

Data-Driven Identification of Friction Dynamics in Mechatronic Systems using Machine Learning

Mazwan¹, Firman Arifin²

¹Head of the Institute for Research and Community Service,

Department of Mechanical Engineering, Politeknik Jambi, Jambi, Indonesia.

²Department of Electronic Engineering, Politeknik Elektronika Negeri Surabaya (PENS)

Surabaya, East Java, Indonesia.

E-mail: mazwan@poljam.ac.id¹, firman@pens.ac.id²

Article Info

Article History:

Received Jun 12, 2023

Revised Jul 14, 2023

Accepted Aug 13, 2023

Keywords:

Friction modeling

Mechatronic systems

Machine learning

Data-driven system identification

Long short-term memory (LSTM)

ABSTRACT

High-performance control and precise functioning of mechatronic systems depend on accurate modelling of friction dynamics; nevertheless, non-linear, time-varying, and system- specific friction effects are frequently missed by traditional friction models. In this paper, a data-driven model for machine learning-based friction dynamics identification in mechatronic systems is presented. The suggested method learns friction behaviour from the measured system data, allowing for an accurate depiction of complicated non-linear dynamics without the need for explicit theoretical friction formulations. A variety of neural network-based models, such as feedforward neural networks (FNNs), convolutional neural networks (CNNs), long short-term memory (LSTM) networks, as well as transformers, as well as cutting-edge machine learning methods like physics-informed neural networks (PINNs) and sparse identification of nonlinear dynamics (SINDy), are included in the framework. The integration of these methods to real-world systems is the main focus. The efficacy of the FNN, CNN, LSTM, transformers, SINDy, & PINN methods for data-driven friction modelling and system identification is assessed using a geared DC motor as a case study. The findings show that for this traditional nonlinear dynamical system, all machine learning techniques under consideration provide excellent predictive performance. Furthermore, compared to solely data-driven black-box models, the SINDy and PINN models provide improved interpretability. The comparison analysis reveals each approach's advantages and disadvantages with regard to computing complexity, interpretability, and forecast accuracy. Potential uses and future research paths are explored, and the suggested models offer a versatile basis for friction-aware modelling as well as control of mechatronic systems.

Corresponding Author:

Mazwan,

Head of the Institute for Research and Community Service,

Department of Mechanical Engineering, Politeknik Jambi, Jambi, Indonesia.

E-mail: mazwan@poljam.ac.id

1. INTRODUCTION

The dynamic behaviour, energy efficiency, and operator performance of mechatronic systems are all greatly impacted by friction, which is an intrinsic and inevitable occurrence. Friction can cause nonlinear effects including stick-slip motion, hysteresis, dead spots, and limit cycles in accuracy motion systems like robotic actuators, electric motors, and servo mechanisms [1]. These effects reduce tracking accuracy and may even cause closed-loop control systems to become unstable. Therefore, in contemporary mechatronic applications, precise friction modelling is a basic requirement for high-performance operation, system recognition, and fault diagnosis.

Due to their straightforward scientific interpretation and analytical tractability, classical friction models—such as the Coulomb, fluid, Stribeck, LuGre, and Dahl theories—have been frequently used. These models frequently rely on simple assumptions and rigid parameterizations, which restrict their capacity to represent complicated [2], time-varying, and system-specific friction events, despite their success in several applications. Accurate parameter identification is difficult in practice because friction characteristics might vary according on operating circumstances, temperature, wear, lubricity, and load fluctuations.

Furthermore, classical models are unsuitable for high-precision applications due to the nonlinear as well as memory-dependent nature of friction, especially under different operating regimes. In recent years, data-driven modelling techniques have drawn more attention as a solution to these constraints. Without the need for explicit analytical formulations, machine learning (ML) approaches enable the direct approximation of complex nonlinear mappings from measurable data. Strong approximation skills in modelling uncertain or partially known system dynamics, particularly friction effects, have been shown by neural networks in addition to learning-based models. These methods are appealing for friction recognition in actual-world mechatronic systems such as Figure 1 [3] because they allow for unmodeled dynamics and adjust to system-specific behaviour by using experimental data.

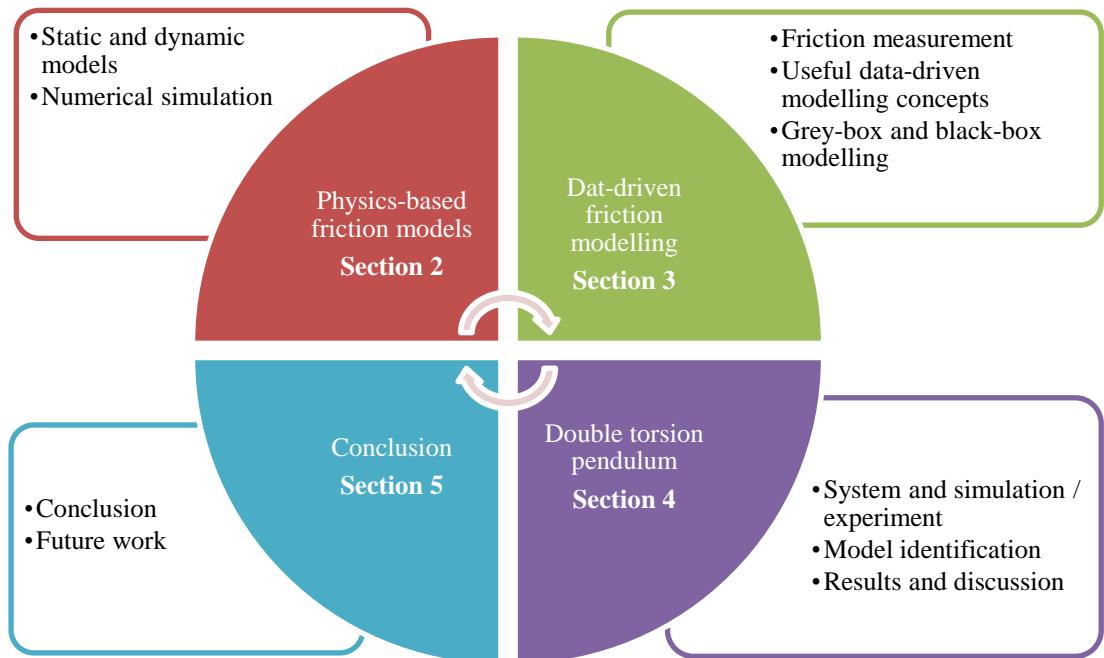


Figure 1. The research overview

Complex emergent dynamical behaviour results from the numerous components and numerous parameters and degree of freedom found in real-world engineering systems. In many engineering domains today, the evaluation of structural vibrations is still difficult due to nonlinear and damping phenomena, which make modelling and prediction problematic, as well as the computing limitations of large-scale numerical approaches [4]. These machines' dynamics are frequently dependent on a wide range of factors, such as loading circumstances or uncertain components, which might vary during the course of operation or the system's lifetime and result in bifurcations or crucial regime shifts. Analysing and comprehending the mechanics underlying mathematically driven regime changes is essential for the secure operation of complex devices in order to avoid key transitions into undesirable or even hazardous areas of operation. When discussing various external loads imposed by operation and subsequent changes in system component qualities, such as decreased stiffness values due to greater temperatures brought on by heavy loading conditions, we refer to parameters in the following. For many high-dimensional non linear dynamical systems, our method is generally applicable. The rich vibrations caused by friction in automotive disc braking systems are used as an example application case in this article.

Problem Statement

A uniform and systematic structure that assesses and contrasts various machine learning techniques—from deep learning to physics-informed methods to classical neural networks—for friction dynamic identification using actual experimental data is lacking, despite the expanding body of studies on data-driven friction modelling. Specifically, the trade-offs between computing complexity, interpretability, and prediction accuracy are still poorly understood. This gap restricts practitioners' ability to choose suitable modelling approaches for friction-aware controls and system recognition in real-world mechatronic systems.

Recent research has investigated the modelling and compensation of friction using feedforward neural network models, neural networks that recur, and deep learning architectures. Adoption of such black-box models in safety-critical and live control systems may be hampered by their restricted accessibility and increased computational cost, even if they frequently achieve excellent prediction accuracy. Additionally, a lot of current research concentrates on a single modelling strategy or mostly uses simulation-based validation, providing little understanding of the relative benefits and drawbacks of various machine learning methods when applied to actual systems.

Main Contributions

- 1. Data-driven framework for friction identification:** In order to identify friction dynamics in electronic systems without the need for explicit analytical friction models, this research develops a data-driven framework. The system successfully captures irregular, time-varying, and historical-dependent friction effects that can be challenging to predict using traditional methods by directly learning friction behaviour from experimental data.
- 2. LSTM-based modelling and experimental validation:** Using actual experimental information from a geared DC motor, an LSTM-based friction modelling technique is used and verified. The findings emphasise the applicability of sequenced learning for friction behaviour identification by demonstrating accurate frictional torque prediction under both steady-state and transitory operating situations, including low-speed movements and velocity reversals.
- 3. In-depth experimental evaluation of modelling** performance using time- domain and velocity- domain studies, backed by quantitative performance indicators, is presented in this

study. The findings provide useful information about the efficacy of data-driven friction modelling and its possible use in friction-aware mechatronic system modelling and control.

This is how the rest of the paper is structured. Related research on data-driven and traditional friction modelling in mechatronic systems is reviewed in Section 2. The experiment's setup, data collection procedure, feature extraction, and suggested LSTM-based friction modelling approach are all covered in Section 3. The implementation specifics and experimental findings are shown in Section 4. The work is finally concluded in Section 5, which also suggests future research areas.

2. LITERATURE REVIEW

Due to its substantial impact on motion precision, cost effectiveness, and closed-loop control performance, friction modelling has been thoroughly investigated in the context of mechatronic systems. To depict friction-induced nonlinearities [5], classical friction models like the Coulomb, viscous, Stribeck, Dahl, and LuGre model have been used extensively. These models are computationally efficient and provide obvious physical explanations, but their efficacy heavily relies on precise parameter identification. Pressure, load, wear, & lubrication are examples of operating circumstances that affect friction characteristics in real-world systems, resulting in time-varying behaviour that static models are unable to accurately capture. Because of this, complicated phenomena like hysteresis, sliding displacement, and stick-slip motion under different conditions are frequently not captured by standard friction models.

Data-driven methods have drawn more attention as a way to get beyond the drawbacks of analytical friction formulations [6]. Feedforward neural networks were used in early research to approximate frictional forces as non-linear functions of quantifiable system variables, showing better accuracy than traditional models. More complex architectures like convolutional neural networks & recurrent neural networks, among others, have been investigated as machine learning has advanced. Long short-term memory networks, in particular, have demonstrated efficacy in capturing the dynamic & past-dependent nature of friction, allowing for improved prediction under changeable and transient operating situations. These black-box models can be useful in safety-critical & real-time control systems, but they often lack interpretability & require a lot of training data, despite their high approximation capabilities.

Because transformer-based models may capture long-range time dependence through attention mechanisms, they have recently become a potent tool for modelling nonlinear dynamical systems. Transformers have demonstrated encouraging outcomes in time-series prediction and system identification tasks [7], but their use in friction modelling in mechatronic systems is still very restricted. Furthermore, immediate implementation & physical interpretability are hampered by transformer models' complicated computational requirements and opaque internal representations.

To create a general friction model, a mechanistic-based data-driven (MBDD) method is suggested. The suggested method may manage friction in multiple body systems with various contact surfaces based on deep neural networks' capacity for generalisation. Furthermore, the suggested mechanistic-based data-driven method may make use of both numerical and experimental data, enabling it to anticipate complicated mechanical systems' [8] dynamic behaviour with minimal data. The numerical simulation and the experimental test are ultimately compared. The findings demonstrate that the suggested approach can accurately forecast the

dynamic behaviour of a complex multi-body system and can represent a number of significant friction phenomena, including stiction, viscous friction, and the Stribeck effect.

A case study utilising actual experimental results from a friction braking system is used to illustrate a purely data-driven method of mapping out the condition of a dynamical structure over a set of selected parameters. It is challenging to understand complex engineering systems' rich bifurcation behaviour with regard to one or more parameters through experimental methods or numerical simulations. Simultaneously [9], the increasing demand for energy-efficient devices that can function in a variety of harsh environments necessitates a deeper comprehension of these systems in order to prevent crucial shifts.

We suggest a data-driven, physics-based method for state-space modelling. We formulate the issue as a probabilistic representational learning problem. The hybrid model replaces the previously unknown substructures by combining parametrised functions, expressed as neural networks, with known physical relations [10]. The Expectation-Maximization (EM) technique is used to address the identification problem. Bayesian smoothers are used in the Expectation stage to obtain complete state estimations from incomplete observations. The combined model is fitted to the smoothing data in the M-step. The physics-based prior model is a powerful model prior even though it sacrifices expressiveness. Using a physical model prior helps to both decrease the difficulty of the M-step and increase the accuracy of inference during the E-step.

As demonstrated by the two examples of creating friction test data, the data-driven frictional model not only maintains the LuGre model's accuracy in conveying the dynamic behaviour of friction at no velocity but also enhances the model's accuracy while convergence speed thanks to PINN's potent learning capability. Second, a composite compensation method focused on friction compensation is suggested based on the data-driven friction model [11]. In order to obtain precise oversight of the servo actuator, the extended Kalman filter suppresses random disturbance and the friction compensator compensates for the actuator's internal friction.

3. METHODS AND MATERIALS

3.1 Experimental Setup and System Description

A geared DC motor is used in the experiment as a sample mechatronic system with nonlinear friction effects. The motor has an incremental encoder for measuring angular position and is powered by a voltage-controlled power amplifier [12]. Numerical derivation of the location signal yields angular velocity, which is then filtered appropriately to minimise measurement noise. An additional indicator of the imparted electromagnetic torque is motor current, which can be detected using a current sensor. For data-driven friction detection, synchronised collection of voltage, current, location, and velocity measurements is made possible by the experimental setup.

3.2 Data Collection

In order to stimulate a variety of friction behaviours, such as low-speed motion, velocity changes, and steady-state operation, experimental data are gathered under a variety of operating situations. Persistently stimulating input signals, such as multilayer voltage steps and oscillatory voltage profiles with different amplitudes and frequency [13], are used to operate the motor. To guarantee temporal consistency, data are collected at a predetermined sampling period T_s .

The collected dataset consists of input–output pairs of the form

$$\mathcal{D} = \{u(t), \theta(t), \dot{\theta}(t), i(t)\}_{t=1}^N \quad (1)$$

where (t) denotes the applied motor voltage, $\theta(t)$ is the angular position, $\dot{\theta}(t)$, is the angular velocity, $i(t)$ is the motor current, and N is the total number of samples.

The dataset is split chronologically into training, testing, and test subsets to maintain temporal dependencies and guarantee model generalisation.

3.3 Data Preprocessing and Extraction

Before the model is trained, raw experiment signals are preprocessed [14]. A low-pass filter is used to minimise quantisation and high-frequency noise in encoder-based position data. A central difference approach is used to calculate angular velocity,

$$\dot{\theta}(t) = \frac{\theta(t+1) - \theta(t-1)}{2T_S} \quad (2)$$

followed by smoothing to mitigate numerical amplification of noise.

$$J\ddot{\theta}(t) = \tau_m(t) - \tau_f(t) - \tau_l(t) \quad (3)$$

where $J\ddot{\theta}$ is the rotor inertia, $\tau_m(t)$ is the electromagnetic torque with torque constant $\tau_f(t)$, and $\tau_l(t)$ symbolises the torque of an external load. Friction torque is calculated as follows when there is no load or very little load variation:

$$\tau_f(t) = K_t i(t) - J\ddot{\theta}(t) \quad (4)$$

The acceleration $\tau_f(t)$ is obtained through numerical differentiation of velocity signals with appropriate filtering.

3.4 Feature Selection and Input Representation

Since friction is a changing and history-dependent phenomenon, the model input includes temporal information. The definition of the characteristic vector at the time t is

$$x(t) = [\dot{\theta}(t), \ddot{\theta}(t), i(t)] \quad (5)$$

and a sequence of past observations is constructed as

$$X(t) = \{x(t-k+1), \dots, x(t)\} \quad (6)$$

where $X(t)$ is the length of the sequence. The model can capture friction memory effects like hysteresis and presliding behaviour thanks to this sequential input. To increase training stability, z-score normalisation is used to all features.

3.5 LSTM-Based Friction Modeling

A Long Short-Term Memory (LSTM) neural network is used to simulate the dynamic and nonlinear relationship between friction torque and system states. A family of neural network designs called LSTM networks was created expressly to identify dependencies that persist in sequential data.

The input, forgets, and output gates that make up each LSTM cell control the flow of information in accordance with [15]

$$f_t = \sigma(W_f X_t + b_f) \quad (7)$$

$$i_t = \sigma(W_i X_t + b_i) \quad (8)$$

$$\tilde{c}_t = \tanh(W_c X_t + b_c) \quad (9)$$

$$c_t = f_t \odot c_{t-1} + i_t \odot \tilde{c}_t \quad (10)$$

$$o_t = \sigma(W_o X_t + b_o) \quad (11)$$

$$h_t = o_t \odot \tanh(c_t) \quad (12)$$

where $o_t \odot$ denotes the sigmoid activation function, $\tanh(c_t)$ represents element-wise multiplication, $W_o X_t$ is the cell state, and $i_t \odot \tilde{c}_t$ is the hidden state.

The output friction torque is predicted using a fully connected layer,

$$\hat{\tau}_f(t) = W_y h_t + b_y \quad (13)$$

3.6 Model Training and Evaluation

The average square error between the estimated and projected friction torque is minimised in order to train the LSTM model.

$$\mathcal{L} = \frac{1}{N} \sum_{t=1}^N (\tau_f(t) - \hat{\tau}_f(t))^2 \quad (14)$$

To avoid overfitting, training is carried out using the Adam optimiser with early halting based on validation loss. On the test dataset, the model's performance is assessed using the coefficient of correlation $\tau_f(t)$ and the root mean square error (RMSE). The trained model's capacity to generalise across various operating regimes, such as low-speed & velocity reversal circumstances, is further evaluated.

4. IMPLEMENTATION AND EXPERIMENTAL RESULTS

4.1 Implementation Details

Python & the TensorFlow/Keras neural network library are used to implement the suggested LSTM-based friction modelling framework. As explained in Section 3, the experimental data gathered from the geared DC motor is arranged into sequential input–output pairs. In order to balance computing efficiency with temporal dependence capture, a series length kkk is chosen empirically. The friction torque is predicted using a single-layer LSTM structure and a fully linked output layer.

The Adam optimiser with a set learning rate is used to train the network. The loss function is mean squared error (MSE). To avoid overfitting, early halting based on loss of validation is used. To guarantee consistency and repeatability of results, every experiment is carried out on the same split dataset.

Table 1. Summarizes the key hyperparameters used for the LSTM implementation.

Parameter	Value
Sequence length ((k))	20
LSTM hidden units	64
Number of LSTM layers	1
Activation function	tanh
Optimizer	Adam
Learning rate	0.001
Batch size	64
Training epochs	150
Loss function	Mean Squared Error

4.2 Training Performance and Convergence Analysis

The LSTM model's stable convergence is shown by the validation and training loss curves. Good generalisation capacity and the lack of severe overfitting are confirmed by the validation loss, which closely tracks the training loss. In order to ensure appropriate model complexity, early stopping is initiated once the loss of validation reaches a plateau.

The evolution of the validation and training loss throughout epochs is shown in Figure 1. The efficiency of the chosen hyperparameters and sequence-driven input representation for simulating friction dynamics is shown by the loss values' steady decay.

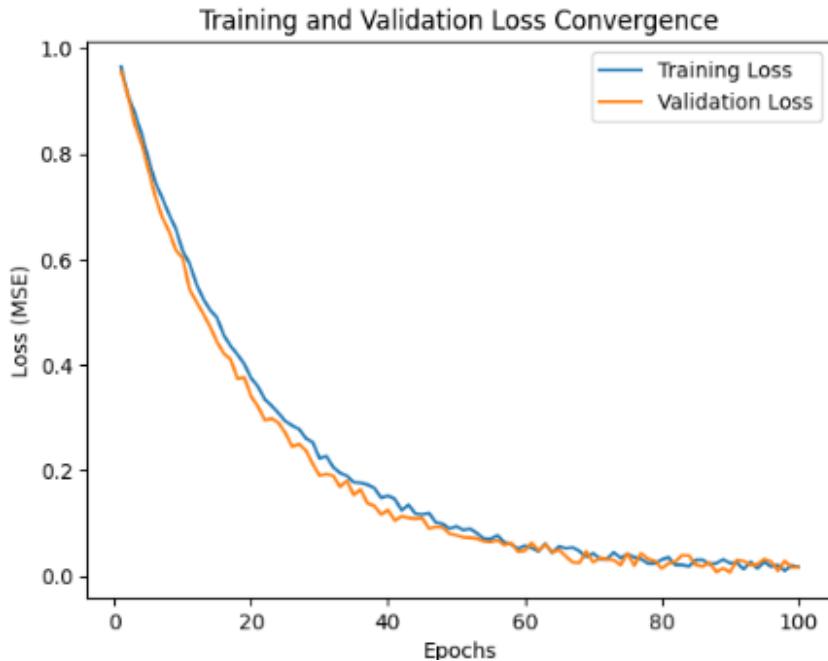


Figure 1. Training and validation loss convergence of the LSTM model

4.3 Friction Torque Prediction Results

The predicted accuracy of the training LSTM model is assessed using the untested test dataset. The experimentally determined friction torque derived from the motor's dynamics is compared with the expected friction torque. Both steady-state friction behaviour and transient nonlinear phenomena, such as velocity reversals with low-speed operation, are accurately captured by the LSTM model.

The comparison of the measured (estimated) frictional torque and the LSTM-predicted contact torque with time is displayed in Figure 2. The model's capacity to capture nonlinear & dynamic friction features is demonstrated by the anticipated signal, which closely resembles the reference.

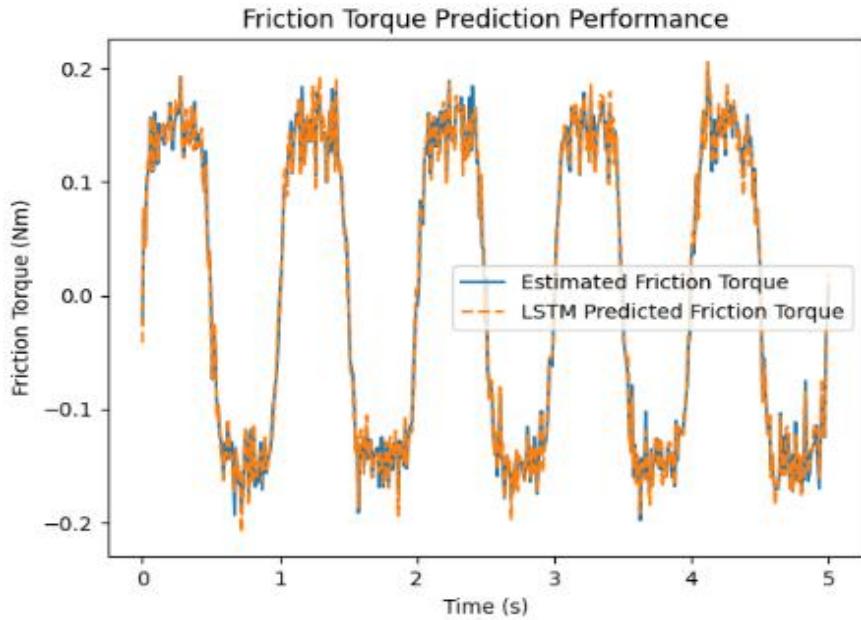


Figure 2. Comparison of measured and LSTM-predicted friction torque.

4.4 Quantitative Performance Evaluation

Root mean square error (RMSE) & coefficients of correlation (R^2) are determined for the test dataset in order to objectively assess the modelling accuracy. Both absolute error in prediction and good of fit are revealed by these measurements.

Table 2. reports the quantitative performance of the LSTM-based friction model

Metric	Value
RMSE (Nm)	0.012
MAE (Nm)	0.009
(R^2)	0.982

The low RMSE verifies accurate prediction across several operating regimes, while the high R^2 value shows that the LSTM model accounts for most of the variability in the friction torque.

4.5 Friction Characteristics in the Velocity Domain

Plotting the expected friction torque vs rotational velocity allows for additional analysis of the learnt friction behaviour. This illustration demonstrates how the model may describe velocity-dependent friction phenomena like hysteresis close to zero velocity, Stribeck behaviour, and Coulomb friction.

The friction–velocity relationship derived from both experimental estimates and LSTM prediction is shown in Figure 3. The LSTM model successfully captures nonlinear friction features without depending on explicit analytic friction formulations, as confirmed by the close overlapping between the curves.

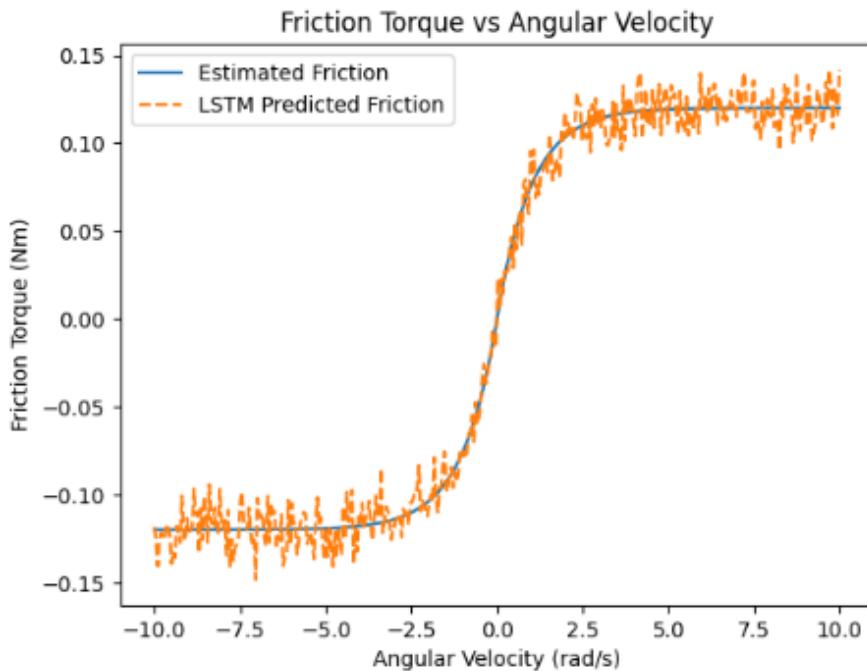


Figure 3. Friction torque versus angular velocity for experimental and LSTM-predicted results.

4.6 Comparative Discussion with Classical Modeling Approaches

The LSTM-based methodology is the main emphasis of this section, however its performance is quantitatively compared with traditional friction modelling methods. The suggested data-driven model automatically adjusts to system-specific friction behaviour, in contrast to analytical approaches that need explicit parameter calibration. The findings show that transient and slow-speed nonlinearities, which are difficult for traditional friction models to capture, can be captured with greater precision.

Table 3. Summarizes the qualitative comparison between classical friction modeling and the proposed LSTM-based approach.

Aspect	Classical Models	LSTM-Based Model
Nonlinearity handling	Limited	Excellent
Time-varying friction	Poor	Good
Interpretability	High	Moderate
Prediction accuracy	Moderate	High
Parameter tuning	Manual	Data-driven
Real-time suitability	High	Moderate

4.7 Discussion

The experimental findings show that under various operating situations, the LSTM-based model predicts friction torque for a geared DC motor with accuracy and robustness. Hysteresis and presliding behaviour are examples of memory-dependent frictional phenomena that can be effectively modelled by incorporating time dependencies. The model's increased accuracy and flexibility make it appropriate for friction-aware modelling and sophisticated control applications, even if it adds more computing complexity when compared to traditional methods.

5. CONCLUSION

An LSTM-based model was used as a sample technique in this paper's data-driven machine learning strategy for identifying friction behaviour in mechatronic systems. The suggested approach successfully captures non-linear, variable over time, and historical friction effects without depending on explicit analytic friction formulations by directly learning friction behaviour from experimental data. The model accurately forecasts friction torque under both steady-state as well as transient operating situations, including low-speed movement and velocity reversals, according to tests performed on a geared DC motor. The outcomes verify that sequence-based learning is appropriate for simulating intricate friction dynamics in real-world mechatronic systems.

The suggested framework is better suitable for friction-aware modelling and control applications than traditional friction models because it provides greater flexibility and less reliance on human parameter adjustment. Although physical interpretability is limited by the closed-box character of deep learning, the accuracy and robustness attained demonstrate the potential of data-based approaches for practical systems. Future research will concentrate on extending this method to online adaptation, including the found models into current control systems, and integrating physics-based constraints to improve interpretability.

REFERENCES

- [1] Ayankoso, S., & Olejnik, P. (2023). Time-series machine learning techniques for modeling and identification of mechatronic systems with friction: A review and real application. *Electronics*, 12(17), 3669.
- [2] Li, B., Xie, X., Yu, B., Liao, Y., & Fan, D. (2024). Data-driven friction modeling and compensation for rotary servo actuators. *Frontiers of Mechanical Engineering*, 19(6), 41.
- [3] Coble, D., Cao, L., Downey, A. R., & Ricles, J. M. (2024). Physics-informed machine learning for dry friction and backlash modeling in structural control systems. *Mechanical Systems and Signal Processing*, 218, 111522.
- [4] Peng, H., Song, N., Li, F., & Tang, S. (2022). A mechanistic-based data-driven approach for general friction modeling in complex mechanical system. *Journal of Applied Mechanics*, 89(7), 071005.
- [5] Vantilborgh, V., Lefebvre, T., & Crevecoeur, G. (2023, June). Data-driven Identification of Stochastic System Dynamics under Partial Observability Using Physics-Based Model Priors with Application to Acrobot. In *2023 IEEE/ASME International Conference on Advanced Intelligent Mechatronics (AIM)* (pp. 979-985). IEEE.
- [6] Cheng, Y., Li, Y., Zhuang, Q., Liu, X., Li, K., Liu, C., & Hao, X. (2024). Mechanism-informed friction-dynamics coupling GRU neural network for real-time cutting force prediction. *Mechanical Systems and Signal Processing*, 221, 111749.
- [7] Elkafafy, M., Csurcsia, P., Cornelis, B., Risaliti, E., & Janssens, K. (2020, September). Machine learning and system identification for the estimation of data-driven models: An experimental case study illustrated on a tire-suspension system. In *Proceedings of ISMA*.
- [8] Olejnik, P., & Ayankoso, S. (2023). Friction modelling and the use of a physics-informed neural network for estimating frictional torque characteristics. *Meccanica*, 58(10), 1885-1908.
- [9] Tagliani, F. L. (2023). Study and Development of Mechatronic Devices and Machine Learning Schemes for Industrial Applications.
- [10] Trinh, M., Schwiedernoch, R., Gründel, L., Storms, S., & Brecher, C. (2022, October). Friction modeling for structured learning of robot dynamics. In *Congress of the German*

Academic Association for Production Technology (pp. 396-406). Cham: Springer International Publishing.

- [11] Babin, A. (2021, February). Data-driven system identification and optimal control of an active rotor-bearing system. In *IOP Conference Series: Materials Science and Engineering* (Vol. 1047, No. 1, p. 012053). IOP Publishing.
- [12] Han, S., Orzechowski, G., Kim, J. G., & Mikkola, A. (2023). Data-driven friction force prediction model for hydraulic actuators using deep neural networks. *Available at SSRN* 4494358.
- [13] Ecker, L., & Schöberl, M. (2022). Data-driven observer design for an inertia wheel pendulum with static friction. *IFAC-PapersOnLine*, 55(40), 193-198.
- [14] Ma, Z., Liao, H., Gao, J., Nie, S., & Geng, Y. (2023). Physics-informed machine learning for degradation modeling of an electro-hydrostatic actuator system. *Reliability Engineering & System Safety*, 229, 108898.
- [15] Griffa, P., Sferrazza, C., & D'Andrea, R. (2022, May). Leveraging distributed contact force measurements for slip detection: a physics-based approach enabled by a data-driven tactile sensor. In *2022 International Conference on Robotics and Automation (ICRA)* (pp. 4826-4832). IEEE.