Data-Driven Security Assessment of Power Grids Based on Machine Learning Approach

Preprint

H. Xiao,[1] S. Fabus,[1] Y. Su,[1] S. You,[1] Y. Zhao,[1] H. Li,[1] C. Zhang,[1] Y. Liu[1,2] H. Yuan,[3] Y. Zhang,[3] and J. Tan[3]

[1] *University of Tennessee, Knoxville*
[2] *Oak Ridge National Laboratory*
[3] *National Renewable Energy Laboratory*

Presented at the 2019 CIGRE Grid of the Future Symposium
Atlanta, Georgia
November 3–6, 2019

NREL is a national laboratory of the U.S. Department of Energy
Office of Energy Efficiency & Renewable Energy
Operated by the Alliance for Sustainable Energy, LLC

This report is available at no cost from the National Renewable Energy Laboratory (NREL) at www.nrel.gov/publications.

Contract No. DE-AC36-08GO28308

Conference Paper
NREL/CP-5D00-74256
March 2020

Data-Driven Security Assessment of Power Grids Based on Machine Learning Approach

Preprint

H. Xiao,[1] S. Fabus,[1] Y. Su,[1] S. You,[1] Y. Zhao,[1] H. Li,[1] C. Zhang,[1] Y. Liu[1,2] H. Yuan,[3] Y. Zhang,[3] and J. Tan[3]

[1] *University of Tennessee, Knoxville*
[2] *Oak Ridge National Laboratory*
[3] *National Renewable Energy Laboratory*

Suggested Citation
Xiao, H., S. Fabus, Y. Su, S. You, Y. Zhao, H. Li, C. Zhang, Y. Liu, H. Yuan, Y. Zhang, and J. Tan. 2020. *Data-Driven Security Assessment of Power Grids Based on Machine Learning Approach: Preprint*. Golden, CO: National Renewable Energy Laboratory. NREL/CP-5D00-74256. https://www.nrel.gov/docs/fy20osti/74256.pdf.

Contract No. DE-AC36-08GO28308

Conference Paper
NREL/CP-5D00-74256
March 2020

National Renewable Energy Laboratory
15013 Denver West Parkway
Golden, CO 80401
303-275-3000 • www.nrel.gov

NOTICE

This work was authored in part by the National Renewable Energy Laboratory, operated by Alliance for Sustainable Energy, LLC, for the U.S. Department of Energy (DOE) under Contract No. DE-AC36-08GO28308.Funding provided by U.S. Department of Energy Office of Energy Efficiency and Renewable Energy Solar Energy Technologies Office. The views expressed herein do not necessarily represent the views of the DOE or the U.S. Government. The U.S. Government retains and the publisher, by accepting the article for publication, acknowledges that the U.S. Government retains a nonexclusive, paid-up, irrevocable, worldwide license to publish or reproduce the published form of this work, or allow others to do so, for U.S. Government purposes.

This report is available at no cost from the National Renewable Energy Laboratory (NREL) at www.nrel.gov/publications.

U.S. Department of Energy (DOE) reports produced after 1991 and a growing number of pre-1991 documents are available free via www.OSTI.gov.

Cover Photos by Dennis Schroeder: (clockwise, left to right) NREL 51934, NREL 45897, NREL 42160, NREL 45891, NREL 48097, NREL 46526.

NREL prints on paper that contains recycled content.

21, rue d'Artois, F-75008 PARIS
http://www.cigre.org

CIGRE US National Committee
2019 Grid of the Future Symposium

Data-Driven Security Assessment of Power Grids Based on Machine Learning Approach

H. XIAO, S. FABUS, Y. SU, S. YOU, Y. ZHAO, H. LI, C. ZHANG
University of Tennessee Knoxville
USA

Y. LIU
University of Tennessee Knoxville and Oak Ridge National Laboratory
USA

H. YUAN, Y. ZHANG, J. TAN
National Renewable Energy Laboratory
USA

SUMMARY

Security assessment is a fundamental function for both short-term and long-term power system operations. The data-driven security assessment (DSA) criteria will help determine when it is necessary to trigger a dynamic simulation. The DSA criteria will provide a key indicator of switching the simulation method. It will be the link between the traditional isolated dynamic simulation and scheduling simulation.

This paper investigates a data-driven security assessment of electric power grids based on machine learning. Multivariate random forest regression is used as the machine learning algorithm because of its high robustness to the input data. Three stability issues are analyzed using the proposed machine learning tool: transient stability, frequency stability, and small-signal stability. The estimation values from the machine learning tool are compared with those from dynamic simulations. Results show that the proposed machine learning tool can effectively predict the stability margins for the three stability metrics, i.e. transient stability, frequency stability and small-signal stability.

KEYWORDS

Data-driven security assessment (DSA), frequency stability, machine learning, multivariate random forest regression (MRFR), small-signal stability, transient stability.

1. INTRODUCTION

In modern power systems, the penetration level of renewable generation is increasing [1], [2]. Renewable generators could become not only an energy producer but also a necessary provider of ancillary services at multiple timescales [3], [4]. Conventional methods to simulate power system operations—such as short-term transient studies and long-term production simulation—are not sufficient for studying the multi-timescale variation of renewable generation and its impact on system reliability.

Security assessment is a fundamental function for both short-term and long-term power system operations [5]. The data-driven security assessment (DSA) criteria will help determine when it is necessary to trigger a dynamic simulation. The DSA criteria will provide a key indicator of switching the simulation method. It will be the link between the traditional isolated dynamic simulation and scheduling simulation. The stability margins associated with security assessments include small-signal stability, frequency stability, and transient stability.

Though previous studies [6]-[8] have used machine learning to predict stability, they usually used specific power flow solution parameters as input to a machine learning tool, such as voltage and angle at particular buses, power flow of individual branches, along with individual machine and load data. Because the model for simulation is supposed to integrate long-term scheduling with short-term dynamics, only scheduling data, or generator dispatch, is available as an input into the tool. Additionally, each of these past studies analyzed only a single stability issue, most often transient stability.

The machine learning approach proposed in this paper analyzes three stability issues: transient, frequency, and small-signal stability. With the proposed DSA approach, the real-time computational burden from conducting dynamic simulations can be transferred to offline training of the machine learning algorithm, which provides a promising path for real-time applications. The developed machine learning tool can accurately predict the system stability margins and has high robustness to the input data.

2. FRAMEWORK OF MACHINE LEARNING-BASED SYSTEM STABILITY ASSESSMENT

Fig. 1 shows the detailed DSA flowchart based on the machine learning approach proposed in this paper. In general, the DSA process includes three steps:

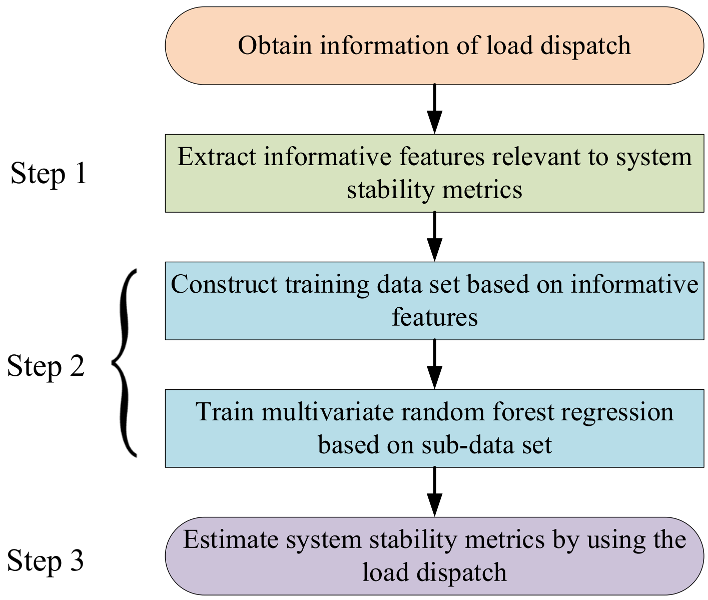

Fig. 1. Flowchart of machine learning-based DSA.

Step 1: Extract informative features relevant to the system stability metrics from the load dispatch.

Step 2: The extracted features are used to train the machine learning model to learn the underlying relationship between the informative features and the system stability metrics. The multivariate random forest regression (MRFR) algorithm is used for the machine learning tool in this paper.

Step 3: The trained MRFR algorithm is further implemented to estimate the system stability metrics using the load dispatch.

The MRFR is an ensemble of decision trees trained by bootstrap sampling and random feature selection. It aims to build a large collection of regression trees and average the output of each tree to reduce the variance of the prediction results and boost the performance of the final model. Fig. 2 shows a diagram of the MRFR algorithm. Considering a training data set $\mathbf{X} = [X_1, X_2 \ldots X_n]$ and the corresponding response $\mathbf{Y} = [Y_1, Y_2 \ldots Y_n]$, MRFR first uses a bootstrap to draw a set of samples with size m from the training data set. Then it establishes a regression Tree-i based on these bootstrapped data. The following steps are recursively repeated for each terminal node of the tree, until the minimum node size is attained.

1) Randomly select D dimension features of each bootstrapped sample in the training data set.

2) Split the parent node into two children nodes based on the information gain ratio criterion.

This iterative procedure is repeated for k times, and the output of the forest is the average of the predictions from each regression tree, as shown in (1)

$$\hat{Y} = \frac{1}{k}\sum_{i=1}^{k}\hat{Y}_i(\mathrm{X}) \tag{1}$$

We use MRFR in this paper because of its ease of implementation, high robustness to the input data, and capability to avoid overfitting during the training process. In this paper, the input of the DSA machine learning tool is the dispatch data updated every 5 minutes. Its outputs are system stability metrics, including critical clearing time, frequency nadir, and damping ratio of oscillation modes.

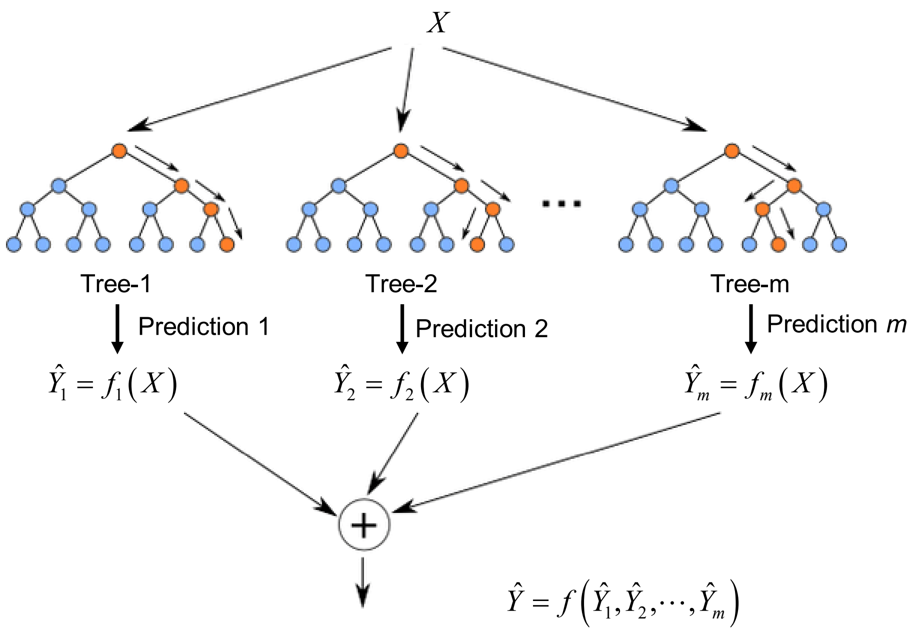

Fig. 2. Diagram of multivariate random forest algorithm.

3. DATA-DRIVEN SECURITY ASSESSMENT: CASE STUDY

In this section, three system stability issues are performed with the proposed machine learning tool: transient stability, frequency stability, and small-signal stability. An 18-bus test system extracted from the case library of the PSLF simulation tool [9] is used to demonstrate the performance of the proposed machine learning tool. The single-line diagram of the 18-bus system is shown in Fig. 3, which includes 18 buses, 24 branches, 5 generators (including 1 photovoltaic power plant), and 7 loads. The dispatch load and generator data are given for every 5 minutes of the 24-hour period. Therefore, there are 288 cases in total.

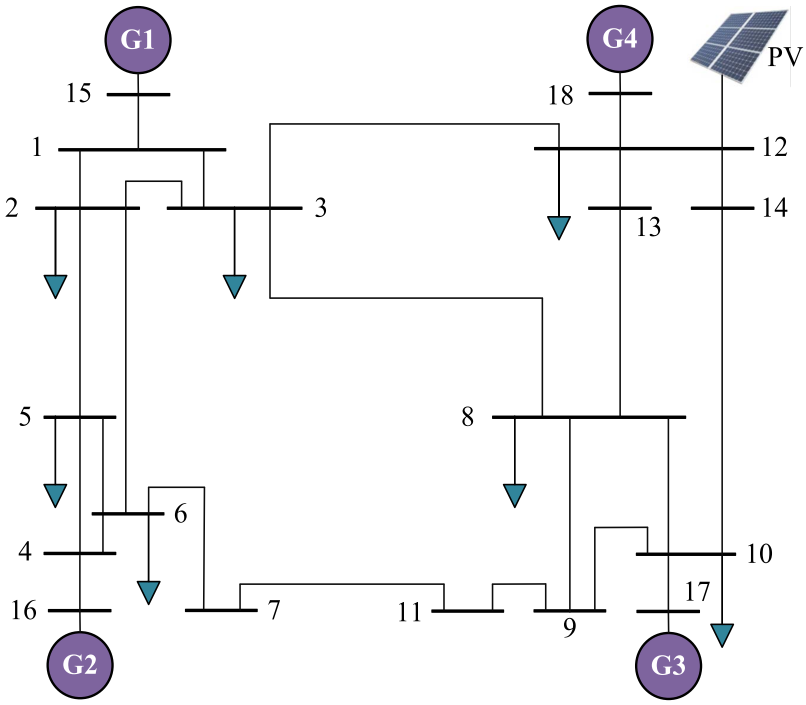

Fig. 3. 18-bus test system.

3.1 DATA-DRIVEN TRANSIENT STABILITY ASSESSMENT

To predict the transient stability margin, a three-phase fault on a transmission line is applied in the system. Although the location of the fault is fixed during all simulations, the fault clearing time is varied. In this example, the fault is placed at the branch between North-01 (Bus 1) and North-02 (Bus 2) in the 18-bus system. The same set of fault clearing times are tested at each time step, or every 5 minutes of the 24-hour data period. At each 5-minute time step, the fault clearing time is adjusted from 60 ms–720 ms in intervals of 20 ms. Therefore, 9,792 total test cases are created. Critical clearing time (CCT) is selected as the metric of transient stability. CCT is defined as the maximum time allowed to remove the disturbance without interrupting the system's performance. The system is stable if the disturbance can be cleared before the time allowed. The purpose of adjusting the fault clearing time in small intervals is to create a sufficient number of stable and unstable cases to determine a more exact CCT. During each simulation, rotor angles of all machines are monitored. If the rotor angle deviation of any two generators exceeds 180 degrees, the case is considered unstable.

The capability of the machine learning tool in predicting the transient stability margin is then validated by using the power flow solution. The input to the machine learning tool is individual real power output of all generators. During testing, the output is the transient stability margin prediction, which is determined by computing estimated fault clearing time as a function of dispatch. Among the 288 cases, 200 cases are randomly selected for training, and the remaining cases are used for testing. Fig. 4 compares the simulated CCT and estimated CCT from the test results of the machine learning tool. The errors of the two results are illustrated in Fig. 5. As shown in Fig. 5, the error of the CCT is less than 40 ms, which means that the proposed machine learning tool has very high accuracy.

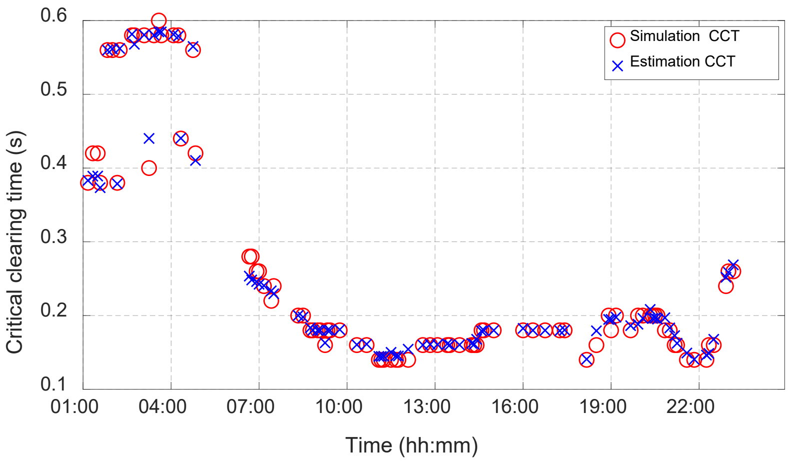

Fig. 4. Simulation and estimation CCT in testing data set.

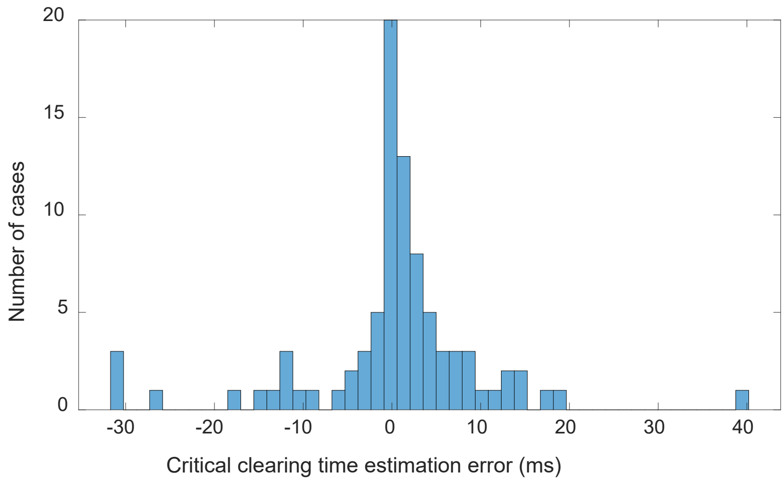

Fig. 5. CCT estimation error distribution.

3.2 DATA-DRIVEN FREQUENCY STABILITY ASSESSMENT

To predict the frequency stability, frequency nadir is selected as the metric. The contingency is selected by tripping the maximum amount of generation (70 MW) the system can handle while still converging. The inertia level is forced to be a percentage of the base inertia depending on total generation output. The purpose of adjusting the inertia at each generation level is to create a sufficient number of stable and unstable cases to train and test the machine learning tool. Though the total generation output over the 24-hour period does not change enough to alter unit commitment and dispatch by large amounts, the inertia is forced to change linearly, even with a smaller range of generation output variation. A linear function is created to correlate generator output and inertia level. The percentages range from 20% to 100% of the base inertia, corresponding to the minimum and maximum total generation output.

For machine learning model development, stability is determined by examining the frequency nadir of each case. Real power dispatch of all generators is provided as the input to the machine learning tool. The frequency nadir is estimated, subjected to the generation trip. Similar to the transient stability assessment, 200 cases are randomly selected for training, and the rest of the 88 cases are for testing.

The capability of the machine learning model in predicting the frequency stability margin is then validated by using the power flow solution. During testing, the input to the machine learning tool is the real power output of all machines. The output is frequency stability margin prediction, based on estimated frequency nadir. The results of the power flow simulation and machine learning estimation are compared and shown in Fig. 6. The distribution of the frequency nadir estimation errors is plotted in Fig. 7. The figures show that the estimation error of the frequency nadir of the generators is less than 7 mHz.

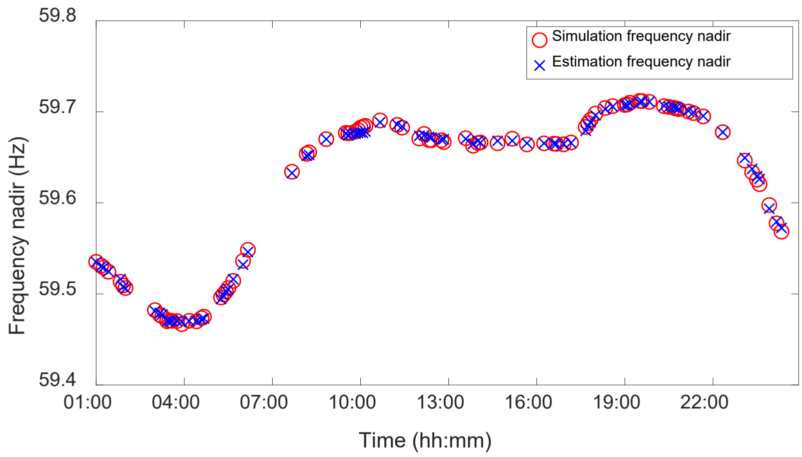

Fig. 6. Simulation and estimation frequency nadir in testing data set.

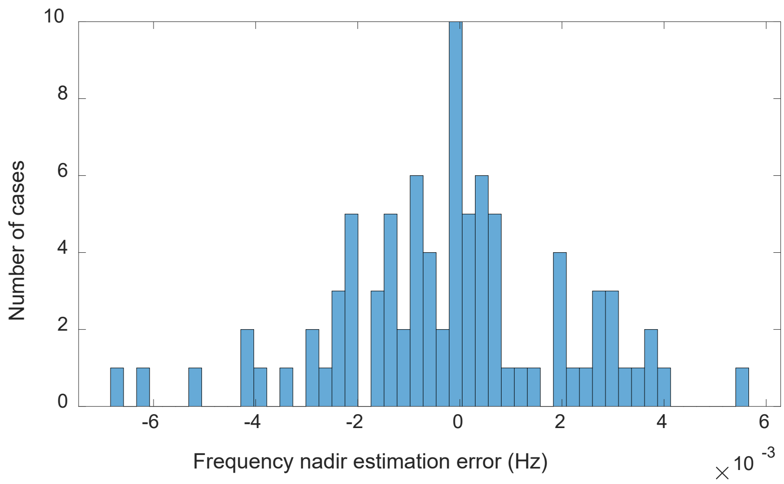

Fig. 7. Frequency nadir estimation error distribution.

3.3 DATA-DRIVEN SMALL-SIGNAL STABILITY ASSESSMENT

The data-driven small-signal stability assessment is performed on the 18-bus system. Damping ratio is selected as the metric of small-signal stability. The damping ratio is a dimensionless measure describing how oscillations in power systems decay after a disturbance. Ideally, the system is stable if the damping ratio is positive and unstable if the damping ratio is negative. The oscillation mode with smallest damping ratio affect the system performance most, so the small-signal stability assessment mainly focuses on this mode. To train and test the model, a specific mode is selected at a time. For training, the damping ratio and generator dispatch at each 5-minute interval are provided as inputs to the model, with 200 of the 288 cases randomly selected to train. The rest of the cases are used to test the accuracy of the model.

The results of the small-signal stability assessment are illustrated in Fig. 8 and Fig. 9. The oscillation frequency of this mode is 0.73 Hz, and the average damping ratio is 1.11%. As shown in Fig. 9, the damping ratio estimation error is less than 0.15%, which means that the machine learning estimated results agree well with the generated simulations.

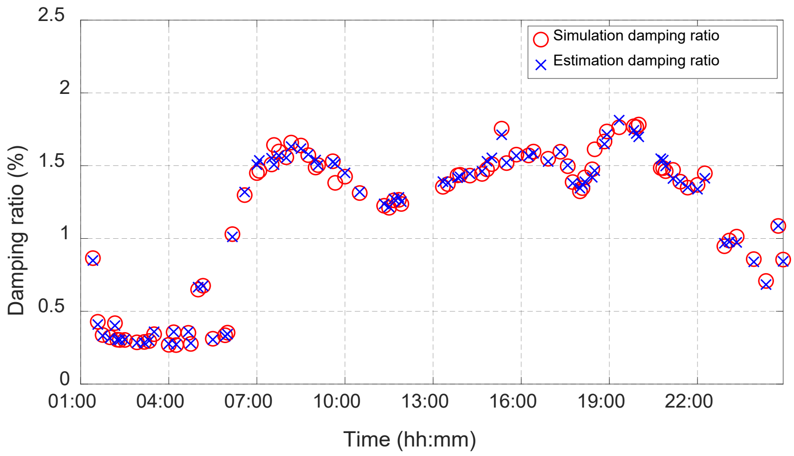

Fig. 8. Simulation and estimation damping ratio in testing data set.

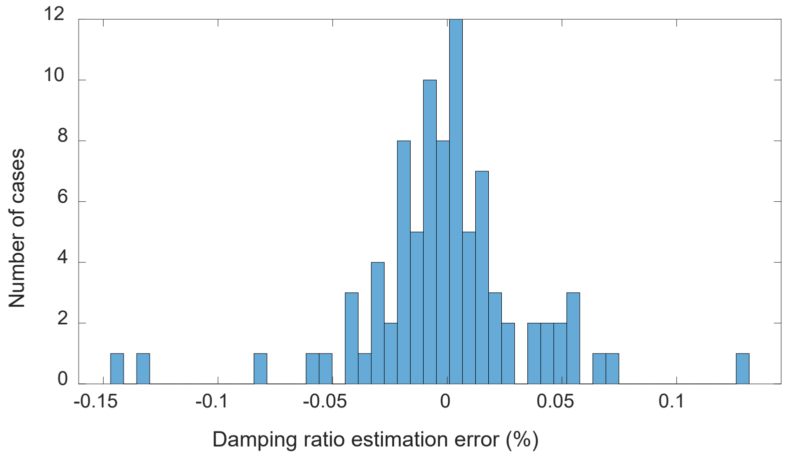

Fig. 9. Damping ratio estimation error distribution.

4. CONCLUSION

This paper proposed a unified machine learning approach for DSA of electric power grids. We used the MRFR algorithm for the machine learning tool because of its ease of implementation, high robustness to the input data, and capability to avoid overfitting during the training process. Three system stability metrics were performed on an 18-bus test system with the proposed machine learning tool, including critical clearing time, frequency nadir, and damping ratio of oscillation mode. The proposed DSA approach transferred the real-time computational burden from conducting dynamic simulations to offline training of the machine learning algorithm, thus providing a promising path for real-time applications. Results show that the developed machine learning tool can accurately predict the system stability margins and has high robustness to the input data.

DISCLAIMER

This work was authored in part by the National Renewable Energy Laboratory, operated by Alliance for Sustainable Energy, LLC, for the U.S. Department of Energy (DOE) under Contract No. DE-AC36-08GO28308. Funding provided by U.S. Department of Energy Office of Energy Efficiency and Renewable Energy Solar Energy Technologies Office. The views expressed in the article do not necessarily represent the views of the DOE or the U.S. Government.

BIBLIOGRAPHY

[1] L. Vimmerstedt *et al.*, "High Penetration of Renewable Energy in the Transportation Sector: Scenarios, Barriers, and Enablers," National Renewable Energy Lab.(NREL), Golden, CO (United States), May 2012.

[2] E. Du *et al.*, "The Role of Concentrating Solar Power Toward High Renewable Energy Penetrated Power Systems," *IEEE Trans. Power Syst.,* vol. 33, no. 6, pp. 6630-6641, Nov. 2018.

[3] V. Gevorgian and B. O'Neill, "Advanced grid-friendly controls demonstration project for utility-scale PV power plants," National Renewable Energy Lab.(NREL), Golden, CO (United States), Jan. 2016.

[4] J. Tan and Y. Zhang, "Coordinated Control Strategy of a Battery Energy Storage System to Support a Wind Power Plant Providing Multi-Timescale Frequency Ancillary Services," *IEEE Transactions on Sustainable Energy,* vol. 8, no. 3, pp. 1140-1153, Jul. 2017.

[5] P. Kundur *et al.*, "Definition and Classification of Power System Stability," *IEEE Trans. Power Syst.,* vol. 19, no. 3, pp. 1387-1401, May 2004.

[6] L. S. Moulin, A. P. A. d. Silva, M. A. El-Sharkawi, and R. J. Marks, "Support Vector Machines for Transient Stability Analysis of Large-Scale Power Systems," *IEEE Trans. Power Syst.,* vol. 19, no. 2, pp. 818-825, May 2004.

[7] Y. Xu, Z. Y. Dong, Z. Xu, K. Meng, and K. P. Wong, "An Intelligent Dynamic Security Assessment Framework for Power Systems With Wind Power," *IEEE Trans. Ind. Inf. ,* vol. 8, no. 4, pp. 995-1003, Nov. 2012.

[8] S. P. Teeuwsen, I. Erlich, M. A. El-Sharkawi, and U. Bachmann, "Genetic Algorithm and Decision Tree-Based Oscillatory Stability Assessment," *IEEE Trans. Power Syst.,* vol. 21, no. 2, pp. 746-753, May 2006.

[9] Q. Wang, H. Wu, J. Tan, and B.-M. Hodge, "Analyzing the Impacts of Increased Wind Power on Generation Revenue Sufficiency," in *2016 IEEE Power and Energy Society General Meeting*, Boston, Massachusetts, Jul. 2016.