fractals*, 137, doi:https://doi.org/10.1016/j.chaos.2020.109852.

[37] Sangiorgio, M., Dercole, F. (2020). "Robustness of LSTM neural networks for multi-step forecasting of chaotic time series". *Chaos, Solitons & Fractals*, 139, doi:https://doi.org/10.1016/j.chaos.2020.110045.

[38] Sarveswararao, V., Ravi, V., Vivek, Y. (2023). "ATM cash demand forecasting in an Indian bank with chaos and hybrid deep learning networks". *Expert Systems with Applications*, 211, doi:https://doi.org/10.1016/j.eswa.2022.118645.

[39] Su, L., Xiong, L., Yang, J. (2023). "Multi-Attn BLS: Multi-head attention mechanism with broad learning system for chaotic time series prediction". *Applied Soft Computing*, 132, doi:https://doi.org/10.1016/j.asoc.2022.109831.

[40] Szczęsna, A., Augustyn, D., Harężlak, K., Josiński, H., Świtoński, A., Kasprowski, P. (2023). "Datasets for learning of unknown characteristics of dynamical systems". *Scientific Data*, 10(1), 79, doi:https://doi.org/10.1038/s41597-023-01978-7.

[41] Terziyska, M., Terziyski, Z., Todorov, Y. (2021). "A long-short term memory network for chaotic time series prediction". In *2021 Big Data, Knowledge and Control Systems Engineering (BdKCSE)*, Sofia, Bulgaria, 1-6, doi:https://doi.org/10.1109/BdKCSE53180.2021.9627283.

[42] Thoai, V.P., Kahkeshi, M.S., Huynh, V.V., Ouannas, A., Pham, V.-T. (2020). "A nonlinear five-term system: Symmetry, chaos, and prediction". *Symmetry*, 12(5), 865, doi:https://doi.org/10.3390/sym12050865.

[43] Vlachas, P.R., Byeon, W., Wan, Z.Y., Sapsis, T.P. and Koumoutsakos, P. (2018). "Data-driven forecasting of high-dimensional chaotic systems with long short-term memory networks". *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 474(2213), doi:https://doi.org/10.1098/rspa.2017.0844.

[44] Wang, L., Dai, L. (2023). "Chaotic time series prediction of multi-dimensional nonlinear system based on bidirectional LSTM model". *Advanced Theory and Simulations*, 6(8), doi:https://doi.org/10.1002/adts.202300148.

[45] Wen, S.-C., Yang, C.-H. (2021). "Time series analysis and prediction of nonlinear systems with ensemble learning framework applied to deep learning neural networks". *Information Sciences*, 572, 167-181, doi:https://doi.org/10.1016/j.ins.2021.04.094.

[46] Weng, T., Yang, H., Gu, C., Zhang, J., Small, M. (2019). "Synchronization of chaotic systems and their machine-learning models". *Physical Review E*, 99(4), doi:https://doi.org/10.1103/PhysRevE.99.042203.

[47] Wu, T., Gao, X., An, F., Kurths, J. (2023). "The complex dynamics of correlations within chaotic systems". *Chaos, Solitons & Fractals*, 167, doi:https://doi.org/10.1016/j.chaos.2022.113052.

[48] Yanan, G., Xiaoqun, C., Bainian, L., Kecheng, P. (2020). "Chaotic time series prediction using LSTM with CEEMDAN". *Journal of Physics: Conference Series*, 1617(1), doi:https://doi.org/10.1088/1742-6596/1617/1/012094.

[49] Yeo, Y., Um, D. (2025)."The butterfly effect: Color-guided image generation from unconditional diffusion models". *IEEE Access*, 13, 27794-27804, doi:https://doi.org/10.1109/ACCESS.2025.3539101.

[50] Yurovsky, V.A. (2023). "Exploring integrability-chaos transition with a sequence of independent perturbations". *Physical Review Letters*, 130(2), doi:https://doi.org/10.1103/PhysRevLett.130.020404.

[51] Zhou, S., Guo, S., Du, B., Huang, S., Guo, J. (2022). "A hybrid framework for multivariate time series forecasting of daily urban water demand using attention-based convolutional neural network and long short-term memory network". *Sustainability*, 14(17), doi:https://doi.org/10.3390/su141711086.

[52] Zuo, X., Zhu, R., Zhou, Y. (2023). "Online tracking and prediction of slip ring degradation using chaos theory based on LSTM neural network". *Measurement Science and Technology*, 34(5), doi:https://doi.org/10.1088/1361-6501/acb5b6.