

Received January 7, 2022, accepted January 31, 2022, date of publication February 7, 2022, date of current version February 15, 2022.

Digital Object Identifier 10.1109/ACCESS.2022.3149472

Identification of Oscillatory Modes in Power System Using Deep Learning Approach

RAHUL SATHEESH¹, (Member, IEEE), **NASHIDA CHAKKUNGAL**¹,
SUNITHA RAJAN¹, (Senior Member, IEEE), **MANU MADHAVAN**²,
AND HASSAN HAES ALHELOU³, (Senior Member, IEEE)

¹Department of Electrical Engineering, National Institute of Technology Calicut, Calicut, Kerala 673601, India

²Department of Computer Science and Engineering, Amrita School of Engineering, Amrita Vishwa Vidyapeetham, Coimbatore 641112, India

³Department of Electrical Power Engineering, Faculty of Mechanical and Electrical Engineering, Tishreen University, Lattakia 2230, Syria

Corresponding author: Hassan Haes Alhelou (alhelou@ieee.org)

ABSTRACT The increase in electric power demand pushes the modern power system for more interconnected networks. It leads to a lack of inertia and creates more critical disturbances in the power system. When this oscillation isn't damped out, it results in cascade tripping. Immediate detection of low-frequency oscillatory modes and their parameters will help the power system operator to act on a particular event without consuming much time. This research paper proposes novel strategies for identifying low-frequency modes using deep learning techniques, and the model can predict the LFO modes in different topologies. This work presents the Long Short-Term Memory Recurrent Neural Network (LSTM-RNN) approach to predict the instantaneous mode oscillatory parameters in the power system. Once the LSTM-RNN model is trained for different power disturbance situations, it can be used for any events associated with the system. Simulation results are verified using two area Kundur systems at various disturbance conditions. The simulations are performed using MATLAB software and python tensor flow library. The results are validated using statistical methods, and it confirms the superior viability and adaptability of the proposed approach in predicting the instantaneous mode parameters.

INDEX TERMS ANN, inter-area oscillation, low-frequency oscillation, LSTM, Teager Kaiser energy operator, variational mode decomposition.

LIST OF ABBREVIATIONS

ANN	Artificial neural network.	IF	Instantaneous frequency.
ARMA	Autoregressive moving average.	LFO	Low frequency oscillation.
CEEMDAN	Complete ensemble empirical mode decomposition with adaptive noise.	LSTM	Long short term memory.
DR	Damping ratio.	MAE	Mean absolute error.
EMD	Empirical mode decomposition.	MAPE	Mean absolute percentage error.
ERA	Eigen realization algorithm.	MPM	Matrix pencil method.
ESPRIT	Estimation of signal parameters via rotational invariant techniques.	MSE	Mean square error.
EWT	Empirical wavelet transform.	PDC	Phasor data concentrator.
BPNN	Back propagation neural network.	PMU	Phasor measurement unit.
GRU	Gated recurrent unit.	RBFNN	Radial basis function neural network.
HHT	Hilbert Huang transform.	RNN	Recurrent neural network.
HT	Hilbert transform.	TKEO	Teager Kaiser Energy operator.
IA	Instantaneous amplitude.	VMD	Variational mode decomposition.

The associate editor coordinating the review of this manuscript and approving it for publication was Ali Raza ^{id}.

I. INTRODUCTION

Due to the rapidly changing electrical dynamics, the modern interconnected power system is inevitably disturbed by various oscillation events. Electromechanical oscillations may

occur in different frequencies, and it is not dangerous if they decay quickly. The stability of the power system will be wrecked if there is no proper damping [1], [2]. The analysis of low-frequency oscillatory modes and their characteristics lead to an adequate understanding of the dynamic performance of the power system. It will give productive inputs to the operator for prevention and control. Due to the issues impacted by LFO, the capability of monitoring grid operations in real-time is critical for the safe and reliable operation of the grid. Underdamped oscillations lead to significant power swings and tripping of protective relays, resulting in the disconnection of loads [3].

Advanced wide-area monitoring of power grid systems associated with phasor measurements units (PMU) can continuously evaluate the health of the power system. Dynamic monitoring of the power system for real-time operation and control has risen in the last two decades. Researchers have proposed various linear and nonlinear approaches to assess the dynamic responses and estimate the parameters of dominant low-frequency oscillatory modes. Power system modes are evaluated using two methods: Modal based approach and Measurement-based approach. The former works on linearizing the governing equation about an operating point [4], and the latter purely follow a data-driven analysis on the system measurement data [5]. IEEE task force on identifying electromechanical oscillatory modes summarizes different techniques used in modal and data-driven approaches [6]. The applicability of model-based techniques in a large-scale power system is constrained, and computation time is too long. Simultaneously, measurement-based methods are used widely to identify the low-frequency modes with synchrophasor technology. The measurement-based techniques are found in many works of literature. Some of them are Prony analysis [7], matrix pencil method (MPM) [8], estimation of signal parameters via rotational invariant techniques (ESPRIT) [9], autoregressive moving average (ARMA) technique [10], and eigenvalue realization algorithm (ERA) [11]. The above mentioned methods are used in ring-down oscillation studies. The methods used for ambient oscillation studies include transfer function methods [12], [13] and subspace methods [14], [15]. The subspace approach gives better results in terms of accuracy, but transfer function methods are preferred in terms of computational time [16].

The majority of power oscillation analysis follows a stationary nature. However, nonstationary analysis is preferred for power oscillation studies and parameter estimation. Hilbert Huang transform (HHT) is very popular in this class, and it comprises of two techniques, Empirical mode decomposition (EMD) and Hilbert transform (HT) [17], [18]. EMD is a well-known signal decomposition technique and HT is used for instantaneous parameter estimation. EMD is modified into ensemble EMD and complete ensemble EMD with adaptive noise (CEEMDAN) [19] based on their denoising power. The drawbacks of EMD are overcome in Variational Mode Decomposition (VMD), a signal decomposition technique with non-recursive nature suggested by

Dragomiretskiy *et al.* for modal parameter estimation [21]. It exhibits better denoising properties and has a credible theoretical foundation. An adaptive Wiener filter bank is employed to efficiently decompose the center frequency test signal into a restricted bandwidth. The work presented in this paper uses VMD with Teager Kaiser Energy Operator (TKEO) to estimate instantaneous parameters [22]–[24]. TKEO is an estimation method that is highly robust than the HT method [25]. This approach is used in a supervised manner for the deep learning techniques adopted in this paper.

Recent trends show that Artificial Neural Network (ANN) methods are more powerful and reliable and provide better results in real-time applications. Besides, the proposed method should track LFOs in real-time and thus entail wide-area PMU data. Therefore measurement-based estimation techniques combined with ANN-based methods extract useful information from the power system. ANN-based power oscillation damping control is suggested in a few pieces of literature [26], [27]. Because of massive data from synchrophasors, initial data is appropriately subjected to pre-processing stage and dimensionality reduction. It results in a marginalized reduction in input data and network size, and hence computation time for the offline training can be reduced. In recent years deep learning techniques have become one of the most efficient tools in many research areas. It is successfully implemented in power systems, particularly short-term load forecasting problems and disturbance classification [28]. Other works include the deep learning approach of cost loss function for transient stability assessment proposed by Zhou *et al.* [29] and prediction of load demand in the smart grid using the LSTM network presented by Cheng *et al.* [30].

This paper presents a deep learning method using LSTM-RNN to estimate low frequency mode parameters. This method is compared with the results obtained using conventional back propagation neural network (BPNN), radial basis function neural network (RBFNN), and gated recurrent unit (GRU) architecture. The results of the Kundur two-area system is verified using MATLAB and python tensor flow library. Here the analysis pattern follows a feature imperative strategy method that enhances the determination of dominant modes in the power system oscillation. Hence an operator can quickly determine the critical modes in the power system. Also, Deep learning methodologies enhance the data visualization possibilities of the dominant mode situation.

The significant contributions of this study are as follows (1) The identification of oscillatory modes in power systems using deep learning techniques are discussed

(2) The LSTM-RNN model is trained using the advanced signal processing methods of VMD and TKEO method.

(3) The proposed method of LSTM-RNN offers lesser computation time and memory storage than the existing learning methods.

The rest of the paper is organized as follows. Section II describes the methodology of synchrophasor technology, VMD approach, TKEO technique, and LSTM architecture.

Section III describes results and discussion, and finally, the concluding remarks are given in Section IV.

II. METHODOLOGY

This section introduces the standard synchrophasor data acquisition procedure, signal decomposition using VMD, estimation of instantaneous parameters using TKEO and LSTM architecture. It also states the importance of LSTM over other learning approaches.

A. SYNCHROPHASOR DATA ACQUISITION

The PMUs can measure time-stamped measurements of positive-sequence voltages and currents of all buses and feeders where it is installed in addition to frequency and rotor angle. The system voltage is collected using PMU connected to the buses. Fig.1. represents the process of data acquisition using PMUs. The measured data sets are stored in a phasor data concentrator (PDC) which rejects the bad data and stores all the information for further analysis. Another advantage of synchrophasor is that it can correlate data measured from the different substations to the common time reference, and hence a wide-area network status can be accessed. When a disturbance occurs, the data is collected and undergoes a post-disturbance analysis. One such method consisting of VMD and TKEO is discussed in the following subsection.

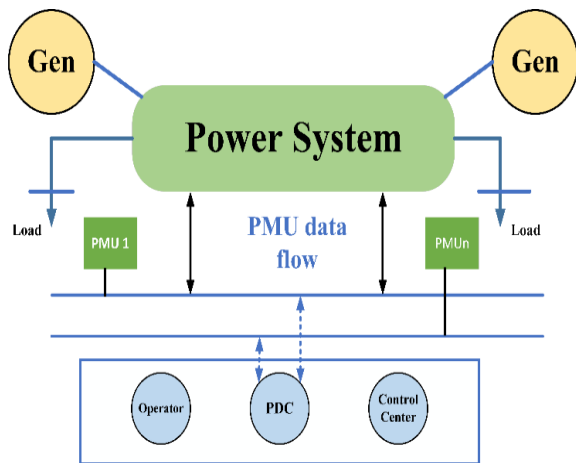


FIGURE 1. The process of collecting data using PMU.

B. VARIATIONAL MODE DECOMPOSITION

The preprocessed data is decomposed using the VMD technique, and the instantaneous parameters like instantaneous amplitude (IA), instantaneous frequency (IF), and damping ratio (DR) were estimated using the HT method. This section will first discuss the VMD technique and demonstrate the HT platform in the later session. VMD is a multi-resolution analytical signal decomposition method based on the concepts of adaptive Wiener filtering, one-dimensional Hilbert transform, and Heterodyne demodulation [21]. VMD's motive is to decompose a real-valued nonlinear nonstationary signal $f(t)$ into a set of discrete sets of quasi-orthogonal intrinsic

mode functions (IMF) represented as u_k , where K denotes the mode number. This set of IMF signals are regarded as modulated amplitude, and frequency signals with a center frequency of ω_k . VMD require the subsequent computational processes as follows:

- (1) Hilbert Transform is applied to the one-sided spectrum of each of the IMFs to compute its signal characteristics.
- (2) A multiplication factor of $e^{-j\omega_k t}$ is considered to shift the frequency spectrum of mode to baseband.
- (3) The estimation of bandwidth using gradient of modulated signal based on the L_2 norm.

The VMD method is assumed as a constrained optimization as in (1)

$$\min (\omega_k, u_k) \left\langle \sum_{k=1}^K \left\| \delta_t \left[\left(\delta(t) + \frac{i}{\pi t} \right) * u_k(t) \right] e^{-i\omega_k t} \right\|_2 \right\rangle \quad (1)$$

such that $\sum_{k=1}^K u_k(t) = f(t)$.

The objective function is modified into an unconstrained optimization problem as in (2)

$$\begin{aligned} L(\{u_k\}, \{\omega_k\}, \{\lambda\}) &= \alpha \sum_{k=1}^K \left\| \delta_t \left[\left(\delta(t) + \frac{i}{\pi t} \right) * u_k(t) \right] e^{-i\omega_k t} \right\|_2^2 \\ &+ \left\| f(t) - \sum_{k=1}^K u_k(t) \right\|_2^2 \\ &+ \left\langle \lambda, f(t) - \sum_{k=1}^K u_k(t) \right\rangle \end{aligned} \quad (2)$$

VMD measures these central frequencies and IMFs at these frequencies concurrently using an optimization technique called the alternate direction method of multipliers [36]. The precise framework of the optimization problem in the time domain is continuous. The various modes are determined by updating the previous mode and center frequency, using equations (3) and (4).

$$\hat{u}_k^{n+1} = \frac{f - \sum_{i < k} \hat{u}_i^{n+1} - \sum_{i > k} \hat{u}_i^n + \frac{\lambda^n}{2}}{1 + 2\alpha(\omega - \omega_k^n)^2} \quad (3)$$

$$\hat{\omega}_k^{n+1} = \frac{\int_0^\omega \omega |\hat{u}_k^{n+1}(\omega)|^2 d\omega}{\int_0^\omega |\hat{u}_k^{n+1}(\omega)|^2 d\omega} \quad (4)$$

The updation of the modes and center frequencies are made, and the Lagrangian multiplier is also restructured as in equation (5).

$$\hat{\lambda}^{n+1} = \hat{\lambda}^n + \left(f - \sum_k \hat{u}_k^{n+1} \right) \quad (5)$$

The mode update process is performed until it converges to a tolerance factor using equation (6), where ε is the tolerance factor.

$$\sum_k \frac{\|\hat{u}_k^{n+1} - \hat{u}_k^n\|_2^2}{\|\hat{u}_k^n\|_2^2} < \varepsilon \quad (6)$$

Based on the mode number, IMFs are extracted. Parameters like fidelity factor (α) and mode number (K) are needed to initialize the VMD operation. These two parameter values are randomly allocated, leading to needless decomposition stages. In this work, the data samples from the PMU have undergone Fourier spectra [25]. The number of peaks in spectra was identified and assigned as the mode number. In the fidelity factor, typically for low-frequency extraction, higher values of α are preferred, and it is assumed as 8000 for the VMD process.

C. TEAGER KAISER ENERGY OPERATOR

TKEO monitors the modulation energy and determines the instantaneous amplitude and frequency. A highly nonlinear energy operator is considered here and the TKEO $\psi(\cdot)$ is one of the finest options. A continuous signal $c(t)$ is defined as

$$\psi(c(t)) = [\dot{c}(t)]^2 - c(t)\ddot{c}(t) \quad (7)$$

The energy operator in discrete form is given by

$$\psi(c(n)) = [c(n)]^2 - c(n+1)c(n-1), \quad (8)$$

where $c(n)$ represents the discrete-time signal. The operator has a better time resolution in seizing energy fluctuations. The instantaneous amplitude (IA) and frequency (IF) are shown as

$$IA(n) = \frac{2\psi[c(n)]}{\sqrt{\psi[c(n+1) - c(n-1)]}} \quad (9)$$

$$IF(n) = \frac{1}{4} \arccos \left\{ 1 - \frac{\psi[c(n+1) - c(n-1)]}{2\psi[c(n)]} \right\} \quad (10)$$

Instantaneous values of damping ratio (DR) are calculated by (15)

$$DR(n) = \frac{IF(n)}{\sqrt{IF^2(n) + \omega_n^2}} \quad (11)$$

where ω_n be the natural frequency. The TKEO method gives a quality estimate of instantaneous parameters with less computational complexity.

In all the cases analyzed in this paper, the properly pre-processed input data [31], [32] is fed to the VMD process, and the mode parameters are estimated through the TKEO. This kind of instantaneous parameter estimation is described in the literature in [25]. These results were taken as the actual values for comparison with learning algorithms. LSTM-RNN architecture is described in the following subsection of the methodology part.

D. LSTM ARCHITECTURE

A normal BPNN assumes the data instances are independent and cannot handle sequential data such as text and time-series data. RNN is a class of neural networks, and its cyclic nature gives the ability to work with temporal data. Two of them are famous; one is from Jordan and Elman. Jordan network is a simple neural network and is formulated as

$$h^t = \sigma(W_h X + W_r y^{t-1}) \quad (12)$$

$$y = \sigma(W_y h^t) \quad (13)$$

Here in this equation X is the input, h is the hidden representation, y be the output, σ be the activation function, W_h is the weight of hidden layers, W_y is the weight of the output layer, and W_r weight of the recurrent computation. Elman proposed a slight modification in information from the previous time step provided by the previously hidden layer, and selection of W_r is also different from the Jordan technique [33], [34].

After these inventions, bidirectional RNN was invented by Schuster and Paliwal [35]. This first hidden layer is unfolded using basic RNN, and the second hidden layer is simulated in a reversed connection. Then the backpropagation can be applied on a time scale, and the weights are updated. This is the basic concept of LSTM-RNN. The disadvantage of this technique is the vanishing gradient problem stated by *et al.* [36]. It deals with the traditional activation functions in which the gradient is bounded. Usually, the gradients are calculated by backpropagation, and its error values decrease exponentially within the time steps, which eventually tends to long-term dependency loss. LSTM has been introduced and incorporated by a specially designed memory cell unit called an LSTM cell to overcome the vanishing gradient problem. Also, LSTM avoids the long-term dependencies of RNN.

LSTM has different critical parameters in terms of states and gates, in which states are the values that offer information for output, and the gates decide the information track of states. Input state I is defined as

$$I^t = \sigma(W_{ix}x^t + W_{ih}h^{t-1}) \quad (14)$$

where h for the value of the hidden layer otherwise hidden state and x is the input data.

One more state is there rather than the input state is an internal state, and it serves as a memory, denoted as m . Gates are mainly of three types: input gate, forget gate, and output gate. The input gate decides whether the input state enters the internal state and is denoted as g and represented as

$$g^t = \sigma(W_{gi}I^t) \quad (15)$$

The forget gate adopts whether the internal state forgets the previous internal state and is denoted as f' and represented as

$$f^t = \sigma(W_{fi}I^t) \quad (16)$$

The output gate determines if the internal state passes its value to the output and to the next hidden state. It is denoted as o and represented as

$$o^t = \sigma(W_{oi}I^t) \quad (17)$$

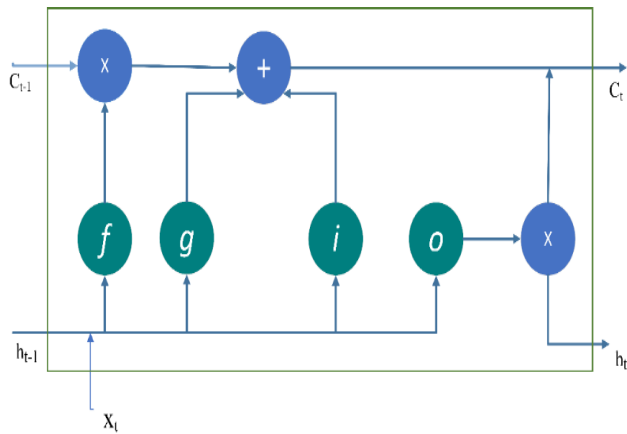


FIGURE 2. LSTM cell.

Generally, the complete formulation of the LSTM network is represented as

$$m^t = g^t * I^t + (f^t m^{t-1}) \tag{18}$$

$$h^t = o^t * m^t \tag{19}$$

The symbol * represents the element-wise multiplication of parameters, and the whole LSTM architecture is represented symbolically, as in Fig.2.

Hyperparameters of deep neural networks are customized by their optimization or tuning techniques. Searching for hyperparameters that result in the best model performance on a specified set of data is called hyperparameter optimization. It involves defining a search space with the volume to be searched in which each dimension represents a hyperparameter, and each point represents a model configuration [37]. First tuned for batch size and epoch, and take the batch size of 10, 20, 40, 60,80, and 100, epochs of 100,150, 200, and 250, and tuned different combinations of this batch size and epochs finally get 20 batch size, and 200 epochs are the best parameters. Next is the grid search for the number of hidden layers and the number of neurons in the hidden layers. Search number of neurons in the dictionary of [10], [10, 10], [10, 20], [10, 20, 30]. From which [10] shows the best results that are the bidirectional neural network. It has one hidden layer and ten neurons.

As a result of an activation function, a node or nodes within a layer of a neural network are transformed into an output based on the weighted sum of the input. A rectified Linear activation (*ReLU*) function is used, and it gives an output x , if x is positive and otherwise zero. Thus $A(x) = \max(0, x)$. Due to simpler mathematical operations, *ReLU* is a less computationally expensive option than *tanh* or *sigmoid*. Only a few neurons are activated, making the network sparse and efficient, leading to easy computation.

The assessment of LSTM-RNN architecture towards PMU data is as follows. The datasets from the PMU units are collected and evaluated; these are called testing datasets.

The next step is the preprocessing of data in which the data is subjected to noise filtering, checking for missing values, and averaging to the mean value. This phase is necessary to make the data set comfortable with the algorithm and reduce computation time. Training and testing data can be calculated as

$$\text{Training data} = PQ * 80\%$$

$$\text{Testing data} = PQ * 20\%$$

Here the P represents the number of samples Q is the input variable. 80% is the percentage target of the training data, and 20% defines the percentage target of testing data.

Using the Python tensor flow library is used to evaluate the LSTM network. The decision for each weight of the LSTM is apprehended in quantity to the derivative of the error. There are three hidden gate layers in each hidden node. A comparison between the estimated accuracy of the algorithm and the actual response is made to determine the learning rate and activation. Then the performance of LSTM accuracy is an evaluation of Mean Absolute Error (MAE), Mean Square Error (MSE), and Mean Absolute Percentage Error (MAPE).

$$MAE = \frac{\sum |\bar{\lambda}_t - \lambda_t|}{S} \tag{20}$$

$$MSE = \frac{\sum |\bar{\lambda}_t - \lambda_t|^2}{S} \tag{21}$$

$$MAPE = \frac{1}{S} \sum \left| \frac{\bar{\lambda}_t - \lambda_t}{\lambda_t} \right| \times 100\% \tag{22}$$

where S represents the number of testing data, $\bar{\lambda}_t$ represents identified data, and λ_t is the target data. The complete algorithm is programmed and implemented in Keras with tensor flow backend on a desktop computer with 3.2 GHz.

The raw frequency data from various PMU units over different nodes should be preprocessed adequately in the first stage. It should filter out noise from the measurements, detrend the data, and set up a proper window frame for offline training. After the preprocessing stage, different power disturbance cases were analyzed and fed to ANN or LSTM approach in a sliding window manner. The estimated instantaneous values of amplitude, frequency, and damping ratio are the results for each mode and it is the output. Fig.3. represents the overall approach for implementing training and testing of the learning algorithm. The present study focuses only on applying the learning algorithms to estimate parameters in the power oscillations. The results should be verified with actual values. ANN approaches of backpropagation neural network and radial basis function are tested. Also, the GRU technique is tested with the dataset, and finally, the LSTM-RNN approach is implemented for the instantaneous parameter prediction with a good fit LSTM model. The results and discussion section describe the dataset generation and application of various learning approaches in parameter estimation.

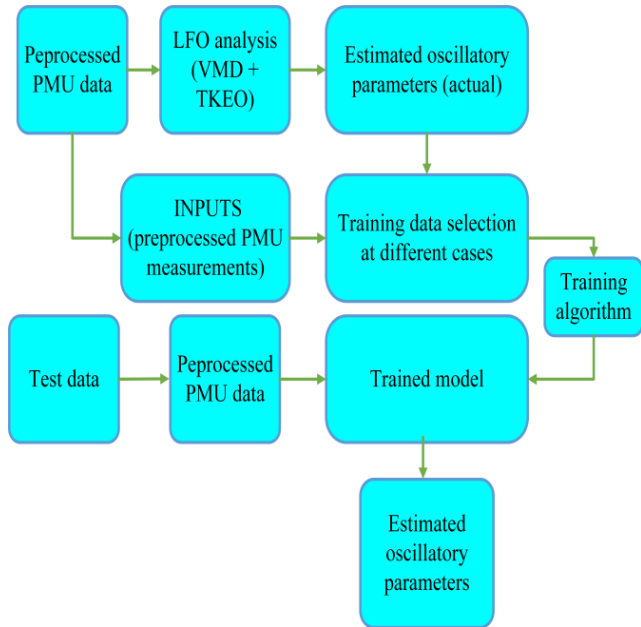


FIGURE 3. Scheme for training using a learning approach.

III. RESULTS AND DISCUSSION

A. DATA SET CREATION

The application of the proposed approach to a two-area Kundur system is discussed in this section and is shown in Fig.4. A post-event analysis is carried out based on the measurements obtained via the standard system. The system consists of 11 buses, four generators, and two areas. The two areas are connected through a weak tie-line between buses 7 and 9. A rotor speed signal is recorded during the 20-second simulation period at 60 samples per second. Generator three is considered for reference, and a few disturbances are applied to the system to create a suitable training model. VMD technique is applied to the measurements, and the predefined parameters are assigned. The mode number is assumed to be six and the fidelity factor to be 8000. The measurements are decomposed into six modes by applying VMD, and the IMF is selected by using the correlation coefficient. Based on the appropriate IMF, instantaneous parameters are estimated using the TKEO method. It includes instantaneous amplitude, frequency, and damping ratio. This method is trained using various disturbances like three-phase fault, line removal, varying the load, etc., and is shown in Fig.5. Similarly, so many disturbance cases can be created for the data generation, and the result is used for the training purpose of learning techniques. IMF3 is preferred based on the excellent correlation coefficient and is used to estimate the instantaneous parameters.

B. APPLICATION OF VARIOUS LEARNING APPROACHES

Based on the generated sample set, various learning algorithms and their application on parameter estimation are verified in this subsection. The parameter summary of various

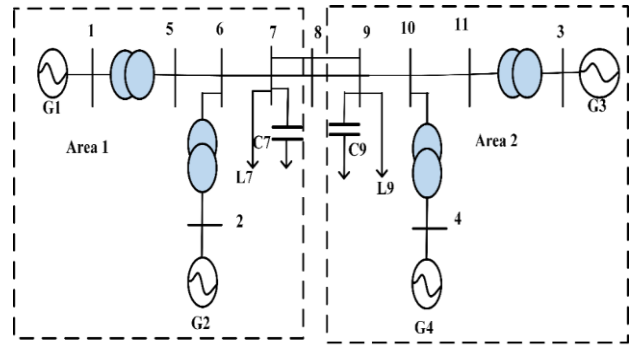


FIGURE 4. Two area interconnected system.

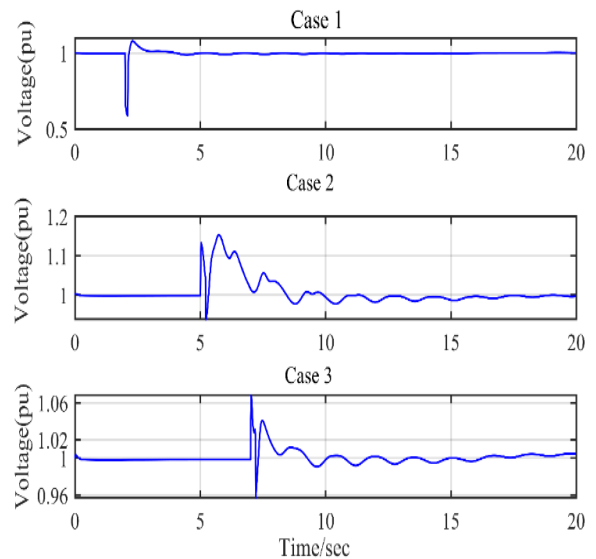


FIGURE 5. Various cases of disturbances of two area system.

learning algorithms and the statistical error parameters are shown in TABLE 1 and TABLE 2 respectively.

ANN models use a feed-forward backpropagation neural network that sets the maximum number of epochs to 500. After each epoch, the network weight is adjusted and biased to the minimum error value. TANSIG transfer function gives lower MSE. The model is trained with different numbers of hidden layer neurons 1, 2, 5, 8, and 10, 12, 15. Out of these hidden layers, 20 neurons show a better least mean square error of 0.0010, and the TRAIN-LM algorithm gives faster results. It is verified with both 10 and 60 time steps, and when we increase the time steps, the errors become smaller. The ANN model shows good performance and an overall regression value of 0.96574 for mode 3, as shown in Fig.6. Regression means the correlation between output and target and the R-value estimates as one indicates a close relationship. The trained models are tested using another disturbance within the system by applying a three-phase fault at 6 seconds between bus 7 and 8, as shown in Fig.7. Comparison of BPNN estimate with actual values is visualized in Fig.8. The RBFNN is also adapted with different time steps.

The results show that the MAE parameter reduces much better than backpropagation. It still shows performance difficulties in the nonlinear problems and needs to catch this by upgrading to another learning technique. However, it is not easy to follow the peak when the amplitude varies dramatically. The ANN techniques are unsuitable for nonlinear systems, indicating that more accurate approaches are necessary to estimate instantaneous parameters [38].

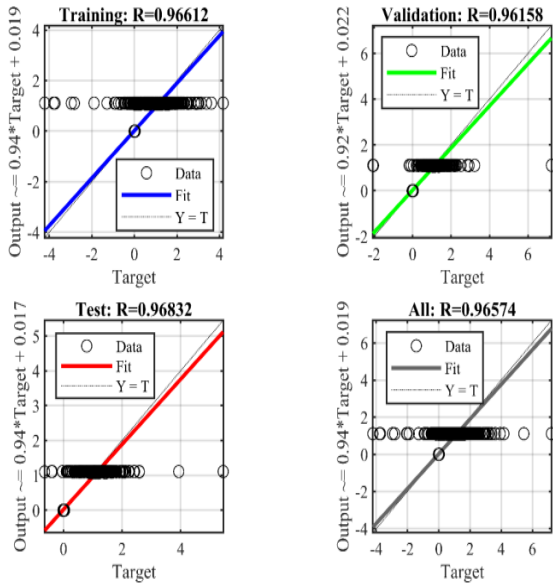


FIGURE 6. Regression curve for the trained model using BPNN.

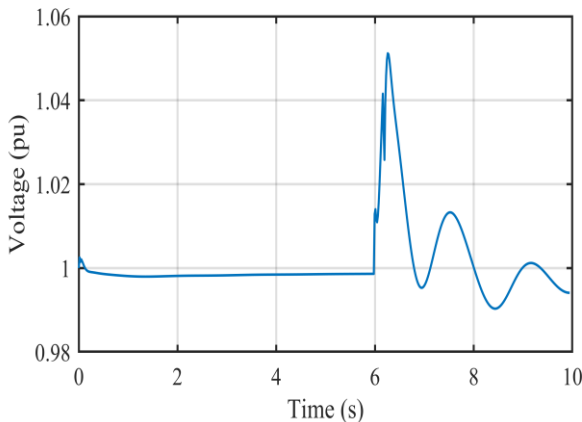


FIGURE 7. Test data for the assessment.

In the case of GRU can capture the dependencies of input and output parameters. It achieves better performance than ANN methods regarding statistical parameters shown in TABLE 2. GRU with 60 time steps offer better performance than with 10 or 1 time step. Its MAE values are smaller than BPNN and more than the RBFNN technique. The identified model obtained through GRU can be modified using better learning techniques.

TABLE 1. Parameter summary.

Technique	layers	Neurons	Epoch
BPNN	3	20	500
RBFNN	3	varying	200
GRU	3	15	100
LSTM	3	10	100

TABLE 2. Estimated MAE, MSE, and MAPE values of parameters (IF).

BPNN		
MAE	MSE (x10 ²)	MAPE (%)
0.1956	325.63	26.85
RBFNN		
MAE	MSE (x10 ²)	MAPE (%)
0.136	152.64	22.57
GRU		
MAE	MSE (x10 ²)	MAPE (%)
0.152	86.364	18.63
LSTM		
MAE	MSE (x10 ²)	MAPE (%)
0.00153	46.25	8.58

TABLE 3. Comparison of computational strength.

Parameter	LSTM	GRU	RBFNN	BPNN
CPU time (ms)	42	58	72	86
Memory use (MB)	215	326	385	390

LSTM-RNN models are also trained using the mode estimation results. The instantaneous parameters are taken as target data and input as preprocessed frequency signals from PMU. Optimization helps the model to reduce losses and provide the most accurate results. This model uses adaptive moment estimation because it is too fast and converges rapidly. Out of the total samples, 80% is used for training and 20% for validation. The convergence characteristic of various algorithms of BPNN, RBFNN, and GRU is compared with the LSTM network. LSTM ensures a good fit model for the estimation of parameters, and it reflects in the convergence characteristics.

Overfitting and underfitting analysis are the primary purposes of investigating training and validation errors. The underfitting analysis should be considered if training performance is lower than validation. Also, if it is vice versa, the overfitting analysis should be considered. Training loss and validation loss are close for a properly trained model, with validation loss being slightly greater than the training loss. Comparison of LSTM estimate with actual values is visualized in Fig.9. The convergence characteristics in Fig.10 show the better performance of LSTM. It is observed that LSTM shows better convergence and minimum error compared with

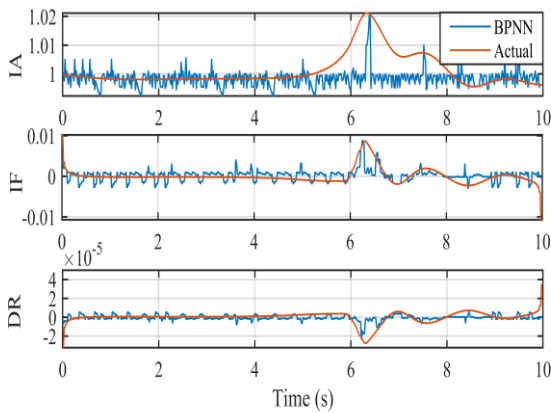


FIGURE 8. Comparison of BPNN estimate with actual values.

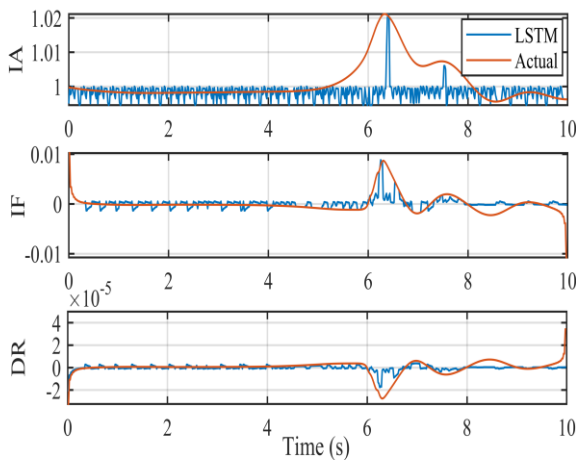


FIGURE 9. Comparison of LSTM estimate with actual values.

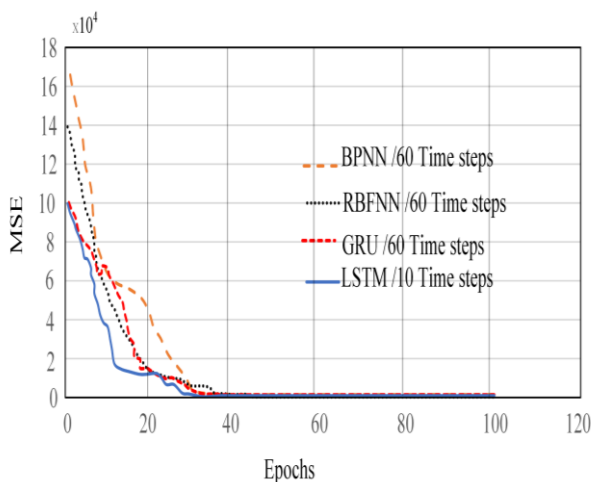


FIGURE 10. Convergence characteristics of various algorithms.

the ANN techniques and GRU. The increase in the time steps of LSTM will improve the convergence as per the studies.

For a well-trained LSTM, the identifying performance is independent of the previous system configuration

and parameters. The minimum values of MAE, MAPE, and MSE values of LSTM in TABLE 2 reflect the quality of the deep learning approach in LSTM in applying power oscillation analysis. As per the convergence characteristics, LSTM takes 26 epochs with 10 time steps for converging, and it implies the higher training quality in the results.

Overall CPU time and memory storage are verified for different methods are shown in TABLE 3. LSTM approach dominates in this analysis too, with its more incredible converging speed. From the final results of convergence and statistical analysis, it is observed that Ann techniques and GRU sacrifice more input characteristics and time steps to get a better result. Meanwhile, the LSTM method always follows the peak and uses fewer neurons and time steps.

IV. CONCLUSION

This research paper is a preliminary step towards a more extended project of a complex real-time power system environment. In this LSTM architecture is used to predict the dominant modes or the instantaneous behaviour of the system. The proposed technique is based on the offline training of the LSTM technique and is compared with conventional approaches like BPNN, RBFNN, and GRU techniques. The development of the training set is done through a preprocessing stage, VMD and TKEO combination is used for the mode estimation. The proposed technique is validated using two area Kundur system with a simulated test signal, and its effectiveness is verified using statistical parameters like MAE, MAPE, and RMSE values. Convergence characteristics were also plotted based on the learning techniques used in this article. LSTM results offer the lowest error values than the conventional learning techniques. Through this research, paper authors are trying to establish the importance of the deep learning approach in the estimation of instantaneous oscillatory parameters in power systems.

ACKNOWLEDGMENT

The authors gratefully acknowledge the contributions of the Power System Operation Corporation, Ltd., Bangalore, India, for their technical assistance.

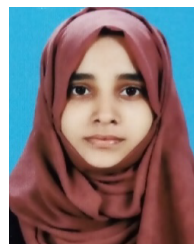
REFERENCES

- [1] D. Gautam, V. Vittal, and T. Harbour, "Impact of increased penetration of DFIG based wind turbine generators on transient and small signal stability of power system," *IEEE Trans. Power Syst.*, vol. 24, no. 3, pp. 1426–1434, Aug. 2009.
- [2] G. Rogers, *Power System Oscillations*. Boston, MA, USA: Springer, 2000, doi: 10.1007/978-1-4615-4561-3.
- [3] P. Kundur, *Power System Stability and Control*, 12th ed. New Delhi, India: McGraw-Hill, 2011.
- [4] S. Okubo, H. Suzuki, and K. Uemura, "Modal analysis for power system dynamic stability," *IEEE Trans. Power App. Syst. (through 1985)*, vol. PAS-97, no. 4, pp. 1313–1318, Jul. 1978.
- [5] P. Ray, "Power system low frequency oscillation mode estimation using wide area measurement systems," *Eng. Sci. Technol., Int. J.*, vol. 20, no. 2, pp. 598–615, Apr. 2017.
- [6] J. Sanchez-Gasca et al., "Identification of electromechanical modes in power system," IEEE Task Force Rep., IEEE PES, Piscataway, NJ, USA, Tech. Rep. TP462, 2012. [Online]. Available: <https://resourcecenter.ieee-pes.org/publications/technical-reports/PESTR15.html>

- [7] K. S. Shim, H. K. Nam, and Y. C. Lim, "Use of Prony analysis to extract sync information of low frequency oscillation from measured data," *Eur. Trans. Electr. Power*, vol. 21, no. 5, pp. 1746–1762, Jul. 2011.
- [8] T. Sarkar and O. Pereira, "Using the matrix pencil method to estimate the parameters of a sum of complex exponentials," *IEEE Antennas Propag. Mag.*, vol. 37, no. 1, pp. 48–55, Feb. 1995.
- [9] J. G. Philip and T. Jain, "Analysis of low frequency oscillations in power system using EMO ESPRIT," *Int. J. Electr. Power Energy Syst.*, vol. 95, pp. 499–506, Feb. 2018.
- [10] C. Jakkattanajit, N. Hoonchareon, and A. Yokoyama, "On-line estimation of power system low frequency oscillatory modes in large power systems," *J. Int. Council Electr. Eng.*, vol. 1, no. 3, pp. 352–358, Jul. 2011.
- [11] A. Almunif, L. Fan, and Z. Miao, "A tutorial on data-driven eigenvalue identification: Prony analysis, matrix pencil, and eigensystem realization algorithm," *Int. Trans. Electr. Energy Syst.*, vol. 30, no. 4, Apr. 2020, Art. no. e12283.
- [12] J. R. Smith, F. Fatehi, C. S. Woods, J. F. Hauer, and D. J. Trudnowski, "Transfer function identification in power system applications," *IEEE Trans. Power Syst.*, vol. 8, no. 3, pp. 1282–1290, Aug. 1993.
- [13] H. Khalilinia, L. Zhang, and V. Venkatasubramanian, "Fast frequency-domain decomposition for ambient oscillation monitoring," *IEEE Trans. Power Del.*, vol. 30, no. 3, pp. 1631–1633, Jun. 2015.
- [14] T. Jiang, H. Yuan, H. Jia, N. Zhou, and F. Li, "Stochastic subspace identification-based approach for tracking inter-area oscillatory modes in bulk power system utilising synchrophasor measurements," *IET Gener., Transmiss. Distrib.*, vol. 9, no. 15, pp. 2409–2418, Nov. 2015.
- [15] S. A. N. Sarmadi and V. Venkatasubramanian, "Electromechanical mode estimation using recursive adaptive stochastic subspace identification," *IEEE Trans. Power Syst.*, vol. 29, no. 1, pp. 349–358, Jan. 2014.
- [16] L. Dosiek, N. Zhou, J. W. Pierre, Z. Huang, and D. J. Trudnowski, "Mode shape estimation algorithms under ambient conditions: A comparative review," *IEEE Trans. Power Syst.*, vol. 28, no. 2, pp. 779–787, May 2013.
- [17] D. S. Laila, A. Messina, and B. C. Pal, "A refined Hilbert–Huang transform with applications to inter-area oscillation monitoring," *IEEE Trans. Power Syst.*, vol. 24, no. 2, pp. 610–619, May 2013.
- [18] A. Prince, N. Senroy, and N. Balasubramanian, "Targeted approach to apply masking signal-based empirical mode decomposition for mode identification from dynamic power system wide area measurement signal data," *IET Gener., Transmiss. Distrib.*, vol. 5, no. 10, pp. 1025–1032, 2011.
- [19] M. A. Colominas, G. Schlotthauer, and M. E. Torres, "Improved complete ensemble EMD: A suitable tool for biomedical signal processing," *Biomed. Signal Process. Control*, vol. 14, pp. 19–29, Nov. 2014.
- [20] S. Dutta, J. G. Philip, and T. Jain, "A Fourier-EWT approach for identifying the low frequency modes present in the power system," in *Proc. 15th IEEE India Council Int. Conf. (INDICON)*, Dec. 2018, pp. 1–6.
- [21] K. Dragomiretskiy and D. Zosso, "Variational mode decomposition," *IEEE Trans. Signal Process.*, vol. 62, no. 3, pp. 531–544, Feb. 2014.
- [22] M. R. A. Paternina, R. K. Tripathy, A. Zamora-Mendez, and D. Dotta, "Identification of electromechanical oscillatory modes based on variational mode decomposition," *Electr. Power Syst. Res.*, vol. 167, pp. 71–85, Feb. 2019.
- [23] S. Rahul, S. Koshy, and R. Sunitha, "Electromechanical mode estimation in power system using a novel nonstationary approach," in *Wide Area Power Systems Stability, Protection, and Security (Power Systems)*, H. H. Alhelou, A. Y. Abdelaziz, and P. Siano, Eds. Cham, Switzerland: Springer, 2020.
- [24] S. Rahul and R. Sunitha, "Dominant electromechanical oscillation mode identification using modified variational mode decomposition," *Arabian J. Sci. Eng.*, vol. 46, no. 10, pp. 10007–10021, Oct. 2021.
- [25] H. Xiao, J. Wei, H. Liu, Q. Li, Y. Shi, and T. Zhang, "Identification method for power system low-frequency oscillations based on improved VMD and Teager–Kaiser energy operator," *IET Gener., Transmiss. Distrib.*, vol. 11, no. 16, pp. 4096–4103, Nov. 2017.
- [26] N. I. A. Wahab, A. Mohamed, and A. Hussain, "Fast transient stability assessment of large power system using probabilistic neural network with feature reduction techniques," *Expert Syst. Appl.*, vol. 38, no. 9, pp. 11112–11119, Sep. 2011.
- [27] S.-S. Jhang, H.-J. Lee, C.-N. Kim, C.-H. Song, and W.-K. Yu, "ANN control for damping low-frequency oscillation using deep learning," in *Proc. Australas. Univ. Power Eng. Conf. (AUPEC)*, Nov. 2018, pp. 1–4.
- [28] W. L. Rodrigues, F. A. S. Borges, A. O. de Carvalho Filho, and R. D. A. L. Rabelo, "A deep learning approach for the detection and classification of power quality disturbances with windowed signals," *Soc. Netw. Comput. Sci.*, vol. 2, no. 2, pp. 1–14, Apr. 2021.
- [29] Y. Zhou, W. Zhao, Q. Guo, H. Sun, and L. Hao, "Transient stability assessment of power systems using cost-sensitive deep learning approach," in *Proc. 2nd IEEE Conf. Energy Internet Energy Syst. Integr. (EI2)*, Oct. 2018, pp. 1–6.
- [30] J. Zheng, C. Xu, Z. Zhang, and X. Li, "Electric load forecasting in smart grids using long-short-term-memory based recurrent neural network," in *Proc. 51st Annu. Conf. Inf. Sci. Syst. (CISS)*, Mar. 2017, pp. 1–6.
- [31] L. Vanfretti, S. Bengtsson, and J. O. Gjerde, "Preprocessing synchronized phasor measurement data for spectral analysis of electromechanical oscillations in the Nordic Grid," *Int. Trans. Electr. Energy Syst.*, vol. 25, no. 2, pp. 348–358, Feb. 2015.
- [32] E. A. Felix and S. P. Lee, "Systematic literature review of preprocessing techniques for imbalanced data," *IET Softw.*, vol. 13, no. 6, pp. 479–496, Dec. 2019.
- [33] W. De Mulder, S. Bethard, and M.-F. Moens, "A survey on the application of recurrent neural networks to statistical language modeling," *Comput. Speech Lang.*, vol. 30, no. 1, pp. 61–98, 2015.
- [34] Y. Yu, X. Si, C. Hu, and Z. Jianxun, "A review of recurrent neural networks: LSTM cells and network architectures," *Neural Comput.*, vol. 31, no. 7, pp. 1235–1270, Jul. 2019.
- [35] M. Schuster and K. K. Paliwal, "Bidirectional recurrent neural networks," *IEEE Trans. Signal Process.*, vol. 45, no. 11, pp. 2673–2681, Nov. 1997.
- [36] R. Pascanu, T. Mikolov, and Y. Bengio, "On the difficulty of training recurrent neural networks," in *Proc. ICML*, 2013, vol. 28, no. 3, pp. 1310–1318.
- [37] D. Bouneffouf and E. Claeys, "Online hyper-parameter tuning for the contextual bandit," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process. (ICASSP)*, Jun. 2021, pp. 3445–3450.
- [38] C. Olivieri, F. de Paulis, A. Orlandi, C. Pisani, G. Giannuzzi, R. Salvati, and R. Zaotini, "Estimation of modal parameters for inter-area oscillations analysis by a machine learning approach with offline training," *Energies*, vol. 13, no. 23, p. 6410, Dec. 2020.



RAHUL SATHEESH (Member, IEEE) received the B.Tech. degree in electrical and electronics engineering from the University of Kerala, India, and the M.E. degree in power systems from Anna University, Chennai, India, in 2012 and 2014, respectively. He is currently pursuing the Ph.D. degree with the Department of Electrical Engineering, National Institute of Technology Calicut, Kerala. He has published several research articles in various journals. His research interests include wide-area monitoring systems, power oscillation studies, power grid resilience, smart grids, and optimization techniques. He was honored with the Young Innovator Award 2019 by the Kerala State Government.



NASHIDA CHAKKUNGAL received the B.Tech. degree in electrical and electronics engineering from the NSS College of Engineering, Palakkad, and the M.Tech. degree in power systems engineering from the National Institute of Technology Calicut. Her research interests include power system security, power system dynamics and stability, and machine-learning approaches.



SUNITHA RAJAN (Senior Member, IEEE) received the B.Tech. degree in electrical and electronics, the M.Tech. degree in energetics, and the Ph.D. degree in power systems from the National Institute of Technology Calicut, in 1996, 1999, and 2014, respectively. She is currently working as an Associate Professor with the Department of Electrical Engineering, National Institute of Technology Calicut, Kerala, India. She has authored and coauthored several research papers and has been a referee to several journals and conferences. Her current interests include the applications of wide-area monitoring systems in power systems, power system stability and control, power system security, smart grids, and microgrids.



MANU MADHAVAN received the B.Tech. degree from the Nehru College of Engineering and Research Center, the M.Tech. degree from the Government Engineering College Palakkad, and the Ph.D. degree from the National Institute of Technology Calicut. He is currently working as an Assistant Professor with the Amrita School of Engineering, Coimbatore. He has published several research papers. His research interests include machine learning, natural language processing, and bioinformatics.



HASSAN HAES ALHELOU (Senior Member, IEEE) was with University College Dublin (UCD), Ireland, and the Isfahan University of Technology (IUT), Iran. He is a Faculty Member with Tishreen University, Syria. He is included in the 2018 and 2019 Publons list of the top 1% Best Reviewer and Researchers in the field of engineering. He has published more than 150 research papers in high-quality peer-reviewed journals and international conferences. His research interests include power systems, power system dynamics, power system operation and control, dynamic state estimation, frequency control, smart grids, microgrids, demand response, load shedding, and power system protection. He was a recipient of the Outstanding Reviewer Award from *Energy Conversion and Management*, in 2016, *ISA Transactions*, in 2018, *Applied Energy*, in 2019, and many other awards. He was the recipient of the Best Young Researcher Award by the Arab Creative Student Forum among 61 researchers from 16 countries at Alexandria University, Egypt, in 2011. He has also performed reviews for high prestigious journals, including IEEE TRANSACTIONS ON POWER SYSTEMS, IEEE TRANSACTIONS ON INDUSTRIAL INFORMATICS, IEEE TRANSACTIONS ON INDUSTRIAL ELECTRONICS, ENERGY CONVERSION AND MANAGEMENT, *Applied Energy*, and *International Journal of Electrical Power and Energy Systems*. He has participated in more than 15 industrial projects.

...