Neural network and Sparse Identification of Nonlinear Dynamics Integrated Algorithm for Digital Twin Identification

Jingyi Wang * **Jesús Moreira** ** **Yankai Cao** *
R. Bhushan Gopaluni *

* *Department of Chemical and Biological Engineering, University of British Columbia, Vancouver, BC V6T 1Z3, Canada (e-mail: wjy1129@student.ubc.ca).*
** *Digital, Innovation and Lean Capability Development Team, Imperial Oil Company, Calgary, AB T2C 5N1, Canada (e-mail: jesus.moreira@esso.ca)*

Abstract: Digital twins play a critical role in simulating industrial manufacturing systems to increase productivity and reduce time spent on troubleshooting. Owing to the complexity of real-world industrial systems, automatic sparse identification has emerged as an attractive approach to perform digital twin modelling. The sparse identification of nonlinear dynamics (SINDy) is a machine learning algorithm that performs feature engineering by generating a model term library and then solves a sparse regression problem between the objective outputs and the generated features. By solving a linear-in-parameter sparse regression problem, SINDy provides automatic discovery of system governing equations. However, the performance of SINDy-based algorithms may decline dramatically when applied to identify complex nonlinear relationships, such as implicit relationships. The substantial number of input variables for a real industrial process may further complicate the modelling procedure. We therefore propose the neural network and SINDy integrated algorithm to automatically select the critical features from a model term library and utilize the neural network to capture the process nonlinearity that cannot be captured by a linear-in-parameter model. SINDy performs feature generation considering both numerical methods and first-principles knowledge, making the proposed algorithm a hybrid system identification approach. A diesel hydrotreating unit case study with 37 input variables is analyzed in this paper to demonstrate the advantages of the proposed algorithm for nonlinear digital twin identification. By combining the advantages from both SINDy and neural networks, the proposed algorithm is able to improve the output prediction accuracy for all the three objectives.

Keywords: Digital twin, Feature engineering, Hybrid modelling, Industry 4.0, Nonlinear model reduction, Sparse process modelling and identification.

1. INTRODUCTION

The development of computing technology and the acceleration of data processing enable more efficient system information transfer and system identification, hence facilitating significant production innovations (Brettel et al., 2014). Digital twins have been identified as a promising approach to integrate physical operations and digital simulation for design, optimization, and control of production loops (Wang et al., 2022a). It has the potential to make considerable contributions to industrial processes in the future, opening the way for Industry 4.0. The development and widespread applications of digital twin technologies have facilitated the convergence of physical production systems with virtual simulations. In (Min et al., 2019), the authors integrated the machine learning technique with the internet of things (IoT)-collected, real-time data to develop a digital twin construction framework for the petrochemical industry. In (Park et al., 2019), a digital twin was designed and implemented for a connected micro-smart factory to assist the managers in decision-making and, as a result, lower production costs and enhance manufacturing efficiency. The digital twin technique expedited the design of individual portions of the comprehensive automated flow-shop manufacturing system (Liu et al., 2018). The developed digital twin used grey-box modelling to provide a reliable digital simulation of the flow-shop pre-production system.

Reliable and practical identification methods are critical to the comprehensive digital twin construction and extensive digital twin applications. One of the most pressing issues in digital twin identification is to automatically perform feature engineering and identify the complicated, nonlinear digital twin model with high accuracy (Shah et al., 2020).

⋆ This work was supported in part by Natural Sciences and Engineering Research Council of Canada.
Corresponding author.
Email address: bhushan.gopaluni@ubc.ca (R. B. Gopaluni).

The SINDy preforms feature engineering by creating a model term library containing all the potential digital twin model terms. Incorporating both data-driven terms, such as polynomial terms, and first-principles terms into the construction of the model term library enables a hybrid system identification procedure. Despite its flexibility in feature generation and selection, the SINDy is restricted to identifying linear-in-parameter system dynamics, and is unable to approximate certain complex nonlinear relationships, such as rational relationships (Kaheman et al., 2020).

In the meanwhile, neural networks have made substantial contributions to smart manufacturing and industrial 4.0. In conjunction with fog computing, a deep convolutional neural network was adapted to construct an inspection model to enhance factory productivity (Li et al., 2018). Moreover, a deep neural network (DNN)-based soft sensor was developed in (Villalba-Diez et al., 2019) to implement automatic quality control by comparing the scanned surface to the engraved file. This DNN-based soft sensor was able to provide an automated classification accuracy of 98.4%. As the universal approximator, the deep neural network can use its first few layers to perform feature engineering and then use the remaining layers for regression (Hornik, 1989). However, a deep neural network contains numerous parameters and will require a significant amount of training data. In addition, as a black-box model, the deep neural network is susceptible to overfitting and lacking of interpretability. In the proposed neural network and SINDy integrated algorithm, a sparse-connected, single-layer, feed-forward neural network with the input layer composed of library model terms is used to improve the model efficiency and accuracy. The proposed approach can be applied to general industrial systems, such as petroleum production, biochemical product manufacturing, and wastewater treatment, to automatically perform feature engineering for the various input variables and construct accurate nonlinear digital twin models to improve operational efficiency.

The remaining sections of this paper is organized as follows. Section 2 introduces the SINDy algorithm as well as its current limitations. This section also introduces the research objective. Section 3 introduces the three major steps of the proposed neural network and SINDy integrated algorithm. Afterward, Section 4 uses a diesel hydrotreating unit case study to demonstrate the procedure of applying the proposed algorithm to identify the multi-input multi-output (MIMO) digital twin model. Performance from the generalized SINDy (GSINDy) and a conventional single-layer, feed-forward neural network is used as benchmark. Finally, Section 5 concludes this paper.

2. RELATED WORK AND PROBLEM STATEMENT

Accurately identifying a process model while reducing the time consumption is important for the development and operation of an effective digital twin (Cimino et al., 2022). The SINDy (Brunton et al., 2016) initially established the framework of linear-in-parameter sparse identification to automatically identify parsimonious system governing equations. However, this algorithm focuses on identifying continuous nonlinear state dynamics,

$$\dot{\mathbf{x}}_t = f_x(\mathbf{x}_t), \tag{1}$$

where $\mathbf{x}$ represents the state and f_x indicates the continuous nonlinear state dynamics. In this relationship, the output is the derivative of the input, and the number of variables for input and output is equal. The continuous-time dynamics and restrictions on output variables limit SINDy's applications on the identification of general digital twin models.

In (Wang et al., 2022b), the authors extended SINDy to the GSINDy to identify general MIMO relationships between measurable system inputs and outputs,

$$\mathbf{y}_t = f_y(\mathbf{x}_t), \tag{2}$$

where $\mathbf{y}_t$ represents the general prediction objectives, which is no longer limited to a derivative, and f_y is the nonlinear relationship between system inputs, $\mathbf{x}_t$, and the outputs. The number of output variables can then differ from number of input variables. Following that, a linear-in-parameter sparse regression problem is solved,

$$\mathbf{Y} = \mathbf{\Theta}(\mathbf{X})\mathbf{\Xi}, \tag{3}$$

where $\mathbf{Y}$ is the output matrix, $\mathbf{\Theta}, (\mathbf{X})$ is the model term library, and $\mathbf{\Xi}$ is the sparse parameter matrix with most of its entries equaling zero to promote model sparsity. Usually, the sequential least squares (SLS) regression approach will be used to solve the sparse regression problem in SINDy-based algorithms. When implementing the SLS regression, a thresholding parameter λ is selected to determine the minimum parameters' magnitude of feature selection (Brunton et al., 2016). The regression parameters inside the sparse parameter matrix, $\mathbf{\Xi}$, whose magnitudes are smaller than λ will be forced to be zero, indicating the corresponding model terms are not selected for digital twin modelling.

Both SINDy and GSINDy have successfully identified governing equations for a variety of systems, including fluid dynamics, biology, and petroleum production (Champion et al., 2019; Mangan et al., 2016; Wang et al., 2022a,b). However, both of these methods are restricted to identifying linear-in-parameter relationships between prospective model terms and objective outputs. Consequently, when applied to identify complicated nonlinear relationships, such as rational relationships or implicit dynamics, their performance will decrease (Kaheman et al., 2020). Under these circumstances, the neural network, with its universal approximation capacity, are more applicable to identify the complex nonlinear digital twin models for an industrial process. In this research, we focus on integrating the feature engineering procedure from SINDy with the single-layer, feed-forward neural network to promote model simplicity and the digital twin identification accuracy.

3. THE NEURAL NETWORK AND SPARSE IDENTIFICATION OF NONLINEAR DYNAMICS INTEGRATED ALGORITHM

In this section, the neural network and SINDy integrated algorithm is introduced, which combines the SINDy's feature engineering with a single-layer, feed-forward neural network utilizing sparse connections among the three layers.

3.1 Data Collection

The initial stage of constructing a digital twin model is to collect input and output data. With the help of industrial internet of things (IoT), the input and output data will be collected as follows,

$$\mathbf{X} = [\mathbf{x}_{t,1}\ \mathbf{x}_{t,2}\ \dots\ \mathbf{x}_{t,m}]^T, \tag{4}$$

where $\mathbf{x}$ represents the multivariate input measurements and $\mathbf{x} \in \mathbb{R}^n$, and $t \in \mathbb{R}^m$ is the number of time instants of data collection. Similarly, the output measurements are recorded as a function of time,

$$\mathbf{Y} = [\mathbf{y}_{t,1}\ \mathbf{y}_{t,2}\ \dots\ \mathbf{y}_{t,m}]^T, \tag{5}$$

where $\mathbf{y}$ is the multivariate objective output and $\mathbf{y} \in \mathbb{R}^j$. In this analysis, $\mathbf{x}$ is assumed to be the easy-to-measure process variable, such as temperature and pressure, and $\mathbf{y}$ is the critical process variable, which usually hard-to-measure or available is a slow-rate.

3.2 Generation of a Model Term Library

Compare with general system identification, to improve the model's domain of applicability, first-principles information should be incorporated during the construction of a digital twin model. In the proposed algorithm, the critical step for combining both first-principles knowledge and data-driven techniques is to create an appropriate model term library. If the library is too complicated, even when a proper regularization approach is implemented, the resulting model can still overfit the data. If the model term library is too simple, the sparse nonlinear identification algorithm will not be able to identify the necessary model terms, resulting in inaccurate predictions.

Generally, when developing the model term library for digital twin identification, the library complexity should be gradually increased until the model's performance suddenly decreases. When first-principles information is available, the library construction can begin with involving only first-principles terms and then incorporating data-driven terms progressively. For instance, suppose that we would like to identify the Bernoulli's equation for pressure estimation. Fig.1 shows a graphical illustration of this problem. Use P to represent pressure, v to represent fluid velocity, and h to represent the location height. Then, assign $y = P_2$ and $\mathbf{x} = [P_1\ v_1\ v_2\ h_1\ h_2]$, where the subscripts 1 and 2 correspond to two ends of a flow system. Our objective is to determine the equation that can accurately predict P_2. According to first-principles knowledge about fluid dynamics, we identify, $\rho g h$ as a potential model term, where ρ is the fluid density, g is gravity, and h represents the relative height. Then, we can construct a data-driven and first-principles integrated model term library of the form,

$$\mathbf{\Theta}(\mathbf{X}) = \left[\mathbf{1}\ \mathbf{X}\ \mathbf{X}^{\mathrm{PO}_2}\ \ \rho g(h_2 - h_1)\right], \tag{6}$$

where $\mathbf{1}$ represents the bias term, $\mathbf{X}$ represents all the input variables, and $\mathbf{X}^{\mathrm{PO}_2}$ represents all the second-order polynomial combinations of the inputs. According to the first-principles knowledge, only h_1 and h_2 are used to generate the last model term, representing the hydrostatic pressure.

If no prior knowledge is available for a target process, we can construct a polynomial model term library and

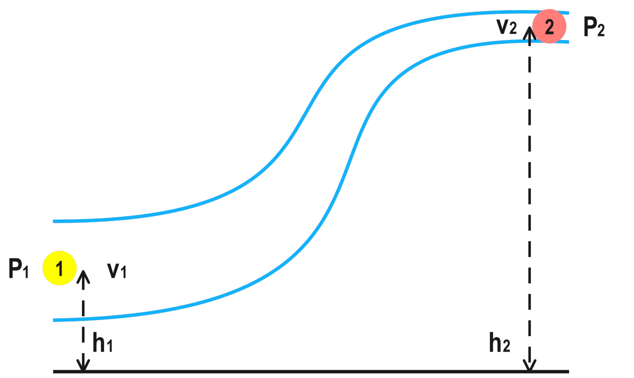

Fig. 1. Fluid dynamics example graphical illustration.

then gradually increase its complexity by including more data-driven terms. A sample data-driven, polynomial and trigonometric library is of the form,

$$\mathbf{\Theta}(\mathbf{X}) = \left[\mathbf{1}\ \mathbf{X}\ \mathbf{X}^{\mathrm{PO}_2} \dots \sin(\alpha\mathbf{X}) \dots \tanh(\beta\mathbf{X})\right]. \tag{7}$$

where α and β are scaling parameters to scale the input values within a data-driven term.

3.3 Integration of the SINDy-based Feature Engineering with the Neural Network

The neural network is capable of capturing more complex nonlinearities, such as rational relationships and implicit relationships, than linear-in-parameter identification approaches. By increasing the number of hidden layers, deep neural networks can perform feature engineering at the expense of increased model complexity and computational cost. To promote model simplicity and lower the computational cost, a single-layer, feed-forward neural network is utilized with the input layer comprised of all the terms from the SINDy-based library. Typically, the number of model terms within the model term library is sizeable. To perform feature selection and promote the model sparsity, sparse connections are utilized among the three layers within the single-layer, feed-forward neural network. The sparse connections can also help to save computational time costs and memory storage (Kepner et al., 2018; Mishra et al., 2021; Kepner et al., 2019).

The overall procedure of the proposed neural network and SINDy integrated algorithm is shown in Fig. 2. The initial step of applying the proposed algorithm to identify a digital twin model is to collect process data. After defining input and output variables, a comprehensive digital twin model term library is constructed by combining first-principles knowledge with data-driven techniques when feasible. Then, the model terms inside the library will constitute the input layer of the neural network. As the number of input variables increases, the size of the library would increase dramatically. Then, sparse connections are applied to perform feature selections. To implement the sparse connections among the three layers, a fully-connected neural network is first trained. Then, the weights whose magnitudes are less than a predetermined threshold will be reduced to zero. Then, the neural network is re-trained on the active links. As a consequence, only critical features will contribute to the digital twin identification. When tuning the threshold of the weight magnitudes, the threshold value is progressively increased until the prediction accuracy decreases sharply.

4. CASE STUDY

In Section 3, the proposed neural network and SINDy integrated digital twin identification algorithm is discussed in detail. In this section, we use the diesel hydrotreating (DHT) unit case study as an example to illustrate the algorithm implementation procedure and demonstrate the advantages of using it to identify the digital twin model.

The DHT unit is a critical component in the petroleum industry to ensure that the product fulfils the certification and commercialization standards (Garcia et al., 2014). Fig. 3 shows a graphical representation of the DHT unit. The overall inlet to the reactor consists of feed streams, recycling streams, and a hydrogen make-up stream. Prior to getting into the reactor, these streams are preheated through the furnaces. The primary diesel hydrotreating reaction happens in the presence of catalysts in the reactor. To regulate the reactor temperature, a portion of unheated hydrogen gas is introduced directly into the reactor as the quench gas. After the output of the reactor has been cooled, it enters the separator for a rough separation of light and heavy products. The light reaction products will enter an absorber for sulfur and ammonia removal. The heavy reaction products will go through the fractionation tower, producing light hydrocarbons, gasoline, jet, diesel, and heavy bottom products (Carelli and da Souza, 2009; Bandyopadhyay et al., 2019; Wang et al., 2022a).

Two data sets are available in this project, including real operational samples collected from real operations and first-principles samples that are calculated using the onsite first-principles software. After data preprocessing, 5713 real operational samples are used for training and 2448 real operational samples are used for testing. In the mean while, 142 first-principles samples are available for analysis. In total 37 input variables are available, including various feed streams' flowrates and densities, recycling stream's flowrate and density, reactor inlet streams' pressures and temperatures, etc. We have three major prediction objectives, including gasoline, diesel, and jet production rates, in the unit of barrels per hour (BPH). All the data have been normalized owing to proprietary reasons. The first-principles equations are unavailable during analysis and only the 142 samples from the first-principles software are accessible.

Compare with the real operational data set, the first-principles data set is less noisy and less complicated. Since the number of input variables is significant. If we put all the second-order polynomial combinations of inputs (666 combinations) to the overall model term library, the library will be too complex. In this case, we first apply the GSINDy algorithm to select the second-order features using the first-principles data set. Afterward, the selected second-order polynomial combinations will be directly involved in the real operational data model term library. From the first-principles data set, ten common second-order terms are identified to predict the objectives. The overall data-driven and first-principles knowledge combined model term library created for the real operational data is as follows,

$$\mathbf{\Theta}(\mathbf{X}) = \left[\mathbf{1} \ \mathbf{X} \ \tanh(0.8\mathbf{X}) \ \mathrm{e}^{(\mathbf{X})} \ \mathbf{FP}(\mathbf{X}^{\mathrm{PO}_2}) \ \frac{1}{\rho} \ \mathrm{e}^{-\frac{\mathrm{E_a}}{\mathrm{RT}}} \right], \tag{8}$$

The trigonometrical term, $\tanh(0.8\mathbf{X})$, and the first exponential term, $e^{(\mathbf{X})}$, are included to increase the nonlinearity of the library, and each of these two terms contains 37 components, one for each input variable. Next, $\mathbf{FP}(\mathbf{X}^{\mathrm{PO}_2})$ represents the ten second-order polynomial components selected from the first-principles data set. The last two terms are generated from diesel hydrotreating first-principles knowledge, where ρ and T are process variables and $\mathrm{E_a}$ and R are constants. Specifically, ρ is the fresh feed stream density in the unit of $\mathrm{kg/m^3}$; T is the reactor temperature in Kelvin; $\mathrm{E_a}$ is the reaction activation energy and equals 21.4 KJ/mol in this case study; R is the gas-law constant and equals 8.314 J/molK. In total, the model term library contains 124 model term components, with 122 components from the data-driven creation and 2 components form the first-principles knowledge. Subsequently, these 124 model terms are used to form the input layer of the single-layer, feed-forward neural network.

The comparative evaluation uses the performance of GSINDy and a fully-connected, single-layer, feed-forward neural network without feature engineering as benchmarks, and the results are presented in Table 1 in terms of mean squared error (MSE). The same model term library is used for both GSINDy and the proposed algorithm. Since the fully-connected, single-layer, feed-forward neural network does not have access to the model term library, its input layer consists only of the 37 individual input variables. After cross-validation, the structure of 37 - 2 - 3 with L_1 regularizers among the three layers, provided optimal performance for the conventional single-layer neural network. The neural network employed in the proposed approach has 124 terms in the input layer. Then, a structure of 124 - 4 - 3 provided its optimal performance. For the neural network parameter optimizations, ADAM optimizer is used in this study. To promote model simplicity and further prevent overfitting, 80% of weights among the three layers are set to zero, leading to sparse connections. Between the input layer and the hidden layer, the linear activation function is used, while the *tanh* activation function is used between the hidden layer and the output layer.

Table 1. Performance comparison among the GSINDy, the conventional single-layer neural network, and the proposed algorithm in terms of MSE.

	Output yields (BPH)		
Methods	**Gasoline**	**Diesel**	**Jet**
GSINDy	0.096	0.237	0.0839
Conventional single-layer neural network	0.068	0.2410	0.110
Proposed algorithm	**0.058**	**0.202**	**0.068**

According to the numerical comparison from the three methods, the conventional single-layer, feed-forward neural network has smaller prediction error in gasoline production, while has inferior prediction accuracy in diesel and jet yields than the GSINDy. This result indicates that even though a neural network better captures the system nonlinearity than the linear-in-parameter sparse

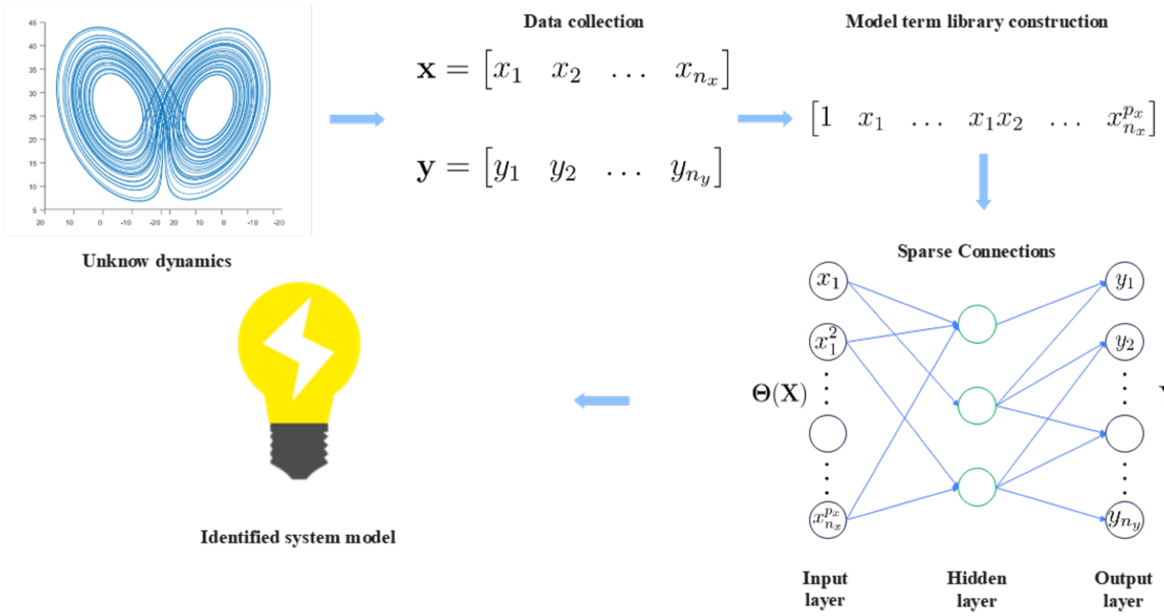

Fig. 2. Graphical illustration of the neural network and SINDy integrated algorithm.

identification, lack of feature engineering reduces the prediction accuracy. By integrating the feature engineering from SINDy with the sparse-connected, single-layer, feed-forward neural network, the proposed algorithm achieves the lowest MSE for all the three predictions, and as a result, provides a more accurate digital twin model.

5. CONCLUSION

In this study, we proposed a neural network and SINDy integrated algorithm for nonlinear digital twin identification combining both first-principles knowledge and data-driven techniques. The feature engineering procedure from the SINDy is utilized to develop a model term library for digital twin identification. In the proposed algorithm, the input layer of the neural network comprised of all the model terms from the library. To ensure only critical model terms are selected to contribute to the digital twin identification, sparse connections are applied among the three layers. The advantages of the proposed approach are demonstrated through a DHT unit case study, with 37 input variables and three major output variables. Compare to the GSINDy approach, the proposed algorithm is applicable to identify digital twin models for more complicated nonlinear systems. Similarly, compare to conventional single-layer, feed-forward neural network, the proposed approach automatically performs feature engineering and considers the first-principles information to improve the models' domain of applicability. As a consequence, the proposed neural network and SINDy integrated digital twin identification approach automatically performs feature engineering for a nonlinear industrial system by creating a model term library considering both first-principles information and data-driven techniques. In addition, the proposed algorithm can not only identify the linear-in-parameter digital twin models but also more complicated nonlinear models, such as implicit relationships. Besides the DHT unit, the proposed approach can be applied to identify digital twin models for general industrial systems, such as those in pulp and paper, automobile manufacturing, mining, as well as mineral and metal processing. In the future, the proposed technique can be further extended to involve model predictive control to achieve automatic identification and control simultaneously.

REFERENCES

Bandyopadhyay, R., Alkilde, O.F., Menjon, I., Meyland, L.H., and Sahlertz, I.V. (2019). Statistical analysis of variation of economic parameters affecting different configurations of diesel hydrotreating unit. *Energy*, 183, 702–715. doi:10.1016/j.energy.2019.06.156.

Brettel, M., Friederichsen, N., Keller, M., and Rosenberg, M. (2014). How virtualization, decentralization and network building change the manufacturing landscape: An industry 4.0 perspective. *International journal of information and communication engineering*, 8(1), 37–44.

Brunton, S.L., Proctor, J.L., and Kutz, J.N. (2016). Discovering governing equations from data by sparse identification of nonlinear dynamical systems. *Proceedings of the National Academy of Sciences*, 113(15), 3932–3937. doi:10.1073/pnas.1517384113. Doi.org/10.1073/pnas.1517384113.

Carelli, A. and da Souza, M. (2009). GPC controller performance monitoring and diagnosis applied to a diesel hydrotreating reactor. *IFAC Proceedings Volumes*, 42(11), 976–981. doi:10.3182/20090712-4-tr-2008.00160.

Champion, K.P., Brunton, S.L., and Kutz, J.N. (2019). Discovery of nonlinear multiscale systems: Sampling strategies and embeddings. *SIAM Journal on Applied Dynamical Systems*, 18(1), 312–333. doi:10.1137/18m1188227.

Cimino, C., Granata, C., Ferretti, G., and Leva, A. (2022). On consistency of physical and DEVS models in control-targeted DTs: an industrial case study. *IFAC-*

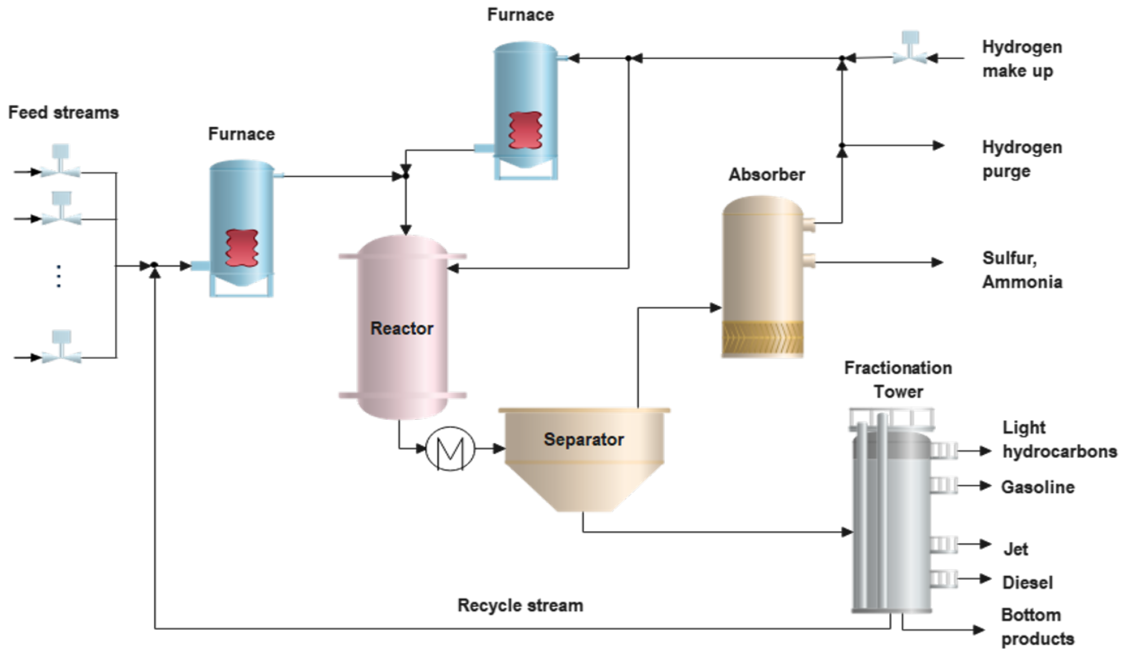

Fig. 3. Graphical illustration of the DHT unit process.

PapersOnLine, 55(2), 7–12. doi:10.1016/j.ifacol.2022.04.161.

Garcia, M.R., Pitta, R.N., Fischer, G.G., and Neto, E.R. (2014). Optimizing diesel production using advanced process control and dynamic simulation. *IFAC Proceedings Volumes*, 47(3), 358–363. doi:10.3182/20140824-6-za-1003.00542.

Hornik, K., S.M..W.H. (1989). Multilayer feedforward networks are universal approximators. *Neural Networks*, 2, 359–366.

Kaheman, K., Brunton, S.L., and Kutz, J.N. (2020). Automatic differentiation to simultaneously identify nonlinear dynamics and extract noise probability distributions from data. doi:10.48550/ARXIV.2009.08810.

Kepner, J., Alford, S., Gadepally, V., Jones, M., Milechin, L., Robinett, R., and Samsi, S. (2019). Sparse deep neural network graph challenge. In *2019 IEEE High Performance Extreme Computing Conference (HPEC)*. IEEE. doi:10.1109/hpec.2019.8916336.

Kepner, J., Gadepally, V., Jananthan, H., Milechin, L., and Samsi, S. (2018). Sparse deep neural network exact solutions. In *2018 IEEE High Performance extreme Computing Conference (HPEC)*. IEEE. doi:10.1109/hpec.2018.8547742.

Li, L., Ota, K., and Dong, M. (2018). Deep learning for smart industry: Efficient manufacture inspection system with fog computing. *IEEE Transactions on Industrial Informatics*, 14(10), 4665–4673. doi:10.1109/tii.2018.2842821.

Liu, Q., Zhang, H., Leng, J., and Chen, X. (2018). Digital twin-driven rapid individualised designing of automated flow-shop manufacturing system. *International Journal of Production Research*, 57(12), 3903–3919. doi:10.1080/00207543.2018.1471243.

Mangan, N.M., Brunton, S.L., Proctor, J.L., and Kutz, J.N. (2016). Inferring biological networks by sparse identification of nonlinear dynamics. *IEEE Transactions on Molecular, Biological and Multi-Scale Communications*, 2(1), 52–63. doi:10.1109/tmbmc.2016.2633265.

Min, Q., Lu, Y., Liu, Z., Su, C., and Wang, B. (2019). Machine learning based digital twin framework for production optimization in petrochemical industry. *International Journal of Information Management*, 49, 502–519. doi:10.1016/j.ijinfomgt.2019.05.020.

Mishra, A., Latorre, J.A., Pool, J., Stosic, D., Stosic, D., Venkatesh, G., Yu, C., and Micikevicius, P. (2021). Accelerating sparse deep neural networks. doi:10.48550/ARXIV.2104.08378.

Park, K.T., Nam, Y.W., Lee, H.S., Im, S.J., Noh, S.D., Son, J.Y., and Kim, H. (2019). Design and implementation of a digital twin application for a connected micro smart factory. *International Journal of Computer Integrated Manufacturing*, 32(6), 596–614. doi:10.1080/0951192x.2019.1599439.

Shah, D., Wang, J., and He, Q.P. (2020). Feature engineering in big data analytics for IoT-enabled smart manufacturing – comparison between deep learning and statistical learning. *Computers &amp Chemical Engineering*, 141, 106970. doi:10.1016/j.compchemeng.2020.106970.

Villalba-Diez, J., Schmidt, D., Gevers, R., Ordieres-Meré, J., Buchwitz, M., and Wellbrock, W. (2019). Deep learning for industrial computer vision quality control in the printing industry 4.0. *Sensors*, 19(18), 3987. doi:10.3390/s19183987.

Wang, J., Moreira, J., Cao, Y., and Gopaluni, B. (2022a). Simultaneous digital twin modeling and signal-noise decomposition through modified generalized sparse identification of nonlinear dynamics.

Wang, J., Moreira, J., Cao, Y., and Gopaluni, B. (2022b). Time-variant digital twin modeling through the kalman-generalized sparse identification of nonlinear dynamics. In *2022 American Control Conference (ACC)*. IEEE. doi:10.23919/acc53348.2022.9867786.