Systematic Evaluation of Deep Neural Network based Dynamic Modeling Method for AC Power Electronic System

Yunlu Li, Guiqing Ma, Junyou Yang, *Member, IEEE*, and Yan Xu

***Abstract*—Since the high penetration of renewable energy complicates the dynamic characteristics of the AC power electronic system (ACPES), it is essential to establish an accurate dynamic model to obtain its dynamic behavior for ensure the safe and stable operation of the system. However, due to the no or limited internal control details, the state-space modeling method cannot be realized. It leads to the ACPES system becoming a black-box dynamic system. The dynamic modeling method based on deep neural network can simulate the dynamic behavior using port data without obtaining internal control details. However, deep neural network modeling methods are rarely systematically evaluated. In practice, the construction of neural network faces the selection of massive data and various network structure parameters. However, different sample distributions make the trained network performance quite different. Different network structure hyperparameters also mean different convergence time. Due to the lack of systematic evaluation and targeted suggestions, neural network modeling with high precision and high training speed cannot be realized quickly and conveniently in practical engineering applications. To fill this gap, this paper systematically evaluates the deep neural network from sample distribution and structural hyperparameter selection. The influence on modeling accuracy is analyzed in detail, then some modeling suggestions are presented. Simulation results under multiple operating points verify the effectiveness of the proposed method.**

***Index Terms*—AC power electronic system, Dynamic behavior, Neural network, Systematic evaluation.**

I. Introduction

AC power electronic systems (ACPESs) are widely applied in power grids, driven by the rapid development of renewable energy generation, electric vehicles, and energy storage systems (ESS). The dynamic behaviors of converter-based power system are complicated, which are affected by the fluctuation of renewable energy and the dynamic interaction of these ac power electronic systems [1]-[3]. Moreover, the integration of large-scale distributed energy may reduce the inertia of the system and make the system vulnerable to interference [4]-[6]. The dynamic behaviors of ACPES are more complex and prone to stability issues. Therefore, it is quite significant to establish the dynamic model to master its dynamic characteristics.

Modeling methods based on state-space modeling method are still the most commonly used methods [7]-[9], such as small signal modeling method, impedance modeling method, etc. To ensure the dynamic behavior of ACPES, the state-space modeling method needs to establish the linear time-invariant (LTI) system in synchronous coordinates, then the system is linearized near equilibrium point to obtain the LTI model [10]. Finally, the system stability is revealed according to the eigenvalues of state matrix. For this modeling method, detailed equipment parameters are indispensable [11]-[13]. However, in practice, it is almost impossible to obtain all the detailed parameters of equipment due to equipment aging, property rights protection and other factors [14]-[15]. Therefore, ACPES becomes a black-box dynamic system. Although the impedance modeling method can establish the impedance model by frequency scanning method, it is effective only at a single operating point. However, the dynamic behavior of ACPES may change significantly at different operating points [16], due to the nonlinear parts in the ACPES such as Phase Lock Loop (PLL) [17]-[18], Park transformation [19], power control loop [20], etc. Therefore, the establishment of a large-scale black box model is the key to ensuring the safe and stable operation of ACPES.

With the rapid development of deep learning in recent years, it provides new solutions for many research fields. In [21], a fault diagnosis strategy based on long short-term memory (LSTM) is proposed. Recursive neural network (RNN) is used to predict power fluctuation in [22]-[23]. The modeling method based on artificial neural network (ANN) is a powerful nonlinear modeling method, which can theoretically fit any complex nonlinear system. In [24]-[25] RNN and LSTM are used to obtain the dynamic model of grid-tied ACPES. In [26], LSTM is used to establish the dynamic equivalent model of active distribution networks. In the field of power systems, deep learning is emerging in more and more solutions.

However, these modeling methods are simple to describe the modeling process. There is no deep neural network modeling guidance for power system and detailed evaluation of the performance under different structural hyperparameters and

Manuscript received July 16, 2022; revised October 09, 2022; accepted October 31, 2022. Date of publication June 25, 2023; Date of current version January 11, 2023.

This work was supported in part by the Science Search Foundation of Liaoning Educational Department. *(Corresponding Author: Junyou Yang)*

Yunlu Li, Guiqing Ma, Junyou Yang and Yan Xu are with the School of Electrical Engineering, Shenyang University of Technology, 110870, China.(e-mail of corresponding author: junyouyang@sut.edu.cn).

Digital Object Identifier 10.30941/CESTEMS.2023.00011

sample distribution. The construction of network models often faces the problems of data screening and network structure determination. Data is the basis of model training. In practice, it is always faced with massive data, different sample distributions lead to extremely different performance after training. Moreover, different network structures also represent different convergence time. These all lead to the inability to quickly and conveniently implemented neural network modeling with high precision and high training speed in practical engineering applications. To fill this gap, this paper evaluated the performance from two aspects of sample distribution and structural hyperparameter selection. The influence of sample distribution and hyperparameter selection on network performance is analyzed in detail, which provides guidance for deep neural network modeling in power systems. The key contributions of this paper are as follows

(1) It is the first time for the neural network modeling in power system is systematically and comprehensively evaluated. The performances of different sample distributions and different structural hyperparameters are described in detail. Design guidelines for general neural network model establishment are also provided.

(2) Discussion and explanation about the reason of LSTM neural as one state variable are originally conducted in detail. Recommendations for the selection of sample distribution and structural hyperparameters in practical engineering applications are presented.

The remainder of this paper is organized as follows. The framework of dynamic modeling based on neural network is presented in Section II. In Section III, then the impact on accuracy of this framework is evaluated in detail from sample distribution and hyperparameter selection. In Section IV, according to the evaluation results, the hyperparameter setting and sample distribution recommended under the framework are provided.

II. Framework of Dynamic Modeling by Using Deep Network

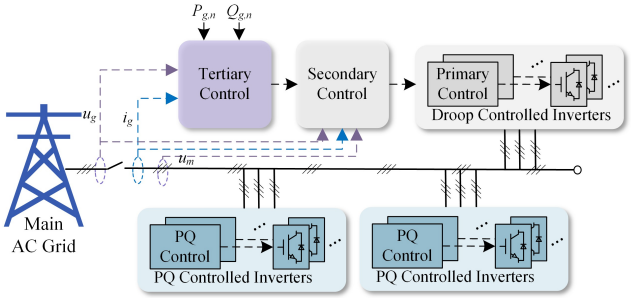


Fig. 1. Typical AC power electronic system diagram.

Fig. 1 shows the typical grid-connected AC power electronic system diagram. Different inverters have different operating points, which may vary with the power distribution plan. This operating point dependent property means that the small signal method of a single operating point cannot express its dynamic behaviors. In addition, all the required detailed modeling parameters cannot be obtained in practice. The modeling method based on deep neural network can fit the dynamic characteristics according to the input and output data of the system. The modeling framework is shown in Fig. 2 This framework shows the procedure from the data preparation to the desired neural network.

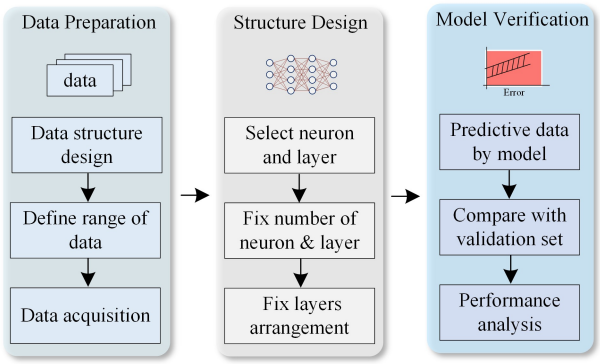


Fig. 2. General framework of the dynamic modeling method based on deep neural network.

From the above analysis, it can be seen that ACPES system in practice is considered as a black-box model so that its output dynamic behavior is only affected by the input. According to the structure shown in Fig. 1, the dynamic behavior of i_g is only affected by grid voltage ug. Therefore, the input data vector is defined as $x(t)= [P_g(t), Q_g(t), u_\alpha(t), u_\beta(t)]$, and the output is defined as $y(t)= [i_\alpha(t), i_\beta(t)]$.

Data preparation is the basis of neural network dynamic modeling. Reasonable sample distribution selection can enable the neural network to learn enough details and make the model achieve satisfactory accuracy. Before collecting data, the first thing to be determined is the sampling rate. Higher sampling rate means more transient details but also brings greater computational burden and sampling cost. According to [27], the sampling rate of kHz is sufficient for millisecond level transient analysis. Then, each inverter integrated in the power system is measured by PMU at different operating points to obtain data sets under multiple operating conditions.

Network structure determination is the most significant step. Different hyperparameter combinations directly affect the final training result. The neuron type selection is the first consideration. Firstly, the selection of neuron type depends on the type of modeling task. Dynamic modeling of ACPES is a regression task, which generates model from measured data. There are two types of neurons often used to deal with regression tasks: RNN and LSTM. Due to the special design of the LSTM hidden layer, it avoids the problems of gradient disappearance and gradient explosion at long time scales. Moreover, the forgetting gates enable neurons to compare input data to avoid recent non-important information occupying large storage space. Therefore, LSTM neurons are selected for modeling in this paper. The formulations of a LSTM neuron are as follows.

$$f_g(t)=\sigma(W_f[x(t),o(t-1)]+b_f) \quad (1)$$

$$i_g(t)=\sigma(W_f[x(t),o(t-1)]+b_i) \quad (2)$$

$$c_g(t)=\tanh(W_c[x(t),o(t-1)]+b_c) \quad (3)$$

$$o_g(t)=\sigma(W_o[x(t),o(t-1)]+b_o) \quad (4)$$

$$s(t)=f_g \cdot s(t-1)+i_g(t)\cdot c_g(t) \quad (5)$$

$$o(t)=\tanh(s(t))\cdot o_g(t) \quad (6)$$

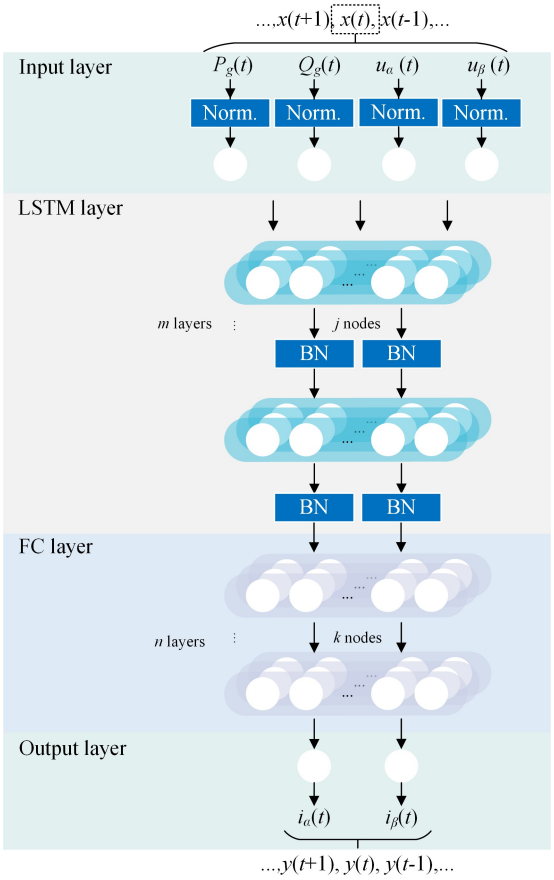


Fig. 3. The proposed generalized neural network.

In addition, adding normalization calculation in the input layer is an effective means to accelerate the network convergence, which can transfer the input data to 0~1 according to the maximum and minimum values of each feature sequence in $x(t)$. Based on the above consideration, the network structure is proposed in Fig. 3. As shown in Fig. 3, the proposed network structure contains an input layer, a LSTM layer, a FC layer and an output layer. The vector is input to network after normalized. In the LSTM layer, each neuron is a processing unit, which iterates the cell state according to the network input value $x(t)$ at the current moment, the previous LSTM output value $h(t\text{-}1)$ and the previous unit state $c(t\text{-}1)$.The result is next processed through the FC layers,which are used to build nonlinear algebraic equations and integrate with differential equations. Finally, the processed vector is output through the output layer.

III. Performance Evaluation

A. Configuration Setup

In order to verify the influence of sample distribution and structural hyperparameters on accuracy, the power system is studied as shown in Fig. 4. The whole system is connected to the grid by two microgrids. The nominal frequency amplitude distribution is 50Hz and 311V. Network and loads parameters are shown in Table I. The cable length of line AB is 4km. The impedance parameters of all lines are with R_{line}=0.16Ω/km and L_{line}=0.26mH/km. The detailed parameters of the inverters are shown in Table II. Training data and test data are collected on bus A. Sampling rate is set to 2kHz to simulate PMU.

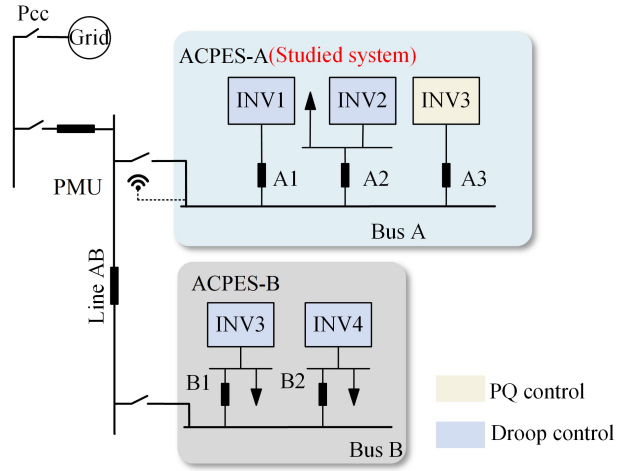


Fig. 4. The framework of testing system.

TABLE I
PARAMETERS OF NETWORK AND LOADS

Bus	Load		Line	Bus	Load		Line
	R(Ω)	L(mH)	(km)		R(Ω)	L(mH)	(km)
A1	29	—	1.2	B1	14.5	231	0.3
A2	14.5	231	0.3	B2	72.6	—	1.6
A3	42	—	0.8	—	—	—	—

TABLE II
PARAMETERS OF INVS

Droop	INV1	INV2	INV4	INV5	PQ	INV3
R_f(Ω)	0.2	0.3	0.2	0.2	R_f(Ω)	0.2
L_f(mH)	1	0.8	1.2	1.2	L_f(mH)	2
L_g(mH)	1.8	1.8	1.4	1.4	k_{cp}	30
C_f(mF)	25e-3	25e-3	0.7	0.7	k_{ci}	10000
R_g(Ω)	0.2	0.2	0.08	0.08	P(W)	2000
k_{vp}	0.07	0.05	1.2	1.2	Q(W)	0
k_{vi}	100	90	200	200	-	-
k_{cp}	1	1	15	15	-	-
k_{ci}	30	30	1000	1000	-	-
m_p(rad/W)	2e-4	2e-4	1e-5	1e-5	-	-
m_q(V/Var)	4e-4	4e-4	1e-4	1e-4	-	-

B. Impact Analysis of Sample Distribution

TABLE III
LIST OF SAMPLE DISTRIBUTION TESTED

Type of samples	Steady-state	Transient	Fault
A	100	0	0
B	90	10	0
C	70	30	0
D	88	10	2
E	85	10	5

This section discusses the impact of sample distribution on accuracy. To evaluate the impact of sample distribution in isolation, the neural network structure is determined. Based on the Fig. 3, the neuron type is selected as LSTM. The overall structure consists of two LSTM layers and two FC layers (m=2, n=2). Each LSTM layers set 64 neurons and each FC layers set 64 neurons (k=64, j=64). The sampling rate is set to 2kHz and the sampling time is 5s. 100 groups of sample data were collected at random operating points, 75% of the data used for

model training and the rest as performance test. The data structure under different sample distributions is shown in Table III. After 3000 iterations, training stops uniformly. In order to facilitate comparison, only the change of α-axis current in $\alpha\beta$-frame is considered.

In practice, steady-state data at different operating points is the most and easiest data to obtain. In first scenario, the fitting performance of type A samples after 3000 iterations as shown in Fig. 5. P_{ref} changed from 5kW to 3kW in 3s and Q_{ref} remained at 3kW. There are obvious errors in i_α in the Fig. 5. The peak error of i_α reaches 23.5%.

Since the data under steady-state cannot reflect the dynamic response characteristics of the system, the high-precision fitting performance cannot be achieved within 3000 iterations only depending on the data under steady state.

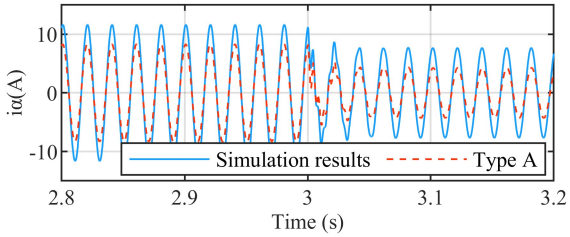


Fig. 5. The transient response of i_α under type A samples.

During the operation of the ACPES, its operating point always drifts according to the demand. The system operating point change process includes the dynamic response characteristics. Therefore, in scenario 2, type B samples replace 10 samples in type A with data samples under operating point changes. The fitting performance is shown in Fig. 6. With the addition of transient samples, the fitting accuracy is significantly improved. the peak error of i_α drops to 11.7%.

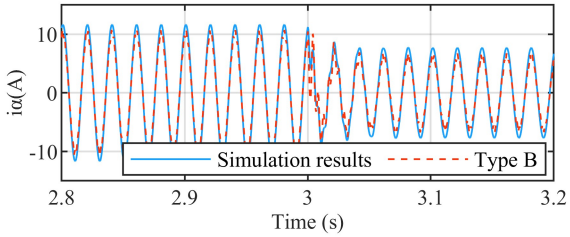


Fig. 6. The transient response of i_α under type B samples.

As the proportion of transient samples increases, the fitting accuracy becomes higher and higher. It achieves a high level of accuracy under type C samples. The peak error of i_α error drops below 4%, as shown in Fig. 7. Fig. 8 shows the estimated transient behavior of two types of samples. During the P_{ref} changing process, the type C samples can provide a more accurate transient behavior than type B samples, which fluctuation converges within two cycles. The dynamic response contained in the transient samples enables the neural network to obtain a large amount of information in a short period. Adding transient samples to the dataset can significantly improve model accuracy.

External short-circuit fault can seriously affect the system dynamic behavior, although it rarely occurs, all the dynamic characteristics of the system under fault are fully exposed. Therefore, the data under fault are added to study.

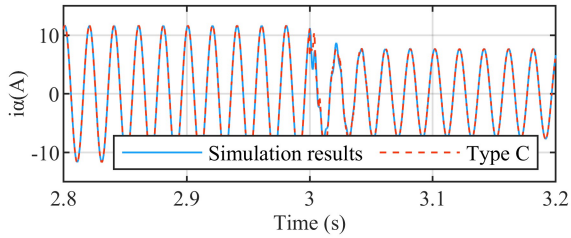


Fig. 7. The transient response of i_α under type C samples.

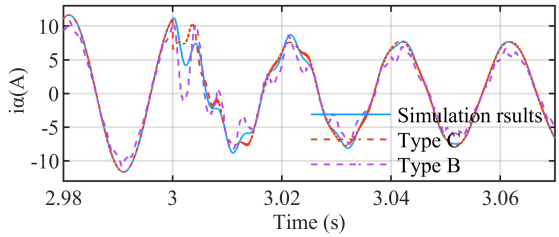
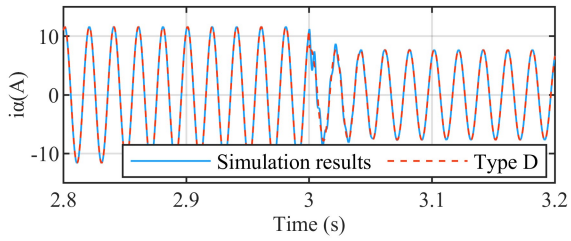


Fig. 8. The comparison of type B and type C when P_{ref} changes from 5kW to 3kW at 3s.

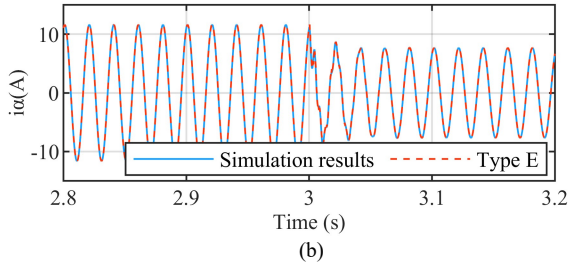
In scenario 3, as shown in Fig. 9(a), the fitting performance of type D samples set is verified under operating point changing conditions. The peak error is 3.2% for i_α. Compared with the type C samples, it can be seen that the performance of adding two fault samples reaches the accuracy of 20 transient samples. Even a small number of fault samples can greatly improve the accuracy. With the addition of type E samples, the accuracy reaches a very high level. The peak error is 0.06% for i_α as shown in Fig. 9(b). From the comparison in Fig. 9(c), it can be seen that the samples under fault are significant helpful to the transient response, the accuracy of E-type samples achieved far more accuracy than others. By adding the data samples under external short-circuit fault, the model better fits the dynamic behavior of the system.

C. *Impact Analysis of Structural Hyperparameter*

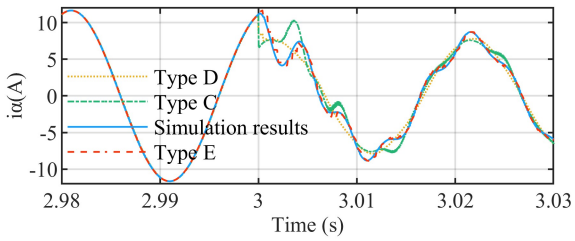
Hyperparameter determination is also a crucial step. Different hyperparameters can lead to huge differences in network performance. In order to verify the influence of hyperparameters on network performance, this section studies under the fixed samples distribution, according to the analysis



(a)



(b)



(c)

Fig. 9. The transient response at terminal when P_{ref} changes from 5kW to 3kW at 3s.(a) Transient response of i_α under type D samples. (b) Transient response of i_α under type E samples. (c) The comparison among type C, type D and type E.

of samples type in the above section, it is determined as type E samples. Similarly, the training stopped after 3000 iterations.

Generally, after determining the network framework, the width and depth of the network need to be tested repeatedly. It can be seen from Eq. (1) to Eq. (6) that the current output of LSTM neurons is also affected by the previous state variable, so it can also be regarded as a discrete dynamic system, which can be regarded as DAEs. Therefore, the calculation process of LSTM can be expressed as state space model as.

$$h(t)=F(h(t-1),x(t),W_{f,i,c,o},b_{f,i,c,o}) \tag{7}$$

$$o(t)=G(h(t),x(t),W_{f,i,c,o},b_{f,i,c,o}) \tag{8}$$

Where $h(t)=[o(t), s(t)]$ is considered as the state variables. From the mathematical level, the whole ACPES system is a set of DAEs, which can be described as.

$$\dot{x}_{acpes}=f(x_{acpes},u,p) \tag{9}$$

$$y=g(x_{acpes},u,p) \tag{10}$$

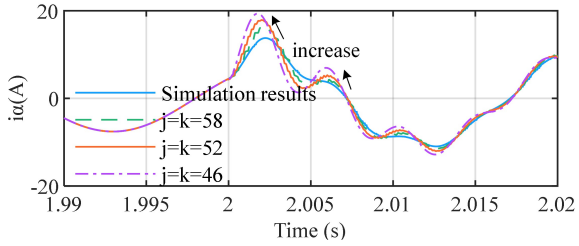
Where x_{acpes} is considered as the state vector, u, p and y represent the input and output and the parameters in the system, respectively. Comparing Eq. (7) to Eq. (8) and Eq. (9) to Eq. (10), it can be seen that theoretically a neuron can represent a system state variable.

TABLE IV
LIST OF HYPERPARAMETER TESTED

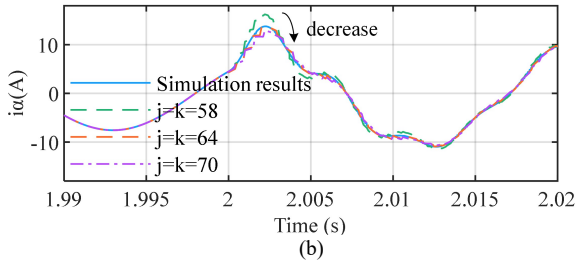
net (j, k)	Network depth (m, n)
j=k=46	m=n=2
j=k=52	m=n=2
j=k=58	m=n=2
j=k=64	m=n=2
j=k=70	m=n=2

For the studied system, there are a total of 58 system variables, so we preliminarily set the network width to 58 and verify it nearby. The depth of the network means the complexity of the fitting function. The more layers, the more mapping of the original data, and the deeper information can be obtained. However, more layers also mean heavier training burden and more information attenuation. Therefore, the number of network layers is generally selected based on experience or repeated tests. The number of network layers is selected as 4 layers by repeated testing. Based on Fig. 3, the detailed parameters of the test are shown in Table IV.

A comparative study is implemented among the different network width, as shown in Fig. 10, P_{ref} changed from 2kW to 4kW in 2s and Q_{ref} remained at 3kW. It can be seen from Fig. 10 that the error increases with the decrease of network width. With the decrease of network width, the coupling between neurons decreases and the fitting ability decreases. In contrast, with the increase in the number of neurons, the fitting ability initially increased, then decreased as the network width continued to increase. This is because too many neurons increase the difficulty of network training, resulting in a longer convergence time. The comparison shows that the best performance is j=k=64, which the network width is slightly larger than the number of system state variables. Although each neuron can represent a state variable in theory, the space of the neuron in the training process is occupied by some other information, so the selection of network width should be slightly larger than the number of state variables.



(a)



(b)

Fig. 10. The transient response at terminal when P_{ref} changes from 2kW to 4kW at 2s. (a) The comparison of network width among 46, 52 and 58. (b) The comparison of network width among 58, 64 and 70.

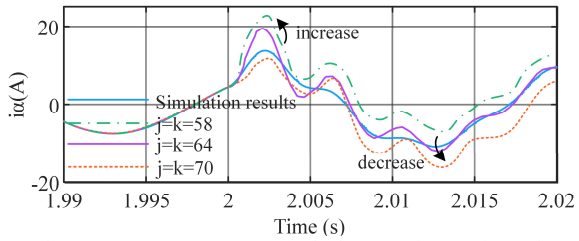


Fig. 11. The transient response at terminal when P_{ref} changes from 2kW to 4kW at 2s by using ANN.

A comparative study is implemented between the proposed LSTM-based method and ANN-based model [28]. All these methods are developed for black-box nonlinear modeling. In order to fairly compare the performances, the ANN is constructed with same structure as the proposed model. Fig. 11 depicts the ANN network width among 58, 64 and 70. It can be seen that the best performance is j=k=64.

By comparing Fig. 10 to 11, it can be seen that the accuracy of the proposed model is much higher than the ANN-based model even when the worst performance (j=k=46). When j=k=46, the error peak of the proposed model is 2.5A and rapidly converges to 0.1A after 0.1s. In contrast, the error peak of the ANN-based model is 5.2A under the best performance (j=k=64), and it still remains around 2A after 0.1s. The comparison results demonstrate that although the ANN-based

model can also predict the transient behavior under transient response, its accuracy still lags far behind the proposed LSTM-based model even in the best performance of j=k=64.

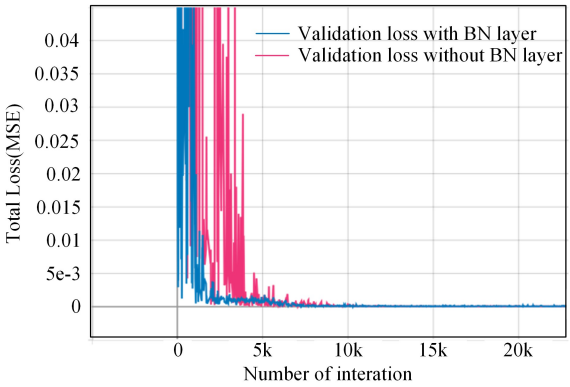


Fig. 12. Total loss between with and without BN layers.

In addition, adding batch normalization (BN) calculation in the hidden layer can reduce the saturation effect of activation function and accelerate the network convergence, which is ignored by many people. The performance with and without BN layers is tested under the optimal network width. It can be seen from Fig. 11 that the training speed was significantly improved after adding BN layer.

D. Modeling Suggestions

From the evaluation of the previous section, it can be seen that the modeling performance relying solely on steady-state samples is inaccurate. Adding additional transient and fault samples can significantly improve the accuracy of the model, especially the fault samples. On the other hand, the network width seriously affects the training efficiency. The best efficiency is achieved when the number of neurons is slightly larger than the system state variables. In addition, the addition of BN layer also significantly improves the training efficiency. Therefore, the suggestions are as follows.

(1) In data acquisition, as far as possible to collect transient samples, especially under fault.

(2) The network width should be slightly larger than the number of system state variables. In addition, the addition of BN layer is necessary to improve training speed.

IV. Conclusion

In this paper, the dynamic modeling method based on neural network is systematically verified from the sample distribution and structural hyperparameter selection level. The modeling framework and design guidelines are presented in detail. Suggestions for the selection of sample distribution and hyperparameter selection are given by simulation. Firstly, transient samples are collected as far as possible in data collection, especially those under fault. Secondly, the network width should be slightly larger than the number of system state variables. Finally, adding BN layers can significantly improve the training speed.

References

[1] T. Wang, C. Li, D. Mi, Z. Wang, and Y. Xiang, "Coordinated modulation strategy considering multi-HVDC emergency for enhancing transient stability of hybrid AC/DC power systems," *CSEE Journal of Power and Energy Systems*, vol. 6, no. 4, pp. 806-815, Dec. 2020.

[2] B. Sahoo, S. Keshari Routray, P. Kumar Rout and M. M. Alhaider, "Neural network and fuzzy control based 11-level cascaded inverter operation," *Computers, Materials & Continua*, vol. 70, no.2, pp. 2319–2346, 2022.

[3] J. Sun, "AC power electronic systems: Stability and power quality," in *Proc. WCMPE*, Zurich, Switzerland, pp. 1093-5142, 2008.

[4] Z. Shuai, Y. Sun, Z. J. Shen, W. Tian, Y. Li and X. Yin, "Microgrid stability: classification and a review, " *Renewable and Sustainable Energy Reviews*, vol. 58, pp. 167-179, 2016.

[5] Z. Zhao, P. Yang, Y. Wang, Z. Xu and J. M. Guerrero, "Dynamic characteristics analysis and stabilization of PV-based multiple microgrid clusters," *IEEE Transactions on Smart Grid*, vol. 10, no. 1, pp. 805-818, 2019.

[6] P. Vorobev, P. Huang, M. Al Hosani, J. L. Kirtley and K. Turitsyn, "High-Fidelity Model Order Reduction for Microgrids Stability Assessment," *IEEE Transactions on Power Systems*, vol. 33, no. 1, pp. 874-887, 2018.

[7] Z. Shuai, C. Shen, X. Yin, X. Liu and Z. J. Shen, "Fault analysis of inverter-interfaced distributed generators with different control schemes," *IEEE Transactions on Power Delivery*, vol. 33, no. 3, pp. 1223-1235, 2018.

[8] Z. Shuai, Y. Peng, X. Liu, Z. Li, J. M. Guerrero and Z. J. Shen, "Dynamic equivalent modeling for multi-microgrid based on structure preservation method," *IEEE Transactions on Smart Grid*, vol. 10, no. 4, pp. 3929-3942, 2019.

[9] H. Alrajhi, "A generalized state space average model for parallel dc-to-dc converters," *Computer Systems Science and Engineering*, vol. 41, no. 2, pp. 717–734, 2022.

[10] Y. Liao and X. Wang, "Small-signal modeling of AC power electronic systems: critical review and unified modeling," *IEEE Open Journal of Power Electronics*, vol. 2, pp. 424-439, 2021.

[11] S. Sen, V. Kumar, "Microgrid modeling: a comprehensive survey," *Annual Reviews in Control*, vol.46, pp. 216-250, 2018.

[12] M. Zhang, X. Wang, D. Yang and M. G. Christensen, "Artificial Neural Network Based Identification of Multi-Operating-Point Impedance Model," *IEEE Transactions on Power Electronics*, vol. 36, no. 2, pp. 1231-1235, 2021.

[13] A. Francés, R. Asensi, Ó. García, R. Prieto and J. Uceda, "Modeling Electronic Power Converters in Smart DC Microgrids—An Overview," *IEEE Transactions on Smart Grid*, vol. 9, no. 6, pp. 6274-6287, 2018.

[14] M. Zhang, X. Wang and Q. Xu, "Data-Driven Modeling of Power-Electronics-Based Power System Considering the Operating Point Variation," *in Proc. ECCE*, Vancouver, BC, Canada, pp.3513-3517, 2021.

[15] A. Francés, R. Asensi, Ó. García, R. Prieto and J. Uceda, "Modeling Electronic Power Converters in Smart DC Microgrids—An Overview," *IEEE Transactions on Smart Grid*, vol. 9, no. 6, pp. 6274-6287, 2018.

[16] B. Li, Y. Liu, B. Li and Y. Xue, "Research on the Coordinated Control of the True Bipolar VSC-HVdc Grid Based on Operating Point Optimization," *IEEE Transactions on Industrial Electronics*, vol. 66, no. 9, pp. 6692-6702, 2019.

[17] Y. Han, M. Luo, X. Zhao, J. M. Guerrero and L. Xu, "Comparative Performance Evaluation of Orthogonal-Signal-Generators-Based Single-Phase PLL Algorithms—A Survey," *IEEE Transactions on Power Electronics*, vol. 31, no. 5, pp. 3932-3944, 2016.

[18] S. Golestan, J. M. Guerrero and J. C. Vasquez, "Three-Phase PLLs: A Review of Recent Advances," *IEEE Transactions on Power Electronics*, vol. 32, no. 3, pp. 1894-1907, 2017.

[19] C. J. O'Rourke, M. M. Qasim, M. R. Overlin and J. L. Kirtley, "A Geometric Interpretation of Reference Frames and Transformations: dq0, Clarke, and Park," *IEEE Transactions on Energy Conversion*, vol. 34, no. 4, pp. 2070-2083, 2019.

[20] Y. Huang and D. Wang, "Effect of Control-Loops Interactions on Power Stability Limits of VSC Integrated to AC System," *IEEE Transactions on Power Delivery*, vol. 33, no. 1, pp. 301-310, 2018.

[21] B. Y. Vyas, B. Das and R. P. Maheshwari, "Improved Fault Classification

in Series Compensated Transmission Line: Comparative Evaluation of Chebyshev Neural Network Training Algorithms," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 27, no. 8, pp. 1631-1642, 2016.

[22] H. Shi, M. Xu and R. Li, "Deep Learning for Household Load Forecasting—A Novel Pooling Deep RNN," *IEEE Transactions on Smart Grid*, vol. 9, no. 5, pp. 5271-5280, 2018.

[23] A. Almalaq and G. Edwards, "A Review of Deep Learning Methods Applied on Load Forecasting," *in Proc. ICMLA*, Cancun, Mexico, pp. 511-516, 2017.

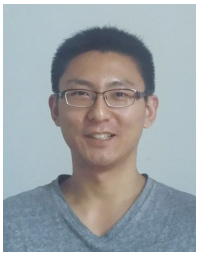
[24] C. Zheng, S. Wang, Y. Liu, C. Liu, W. Xie, C. Fang and S. Liu, "A novel equivalent model of active distribution networks based on LSTM," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 30, no. 9, pp. 2611-2624, 2019.

[25] S. Hochreiter, Y. Bengio, P. Frasconi, J. Schmidhuber, S. C. Kremer, and J. F. Kolen, "Gradient flow in recurrent nets: The difficulty of learning long-term dependencies," *in Proc. A Field Guide to Dynamical Recurrent Networks.* Piscataway, NJ, USA, pp.237-243, 2001.

[26] K. Greff, R. K. Srivastava, J. Koutnik, B. R. Steunebrink, and J. Schmidhuber, "LSTM: A search space odyssey," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 28, no. 10, pp. 2222-2232, 2017.

[27] Y. Li, Z. Wang, J. Yang, X. Wang and J. Feng, "Dynamic Equivalence Modeling for Microgrid Cluster by Using Physical-Data-Driven Method," in *IEEE Transactions on Applied Superconductivity*, vol. 31, no. 8, pp. 1-4, 2021.

[28] A. M. Azmy, I. Erlich, and P. Sowa, "Artifificial neural network-based dynamic equivalents for distribution systems containing active sources," *IEE Proc.-Generat., Transmiss. Distrib.,* vol. 151, no. 6, pp. 681-688, Nov. 2004.



Yunlu Li received the B.S. in electronic information engineering from Shenyang University of Technology in 2009, the M.S. degree in control engineering in 2011 and Ph.D degree in power electronics and drives from Northeastern University in 2017, Shenyang, China. He is currently an associate professor in school of electrical engineering with Shenyang University of Technology, Shenyang, China. His research interests include data-driven based modeling technique for renewable energy system and nonlinear control theory for complex dynamic system.

Guiqing Ma was born in Nanyang, Henan Province, in 1997. He received the B.S. degree in electric automatization from Luoyang Institute of Science and Technology, Luoyang, China, in 2020. He is currently pursuing the M.S. degree in electric engineering at Shenyang University of Technology, Shenyang, China. His main research interests include microgrid modeling, distributed control, and simulation of power systems.

Junyou Yang received the B.Eng. degree from the Jilin University of Technology, Jilin, China, the M.Sc. degree from the Shenyang University of Technology, Shenyang, China, and the Ph.D. degree from the Harbin Institute of Technology, Harbin, China. He was a Visiting Scholar with Department of Electrical Engineering and Computer Science, University of Toronto, Canada, from 1999 to 2020. He is currently the Head of the School of Electrical Engineering, Shenyang University of Technology. He is also a Distinguished Professor of Liaoning Province. His research interests include wind energy, special motor and its control.

Yan Xu was born in Linfen, Shanxi Province, in 1998. She received the B.Eng. degree in electrical engineering and automation from Shanxi Institute of Energy, Jinzhong, China, in 2020. She is currently pursuing the M.S. degree in electrical engineering at Shenyang University of Technology, Shenyang, China. Her research interests include microgrid modeling and simulation of power systems.