**Direct Current Servo Motor Machine Learning-Based System Identification**

Eka Sugiarto\*

\*School of Electrical Engineering, Telkom University, Bandung, 40287, Indonesia

\*ekasugiartoo@student.telkomuniversity.ac.id

**Abstract**

**In the era of Industry 4.0, automation technologies are essential, particularly in manufacturing and logistics. This study presents a novel application of machine learning for system identification of direct current servo motors (DCSMs), aimed at enhancing the precision and efficiency of automated guided vehicles (AGVs) operating under varying load conditions. We compare three methodologies: Autoregressive-Moving Average with Exogenous Inputs (ARMAX), Random Forest Regressor, and Nonlinear Autoregressive with Exogenous Inputs (NLARX) models. Simulated data from Simulink was utilized for training and validation, with analyses conducted in MATLAB and Python. Our results indicate that the ARMAX model achieves superior accuracy, with a Mean Absolute Error (MAE), Mean Squared Error (MSE), and a R-squared (R²), outperforming the NLARX and Random Forest models. These findings suggest that ARMAX and NLARX models hold significant promise for real-time system identification in industrial applications. This research provides critical insights for the development of more efficient AGV control systems and could be extended to other nonlinear and time-varying systems in various engineering domains.**

*Keywords*: System Identification; Machine Learning; ARMAX; AGV; DCSM; Simulink; MATLAB.

1. **Introduction**

Automation and mobile devices have been developing rapidly since the start of Industry 4.0, including in the fields of manufacturing and engineering [1]. One application of automation systems is in logistics vehicles. To reduce operational costs and improve process efficiency, smart logistics equipment is needed [2]. Automated guided vehicle (AGV) technology has been widely developed in recent years. The workflow of AGVs is repetitive, where AGVs are programmed to move from one point to another according to a route. However, the load carried by AGVs is often inconsistent due to the different shapes and sizes of the load. The difference in load can affect the amount of friction between the wheels and the ground [3]. This can affect the accuracy of the AGV in reaching its destination. Therefore, this problem motivates the development of a real-time identification system so that the system can be controlled according to changing conditions.

System identification is a method for improving the understanding of a system and enables the simulation of complex systems for making predictions over time [4]. Advanced machine learning algorithms, big data, and modern computing devices are revolutionizing the complex system modeling and control based on data driven. The dominance of first principle models in control methods is giving way to data-driven models powered by machine learning. This trend is impacting diverse fields like turbulence simulation, disease forecasting, and industrial automation [5].

Machine learning has permeated various fields, from science and engineering to economics, Industry 4.0, geography, and healthcare [6]. In the context of system identification, research has demonstrated the promise of machine learning as a tool, with algorithms like neural networks effectively estimating nonlinear systems [7]. Machine learning also finds applications in weather prediction and climate change analysis, serving as a decision-support tool due to the inherent uncertainties and lack of clear quantitative data in these domains [8].

This research will explore the application of machine learning as a tool for identifying physical systems using data-driven approaches. System dynamics equations will serve exclusively as a source of datasets for testing purposes.

1. **Research Methodology**

This research involves a series of systematic steps. It begins with an extensive review of literature on system identification techniques. Following this, a suitable physical model is chosen for simulation, and an appropriate machine learning model is identified. Rigorous model testing is then conducted, and the results are carefully analyzed to draw meaningful conclusions. Finally, the findings are thoroughly documented and evaluated. The research flow is illustrated in Figure 1.

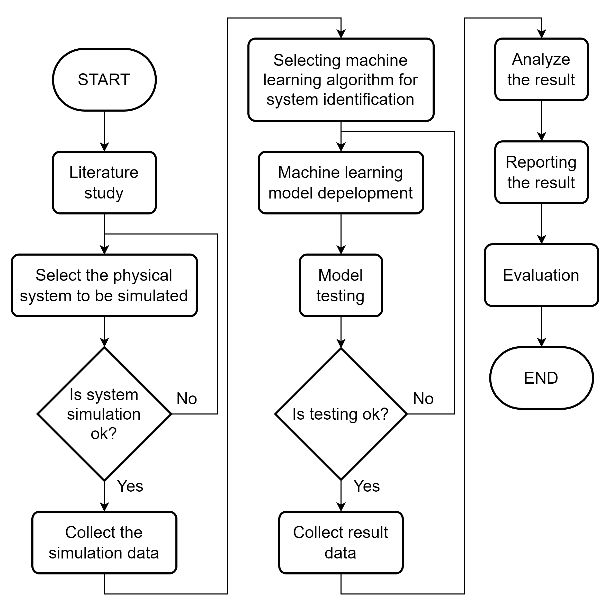


Figure 1. Research Process Flowchart

* 1. System Model

This paper utilizes a DC servo motor (DCSM) as the physical system to develop a system identification model. DCSMs are widely employed in industrial applications as actuators that deliver high torque. DCSMs are widely employed in devices that required precise motion, such as robotic arms, fluid valve control, and following a position trajectory under variable load conditions [9]. The DCSM model will utilize the following paper [10] and [11] as a reference for the physical equations, block diagram, and physical parameters of the system. The details of the DCSM model are presented below.

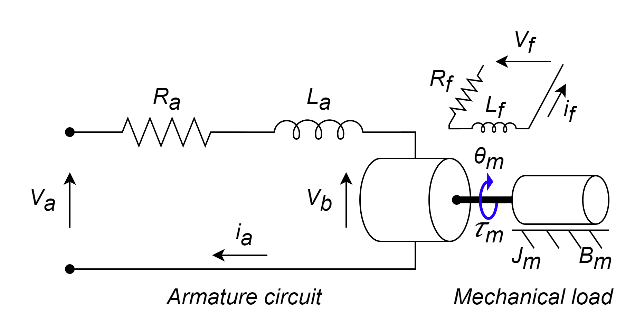


Figure 2. DCSM Model Representation

Figure 1 depicts the model of the DCSM with = applied armature voltage; = armature circuit resistance; = armature circuit inductance; = currect flow through the armature; = back electromotive force (EMF) voltage; = angular position of rotor; = torque produced by the motor; = inertia of rotor; and = viscous friction coefficient. According to reference [11], the transfer function of the DCMS is represented by Equation ( 1 ) below:

Equation ( 1 )

In Equation ( 1 ), given represent DCSM transfer function, where denotes the angular position and is the applied armature voltage. Also = and = . Table 1 below presents the values for each parameter obtained from reference [10] and [11].

Table 1. Parameter of DCSM

|  |  |  |
| --- | --- | --- |
| **Symbol** | **Significance** | **Value** |
|  | Armature resistance |  |
|  | Armature inductance |  |
|  | Moment of inertia |  |
|  | Friction coefficient |  |
|  | Back EMF constant |  |
|  | Torque constant |  |

* 1. Dataset Generation

The dataset for training and testing the model was obtained by simulating the transfer function Equation ( 1 ) in Simulink. The resulting data was then exported to an Excel file for use on platforms other than MATLAB. The Simulink block diagram is shown below.

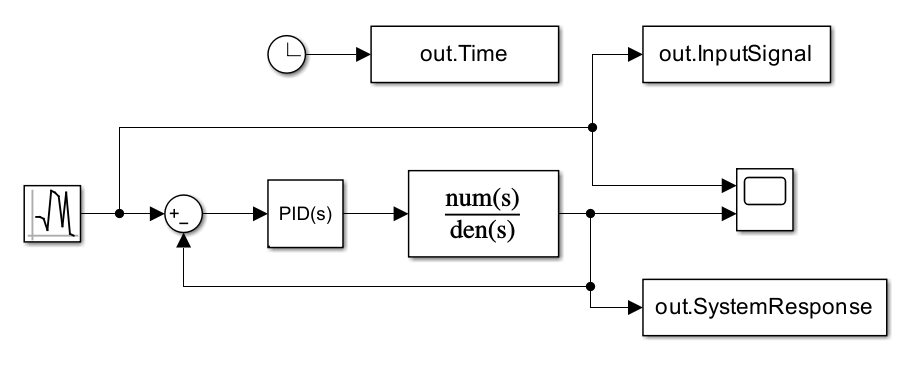


Figure 3. System Diagram Block in Simulink

The Simulink model includes a PID controller, whose parameters (Kp, Ki, Kd) were optimized through a trial-and-error process to achieve minimal error of the system response.

* 1. System Identification Model

This paper employs three system identification methods: autoregressive-moving average with exogenous inputs (ARMAX), Random Forest Regressor, and nonlinear autoregressive exogenous (NLARX). ARMAX is implemented in MATLAB R2024a using the SysId toolbox with customized coding. Random Forest and NLARX are implemented in Python 3.12.4 with updated libraries compatible for this version. The associated files used in this paper are stored in the following repository [12]. Here the explanation of each method.

1. ARMAX

ARMAX model general representation given by Equation ( 2 ) below [13]:

Equation ( 2 )

The = output and = input signals, = white noise, = delay operator, i.e., , and the A, B, and C are polynomials of degrees , , and defined below:

Equation ( 3 )

Common ARMAX model indicates as , where = sampling interval related to transport delay. Consequently, and the polynomial will be given by

1. Random Forest

The Random Forest (RF) algorithm leverages an ensemble of decision trees for prediction tasks within machine learning [14]. This ensemble is constructed through the training of individual decision trees on a training dataset. The resulting predictions from the RF model are derived by aggregating the outputs of the constituent trees, employing the mode for classification problems or the mean prediction for regression problems. A key strength of RF lies in its utilization of bootstrap aggregating, a process that generates multiple training datasets via random sampling with replacement from the original data. Each unique decision tree within the ensemble is subsequently trained on one such bootstrap sample. This strategy injects randomness into the training process, mitigating the risk of overfitting the model to the specific characteristics of the training data and enhancing generalization performance on unseen data [15].

1. Nonlinear ARX

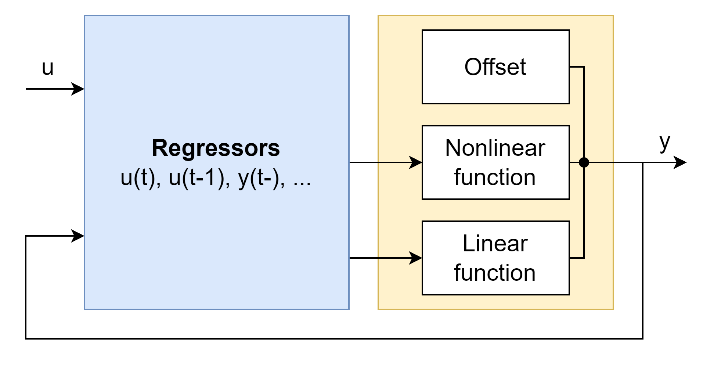


Figure 4. NLARX Diagram [16]

A nonlinear ARX model uses model regressors and an output function to compute outputs in two stages. First, it map a record od observed signals into a finite-dimensional set of regressor. Second, map the set of regressors to an output by using a static nonlinear map [17]. The NLARX equation [18], represented by the Equation ( 4 ).

Equation ( 4 )

where denotes the output at time step , represents the set of regressors, signifies the input, and represents the linear ARX component. The offset term is represented by , while symbolizes the output of the nonlinear function. The matrix plays a crucial role in ensuring well-conditioned computations.

* 1. Model Evaluation

The evaluation for each model were evaluated with these statistical measures [14][15], such as mean absolute error (MAE), mean square error (MSE), root mean square error (RMSE), coefficient of determination (). The selection of these four metrics stems from their established role as benchmarks for evaluating regression analysis across diverse scientific domains [19]. Each equation shown below.

Equation ( 5 )

1. **Result and Discussion**
   1. Simulation Result

Simulation on Simulink was run for 200 seconds with a time sampling of 0.01. The PID values used were Kp = 9, Ki = 3, Kd = 3, and the random number used parameters of mean = 0, variance = 2, seed 0, and sample time = 1. The results of the simulation are shown in Figure 4.

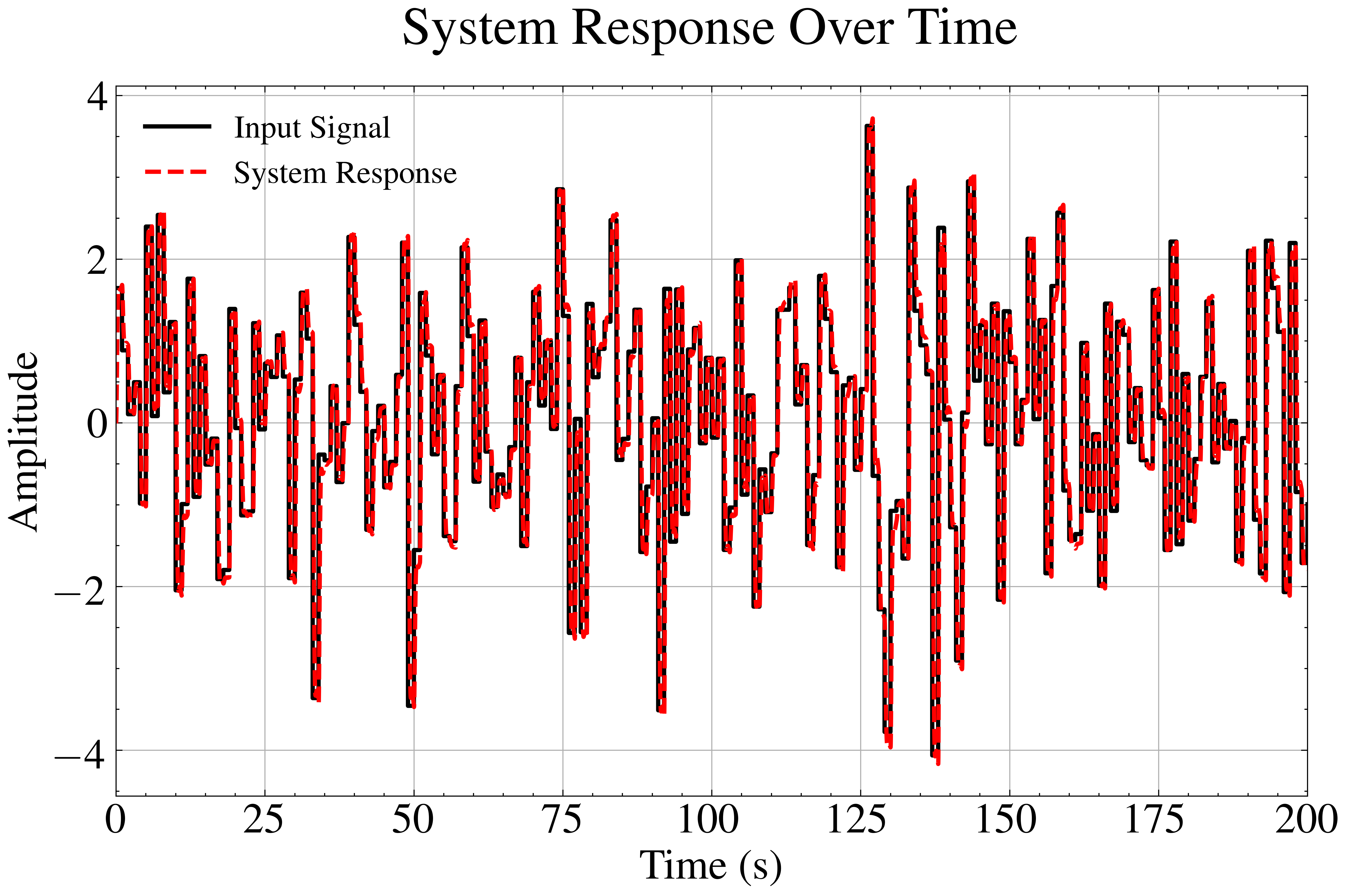


Figure 5. Simulink Simulation Result

Figure 4 shows the random number as the Input Signal and the System Response as the output signal of the system. The simulation results show a close alignment between the input signal and system response, indicating effective control by the PID controller.

* 1. ARMAX Result

ARMAX estimation was performed using MATLAB code, utilizing data from the Simulink simulation stored in the Workspace. The parameters used were = 2, = 2, = 1, and = 1. The estimation results are shown in Figure 5.

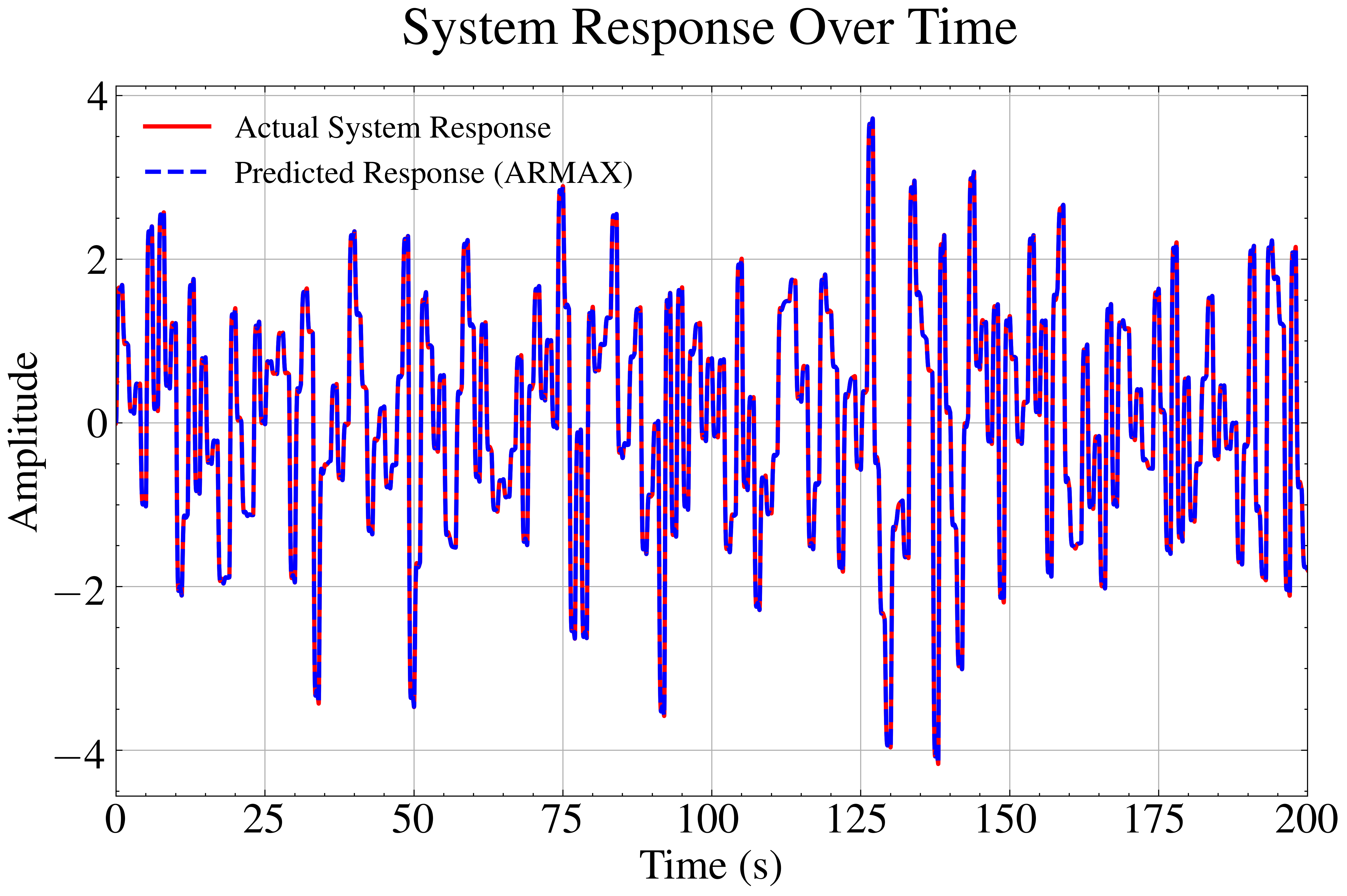


Figure 6. Actual System Response VS ARMAX Estimation

The prediction results indicate MAE = 0.0003, MSE = 0.000, RMSE = 0.0005, and = 0.9999. These values demonstrate the accuracy of the estimation/prediction, with an value close to 1 and an error below 0.01.

* 1. NLARX and Random Forest

In the Random Forest and NLARX testing, the dataset will be divided into two parts: 80% for training and 20% for testing. Figure 6 shows a plot of the Actual System Response and the Predicted System Response using Random Forest and NLARX, along with the ARMAX estimation results. A comparison of the results of the three methods is shown in Table 2.

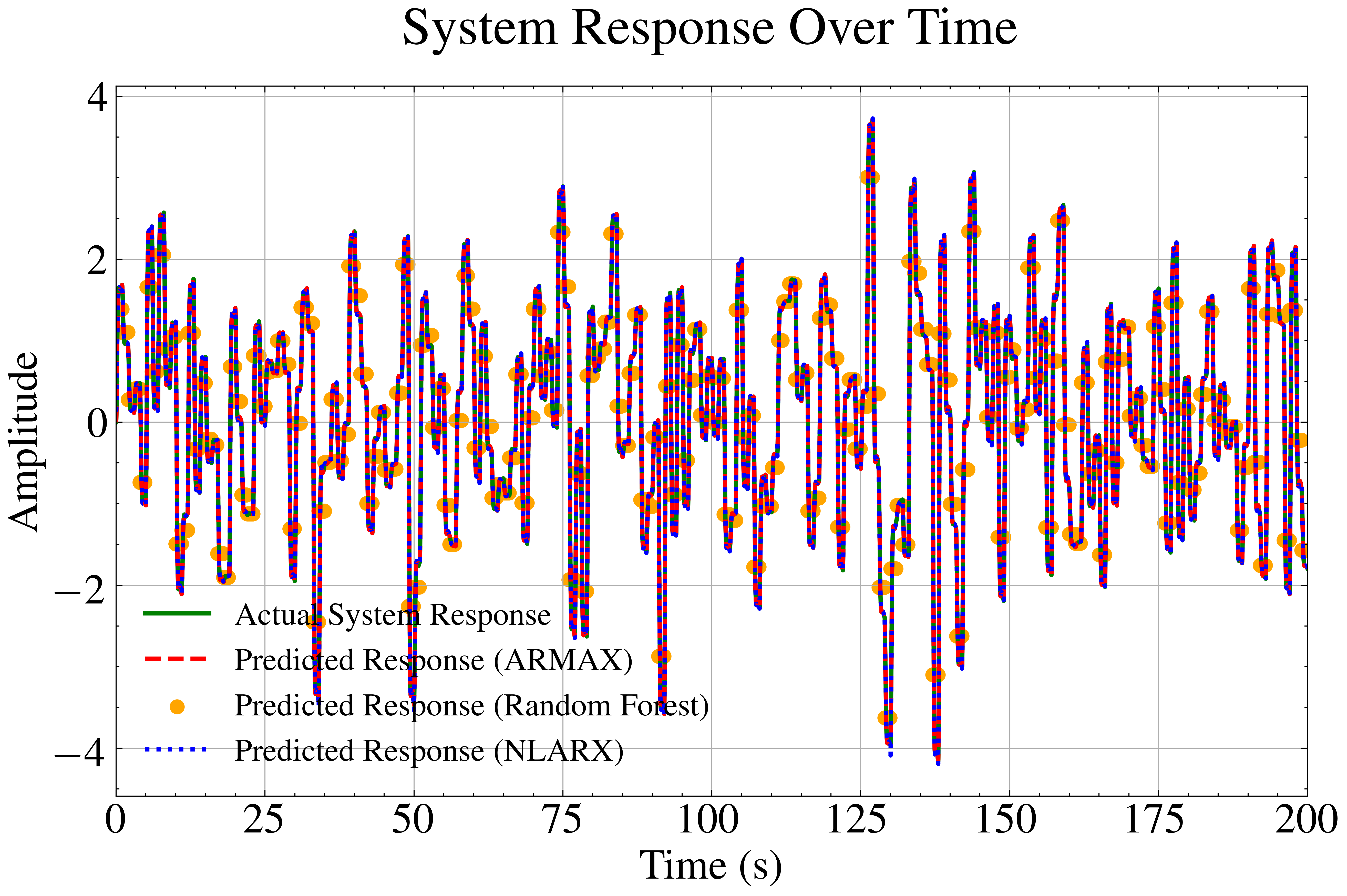


Figure 7. Estimation Plot of ARMAX, NLARX, and Random Forest

Table 2. Metrics Result of Estimation/Prediction

|  |  |  |  |
| --- | --- | --- | --- |
| **Test Parameter** | **ARMAX** | **Random Forest** | **NLARX** |
| MAE | 0.0003 | 0.4190 | 0.0133 |
| MSE | 0.0000 | 0.4275 | 0.0005 |
| RMSE | 0.0005 | 0.6538 | 0.0233 |
|  | 0.9999 | 0.7643 | 0.9996 |

Table 2 presents a comparative analysis of three modeling approaches across four performance metrics. The ARMAX model demonstrates superior predictive accuracy with an exceptionally low MAE, MSE, and RMSE coupled with an value of 0.9999, suggesting an almost perfect fit and minimal prediction error. In contrast, the Random Forest model shows considerably higher error values and a lower value of 0.7643, indicating suboptimal performance and greater prediction inaccuracies. The NLARX model, while not matching the precision of ARMAX, still performs robustly with minimum MAE, MSE, & RMSE, and an value of 0.9996, suggesting it is a highly effective predictive tool but slightly less accurate than ARMAX. These results underscore the exceptional precision of ARMAX, moderate efficacy of NLARX, and relatively lower performance of the Random Forest for system identification.

1. **Conclusions**

This research presents a comprehensive study on the application of machine learning techniques for system identification of DCSM. Through simulation and empirical analysis, three models: ARMAX, random forest, and NLARX were evaluated for their predictive accuracy and robustness. The results clearly indicate that the ARMAX model outperforms the other methods, achieving near-perfect accuracy with minimal prediction errors. The NLARX model also demonstrates strong performance, albeit slightly less precise than ARMAX. Conversely, the Random Forest model, while useful, shows higher error rates and lower R² values, indicating it may be less suitable for this specific application.

These insights provide valuable guidance for selecting appropriate system identification models in the context of AGV operations, emphasizing the importance of model choice in achieving optimal control and efficiency. Future research could explore the application of these models to different types of motors and incorporate real-world operational data for further validation.

**Reference**

[1] Á. Cservenák, “CREATING VOLTAGE, CURRENT AND NAVIGATION MEASUREMENT SYSTEM ON AN AGV FOR MOTION CONTROLLING,” *ACADEMIC JOURNAL OF MANUFACTURING ENGINEERING*, vol. 19, no. 2, p. 31, 2021.

[2] L. Willems, “Understanding the Impacts of Autonomous Vehicles in Logistics,” *The Digital Transformation of Logistics*, pp. 113–127, Mar. 2021, doi: 10.1002/9781119646495.CH8.

[3] W. Jiang, G. Zhu, and Y. Zheng, “Iterative Learning Control for AGV Drive Motor Based on Linear Extended State Observer,” *Machines 2021, Vol. 9, Page 324*, vol. 9, no. 12, p. 324, Nov. 2021, doi: 10.3390/MACHINES9120324.

[4] E. Chee, W. C. Wong, and X. Wang, “An integrated approach for machine-learning-based system identification of dynamical systems under control: application towards the model predictive control of a highly nonlinear reactor system,” *Front Chem Sci Eng*, vol. 16, no. 2, pp. 237–250, Feb. 2022, doi: 10.1007/S11705-021-2058-6/METRICS.

[5] R. Cao, Y. P. Lu, and Z. He, “System identification method based on interpretable machine learning for unknown aircraft dynamics,” *Aerosp Sci Technol*, vol. 126, p. 107593, Jul. 2022, doi: 10.1016/J.AST.2022.107593.

[6] R. H. Jhaveri, A. Revathi, K. Ramana, R. Raut, and R. K. Dhanaraj, “A Review on Machine Learning Strategies for Real-World Engineering Applications,” *Mobile Information Systems*, vol. 2022, no. 1, p. 1833507, Jan. 2022, doi: 10.1155/2022/1833507.

[7] A. Perrusquía and W. Yu, “Identification and optimal control of nonlinear systems using recurrent neural networks and reinforcement learning: An overview,” *Neurocomputing*, vol. 438, pp. 145–154, May 2021, doi: 10.1016/J.NEUCOM.2021.01.096.

[8] B. Bochenek and Z. Ustrnul, “Machine Learning in Weather Prediction and Climate Analyses—Applications and Perspectives,” *Atmosphere (Basel)*, vol. 13, no. 2, p. 180, Feb. 2022, doi: 10.3390/ATMOS13020180/S1.

[9] M. Khalifa, A. H. Amhedb, and M. Al Sharqawi, “Real Time DC Motor Position Control Using PID Controller in LabVIEW,” *Journal of Robotics and Control (JRC)*, vol. 2, no. 5, pp. 342–348, Sep. 2021, doi: 10.18196/jrc.25104.

[10] H. S. Dakheel, Z. B. Abdullah, N. S. Jasim, and S. W. Shneen, “Simulation model of ANN and PID controller for direct current servo motor by using Matlab/Simulink,” *TELKOMNIKA (Telecommunication Computing Electronics and Control)*, vol. 20, no. 4, pp. 922–932, Aug. 2022, doi: 10.12928/TELKOMNIKA.V20I4.23248.

[11] Z. B. Abdullah, S. W. Shneen, and H. S. Dakheel, “Simulation Model of PID Controller for DC Servo Motor at Variable and Constant Speed by Using MATLAB,” *Journal of Robotics and Control (JRC)*, vol. 4, no. 1, pp. 54–59, Feb. 2023, doi: 10.18196/JRC.V4I1.15866.

[12] E. Sugiarto, “GitHub Repository - UAS\_Pemodelan\_Sistem\_Identifikasi.” Accessed: Jun. 18, 2024. [Online]. Available: https://github.com/EkaSugiarto/UAS\_Pemodelan\_Sistem\_Identifikasi

[13] G. C. Silva Neto, D. S. Chui, F. P. R. Martins, A. T. Fleury, F. Furnari, and F. C. Trigo, “Identification of CO 2 and O 2 emissions dynamics in a natural gas furnace through flame images, ARMAX models, and Kalman filtering,” *Journal of the Brazilian Society of Mechanical Sciences and Engineering*, vol. 43, no. 5, pp. 1–13, May 2021, doi: 10.1007/S40430-021-02967-W/METRICS.

[14] L. Xue, Y. Liu, Y. Xiong, Y. Liu, X. Cui, and G. Lei, “A data-driven shale gas production forecasting method based on the multi-objective random forest regression,” *J Pet Sci Eng*, vol. 196, p. 107801, Jan. 2021, doi: 10.1016/J.PETROL.2020.107801.

[15] F. Wang, Y. Wang, K. Zhang, M. Hu, Q. Weng, and H. Zhang, “Spatial heterogeneity modeling of water quality based on random forest regression and model interpretation,” *Environ Res*, vol. 202, p. 111660, Nov. 2021, doi: 10.1016/J.ENVRES.2021.111660.

[16] “What are Nonlinear ARX Models? - MATLAB & Simulink.” Accessed: Jun. 18, 2024. [Online]. Available: https://www.mathworks.com/help/ident/ug/what-are-nonlinear-arx-models.html

[17] R. Singh and M. Sznaier, “On Identification of Nonlinear ARX Models with Sparsity in Regressors and Basis Functions,” *IFAC-PapersOnLine*, vol. 54, no. 7, pp. 720–725, Jan. 2021, doi: 10.1016/J.IFACOL.2021.08.446.

[18] A. E. Osamah, H. B. Hassrizal, and S. M. Othman, “Modeling a new Gimbal System using Nonlinear ARX (NARX) Model,” *J Phys Conf Ser*, vol. 2641, no. 1, p. 012023, Nov. 2023, doi: 10.1088/1742-6596/2641/1/012023.

[19] D. Chicco, M. J. Warrens, and G. Jurman, “The coefficient of determination R-squared is more informative than SMAPE, MAE, MAPE, MSE and RMSE in regression analysis evaluation,” *PeerJ Comput Sci*, vol. 7, pp. 1–24, Jul. 2021, doi: 10.7717/PEERJ-CS.623/SUPP-1.