



# A Comparison of Machine Learning Methods for Frequency Nadir Estimation in Power Systems

## Preprint

Xuebo Liu,<sup>1,2</sup> Jian Xie,<sup>1,3</sup> Xin Fang,<sup>1</sup> Haoyu Yuan,<sup>1</sup>  
Bin Wang,<sup>1</sup> Hongyu Wu,<sup>2</sup> and Jin Tan<sup>1</sup>

*1 National Renewable Energy Laboratory*

*2 Kansas State University*

*3 University of Central Florida*

*Presented at the IEEE Kansas Power & Energy Conference*

*Manhattan, Kansas*

*April 25–26, 2022*

**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](http://www.nrel.gov/publications).

Contract No. DE-AC36-08GO28308

**Conference Paper**  
NREL/CP-6A40-82116  
May 2022



# A Comparison of Machine Learning Methods for Frequency Nadir Estimation in Power Systems

## Preprint

Xuebo Liu,<sup>1,2</sup> Jian Xie,<sup>1,3</sup> Xin Fang,<sup>1</sup> Haoyu Yuan,<sup>1</sup>  
Bin Wang,<sup>1</sup> Hongyu Wu,<sup>2</sup> and Jin Tan<sup>1</sup>

*1 National Renewable Energy Laboratory*

*2 Kansas State University*

*3 University of Central Florida*

### Suggested Citation

Liu, Xuebo, Jian Xie, Xin Fang, Haoyu Yuan, Bin Wang, Hongyu Wu, and Jin Tan. 2022. *A Comparison of Machine Learning Methods for Frequency Nadir Estimation in Power Systems: Preprint*. Golden, CO: National Renewable Energy Laboratory. NREL/CP-6A40-82116. <https://www.nrel.gov/docs/fy22osti/82116.pdf>.

© 2022 IEEE. Personal use of this material is permitted. Permission from IEEE must be obtained for all other uses, in any current or future media, including reprinting/republishing this material for advertising or promotional purposes, creating new collective works, for resale or redistribution to servers or lists, or reuse of any copyrighted component of this work in other works.

**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](http://www.nrel.gov/publications).

Contract No. DE-AC36-08GO28308

**Conference Paper**  
NREL/CP-6A40-82116  
May 2022

National Renewable Energy Laboratory  
15013 Denver West Parkway  
Golden, CO 80401  
303-275-3000 • [www.nrel.gov](http://www.nrel.gov)

## NOTICE

This work was authored 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 the U.S. Department of Energy Office of Renewable Energy and Energy Efficiency Solar Energy Technologies Office Award Numbers 34224 and 37772. The views expressed herein do not necessarily represent the views of the DOE or the U.S. Government.

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

# A Comparison of Machine Learning Methods for Frequency Nadir Estimation in Power Systems

Xuebo Liu<sup>1,2</sup>, Jian Xie<sup>1,3</sup>, Xin Fang<sup>1</sup>, Haoyu Yuan<sup>1</sup>, Bin Wang<sup>1</sup>, Hongyu Wu<sup>2</sup>, and Jin Tan<sup>1</sup>

1-National Renewable Energy Laboratory, Golden, CO, 80401, US

2-Mike Wieggers Department of Electrical and Computer Engineering, Kansas State University, Manhattan, KS, 66502, US

3-Department of Electrical and Computer Engineering, University of Central Florida, Orlando, FL, 32816, US

**Abstract**—An increasing penetration level of inverter-based renewable energy resources changes the inertia of power systems, posing challenges for maintaining the desired system frequency stability. An accurate frequency nadir estimation is crucial for power system operators to prepare preventive actions against large frequency excursions. In this paper, five machine learning methods—linear regression, gradient boosting, support vector regression, an artificial neural network, and XGBoost—are applied to two different datasets, i.e., 1) the unit generation dataset and 2) the system total inertia and headroom dataset, for the prediction of the frequency nadir. The training and testing datasets are generated through extensive generation scheduling simulations using Multi-timescale Integrated Dynamic and Scheduling (MIDAS) toolbox on the Western Electricity Coordinating Council 240-bus system with high renewable penetration levels. Numerical results show that all five machine learning methods perform well in predicting the nadir frequency of the system. Among them, the gradient boosting and the XGBoost are clear winners yielding the best prediction accuracy in terms of four evaluation metrics.

**Index Terms**—data driven, machine learning, frequency nadir estimation, power system stability

## I. INTRODUCTION

To tackle the problem of climate change and fossil fuel depletion, renewable energy sources (RES) and energy storage are considered the most practical solutions in the power system. Many works in the literature have explored uncertain RES in the unit commitment (UC) and economic dispatch problems in the day-ahead market [1]–[5]. References [1], [2] proposed mixed-integer linear programming-based optimization for power system integration high photovoltaic (PV) penetration, and the authors demonstrated that it significantly reduces day-ahead market cost while considering PV and a battery energy storage system as ancillary service providers. References [3], [4] proposed a method to optimize generation under high RES penetration on the distribution locational marginal price components and distributed energy resources' schedules in the day-ahead market model. Reference [5] proposed a method to solve stochastic network-constrained unit commitment with an increasing penetration level of renewable energy by an accelerating technique; however, as more conventional generators are retired and replaced by RES, the power system stability can be adversely affected due to the reduced system inertia [6]. Therefore, a challenge emerges that a lack of system inertia could contribute to fast decline of the grid frequency during a generator trip in the frequency, as

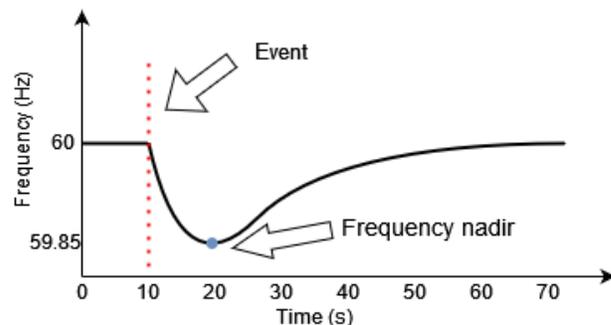


Fig. 1: Illustration of frequency dynamic response after the largest generator trips

shown in Fig. 1, which could lead to cascading failures and eventually a electrical grid blackout [7]. However, none of the above literature consider the frequency of nadir during the day-ahead unit commitment, which may cause server issue in the dynamic results, especially in the island system with high renewable penetration. Our study proposes a straightforward and efficient frequency nadir estimator which helps maintain power system stability in the day-ahead market.

An accurate estimation of frequency nadir can assist power system operators in maintaining sufficient inertia and adequate headroom for primary frequency response and in maintaining system frequency stability [8]. Because of the availability of a large volume of historical and model-based data, machine learning techniques have gained attraction in frequency nadir prediction because they can achieve high accuracy in real-world classification, prediction, and regression problems. They have been used in many practical applications, such as short-term wind speed prediction [9] and temperature prediction [10], [11], unit commitment problems [12], transient stability assessment [8], [13], and power system resilience studies [14]. A multivariate random forest regression-based frequency nadir estimation was proposed in [8] using time-domain simulation data from a realistic power system model to predict the real-time frequency nadir of varying system conditions with the predetermined contingency, given the system inertia and the generation dispatch of conventional generators. Reference [13] proposed state-of-art artificial-intelligence-based applications in power systems including inertia estimation, disturbance

size and location estimation, system stability assessment, and data authentication; thereby providing accurate results and improving efficiency to those applications. A frequency nadir prediction model with high accuracy was proposed in [15] to establish primary frequency response in various scenarios by polynomial fitting. A nonlinear auto-regressive model based on an artificial neural network (ANN) was proposed in [16] to predict the timing of a frequency nadir under various prediction horizons.

All of these works reported that machine learning can achieve high performance for frequency nadir prediction, but a comprehensive comparison among different machine learning methods has not been conducted yet. This paper tries to fill this gap by proposing two different ways to preprocess data and comparing five machine learning methods: linear regression (LR), gradient boosting (GB), support vector regression (SVR), ANN, and XGBoost. A detailed comparison is carried out based on extensive simulations on the Multi-timescale Integrated Dynamic and Scheduling (MIDAS) toolbox [17] developed by the National Renewable Energy Laboratory. The remainder of this paper is structured as follows. The research background and the machine learning methods used are described in Section II. Numerical results are presented and analyzed in Section III. Section IV concludes the paper.

## II. BACKGROUND AND FORMULATION

### A. Background

In this section, we briefly review the five machine learning methods used and the background of the proposed work. In this study, we explore two ways to approach the dataset necessary for the machine learning methods: 1) the unit generation dataset (UGD) and 2) the system total inertia and headroom dataset (TIH). Both datasets are generated by running the PSS/E [18] on the Western Electricity Coordinating Council (WECC) 240-bus system [19], [20]. More details about the dataset generation are discussed in the next section. The UGD is a dataset with hourly generation of all units in the power system. The TIH dataset consists of the hourly total inertia and total headroom of the system, which can be calculated as follows:

$$T_{Inertia} = \sum_{i=0}^N H_i \cdot Mbase_i, \forall i \quad (1)$$

$$T_{HR} = \sum_{i=0}^N (P_i^{max} - P_i), \forall i \in [\mathbf{S}, \mathbf{G}, \mathbf{H}] \quad (2)$$

where  $T_{Inertia}$  and  $T_{HR}$  are the system total inertia and total headroom at each hour;  $H_i$  is the inertia of unit  $i$  in second;  $Mbase_i$  is the rating in MVA of unit  $i$ ;  $P_i^{max}$  is the maximum power generation of unit  $i$ ; and  $\mathbf{S}$ ,  $\mathbf{G}$ , and  $\mathbf{H}$  represent the set of steam, gas, and hydropower generators, respectively. Note that the headroom of the TIH dataset is considered in these three types of generators. The inertia of a unit exists only when the unit is on.

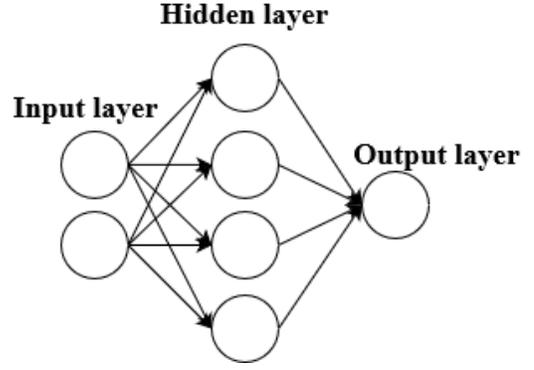


Fig. 2: Simplified ANN model

### B. Machine Learning Methods

#### 1) Linear regression:

$$y_i = \beta x_i^T + \varepsilon_i \quad (3)$$

where  $x_i$  is the input variable, and  $y_i$  is the output variable. The slope of the line is  $\beta$ , and  $\varepsilon_i$  is the intercept (the value of  $y$  when  $x = 0$ ) [21]. Linear regression can be used to construct linear constraint of frequency nadir which is suitable for the day-ahead UC model.

#### 2) Gradient boosting:

$$F_0(x) = \arg \min_{\gamma} \sum_{i=1}^n L(y_i, \gamma) \quad (4)$$

$$F_m(x) = F_{m-1}(x) + \arg \min_{h_m \in H} \sum_{i=1}^n L(y_i, F_{m-1}(x_i) + h_m(x_i)) \quad (5)$$

where  $F_i$ ,  $h_i$  are the objective learning function and the generic function at stage  $m$ , and  $L$  is the loss function. Parameters  $\gamma$ ,  $x$ , and  $y$  are the initialization, input variable, and output variable, respectively [22].

#### 3) Support Vector Regression:

$$\min \frac{1}{2} \|w\|^2 + C \sum_{i=1}^N (\xi_i + \xi_i^*) \quad (6)$$

$$y_i - wx_i - b \leq \epsilon + \xi_i \quad (7)$$

$$-y_i + wx_i + b \leq \epsilon + \xi_i^* \quad (8)$$

$$\xi_i, \xi_i^* \geq 0 \quad (9)$$

where  $w$  refers to the hyperplane;  $\epsilon$  is the boundary line;  $\xi_i$  and  $\xi_i^*$  are the upper-bound error and lower-bound error, respectively;  $x_i$  is the input variable; and  $y_i$  is the output variable. The objective of the SVR is to minimize the error in Equation (6) in the hyperplane, which maximizes the margin as constraints (7)–(9) [23].

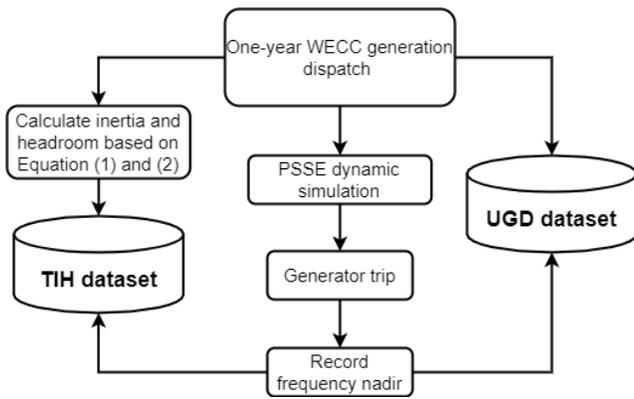


Fig. 3: Flowchart of generating UGD and TIH datasets

#### 4) Artificial Neural Network:

$$W = W - \alpha \nabla J(W) \quad (10)$$

$$\nabla J(W) = \left( \frac{\partial J}{\partial w_1}, \frac{\partial J}{\partial w_2}, \dots, \frac{\partial J}{\partial w_N} \right) \quad (11)$$

where  $W$  is the updated parameters set of the regression function, and  $J$  is the loss function, which is calculated by mean square error (MSE). Fig. 2 shows the graphic architecture of a fully connected neural network with two input neurals, four hidden neurals, and one output neural. The ANN is built based on the gradient descent with activation functions [24].

5) *XGBoost*: *XGBoost* stands for Extreme Gradient Boosting, which is built based on gradient boosting but has more accurate approximations by enabling the regularization and second-order gradients [25]. The objective function of the *XGBoost* is as follows:

$$L^t = \sum_{i=1}^n l(\hat{y}_i, y_i) + \sum_k \Omega(f_k) \quad (12)$$

$$\Omega(f) = \gamma T + \frac{1}{2} \lambda \|w\|^2 \quad (13)$$

where  $l$  is the loss function that measures the difference between the prediction,  $\hat{y}_i$ , and the read value,  $y_i$ . The second term,  $\Omega$ , penalizes the complexity of the model (i.e., the regression tree functions), which helps to smooth the final learnt weights to avoid overfitting [25]. Parameters  $\gamma$ ,  $T$ , and  $w$  are the coefficient, the number of leaves, and the leaf weights for the corresponding regression tree, respectively.

### III. CASE STUDY

The datasets of the proposed study were collected from a 1-year dynamic simulation in the WECC 240-bus system with high RES penetration. The dynamic simulation is conducted in PSS/E. The procedure of obtaining the UGD and TIH datasets is shown in Fig. 3. The hourly generation dispatch is obtained from MIDAS which is an open-source power-system-simulation tool for simulating the electricity market that considers multi-timescale operations with an interface for dynamic simulations. It can address the challenge of operating the grid with extremely high renewable penetrations

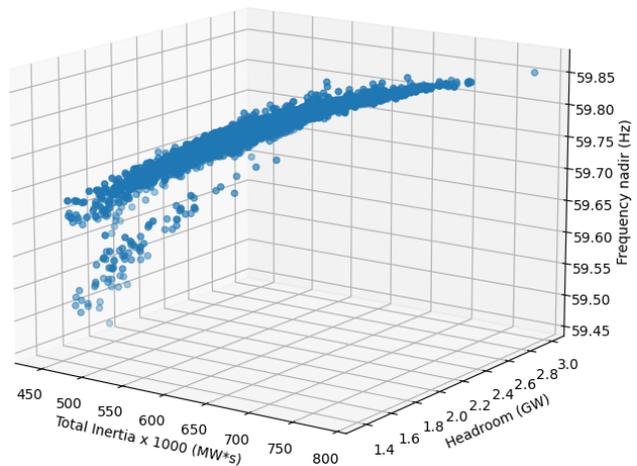


Fig. 4: Frequency nadir versus total inertia and headroom

TABLE I: Hyper-parameters of ANN

	activation	L2 regularization	# hidden units	optimizer
UGD	logistic	0.001	50	Adam
TIH	tanh	5e-05	50	SGD

by bridging the modeling the gaps of different timescales between economics, reliability, and stability of grid operation [26].

Once the generation dispatch is given from the MIDAS scheduling, the PSS/E-based dynamic simulation is processed following the trip of the largest generator. Then, the frequency nadir points are recorded. In this procedure, the frequency nadir in the TIH dataset can be visualized as a 3D plot in Fig. 4. Figure 4 depicts three dimensions—inertia, headroom, and frequency nadir—and the z-axis is the frequency nadir point corresponding to the total inertia and the headroom calculated based on the generator dispatch point according to Equation (1) and (2). The outliers in the Fig. 4 are caused by the various tripped largest generator because the different generator has different effect on frequency. In addition, it is seen that the more inertia and headroom can contribute higher frequency nadir.

We randomly split the dataset into 80% and 20% for the training and testing, respectively. The hyper-parameters and architecture information of the ANN are listed in Table I. Those parameters were finely tuned for an optimal MSE. In addition to the MSE, other metrics—including mean absolute error (MAE), root mean square error (RMSE), and maximum error (ME)—were used on both the UGD and the TIH datasets for training and testing evaluations in the remaining part of this section.

The UGD training results are given in Table II. As shown, the gradient boosting has the best training performance among all five machine learning methods, and the performance of the *XGBoost* is second. Note that a maximum error of 0.02487 from the gradient boosting indicates that the prediction error of the gradient boosting is no greater than 0.02487 Hz. The UGD testing results are shown in Table III. Analogously, the gradient boosting has the best prediction performance in terms of MAE, MSE, and RMSE; however, the maximum error

TABLE II: Evaluation metrics of UGD training

Methods	MAE (Hz)	MSE (Hz)	RMSE (Hz)	ME (Hz)
LR	0.01014	0.00027	0.01646	0.16097
GB	<b>0.00231</b>	<b>0.00001</b>	<b>0.00295</b>	<b>0.02487</b>
SVR	0.00461	0.00250	0.04852	0.10047
ANN	0.02741	0.00161	0.04068	0.27947
XGBoost	0.00366	0.00003	0.00557	0.11243

TABLE III: Evaluation metrics of UGD testing

Methods	MAE (Hz)	MSE (Hz)	RMSE (Hz)	ME (Hz)
LR	0.01028	0.00035	0.01813	0.17674
GB	<b>0.00345</b>	<b>0.00005</b>	<b>0.00682</b>	0.10395
SVR	0.04674	0.00256	0.04886	0.1351
ANN	0.02494	0.00149	0.03675	0.30963
XGBoost	0.00428	0.00006	0.00762	<b>0.10137</b>

of the XGBoost is slightly better than that of the gradient boosting. Additionally, the SVR leads to a smaller maximum error but larger MSE and RMSE than the ANN.

The TIH training results are provided in Table IV. Different from the UGD dataset, the XGBoost has the best MAE, MSE, and RMSE among all five methods, whereas the gradient boosting outperforms the other methods in terms of maximum error. In addition, the evaluation metrics for both the gradient boosting and the XGBoost are quite similar to the TIH training dataset. Table V lists the TIH testing results of all the methods. Again, the XGBoost leads to the best performance; however, the ANN’s performance in this dataset is worse than that of the UGD dataset, whereas the other methods perform quite similarly between the two datasets.

Figure 5 shows the error distribution for the (top) UGD and (bottom) TIH datasets. In general, the gradient boosting achieves the best error distribution in both datasets, whereas the XGBoost performs nearly as well. In contrast, the ANN and the SVR are not suitable for the proposed work because of the high variance distribution. In addition, both methods perform worse in the TIH testing dataset. Note that the performance of the linear regressing ranks third in both datasets.

#### IV. CONCLUSION

In this paper, we propose to use machine-learning method, rather than dynamic simulation, to predict frequency nadir of the largest N-1 contingency analysis based on steady-state scheduling results. Five methods—linear regression, gradient boosting, SVR, ANN, and XGBoost—are compared by testing

TABLE IV: Evaluation metrics of TIH training

Methods	MAE (Hz)	MSE (Hz)	RMSE (Hz)	ME (Hz)
LR	0.01023	0.00028	0.0167	0.16213
GB	0.00461	0.00006	0.00753	<b>0.07957</b>
SVR	0.05291	0.00309	0.05557	0.10048
ANN	0.03251	0.02101	0.14496	3.87351
XGBoost	0.00403	<b>0.00004</b>	<b>0.00623</b>	0.10039

TABLE V: Evaluation metrics of TIH testing

Methods	MAE (Hz)	MSE (Hz)	RMSE (Hz)	ME (Hz)
LR	0.01056	0.00030	0.01733	0.17674
GB	0.00499	0.00007	0.00861	0.09003
SVR	0.05383	0.00318	0.05636	0.11954
ANN	0.04853	0.08453	0.29074	4.00111
XGBoost	<b>0.00469</b>	<b>0.00006</b>	<b>0.00781</b>	<b>0.07819</b>

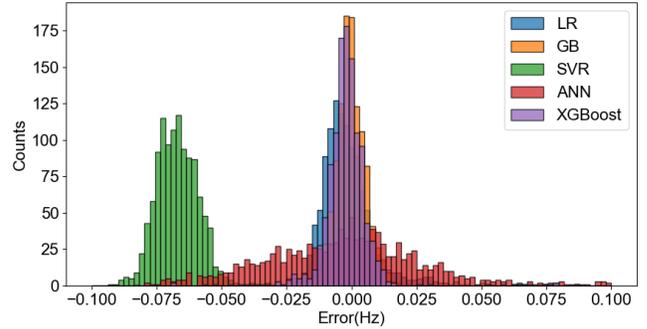
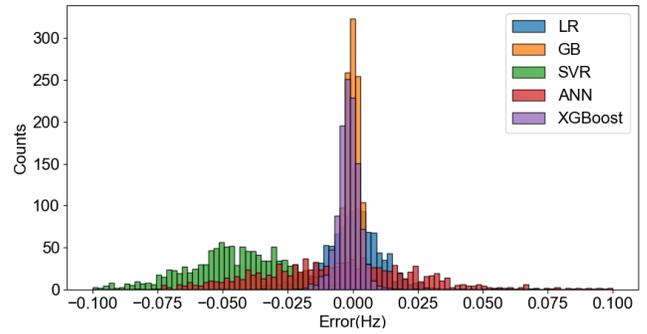


Fig. 5: Error distribution of UGD (top) and TIH (bottom) testing

the reduced WECC 240-bus system with high RES penetration. Results show that these methods can achieve high prediction accuracy in general. The prediction performance of the gradient boosting and the XGBoost are the best. In addition, both gradient boosting and XGBoost achieve high accuracy in the two method’s dataset which are proposed in this paper 1) the unit generation dataset; 2) the system total inertia and headroom dataset. In the future, we will propose a machine learning-based security constrained unit commitment model and analyze the dynamic simulation of the proposed work.

#### ACKNOWLEDGEMENT

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 Award Numbers 34224 and 37772. The views expressed in the article 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.

#### REFERENCES

- [1] S. Yin, J. Wang, Y. Lin, X. Fang, J. Tan, and H. Yuan, “Practical operations of energy storage providing ancillary services: From day-ahead to real-time; practical operations of energy storage providing ancillary services: From day-ahead to real-time,” 2020.

- [2] X. Fang, J. Tan, H. Yuan, S. Yin, and J. Wang, "Providing ancillary services with photovoltaic generation in multi-timescale grid operation." Institute of Electrical and Electronics Engineers Inc., 4 2021.
- [3] M. N. Faqiry, L. Edmonds, H. Wu, and A. Pahwa, "Distribution locational marginal price-based transactive day-ahead market with variable renewable generation," *Applied Energy*, vol. 259, p. 114103, 2 2020.
- [4] L. Edmonds, M. Derby, M. Hill, and H. Wu, "Coordinated operation of water and electricity distribution networks with variable renewable energy and distribution locational marginal pricing," *Renewable Energy*, vol. 177, 06 2021.
- [5] A. M. Palani, H. Wu, and M. M. Morcos, "A fast penalty-based gauss-seidel method for stochastic unit commitment with uncertain load and wind generation," *IEEE Open Access Journal of Power and Energy*, vol. 8, pp. 211–222, 2021.
- [6] A. Fernandez-Guillamon, E. Gomez-Lazaro, E. Muljadi, and Molina-Garcia, "Power systems with high renewable energy sources: A review of inertia and frequency control strategies over time," *Renewable and Sustainable Energy Reviews*, vol. 115, p. 109369, 2019.
- [7] R. Eriksson, N. Modig, and K. Elkington, "Synthetic inertia versus fast frequency response: a definition; synthetic inertia versus fast frequency response: a definition," *IET Renewable Power Generation*, vol. 12, pp. 507–514, 2017.
- [8] Y. Su, H. Li, Y. Cui, S. You, Y. Ma, J. Wang, and Y. Liu, "An adaptive pv frequency control strategy based on real-time inertia estimation," *IEEE Transactions on Smart Grid*, vol. 12, pp. 2355–2364, 2021.
- [9] D. Geng, H. Zhang, and H. Wu, "Short-term wind speed prediction based on principal component analysis and lstm," *Applied Sciences*, vol. 10, no. 13, 2020.
- [10] X. Liu, Y. Wu, and H. Wu, "PV-EV integrated home energy management considering residential occupant behaviors," *Sustainability (Switzerland)*, vol. 13, 2021.
- [11] X. Liu, Y. Wu, H. Zhang, and H. Wu, "Hourly occupant clothing decisions in residential hvac energy management," *Journal of Building Engineering*, vol. 40, 2021.
- [12] L. W. Yafei Yang, "Machine learning approaches to the unit commitment problem: Current trends, emerging challenges, and new strategies," *The Electricity Journal*, vol. 34, p. 106889, 1 2021.
- [13] L. Zhu, Y. Zhao, Y. Cui, S. You, W. Yu, S. Liu, H. Yin, C. Chen, Y. Wu, W. Qiu, M. Mandich, H. Li, A. Ademola, C. Zhang, C. Zeng, X. Jia, W. Wang, H. Yuan, H. Jiang, J. Tan, and Y. Liu, "Adding power of artificial intelligence to situational awareness of large interconnections dominated by inverter-based resources," *High Voltage*, vol. 6, pp. 924–937, 12 2021.
- [14] J. Xie, I. Alvarez-Fernandez, and W. Sun, "A review of machine learning applications in power system resilience," in *2020 IEEE Power Energy Society General Meeting (PESGM)*, 2020, pp. 1–5.
- [15] L. Liu, W. Li, Y. Ba, J. Shen, C. Jin, and K. Wen, "An analytical model for frequency nadir prediction following a major disturbance," *IEEE Transactions on Power Systems*, vol. 35, pp. 2527–2536, 7 2020.
- [16] H. R. Chamorro, A. D. Orjuela-Canon, D. Ganger, M. Persson, F. Gonzalez-Longatt, V. K. Sood, and W. Martinez, "Nadir frequency estimation in low-inertia power systems," *IEEE International Symposium on Industrial Electronics*, vol. 2020-June, pp. 918–922, 6 2020.
- [17] J. Tan, H. Yuan, Y. Zhang, X. Fang, X. Liu, S. Yin, J. Wang, M. Parag, E. Erik, and V. Singhvi, "Multi-timescale integrated dynamic and scheduling model," 2021, NREL SWR-21-97.
- [18] PSS@E — high-performance transmission planning and analysis software. [Online]. Available: <https://new.siemens.com/global/en/products/energy/energy-automation-and-smart-grid/pss-software/pss-e.html>
- [19] H. Yuan, R. S. Biswas, J. Tan, and Y. Zhang, "Developing a reduced 240-bus wecc dynamic model for frequency response study of high renewable integration," in *2020 IEEE/PES Transmission and Distribution Conference and Exposition (T D)*, 2020, pp. 1–5.
- [20] "test case repository for high renewable study". [Online]. Available: <https://www.nrel.gov/grid/test-case-repository.html>
- [21] G. Seber and A. Lee, *Linear Regression Analysis*, ser. Wiley Series in Probability and Statistics. Wiley, 2012.
- [22] J. H. Friedman, "Greedy function approximation: A gradient boosting machine." *The Annals of Statistics*, vol. 29, no. 5, pp. 1189 – 1232, 2001.
- [23] M. Awad and R. Khanna, *Efficient Learning Machines: Theories, Concepts, and Applications for Engineers and System Designers*. Berkeley, CA: Apress, 2015, pp. 67–80.
- [24] I. Goodfellow, Y. Bengio, and A. Courville, *Deep Learning*. MIT Press, 2016, <http://www.deeplearningbook.org>.
- [25] T. Chen and C. Guestrin, "Xgboost," *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, Aug 2016.
- [26] H. Yuan and J. Tan, "Multi-timescale integrated dynamic and scheduling model (midas-solar)," 2020.