Available online at www.eccomasproceedia.org
Eccomas Proceedia UNCECOMP (2023) 787-797

ECCOMAS
Proceedia

UNCECOMP 2023
5th ECCOMAS Thematic Conference on
Uncertainty Quantification in Computational Science and Engineering
M. Papadrakakis, V. Papadopoulos, G. Stefanou (eds.)
Athens, Greece, 12-14 June 2023

COMPARISON OF DATA DRIVEN BASED SYSTEM IDENTIFICATION TECHNIQUES FOR DIFFERENT NON-LINEAR DYNAMICAL SYSTEMS

Pushpa Pandey[1], Hamed Haddad Khodaparast[1], Michael Ian Friswell[1], Tanmoy Chatterjee[2], and Tom Deighan[3]

[1]Swansea University
Swansea University, Bay Campus, Fabian Way, Crymlyn Burrows, Swansea SA1 8EN, UK
e-mail: {pushpa.pandey, h.haddadkhodaparast, m.i.friswell}@swansea.ac.uk

[2] University of Surrey
Centre for Engineering Materials, School of Mechanical Engineering Sciences, Faculty of Engineering and Physical Sciences, Guildford GU2 7XH, UK
e-mail: t.chatterjee@surrey.ac.uk

[3]United Kingdom of Atomic Energy Authority
e-mail: tom.deighan@ukaea.uk

Abstract. *The present work focuses on comparing two different system identification methods for the identification of known nonlinearities, including cubic stiffness, quadratic damping with cubic stiffness, and dry friction. The methods compared are SINDy and PiSL. SINDy is a sparse regression method that identifies governing equations from time-series data, and PiSL is a physics-informed spline learning method that uses prior knowledge of the underlying physical laws of the system to improve the accuracy of the identification. The methods are evaluated using simulated data from known nonlinearities. The results show that the two methods can accurately identify the nonlinearities, but each method has its strengths and weaknesses. This study provides valuable insights into the performance of these system identification methods for nonlinear systems, which can help researchers choose the most appropriate method for their specific application.*

Keywords: System identification, SINDy, PiSL, sparse regression, physics-informed spline learning

doi: 10.7712/120223.10375.19997

1 INTRODUCTION

System identification is the process of identifying a mathematical model that, depending on input and output data, represents the behaviour of a dynamic system. This often entails calculating the state-space model or transfer function of the system in the context of dynamical systems. System identification can be done using time-domain and frequency-domain methods. Input-output information is utilised in time-domain identification to estimate the system's impulse response or step response, which may subsequently be used to determine the transfer function. In order to identify a system in the frequency domain, its input-output behaviour is examined to ascertain its frequency response. Both the frequency and time domains can be used to identify systems, and each technique has its merits and restrictions.

The input-output behaviour of the system is examined in the frequency domain during frequency-domain system identification using techniques like Fourier analysis or the Fast Fourier Transform (FFT). This method works well with linear time-invariant (LTI) systems because they have transfer functions that are constant at all frequencies. It is possible to ascertain the system's transfer function by investigating the frequency response of the system, which can then be used to predict how the system would respond to various input scenarios. The ability to easily study a system's behaviour over a broad frequency range is one benefit of frequency-domain system identification. The analysis of systems that show significant resonance or frequency-dependent behaviour can benefit from this method.

In time-domain system identification, methodologies like impulse response or step response analysis are used to study the system's input-output behaviour in time. This method offers a clear indication of the system's behaviour over time, making it suitable for both linear and non-linear systems. The transfer function, state-space model, and other system parameters can be found by examining the system's time-domain behaviour. One benefit of time-domain system identification is that it gives a clear indication of the system's transient reaction, which is useful for comprehending the stability and performance of the system.

Overall, the characteristics of the system being studied and the identification process's objectives will determine whether to use a frequency-domain or time-domain system identification approach. Frequency-domain methods may be the best strategy for linear systems with constant transfer functions, whereas time-domain methods may yield more precise findings for nonlinear or time-varying systems.

The following are some typical methods for system identification:

- Least-squares techniques: They entail employing a least-squares criterion to reduce the discrepancy between model output and measured output. Either the time domain or the frequency domain can be used for this. [1] [2]
- Maximum likelihood techniques: In these techniques, the model's parameters are found to maximise the possibility that the observed data will match the model. [3]
- Prony's method: This time-domain approach entails fitting an accumulation of exponential functions to the output measurements. [4]
- Eigensystem realisation algorithm: This estimates the state-space model of the system using its frequency response data. It is a frequency-domain approach. [5]
- Subspace identification techniques: They rely on a subspace identification algorithm to estimate the system's state-space model from input-output data.[6] [7]

Once the model has been identified, it can be used for analysis and control of the system. This can include predicting the system's behaviour under different conditions, designing controllers to achieve the desired performance, or optimizing the system's behavior for a particular application. Extending system identification methods to nonlinear dynamical systems is more challenging than linear systems due to the complex and often chaotic behavior exhibited by these systems. In general, nonlinear systems are described by differential equations that cannot be solved analytically, and the complexity of their dynamics means that it is not always possible to obtain reliable input-output data for system identification.
System identification approaches include linearization, parametric and nonparametric techniques. Linearization techniques approximate the system by a linear model around a particular operating point, which can be identified using standard techniques for linear systems. However, the accuracy of the resulting model depends on how well the linear model approximates the true behavior of the highly nonlinear systems. Parametric techniques assume a particular functional form for the system's dynamics, but the identification process can be computationally intensive and may require specialized computational resources. Nonparametric techniques estimate the system's input-output behavior using data-driven approaches such as neural networks or support vector machines, and can be highly accurate but require large amounts of data to train and can be difficult to interpret. These methods are well-suited for control or prediction tasks that require high accuracy.
Among different nonparametric techniques, SINDy and PiSL are the few which can handle both static and dynamic nonlinearities and are found to be computationally efficient. These techniques are highly effective for identifying complex nonlinear systems with unknown dynamics, but each method has its strengths and limitations. SINDy is good at identifying sparsity and is computationally efficient, and PiSL is effective for incorporating prior knowledge about the underlying physics of the system. Overall, these techniques provide powerful tools for nonlinear system identification, enabling a deeper understanding of the underlying dynamics of complex systems.
Current work focuses on comparing the two system identification methods i.e. SINDy (Sparse Identification of Nonlinear Dynamics) and PiSL (Physics-Informed Spline Learning) for the known non-linearity cases (Cubic stiffness, quadratic damping with cubic stiffness and dry friction). By comparing these methods for known nonlinearities, a better understanding of the strengths and weaknesses of each method and their suitability for different types of systems and data can be established.

2 DATA-DRIVEN APPROACH FOR SYSTEM IDENTIFICATION

Nonparametric techniques involve estimating the system's input-output behavior using data-driven approaches. The resulting models can be highly accurate when trained on large amounts of data, but they can also be difficult to interpret and may require large amounts of data to train. This makes them well-suited for applications where high accuracy is important, such as in control or prediction tasks. Non-parametric techniques do not require explicit functional forms to describe the system dynamics. Some of the common non-parametric techniques involve the following:

- Neural Networks: Neural networks are a popular non-parametric technique for system identification. They can learn complex input-output relationships without assuming a particular functional form. Neural networks are often used for modeling nonlinear systems, as they can capture complex nonlinear dynamics.[8][9]

- Support Vector Machines (SVM): SVM is another non-parametric technique for system identification. SVM seeks to find a separating hyperplane in high-dimensional feature space that maximizes the margin between the different classes of data. SVM has been applied to system identification problems, where it can be used to learn nonlinear input-output mappings.[10][11]

- Gaussian Processes (GP): GP is a non-parametric probabilistic approach to system identification. GP can be used to estimate the underlying function that generates the observed data, while also providing uncertainty estimates. GP has been used for modeling nonlinear systems, as it can capture complex nonlinear dynamics. [12][13]

- Kernel Methods: Kernel methods are a family of non-parametric techniques that are used for regression, classification, and density estimation. Kernel methods use a kernel function to map the input data into a higher-dimensional feature space, where the data may be more separable. Kernel methods have been applied to system identification problems, where they can be used to learn nonlinear input-output mappings.[14]

- NARMAX (Nonlinear AutoRegressive Moving Average with eXogenous inputs) is a type of mathematical model used in system identification, which involves analyzing input-output data to obtain a mathematical representation of the system. NARMAX models are typically defined as a combination of nonlinear autoregressive (AR) and moving average (MA) terms, where the nonlinearity is captured using a polynomial. This method is popular because it provides a flexible way to model nonlinear systems, while still using a linear model structure.[15][16][18][17]

Some of the recently emerging techniques have shown promising results in identifying the system dynamics.

- SINDy (Sparse Identification of Nonlinear Dynamics) is a data-driven method for identifying nonlinear dynamical systems using sparse regression. The method is based on the principle that many physical systems can be accurately modelled using only a few important terms, even in the presence of nonlinearity. SINDy identifies these important terms by applying sparse regression to the data, allowing for the identification of a parsimonious, interpretable model of the system.[19]

- PiSL (Physics-Informed Spline Learning) is a data-driven method that combines spline interpolation with knowledge of the system's underlying physics. The method uses a spline function to interpolate the system's behavior, while also incorporating known physical constraints into the model. This allows for a more interpretable model, while still providing the flexibility to capture the complex behavior of nonlinear systems. PiSL can capture the highly nonlinear input-output relationship between the input and output signals of the system. Additionally, by incorporating known physical constraints, PiSL can provide a model that is both accurate and interpretable.[20]

The current work focuses on comparing the recently emerging techniques i.e. SINDy and PiSL which have been successfully applied to a wide range of nonlinear systems, including fluid dynamics [21], chemical reactions [22], and biological systems[23].

3 NON-LINEAR CASES CONSIDERED FOR COMPARISON

A Single Degree of Freedom (SDOF) system for different non-linear systems (Eq. 1, 2, 3) is used to describe the dynamic behavior of a structure or mechanical system. This study examines the behavior of single-degree-of-freedom (SDOF) systems with two distinct non-linear systems. To activate the non-linear response, an initial displacement condition is applied. The equations for each non-linear system, provided below, are used to produce the response for each system.

Cubic stiffness

$$m\ddot{x} + C\dot{x} + K_1 x + K_2 x^3 = F(t) \tag{1}$$

where m = 1, C = 0.2, K_1 = 1 , K_2 = 1

Quadratic damping with cubic stiffness

$$m\ddot{x} + C_1\dot{x} + C_2\dot{x}|\dot{x}| + K_1 x + K_2 x^3 = F(t) \tag{2}$$

where m = 1, C_1 = 0.2, $C_2 = 0.5$, K_1 = 1 , K_2 = 1

Dry friction

$$m\ddot{x} + C_1\dot{x} + C_2\text{sign}\dot{x} + K_1 x = F(t) \tag{3}$$

where m = 1, C_1 = 0.2 , C_2 = 0.1, K_1 = 1

To identify these systems, their time response is analyzed using an input signal applied to the system and the resulting output displacement, which serves as the input for the identification method. The sinusoidal excitation force (sin(t)) as the input signal is used in system identification and the whole response is used for the identification of the full range of system dynamics and the estimation of the model parameters over the entire time period. However, it is important to ensure that the identified models are not overfitting the data and that they are able to generalize to different operating conditions or perturbations.

4 PERFORMANCE OF THE IDENTIFICATION TECHNIQUES FOR DIFFERENT NON-LINEAR CASES

In this study, we compared the performance of two identification techniques, SINDy and PiSL, for three nonlinear cases, including cubic stiffness, quadratic damping, and dry friction. The objective was to evaluate the effectiveness of SINDy and PiSL in identifying the system and its parameters accurately and efficiently for these different nonlinear cases. To conduct this study, each nonlinear case was simulated using a mathematical model and generated data from the simulation was used as input for both identification techniques. The accuracy and computational efficiency of the parameter estimates obtained from SINDy and PiSL for each case were compared.

Two evaluation metrics i.e. Pearson correlation and mean squared error (MSE) were used to compare the performance of both identification techniques. The Pearson correlation coefficient measures the linear relationship between two variables, in this case, the predicted values from the model and the actual values. A value of 1 indicates a perfect positive correlation, while a value of -1 indicates a perfect negative correlation. A value of 0 indicates no correlation.

The MSE measures the average of the squared differences between the predicted values and the actual values. A lower MSE indicates better accuracy of the model.

4.1 Cubic stiffness case

One of the first nonlinear cases considered for comparing system identification techniques was the cubic stiffness case. This case involves a system with a cubic stiffness term, which is a common nonlinear characteristic observed in various physical systems. From Figure 1, both PiSL and SINDy appear to be efficient in capturing the true values (Disp_Test). According to the evaluation metric, both PiSL and SINDy show a high correlation with the actual values, with PiSL having a slightly higher correlation coefficient of 0.999988 compared to SINDy's correlation coefficient of 0.998983 (From Table 1). This indicates that both methods are highly accurate in predicting the actual values.

However, when it comes to MSE, PiSL outperforms SINDy with a significantly lower value of 1.789479e-05 compared to SINDy's MSE of 1.618123e-03 (From Table 1). This suggests that PiSL is more accurate than SINDy in predicting the actual values but also indicates the need to assess the robustness of the identified models, and to investigate how well they generalize to different forcing, initial conditions, or measurement noise. This is important to ensure that the identified models are not overfitting the data and that they are able to capture the true underlying dynamics of the system. For the cubic stiffness case, PiSL and SINDy identified the model, number of parameters and their corresponding values correctly with three additional terms in PiSL (quadratic damping, stiffness and a bias) and one additional term in SINDy (bias). SINDy was able to capture the true dynamics correctly through the identified model.

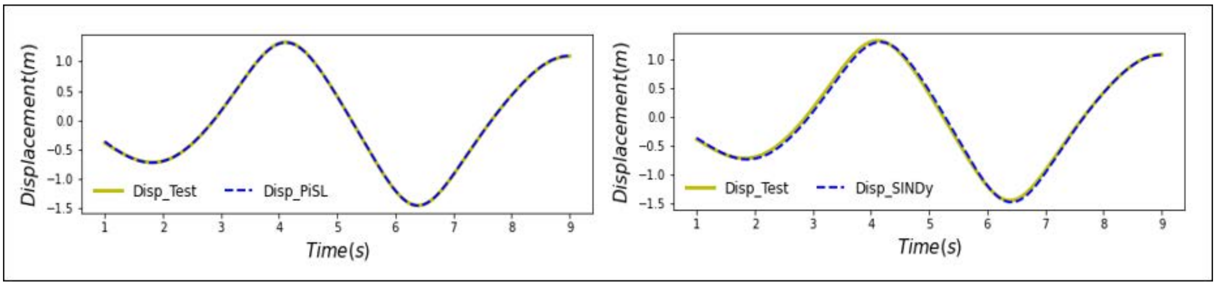

Figure 1: Comparison of SINDy and PiSL performance for Cubic stiffness case.

Metric	PiSL	SINDy
Pearson Correlation	0.999988	0.998983
Mean Squared Error	1.789479e-05	1.618123e-03

Table 1: Comparison metric: Cubic stiffness case.

4.2 Quadratic damping with cubic stiffness case

The second nonlinear case taken into consideration for comparing system identification strategies was quadratic damping with cubic stiffness. Both PiSL and SINDy exhibit a strong correlation with the true values, according to the results presented in Table 2 and Figure 2, with PiSL having a slightly higher correlation coefficient of 0.999996 compared to SINDy's correlation coefficient of 0.998687. This shows that both approaches have a very high degree of accuracy in forecasting the actual values. However, when it comes to MSE, PiSL outperforms SINDy with a significantly lower value of 2.742266e-06 compared to SINDy's MSE of 2.554376e-03. This suggests that PiSL is more accurate than SINDy in predicting the actual values. However, Pisl and SINDy identified the model but with many additional terms of low significance. The number of additional terms in PiSL was found to be less than in SINDy. PiSL was found to be near to capturing the true dynamics through the identified model.

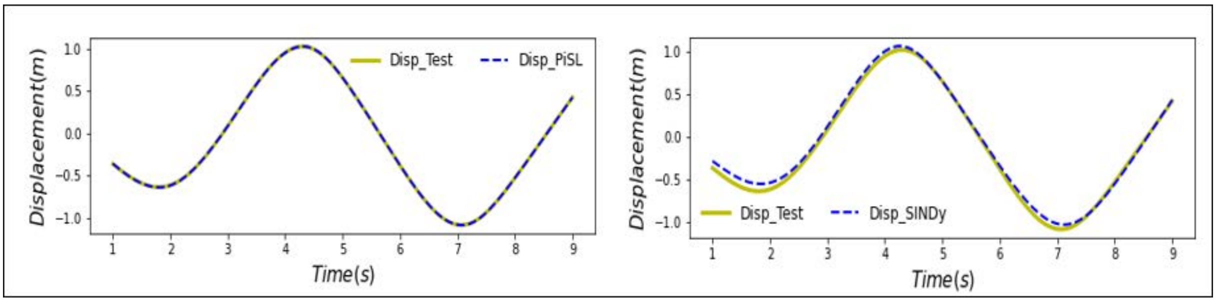

Figure 2: Comparison of SINDy and PiSL performance for Quadratic damping with cubic stiffness case.

Metric	PiSL	SINDy
Pearson Correlation	0.999996	0.998687
Mean Squared Error	2.742266e-06	2.554376e-03

Table 2: Comparison metric: Quadratic damping with cubic stiffness case.

4.3 Dry friction case

The Pearson correlation coefficient between the outputs of the two models (PiSL and SINDy) was very high for both methods, indicating a strong positive linear relationship between the models' outputs. However, the PiSL method had a slightly higher Pearson correlation coefficient than the SINDy method, suggesting a slightly stronger linear relationship between its output and the true values. The PiSL method had a significantly lower MSE (5.948443e-06) than the SINDy method (7.645737e-03) (From Table 3), indicating that PiSL was better at predicting the true values compared to the SINDy method for the cubic stiffness case (Figure 3). Again SINDy had numerous additional terms of low importance than PiSL but both the identification techniques could not identify the sign term which can be attributed to factors such as insufficient data or the probability that the sign term may not have had a large impact on the dynamics of the system during the time period for which data was collected.

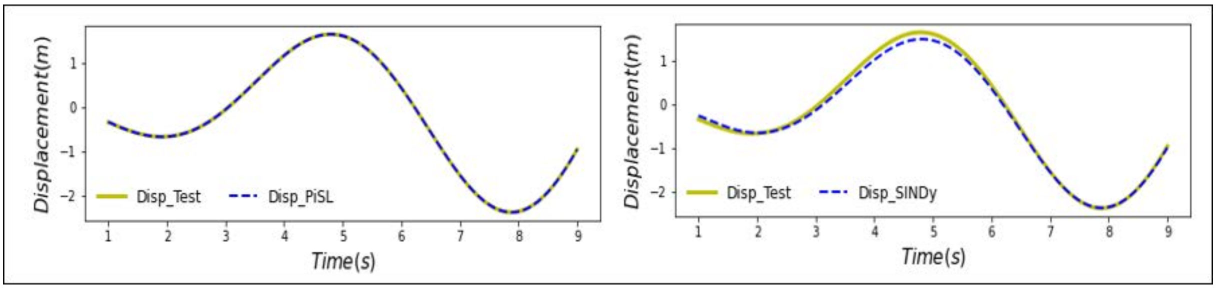

Figure 3: Comparison of SINDy and PiSL performance for Dry friction case.

Metric	PiSL	SINDy
Pearson Correlation	0.999998	0.999416
Mean Squared Error	5.948443e-06	7.645737e-03

Table 3: Comparison metric: Dry friction case.

Overall, based on these evaluation metrics, the PiSL and SINDy method appears to perform well for system identification for all three cases. It should be considered that the success of identification techniques like SINDy and PiSL depends heavily on the quality and quantity of data used as well as the complexity of the system being studied. Also, other factors such as computational efficiency and model interpretability play a major role in choosing a system identification method. For all three cases, PiSL was found to be more computationally expensive than SINDy. The use of spline functions and the incorporation of physics-based constraints may require more computational power than SINDy, which is a simpler approach that relies on sparse regression.

5 CONCLUSIONS

PiSL and SINDy are two different methods for learning the dynamics of a system from data. PiSL uses a physics-informed spline-based approach, where a spline function is used to approximate the unknown underlying dynamics of the system, and the spline coefficients are optimized to minimize the mean squared error between the predicted values and the actual values, subject to the underlying physical laws of the system being modelled.

On the other hand, SINDy uses a sparse regression-based approach, where a set of candidate functions is used to approximate the unknown underlying dynamics of the system, and the coefficients of these functions are optimized to minimize the mean squared error between the predicted values and the actual values. The candidate functions are chosen such that they only depend on a small number of system variables, which leads to a more interpretable model. PiSL appears to perform better than SINDy for all three non-linear scenarios, although it is shown to be computationally intensive. The main reason why PiSL is more computationally expensive than SINDy is due to the use of spline functions. Spline functions are continuous piecewise polynomials that require the optimization of a large number of spline coefficients to approximate the unknown underlying dynamics of the system. This optimization process can be computationally expensive, especially when dealing with high-dimensional systems or large amounts of data. In contrast, the sparse regression-based approach used by SINDy involves optimizing the coefficients of a small set of candidate functions, which is computationally more efficient.

However, the choice of method depends on the specific problem, and both PiSL and SINDy have their own advantages and disadvantages. PiSL may be more accurate in certain cases, especially when the data is noisy or sparse, or when the underlying dynamics of the system are highly nonlinear. Meanwhile, SINDy may be more computationally efficient and easier to interpret, especially when dealing with systems with a small number of variables or when the underlying dynamics are sparse. PiSL can potentially model more complex and highly nonlinear dynamics than SINDy, due to the use of splines, which can approximate a wider range of functions than the sparse set of candidate functions used by SINDy. Both PiSL and SINDy use regularization to enforce physical constraints on the learned models. However, the type and strength of regularization used may differ between the two methods, which can affect the accuracy and interpretability of the resulting models.

In the case of PiSL, using spline functions and incorporating physics-based constraints may require more computational power than SINDy, which is a simpler approach that relies on sparse regression. The computational expense of PiSL may make it less desirable for certain applications where speed and efficiency are critical. However, it's important to note that PiSL's ability to incorporate physical constraints into the model can provide additional benefits in terms of model interpretability and generalizability. In some cases, the additional computational cost may be worth the improved performance and interpretability that PiSL provides.

Ultimately, the choice of system identification method will depend on a variety of factors including the complexity of the system, the availability of computational resources, and the desired balance between model accuracy and computational efficiency. The future work focuses on exploring the efficiency of system identification techniques (PiSL and SINDy) with different excitation levels, noise levels, and non-linear cases to model the complex systems more accurately and efficiently which can be used for a wide range of applications, including optimization, control, and prediction. By comparing different modelling techniques and validating the results with experimental data the target is to improve our understanding of complex systems and develop new approaches for modelling them.

6 ACKNOWLEDGEMENTS

This research work has been partly funded by the RCUK Energy Programme [grant number EP/T012250/1]. The views and options expressed herein do not necessarily reflect those of UKAEA. The support of the Supercomputing Wales project, which is part-funded by the European Regional Development Fund (ERDF) via the Welsh Government is also acknowledged.

REFERENCES

[1] H. Dai and N. K. Sinha, Robust recursive least-squares method with modified weights for bilinear system identification. *IEE Proceedings D Control Theory and Applications, vol. 136, no. 3.* Institution of Engineering and Technology (IET), p. 122, 1989, doi: 10.1049/ip-d.1989.0018.

[2] A. J. Helmicki, C. A. Jacobson, and C. N. Nett, Least squares methods for H/sub infinity / control-oriented system identification. *IEEE Transactions on Automatic Control, vol. 38,*

no. 5. Institute of Electrical and Electronics Engineers (IEEE), pp. 819–826, May 1993. doi: 10.1109/9.277254.

[3] A. Hagenblad, L. Ljung, and A. Wills, Maximum likelihood identification of Wiener models, *Automatica, vol. 44, no. 11.* Elsevier BV, pp. 2697–2705, Nov. 2008. doi: 10.1016/j.automatica.2008.02.016.

[4] S. Foyen, M.-E. Kvammen, and O. B. Fosso, Prony's method as a tool for power system identification in Smart Grids, *International Symposium on Power Electronics, Electrical Drives, Automation and Motion (SPEEDAM).* IEEE, Jun. 2018. doi: 10.1109/speedam.2018.8445308.

[5] J.-N. Juang and R. S. Pappa, An eigensystem realization algorithm for modal parameter identification and model reduction, *Journal of Guidance, Control, and Dynamics, vol. 8, no. 5.* American Institute of Aeronautics and Astronautics (AIAA), pp. 620–627, Sep. 1985. doi: 10.2514/3.20031.

[6] W. Favoreel, B. De Moor, and P. Van Overschee, Subspace state space system identification for industrial processes, *Journal of Process Control, vol. 10, no. 2–3.* Elsevier BV, pp. 149–155, Apr. 2000. doi: 10.1016/s0959-1524(99)00030-x.

[7] A. Hansson, Z. Liu, and L. Vandenberghe, Subspace system identification via weighted nuclear norm optimization, *51st IEEE Conference on Decision and Control (CDC). IEEE,* Dec. 2012. doi: 10.1109/cdc.2012.6426980.

[8] J. Sjöberg, H. Hjalmarsson, L. Ljung, Neural Networks in System Identification, *IFAC Proceedings Volumes, Volume 27, Issue 8,*1994, Pages 359-382, ISSN 1474-6670, https://doi.org/10.1016/S1474-6670(17)47737-8.

[9] S. R. Chu, R. Shoureshi and M. Tenorio, Neural networks for system identification, *IEEE Control Systems Magazine, vol. 10, no. 3,* pp. 31-35, April 1990, doi: 10.1109/37.55121.

[10] P. M. L. Drezet, Support vector machines for system identification, *UKACC International Conference on Control (CONTROL '98). IEE,* 1998. doi: 10.1049/cp:19980312.

[11] V. Laurain, R. Tóth, D. Piga, and W. X. Zheng, An instrumental least squares support vector machine for nonlinear system identification, *Automatica, vol. 54.* Elsevier BV, pp. 340–347, Apr. 2015. doi: 10.1016/j.automatica.2015.02.017.

[12] J. Kocijan, A. Girard, B. Banko, and R. Murray-Smith, Dynamic systems identification with Gaussian processes, *Mathematical and Computer Modelling of Dynamical Systems, vol. 11, no. 4.* Informa UK Limited, pp. 411–424, Dec. 2005. doi: 10.1080/13873950500068567.

[13] W. Ariza Ramirez, Z. Q. Leong, H. Nguyen, and S. G. Jayasinghe, Non-parametric dynamic system identification of ships using multi-output Gaussian Processes, *Ocean Engineering, vol. 166.* Elsevier BV, pp. 26–36, Oct. 2018. doi: 10.1016/j.oceaneng.2018.07.056.

[14] A. Georgiev, Nonparametric system identification by kernel methods, *IEEE Transactions on Automatic Control, vol. 29, no. 4.* Institute of Electrical and Electronics Engineers (IEEE), pp. 356–358, Apr. 1984. doi: 10.1109/tac.1984.1103532.

[15] S. Chen and S. A. Billings, Representations of non-linear systems: the NARMAX model, *International Journal of Control, vol. 49, no. 3.* Informa UK Limited, pp. 1013–1032, Mar. 1989. doi: 10.1080/00207178908559683.

[16] S. Obeid, G. Ahmadi, and R. Jha, NARMAX Identification Based Closed-Loop Control of Flow Separation over NACA 0015 Airfoil, *Fluids, vol. 5, no. 3. MDPI AG, p. 100,* Jun. 29, 2020. doi: 10.3390/fluids5030100.

[17] L. Zhang, Y. Xue, Q. Xie, and Z. Ren, Analysis and neural network prediction of combustion stability for industrial gases, *Fuel, vol. 287. Elsevier BV, p. 119507,* Mar. 2021. doi: 10.1016/j.fuel.2020.119507.

[18] S. L. Kukreja, H. L. Galiana, and R. E. Kearney, NARMAX representation and identification of ankle dynamics, *IEEE Transactions on Biomedical Engineering, vol. 50, no. 1.* Institute of Electrical and Electronics Engineers (IEEE), pp. 70–81, Jan. 2003. doi: 10.1109/tbme.2002.803507.

[19] S. L. Brunton, J. L. Proctor, and J. N. Kutz, Sparse Identification of Nonlinear Dynamics with Control (SINDYc), *IFAC-PapersOnLine, vol. 49, no. 18.* Elsevier BV, pp. 710–715, 2016. doi: 10.1016/j.ifacol.2016.10.249.

[20] F. Sun, Y. Liu, Q. Wang, and H. Sun, PiSL: Physics-informed Spline Learning for data-driven identification of nonlinear dynamical systems, *Mechanical Systems and Signal Processing, vol. 191. Elsevier BV, p. 110165,* May 2023. doi: 10.1016/j.ymssp.2023.110165.

[21] R. Vinuesa and S. L. Brunton, Enhancing computational fluid dynamics with machine learning, *Nature Computational Science, vol. 2, no. 6. Springer Science and Business Media LLC, pp. 358–366,* Jun. 27, 2022. doi: 10.1038/s43588-022-00264-7.

[22] Q. Yang, C. A. Sing-Long, and E. J. Reed, Rapid data-driven model reduction of nonlinear dynamical systems including chemical reaction networks using L1-regularization, *Chaos: An Interdisciplinary Journal of Nonlinear Science, vol. 30, no. 5.* AIP Publishing, p. 053122, May 2020. doi: 10.1063/1.5139463.

[23] N. M. Mangan, S. L. Brunton, J. L. Proctor, and J. N. Kutz, Inferring Biological Networks by Sparse Identification of Nonlinear Dynamics, *IEEE Transactions on Molecular, Biological and Multi-Scale Communications, vol. 2, no. 1.* Institute of Electrical and Electronics Engineers (IEEE), pp. 52–63, Jun. 2016. doi: 10.1109/tmbmc.2016.2633265.