# Feed-Forward and Long Short-Term Neural Network Models for Power System State Estimation

**Tuan-Ho Le**

Faculty of Engineering and Technology, Quy Nhon University
170 An Duong Vuong, 55100, Quy Nhon City, Binh Dinh Province, Vietnam
tuanhole@qnu.edu.vn

*Abstract: The primary objective of this paper is to propose the two new combined approaches based on Feed-Forward and Long Short-Term Memory Neural Network models for Power System State Estimation. First, the Weighted Least Square method and the Generalized Maximum-Likelihood Estimator using the Projection statistics method are used to estimate the voltage magnitude and phase angle. Secondly, the Feed-Forward Neural Network model is proposed to combine the obtained voltages and angles. The optimal structure of the proposed Feed-Forward Neural Network model is defined based on the Akaike Information Criterion. Thirdly, the Long Short-Term Neural Network model is proposed as an alternative hybrid power system state estimation approach. Finally, the different case studies including IEEE 9-bus system and IEEE 14-bus system are used to validate the effectiveness of the proposed approaches. The final results imply that the proposed approaches can provide more effective solutions than the existing approaches according to Mean Absolute Percentage Error and Weighted Average Percentage Error criteria.*

*Keywords: Power System State Estimation; Weighted Least Square; Feed-Forward Neural Network; Long Short-Term Neural Network*

# 1 Introduction

State Estimation (SE) of power systems or Power System State Estimation (PSSE) is an important tool of the Energy Management System to provide a reliable estimation of the states of the electrical power systems. The phasor voltages (voltage magnitude and angle) for all system buses are estimated based on a set of available measurements. In real-time power systems monitoring, the Supervisory Control and Data Acquisition (SCADA) system is responsible for gathering and preprocessing information such as the values of active and reactive power flows, power injections, bus voltage magnitude, and status of the circuit breaker switches. However, the SCADA system is not capable of performing a convenient

treatment of inconsistent information because of gross errors in the measurements, telemetered values, communication noise, etc. [1]. In addition, the collected measurement data may not directly extract the parameters of interest. Besides that telemetering all the data of interest may require large numbers of sensors which are not feasible economically or practically [2]. To handle these issues, PSSE is used to detect and eliminate unreliable data from SCADA to identify the optimal estimate of the operating states consisting of voltage magnitude and angle. SE was first introduced by Gauss and Legendre (around 1800). The main purpose of the SE issue is to identify the voltage magnitude and angle from the various input factors. In engineering modeling, various approaches are proposed for modular robotics and human reasoning [3], cognition processes [4], photovoltaic model [5], biomonitoring studies data [6], tower crane systems modeling [7], and opportunities model [8]. Then, the SE was first applied to electric power systems by Fred Schweppe [9]-[11]. Traditionally, the Weighted Least Square (WLS) method is used to solve the SE. However, several works prove that the conventional WLS method has many disadvantages. The WLS technique is less robust as a single outlier can severely distort the estimation results. Further, the technique has a problem of sluggishness and the likelihood of convergence to local optima [12], [13]. To improve the WLS method, several methods are proposed as Iteratively ReWLS SE through givens rotations [14], robust WLS estimator using Reweighting techniques [15], robust linear-WLS method [16], and linear WLS based Singular Value Decomposition approach [17]. Other methods are also provided to improve the estimate of operating states in comparison with the WLS method, such as Least Absolute Value (LAV) [18], weighted LAV [19], and weighted LAV using Interior Point methods [20]. Recently, the Kalman filter (KF) technique is used in several works to provide efficient results of SE [2]. To increase the efficiency of this method, several improved KF variants are introduced as linear KF [21], the extended KF [22]-[24], the unscented KF [25], [26], the cubature KF [27], the Correntropy KF [28], and the ensemble KF [29]. Other works seek to provide the estimator robustness such as M-estimators [30], the Generalized Maximum-Likelihood (GM) Estimation [31], the H-infinite [32], the Robust Cubature KF [33], and the GM-estimator using Projection statistics [34]-[36]. Another research direction in PSSE is to apply Artificial Intelligence methods, such as the Neural Network (NN) models [37]-[39], Deep Learning [40], [41], Fuzzy Logic [42], [43], and Support Vector Machine [44], [45]. All of these works in the literature show the efficiency of the applied method in a case study. In addition, the optimal estimates of the operating states are normally identified using a single method. Unfortunately, none of the existing works try to combine the advantages of these methods to provide better PSSE solutions.

Therefore, the main contribution of this paper is to combine these single methods using Feed-Forward NN (FFNN) and Long Short-Term Memory NN (LSTM) models to provide better PSSE results. The obtained PSSE results (voltage magnitude and phase angle) using the WLS method and the GM-estimator using Projection statistics are considered from the inputs of FFNN and LSTM models.

These NN models are trained to achieve the optimal solutions. Case studies are conducted to verify the proposed combined approaches in PSSE. To the best of our knowledge, this is the first attempt to provide the combined approaches in this area. The proposed combined approaches for PSSE are represented in Figure 1.

The remainder of this paper is organized as follows. In Section 2, the proposed approaches including the WLS method, the GM-estimator using Projection statistics, the FFNN and LSTM models are presented. In Section 3, the case studies of the IEEE 9-bus system and IEEE 14-bus system are conducted. The conclusions are given in Section 4.

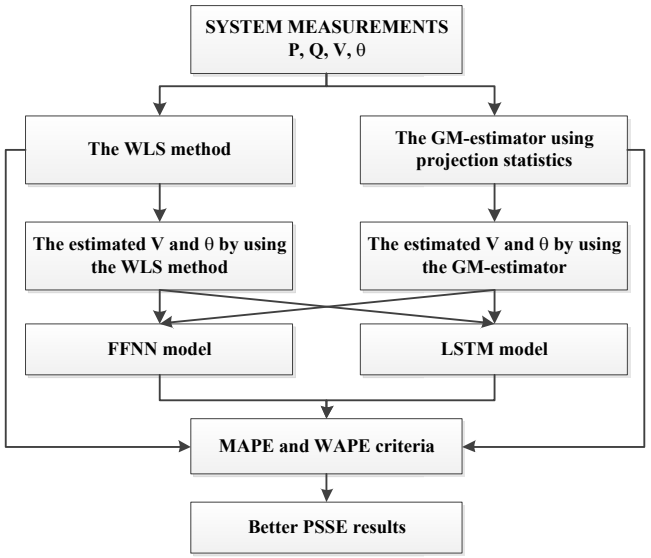


Figure 1
Proposed combined approaches for PSSE

# 2 Proposed Combined Approaches

## 2.1 WLS Method

The WLS method in [46] can be described as follows:

The measurement model of the SE is represented as a set of non-linear equations $\mathbf{h}$ relating measurements $\mathbf{z}$ to state variables $\mathbf{x}$:

$$\mathbf{z} = \mathbf{h}(\mathbf{x}) + \mathbf{e} \tag{1}$$

where vector $\mathbf{x}$ comprises all nodal voltage magnitudes and angles. Vector $\mathbf{z}$ includes active and reactive power injections, active and reactive power flows, and voltage magnitudes. Vector $\mathbf{e}$ is the measurement error that is usually assumed to be independent identically distributed Gaussian with zero mean and diagonal covariance matrix $\mathbf{R} = diag\{\sigma_1^2,...,\sigma_m^2\}$.

The WLS method minimizes the objective function to determine the optimal estimate of $x$:

$$\mathbf{J}(\mathbf{x}) = \left[\mathbf{z} - \mathbf{h}(\mathbf{x})\right]^T \mathbf{R}^{-1} \left[\mathbf{z} - \mathbf{h}(\mathbf{x})\right]. \tag{2}$$

At the minimum, these can be expressed as follows:

$$\mathbf{g}(\mathbf{x}) = \frac{\partial \mathbf{J}(\mathbf{x})}{\partial \mathbf{x}} = -\mathbf{H}(\mathbf{x})^T \mathbf{R}^{-1} \left[\mathbf{z} - \mathbf{h}(\mathbf{x})\right] = 0 \tag{3}$$

where $\mathbf{H}(\mathbf{x}) = \partial \mathbf{h}(\mathbf{x}) / \partial \mathbf{x}$.

Using the Gauss-Newton method, the estimated values of $\hat{\mathbf{x}}$ can be calculated by the following iterative solution

$$\mathbf{G}(\mathbf{x}^k)(\mathbf{x}^{k+1} - \mathbf{x}^k) = \mathbf{H}(\mathbf{x}^k)^T \mathbf{W}\left[\mathbf{z} - \mathbf{h}(\mathbf{x}^k)\right] \tag{4}$$

where $k$ is the iteration index. The measurement weight matrix $\mathbf{W} = \mathbf{R}^{-1}$ is the inverse of the measurement covariance matrix.

## 2.2 GM-estimator using Projection Statistics

This method is represented in [34], [35]. Instead of using Equation (2), the objective function in this method is:

$$\mathbf{J}(\mathbf{x}) = \sum_{i=1}^{m} \omega_i^2 \rho(r_{S_i}) \tag{5}$$

where $\omega_i$ is the weight used to bound the influence of the leverage point. The Huber function can be defined as:

$$\rho(r_{S_i}) = \begin{cases} r_{S_i}^2 / 2 & \text{for } \left|r_{S_i}\right| \le \beta \\ \beta \left|r_{S_i}\right| - \beta^2 / 2 & \text{for } \left|r_{S_i}\right| > \beta \end{cases} \tag{6}$$

where $r_{S_i} = r_i / s\omega_i$ is the standardized residual; $r_i$ is the normalized residual; the parameter $\beta$ is a fixed value; $s$ is the robust scale estimation.

$$s = 1.4826 b_m median_i \left| r_i - median_j (r_j) \right| \tag{7}$$

where $b_m$ is a correction factor for unbiasedness at the Gaussian distribution.

To solve Equation (5), one takes its partial derivative and sets it equal to zero, yielding

$$\frac{\partial \mathbf{J}}{\partial \mathbf{x}} = \sum_{i=1}^{m} -\frac{\omega_i \mathbf{a}_i}{s_i^2} \psi(r_{S_i}) = 0 \tag{8}$$

where $\psi(r_{S_i}) = \partial \rho(r_{S_i}) / \partial r_{S_i}$ ; $\mathbf{a}_i$ is the $i$th row of the Jacobian matrix $\mathbf{H} = \partial \mathbf{h} / \partial \mathbf{x}$.

This set of equations can be solved by using iterated re-WLS algorithm [47].

## 2.3 Proposed FFNN-based Approach

In recent decades, NNs have become a hot topic of research; NNs are now widely used in various fields, including speech recognition, multi-objective optimization, function estimation, and classification. NNs can model linear and nonlinear relationships between inputs and outputs without any assumptions based on the activation function's generalization capacity [48]. A NN comprises one input layer, one output layer, and one or more hidden layers. Among the NNs, the FFNN models are the most popular type for function approximation and multi-objective optimization [49]. In this paper, the proposed FFNN-based approach with one hidden layer is illustrated in Figure 2.

The two different FFNN models, i.e., one for voltage magnitude and one for voltage angle, are used to provide the optimal estimate of the power states. The transfer functions for the hidden layer in the FFNN model are Hyperbolic Tangent Sigmoid (i.e., tansig) and Log-Sigmoid (i.e., logsig). The number of neurons in the hidden layer must be identified carefully since the improper number may lead to overfitting or underfitting. A review that discusses how to fix the number of hidden neurons in NNs was presented in [50]. In this paper, the Akaike Information Criterion (AIC) is used to determine the number of hidden neurons. This criterion is defined as

$$AIC = n \ln\left(\frac{SSE}{n}\right) + 2k \tag{9}$$

where $n$ is the number of data points (observations), $p$ is the number of estimated parameters, and SSE is the Residual Sum of Squares.

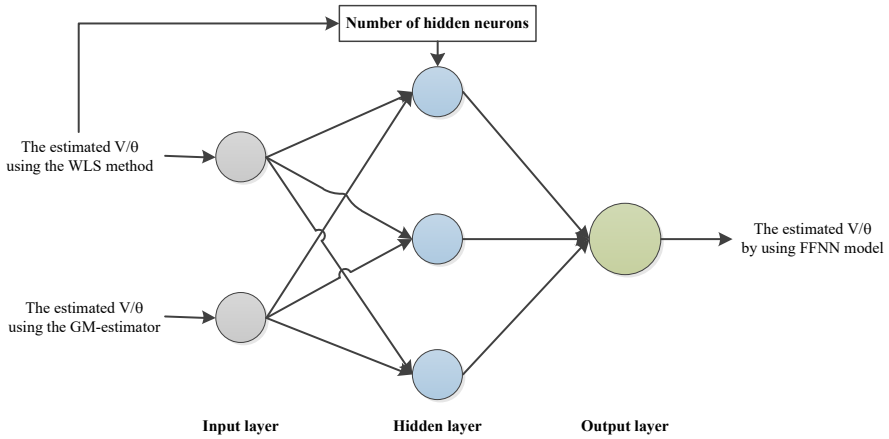


Figure 2
Proposed FFNN-based approach for PSSE

Among several learning algorithms such as Error Correction, Perception Learning, Boltzmann Learning, Hebbian rules, or Back-Propagation (BP), BP is one of the most popular network training algorithms since it is both simple and generally applicable [51]. The difference between actual output and the desired value of the FFNN model is minimized as much as possible by finding the optimal learning rate. The optimization methods for finding the local minimum are Conjugate Gradient such as Gradient Descent with Adaptive Learning Rate (i.e., traingda) and Gradient Descent with Momentum and Adaptive Learning Rate (i.e., traingdx), Steepest Descent such as Resilient BP (i.e., trainrp), and Newton's method such as Levenberg-Marquardt (i.e., trainlm).

Another problem in training NN models is the choice of the number of epochs. The number of epochs determines the number of times that the learning algorithm will work through the entire training dataset. Too many epochs or too few epochs may lead to overfitting or underfitting of the training dataset, respectively. In this case, the early stopping method is often used to solve the generalization issue.

## 2.4 Proposed LSTM-based Approach

LSTM is a deep learning method proposed in [52]. LSTM can be used as a complex nonlinear unit to construct a larger deep NN, which can reflect the effect of long-term memory and has the ability of deep learning [53]. The LSTM model consists of an input layer, an output layer, and several hidden layers. The basic principle of LSTM is shown in Figure 3.

In Figure 3, $x_t$ $t$. $h_{t-1}$ and $C_{t-1}$ $t-1$, respectively. $f$,

$i$, $g$, and $o$ are the forgetting gate, input gate, memory cell, and output gate at time $t$, respectively. $C_t$ and $h_t$ are the updated historical information and the output of the hidden layer at time $t$, respectively.

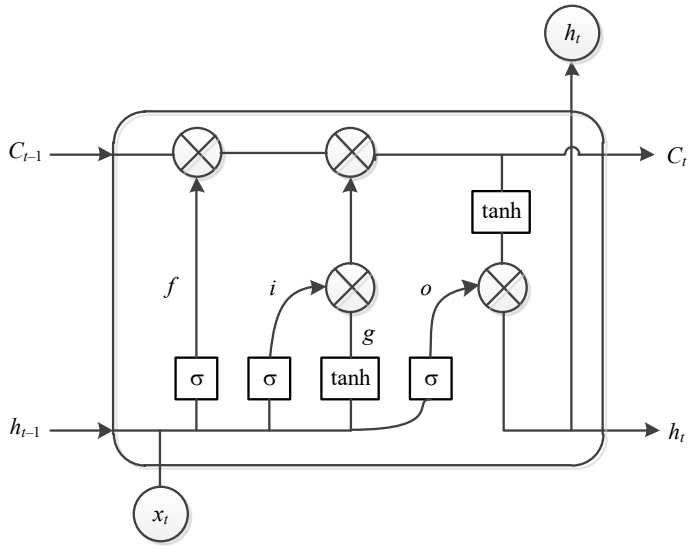


Figure 3
The basic principle of LSTM [52]

The weights and biases to the input gate, forget gate, and output gate control the extent in the cell to compute the output activation of the LSTM block, respectively. Their calculation methods are shown in Equations (10)-(13) as follows:

$$f = \sigma\left(\mathbf{W}_f x_t + \mathbf{U}_f h_{t-1} + \mathbf{b}_f\right), \tag{10}$$

$$i = \sigma\left(\mathbf{W}_i x_t + \mathbf{U}_i h_{t-1} + \mathbf{b}_i\right), \tag{11}$$

$$o = \sigma\left(\mathbf{W}_o x_t + \mathbf{U}_o h_{t-1} + \mathbf{b}_o\right), \tag{12}$$

$$g = \sigma\left(\mathbf{W}_g x_t + \mathbf{U}_g h_{t-1} + \mathbf{b}_g\right) \tag{13}$$

where $\mathbf{W}$, $\mathbf{U}$, $\mathbf{b}$, and $\sigma$ are the parameter matrix from the input layer to the hidden layer, self-recurrent parameter matrix from the hidden layer to the hidden layer, bias parameter vector, and sigmoid function, respectively.

Then, the internal memory cell state $C_t$ is updated. Finally, the output information of the memory cell $h_t$ is obtained. The calculation methods are shown as follows:

$$C_f = f_t \otimes C_{t-1} \otimes i * e, \tag{14}$$

$$h_t = o \otimes \tanh(C_t) \tag{15}$$

where $e$ is the saved new information.

Similar to the FFNN models, two different LSTM models, i.e., one for voltage magnitude and one for voltage angle, are used to provide the optimal estimate of the power states. The estimated voltage magnitudes and angles using the WLS method and GM-estimator are the inputs of the LSTM models.

## 2.5 Evaluation Criteria

To evaluate the performance of the estimation methods, several criteria were proposed. In this paper, two primary evaluation criteria are used as:

Mean Absolute Percentage Error (MAPE) is used to measure the percentage error of the estimate in relation to the actual values. It is represented as

$$MAPE = \frac{1}{n}\sum_{i=1}^{n}\left|\frac{A-\hat{A}}{A}\right|100 \tag{16}$$

where $n$ is the number of observations. $A$ and $\hat{A}$ are the actual and estimated values, respectively.

Weighted Average Percentage Error (WAPE) is similar to MAPE. However, it weighs the estimated error.

$$WAPE = \frac{\sum_{i=1}^{n}\left|A-\hat{A}\right|}{\sum_{i=1}^{n}|A|}100. \tag{17}$$

# 3 Case Studies

## 3.1 IEEE 9-bus System

The proposed approaches along with the conventional WLS method and GM-estimator using Projection statistics have been applied on IEEE 9-bus system. The simulation results of voltage magnitudes and angles are tabulated in Tables 1

and 2, respectively. The proposed FFNN and LSTM models are coded and trained in Matlab. Information about the trained FFNN model for estimating the voltage magnitudes including the transfer function, training function, number of hidden neurons, and number of epochs, is tansig, trainlm, 87, and 4, respectively. Information about the trained FFNN model for estimating the voltage angles including the transfer function, training function, number of hidden neurons, and number of epochs, is logsig, trainlm, 30, and 5, respectively.

Information about the trained LSTM model for estimating the voltage magnitudes including the solver, maximum epochs, gradient threshold, initial learn rate, number of hidden neurons, and dropoutlayer is adam optimizer, 10000, 0.01, 0.001, 6, and 0.1, respectively. Information about the trained LSTM model for estimating the voltage angles including the solver, maximum epochs, gradient threshold, initial learn rate, number of hidden neurons, and dropoutlayer is adam optimizer, 10000, 0.01, 0.001, 35, and 0.1, respectively.

Table 1
Comparison of true and estimated voltage magnitudes for IEEE 9-bus system

| Bus no | True voltage magnitude (p.u) | Estimated voltage magnitude | | | |
|---|---|---|---|---|---|
| | | WLS | GM | FFNN | LSTM |
| 1 | 1.04000 | 0.98871 | 0.99927 | 1.03863 | 1.03529 |
| 2 | 1.02500 | 1.00340 | 1.01731 | 1.02508 | 1.03016 |
| 3 | 1.02500 | 1.02414 | 1.03592 | 1.02500 | 1.02393 |
| 4 | 1.02579 | 0.97384 | 0.98454 | 1.02579 | 1.01423 |
| 5 | 0.99563 | 0.95003 | 0.96177 | 0.99563 | 1.00710 |
| 6 | 1.01265 | 1.04889 | 1.06516 | 1.01265 | 1.01544 |
| 7 | 1.02577 | 1.00357 | 1.01808 | 1.02577 | 1.01884 |
| 8 | 1.01588 | 1.00153 | 1.01468 | 1.01588 | 1.02322 |
| 9 | 1.03235 | 1.03242 | 1.04327 | 1.03235 | 1.03157 |

Table 2
Comparison of true and estimated voltage angles for IEEE 9-bus system

| Bus no | True voltage angle (degree) | Estimated voltage angle | | | |
|---|---|---|---|---|---|
| | | WLS | GM | FFNN | LSTM |
| 1 | 0.00000 | 0.00000 | 0.00000 | -0.00752 | 0.00894 |
| 2 | 9.28001 | 11.17367 | 10.97289 | 9.28386 | 9.28298 |
| 3 | 4.66475 | 7.14438 | 6.99829 | 4.62074 | 4.50812 |
| 4 | -2.21679 | -2.45629 | -2.40388 | -1.72290 | -2.26689 |
| 5 | -3.98881 | -3.93075 | -3.88208 | -3.98727 | -4.00079 |
| 6 | -3.68740 | 0.11674 | 0.22492 | -3.68426 | -3.70708 |
| 7 | 3.71970 | 5.34103 | 5.32414 | 3.72129 | 3.66028 |
| 8 | 0.72754 | 2.58574 | 2.58591 | 0.74213 | 0.66242 |
| 9 | 1.96672 | 4.47395 | 4.36141 | 1.96860 | 1.82292 |

The true and estimated values of voltage magnitudes and angles in this case study are shown in Figures 4 and 5, respectively. In these figures, the true values, estimated values by using WLS, estimated values by using GM, estimated values by using FFNN, and estimated values by using LSTM are demonstrated by the solid line (black color), dash line with * marker (red color), dash-dot line with square marker (yellow color), dash-dot line with + marker (blue color), and dash-dot line with pentagram marker (green color), respectively.

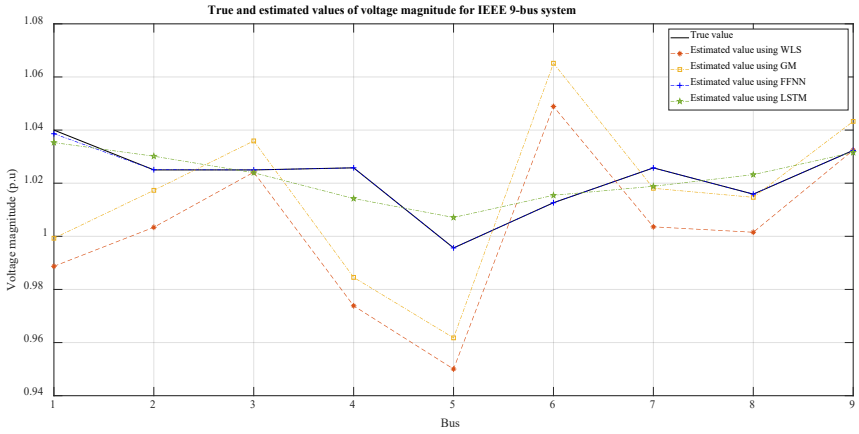


Figure 4

True and estimated values of voltage magnitudes for IEEE 9-bus system

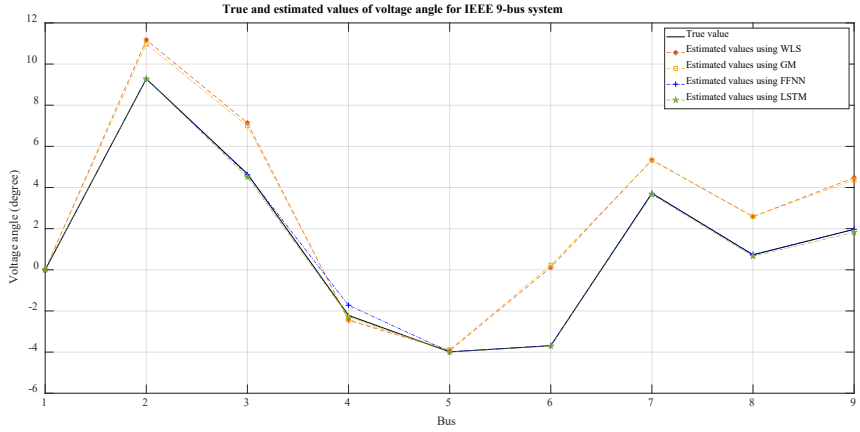


Figure 5

True and estimated values of voltage angles for IEEE 9-bus system

The MAPE and WAPE criteria of the WLS, GM, FFNN, and LSTM methods for estimating voltage magnitudes and angles in this case study are represented in Tables 3 and 4, respectively.

Table 3
Evaluation criteria of the proposed approaches in voltage magnitude estimation for IEEE 9-bus system

| Evaluation criteria | WLS | GM | FFNN | LSTM |
|---|---|---|---|---|
| MAPE (%) | 2.65879 | 2.25172 | 0.01566 | 0.56536 |
| WAPE (%) | 2.65441 | 2.24802 | 0.01592 | 0.56320 |

Table 4
Evaluation criteria of the proposed approaches in voltage angle estimation for IEEE 9-bus system

| Evaluation criteria | WLS | GM | FFNN | LSTM |
|---|---|---|---|---|
| MAPE (%) | 68.38519 | 67.31203 | 2.83688 | 2.70489 |
| WAPE (%) | 47.80473 | 46.57606 | 1.89083 | 1.71443 |

As shown in Table 3, the GM-estimator (MAPE = 2.25172%, WAPE = 2.24802%) can provide better-estimated results compared to the conventional WLS method (MAPE = 2.65879%, WAPE = 2.65441%). Moreover, the proposed FFNN (MAPE = 0.01566%, WAPE = 0.01592%) and LSTM (MAPE = 0.56536%, WAPE = 0.56320%) approaches can provide more accurate estimated voltage magnitudes compared to the two above methods. The proposed FFNN approach is the best estimation method in this case study.

As shown in Table 4, the GM-estimator (MAPE = 67.31203%, WAPE = 46.57606%) can provide better-estimated results compared to the conventional WLS method (MAPE = 68.38519%, WAPE = 47.80473%). Moreover, the proposed FFNN (MAPE = 2.83688%, WAPE = 1.89083%) and LSTM (MAPE = 2.70489%, WAPE = 1.71443%) approaches can provide more accurate estimated voltage angles compared to the two above methods. The proposed FFNN approach is the best estimation method in this case study.

## 3.2 IEEE 14-bus System

Similarly, the estimated results of voltage magnitudes and angles in this case study are tabulated in Tables 5 and 6, respectively. Information about the trained FFNN model for estimating the voltage magnitudes including the transfer function, training function, number of hidden neurons, and number of epochs, is logsig, trainlm, 32, and 6, respectively. Information about the trained FFNN model for estimating the voltage angles including the transfer function, training function, number of hidden neurons, and number of epochs, is tansig, trainlm, 34, and 14, respectively. Information about the trained LSTM model for estimating the voltage magnitudes including the solver, maximum epochs, gradient threshold, initial learn rate, number of hidden neurons, and dropoutlayer is adam optimizer,

10000, 0.01, 0.001, 5, and 0.1, respectively. Information about the trained LSTM model for estimating the voltage angles including the solver, maximum epochs, gradient threshold, initial learn rate, number of hidden neurons, and dropoutlayer is adam optimizer, 10000, 0.01, 0.001, 19, and 0.1, respectively.

Table 5

Comparison of true and estimated voltage magnitudes for IEEE 14-bus system

| Bus no | True voltage magnitude (p.u) | Estimated voltage magnitude | | | |
|---|---|---|---|---|---|
| | | WLS | GM | FFNN | LSTM |
| 1 | 1.06000 | 1.04531 | 1.06873 | 1.06071 | 1.06275 |
| 2 | 1.04500 | 1.02799 | 1.04954 | 1.04529 | 1.03994 |
| 3 | 1.01000 | 0.98940 | 1.01203 | 1.00970 | 1.01431 |
| 4 | 1.01767 | 0.99053 | 1.01540 | 1.01780 | 1.00976 |
| 5 | 1.01951 | 0.99448 | 1.01947 | 1.01975 | 1.02871 |
| 6 | 1.07000 | 1.02631 | 1.05667 | 1.07012 | 1.05692 |
| 7 | 1.06152 | 1.01274 | 1.04099 | 1.06155 | 1.07290 |
| 8 | 1.09000 | 1.04244 | 1.07030 | 1.09009 | 1.07310 |
| 9 | 1.05593 | 0.99548 | 1.02508 | 1.05594 | 1.06250 |
| 10 | 1.05098 | 0.99302 | 1.02325 | 1.05096 | 1.05392 |
| 11 | 1.05691 | 1.00564 | 1.03635 | 1.05335 | 1.05157 |
| 12 | 1.05519 | 1.00915 | 1.04072 | 1.05955 | 1.05139 |
| 13 | 1.05038 | 1.00312 | 1.03457 | 1.04940 | 1.04836 |
| 14 | 1.03553 | 0.97970 | 1.01116 | 1.03543 | 1.04097 |

Table 6

Comparison of true and estimated voltage angles for IEEE 14-bus system

| Bus no | True voltage angle (degree) | Estimated voltage angle | | | |
|---|---|---|---|---|---|
| | | WLS | GM | FFNN | LSTM |
| 1 | 0.00000 | 0.00000 | 0.00002 | 0.00000 | -0.06214 |
| 2 | -4.98259 | -5.12756 | -4.84083 | -4.98258 | -5.02083 |
| 3 | -12.72510 | -13.16706 | -12.50966 | -12.72502 | -12.87459 |
| 4 | -10.31290 | -10.51851 | -10.01290 | -10.31285 | -9.99912 |
| 5 | -8.77385 | -8.96567 | -8.53965 | -8.77382 | -8.69088 |
| 6 | -14.22095 | -14.93620 | -14.17390 | -14.22541 | -14.12077 |
| 7 | -13.35963 | -13.73609 | -13.04024 | -13.35019 | -13.07401 |
| 8 | -13.35963 | -13.73477 | -13.03726 | -13.35929 | -13.13985 |
| 9 | -14.93852 | -15.45928 | -14.65665 | -14.93769 | -14.71624 |
| 10 | -15.09729 | -15.67465 | -14.86375 | -15.09144 | -14.73552 |
| 11 | -14.79062 | -15.43579 | -14.64498 | -14.82734 | -14.53140 |
| 12 | -15.07558 | -15.85483 | -15.04993 | -15.07530 | -14.93506 |
| 13 | -15.15628 | -15.91313 | -15.09867 | -15.17020 | -14.98227 |
| 14 | -16.03364 | -16.76849 | -15.89944 | -16.03184 | -15.53161 |

The true and estimated values of voltage magnitudes and angles in this case study are shown in Figures 6 and 7, respectively. In these figures, the true values, estimated values by using WLS, estimated values by using GM, estimated values by using FFNN, and estimated values by using LSTM are demonstrated by the solid line (black color), dash line with * marker (red color), dash-dot line with square marker (yellow color), dash-dot line with + marker (blue color), and dash-dot line with pentagram marker (green color), respectively.

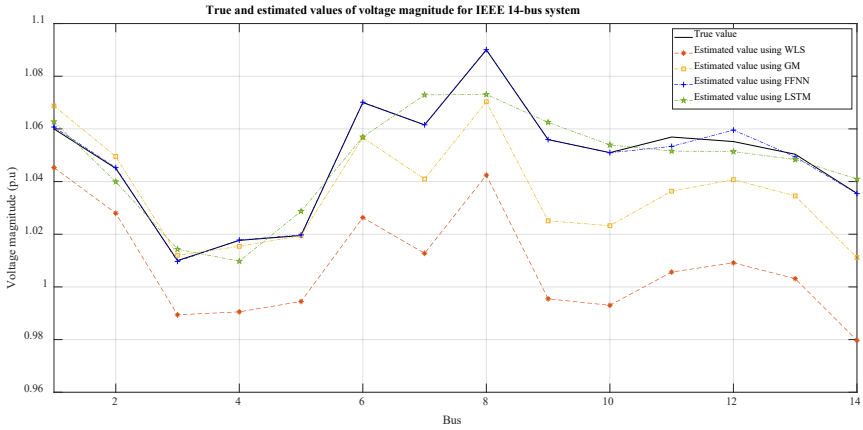


Figure 6
True and estimated values of voltage magnitudes for IEEE 14-bus system

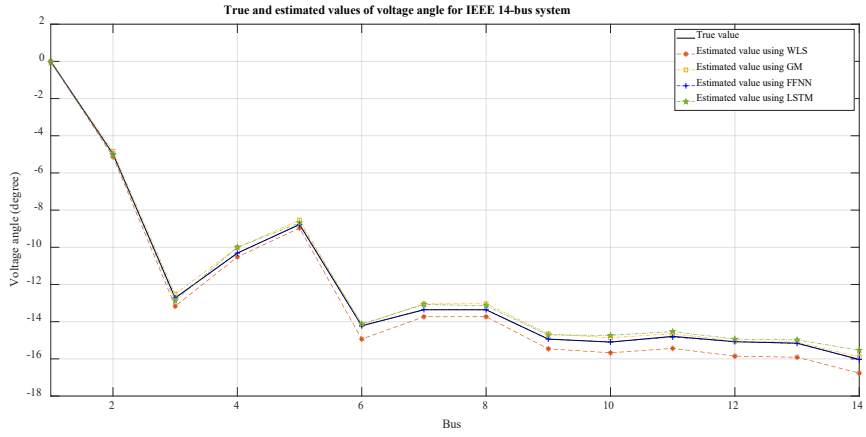


Figure 7
True and estimated values of voltage angles for IEEE 14-bus system

The MAPE and WAPE criteria of the WLS, GM, FFNN, and LSTM approaches for estimating voltage magnitudes and angles in this case study are represented in Tables 7 and 8, respectively.

Table 7

Evaluation criteria of the proposed approaches in voltage magnitude estimation for IEEE 14-bus system

| Evaluation criteria | WLS | GM | FFNN | LSTM |
|---|---|---|---|---|
| MAPE (%) | 3.82581 | 1.38623 | 0.07419 | 0.65567 |
| WAPE (%) | 3.83760 | 1.39623 | 0.07452 | 0.65874 |

Table 8

Evaluation criteria of the proposed approaches in voltage angle estimation for IEEE 14-bus system

| Evaluation criteria | WLS | GM | FFNN | LSTM |
|---|---|---|---|---|
| MAPE (%) | 3.40259 | 1.50406 | 0.03599 | 1.51901 |
| WAPE (%) | 3.82964 | 1.45638 | 0.04373 | 1.72487 |

As shown in Table 7, the GM-estimator (MAPE = 1.38623%, WAPE = 1.39623%) can provide better-estimated results compared to the conventional WLS method (MAPE = 3.82581%, WAPE = 3.83760%). Moreover, the proposed FFNN (MAPE = 0.07419%, WAPE = 0.07452%) and LSTM (MAPE = 0.65567%, WAPE = 0.65874%) approaches can provide more accurate estimated voltage magnitudes compared to the two above methods. The proposed FFNN approach is the best estimation method in this case study.

As shown in Table 8, the GM-estimator (MAPE = 1.50406%, WAPE = 1.45638%) can provide better-estimated results compared to the conventional WLS method (MAPE = 3.40259%, WAPE = 3.82964%). Moreover, the proposed FFNN (MAPE = 0.03599%, WAPE = 0.04373%) and LSTM (MAPE = 1.51901%, WAPE = 1.72487%) approaches can provide more accurate estimated voltage angles compared to the two above methods. The proposed FFNN approach is the best estimation method in this case study.

## Conclusions

The two combined approaches, based on FFNN and LSTM models, for PSSE are proposed. The two proposed approaches are trained to identify the optimal structures. The IEEE 9-bus system and IEEE 14-bus system are used to approve the effectiveness of the proposed approaches in finding the optimal estimate of the states. Two evaluation criteria consisting of MAPE and WAPE are used to specify the better methods. The simulation results indicate that the two proposed combined approaches can provide better solutions compared to the conventional WLS method and the GM-estimator using Projection statistics. For further studies, the weighted Tchebycheff optimization technique and Genetic Algorithm can be applied to solve the PSSE.

## References

[1] M. B. D. C. Filho, A. M. L. da Silva, J. M. C. C. Cantera, and R. A. da Silva, “Information debugging for real-time power systems monitoring,” *IEE Proceedings - Generation, Transmission and Distribution*, 136 (3), 1989, pp. 145-152, doi: 10.1049/ip-c.1989.002

[2] A. Saikia and R. K. Mehta, "Power system static state estimation using Kalman filter algorithm", *International Journal for Simulation and Multidisciplinary Design Optimization*, *7*, A7, 2016

[3] C. Pozna and R.-E. Precup, “Plausible Reasoning in Modular Robotics and Human Reasoning,” *Acta Polytechnica Hungarica*, 4 (4), 2007

[4] C. Pozna and R.-E. Precup, “Aspects concerning the observation process modelling in the framework of cognition processes,” *Acta Polytechnica Hungarica*, 9 (1), 2012

[5] N. Ngoc Son and L. The Vinh, “Parameter estimation of photovoltaic model, using balancing composite motion optimization,” *Acta Polytechnica Hungarica*, 19 (11), 2022, pp. 27-46, doi: 10.12700/APH.19.11.2022.11.2

[6] R.-E. Precup, G. Duca, S. Travin, and I. Zinicovscaia, “Processing, neural network-based modeling of biomonitoring studies data and validation on republic of moldova data”, *Proceedings of the Romanian Academy Series A-Mathematics Physics Technical Sciences Information Science*, 2022, 403-410

[7] E.-L. Hedrea, R.-E. Precup, R.-C. Roman, and E. M. Petriu, “Tensor product-based model transformation approach to tower crane systems modeling,” *Asian Journal of Control*, 23 (3), 2021, pp. 1313-1323, doi: 10.1002/asjc.2494

[8] S. Travin and G. Duca, "New opportunities model for monitoring, analyzing and forecasting the official statistics on Coronavirus disease pandemic", *Romanian Journal of Information Science and Technology*, 1, 2023, pp. 49-64, doi: 10.59277/ROMJIST.2023.1.04

[9] F. C. Schweppe and J. Wildes, "Power system static-state estimation, Part I: Exact model", *IEEE Transactions on Power Apparatus and systems*, (1), 1970, pp. 120-125

[10] F. C. Schweppe and D. B. Rom, “Power System Static-State Estimation, Part II: Approximate Model,” *IEEE Transactions on Power Apparatus and systems*, 89 (1), 1970, pp. 125-130, doi: 10.1109/TPAS.1970.292679

[11] F. C. Schweppe, “Power System Static-State Estimation, Part III: Implementation,” *IEEE Transactions on Power Apparatus and systems*, 89 (1), 1970, pp. 130-135, doi: 10.1109/TPAS.1970.292680

[12] E. Caro and A. J. Conejo, “State estimation via mathematical programming: a comparison of different estimation algorithms,” *IET Generation,*

*Transmission & Distribution*, 6 (6), 2012, pp. 545-553, doi: 10.1049/iet-gtd.2011.0663

[13] E. J. Contreras-Hernandez and J. R. Cedeno-Maldonado, "A self-adaptive evolutionary programming approach for power system state estimation," in *2006 49th IEEE International Midwest Symposium on Circuits and Systems*, Aug. 2006, pp. 571-575. doi: 10.1109/MWSCAS.2006.382127

[14] R. C. Pires, A. Simoes Costa, and L. Mili, "Iteratively reweighted least-squares state estimation through givens rotations," *IEEE Transactions on Power Systems*, 14 (4), 1999, pp. 1499-1507, doi: 10.1109/59.801941

[15] E. Caro, R. Mínguez, and A. J. Conejo, "Robust WLS estimator using reweighting techniques for electric energy systems," *Electric power systems research*, 104, 2013, pp. 9-17, doi: 10.1016/j.epsr.2013.05.021

[16] S. K. Kotha, B. Rajpathak, M. Mallareddy, and R. Bhuvanagiri, "Wide area measurement systems based power system state estimation using a robust linear-weighted least square method," *Energy Reports*, 9, 2023, pp. 23-32, doi: 10.1016/j.egyr.2023.05.046

[17] G. Bei, "Observability analysis for state estimation using Hachtel's augmented matrix method," *Electric power systems research*, 77 (7), 2007, pp. 865-875, doi: 10.1016/j.epsr.2006.07.010

[18] A. Abur and M. K. Celik, "Least absolute value state estimation with equality and inequality constraints," *IEEE Transactions on Power Systems*, 8 (2), 1993, pp. 680-686, doi: 10.1109/59.260812

[19] C. Rakpenthai, S. Uatrongjit, I. Ngamroo, and N. R. Watson, "Weighted least absolute value power system state estimation using rectangular coordinates and equivalent measurement functions," *IEEJ transactions on electrical and electronic engineering*, 6 (6), 2011, pp. 534-539, doi: 10.1002/tee.20692

[20] H. Singh and F. L. Alvarado, "Weighted least absolute value state estimation using interior point methods," *IEEE Transactions on Power Systems*, 9 (3), 1994, pp. 1478-1484, doi: 10.1109/59.336114

[21] S. Sarri, L. Zanni, M. Popovic, J.-Y. Le Boudec, and M. Paolone, "Performance assessment of linear state estimators using synchrophasor measurements," *IEEE Transactions on Instrumentation and Measurement*, 65 (3), 2016, pp. 535-548, doi: 10.1109/TIM.2015.2510598

[22] L. Fan and Y. Wehbe, "Extended Kalman filtering based real-time dynamic state and parameter estimation using PMU data," *Electric Power Systems Research*, 103, 2013, pp. 168-177, doi: 10.1016/j.epsr.2013.05.016

[23] F. Shabani, M. Seyedyazdi, M. Vaziri, M. Zarghami, and S. Vadhva, "State estimation of a distribution system using WLS and EKF Techniques," in

*2015 IEEE International Conference on Information Reuse and Integration*, Aug. 2015, pp. 609-613, doi: 10.1109/IRI.2015.101

[24] E. Ghahremani and I. Kamwa, "Dynamic state estimation in power system by applying the extended Kalman filter with unknown inputs to phasor measurements," *IEEE Transactions on Power Systems*, 26 (4), 2011, pp. 2556-2566, doi: 10.1109/TPWRS.2011.2145396

[25] G. Valverde and V. Terzija, "Unscented Kalman filter for power system dynamic state estimation," *IET generation, transmission & distribution*, 5 (1), 2011, pp. 29-37, doi: 10.1049/iet-gtd.2010.0210

[26] X. Qing, H. R. Karimi, Y. Niu, and X. Wang, "Decentralized unscented Kalman filter based on a consensus algorithm for multi-area dynamic state estimation in power systems," *International Journal of Electrical Power & Energy Systems*, 65, 2015, pp. 26-33, doi: 10.1016/j.ijepes.2014.09.024

[27] A. Sharma, S. C. Srivastava, and S. Chakrabarti, "A cubature Kalman filter based power system dynamic state estimator," *IEEE Transactions on Instrumentation and Measurement*, 66 (8), 2017, pp. 2036-2045, doi: 10.1109/TIM.2017.2677698

[28] J. A. D. Massignan, J. B. A. London, and V. Miranda, "Tracking power system state evolution with maximum-correntropy-based extended Kalman filter," *Journal of Modern Power Systems and Clean Energy*, 8 (4), 2020, pp. 616-626, doi: 10.35833/MPCE.2020.000122

[29] N. Zhou, D. Meng, Z. Huang, and G. Welch, "Dynamic state estimation of a synchronous machine using PMU data: A comparative study", *IEEE Transactions on Smart grid*, 6 (1), 2014, pp. 450-460

[30] G. Durgaprasad and S. S. Thakur, "Robust dynamic state estimation of power systems based on M-estimation and realistic modeling of system dynamics," *IEEE Transactions on Power Systems*, 13 (4), 1998, pp. 1331-1336, doi: 10.1109/59.736273

[31] J. Zhao and L. Mili, "A framework for robust hybrid state estimation with unknown measurement noise statistics", *IEEE Transactions on Industrial Informatics*, 14 (5), 2017, pp. 1866-1875

[32] J. Zhao, "Dynamic state estimation with model uncertainties using H_infinity extended Kalman filter," *IEEE Transactions on power systems*, 33 (1), 2018, pp. 1099-1100, doi: 10.1109/TPWRS.2017.2688131

[33] Y. Wang, Y. Sun, V. Dinavahi, S. Cao, and D. Hou, "Adaptive robust cubature Kalman filter for power system dynamic state estimation against outliers", *IEEE Access*, 7, 2019, pp. 105872-105881

[34] L. Mili, M. G. Cheniae, N. S. Vichare, and P. J. Rousseeuw, "Robust state estimation based on projection statistics [of power systems]," *IEEE*

*Transactions on Power Systems*, 11 (2), 1996, pp. 1118-1127, doi: 10.1109/59.496203

[35] J. Zhao, G. Zhang, and M. La Scala, "A two-stage robust power system state estimation method with unknown measurement noise," in *2016 IEEE Power and Energy Society General Meeting (PESGM)*, Jul. 2016, pp. 1-5, doi: 10.1109/PESGM.2016.7741350

[36] Y. Shi, Y. Hou, Y. Yu, Z. Jin, and M. A. Mohamed, "Robust power system state estimation method based on generalized M-estimation of optimized parameters based on Sampling", *Sustainability*, 15 (3), 2023, pp. 2550

[37] D. M. V. Kumar, S. C. Srivastava, S. Shah, and S. Mathur, "Topology processing and static state estimation using artificial neural networks," *IEE Proceedings-Generation, Transmission and Distribution*, 143 (1), 1996, pp. 99-105

[38] E. Manitsas, R. Singh, B. C. Pal, and G. Strbac, "Distribution system state estimation using an artificial neural network approach for pseudo measurement modeling," *IEEE Transactions on power systems*, 27 (4), 2012, pp. 1888-1896, doi: 10.1109/TPWRS.2012.2187804

[39] H. Salehfar and R. Zhao, "A neural network preestimation filter for bad-data detection and identification in power system state estimation," *Electric power systems research*, 34 (2), 1995, pp. 127-134, doi: 10.1016/0378-7796(95)00966-7

[40] L. Wang, Q. Zhou, and S. Jin, "Physics-guided deep learning for power system state estimation," *Journal of Modern Power Systems and Clean Energy*, 8 (4), 2020, pp. 607-615, doi: 10.35833/MPCE.2019.000565

[41] K. R. Mestav, J. Luengo-Rozas, and L. Tong, "Bayesian state estimation for unobservable distribution systems via deep learning," *IEEE Transactions on Power Systems*, 34 (6), 2019, pp. 4910-4920, doi: 10.1109/TPWRS.2019.2919157

[42] F. Shabani, N. R. Prasad, and H. A. Smolleck, "A fuzzy-logic-supported weighted least squares state estimation," *Electric power systems research*, 39 (1), 1996, pp. 55-60, doi: 10.1016/S0378-7796(96)01107-8

[43] K. E. Holbert and K. Lin, "Reducing state estimation uncertainty through fuzzy logic evaluation of power system measurements," in *2004 International Conference on Probabilistic Methods Applied to Power Systems*, Sep. 2004, pp. 205-211

[44] V. Kirinčić, E. Čeperić, S. Vlahinić, and J. Lerga, "Support vector machine state estimation", *Applied Artificial Intelligence*, 33 (6), 2019, pp. 517-530

[45] A. A. Abod, A. H. Abdullah, and M. K. Abd, "Support vector machine based approach for state estimation of Iraqi super grid network," in *2008*

*Workshop on Power Electronics and Intelligent Transportation System*, Aug. 2008, pp. 252-256, doi: 10.1109/PEITS.2008.102

[46] A. Abur and A. G. Exposito, "*Power system state estimation: theory and implementation*", 2004, CRC press

[47] W. W. Kotiuga and M. Vidyasagar, "Bad data rejection properties of weighted least absolute value techniques applied to static state estimation," *IEEE Transactions on Power apparatus and systems*, 101 (4), 1982, pp. 844-853, doi: 10.1109/TPAS.1982.317150

[48] T. H. Le, L. Dai, H. Jang, and S. Shin, "Robust process parameter design methodology: A new estimation approach by using feed-forward neural network structures and machine learning algorithms", *Applied Sciences*, 12 (6), (2022), pp. 2904

[49] Z. Zainuddin and O. Pauline, "Function approximation using artificial neural networks", *WSEAS Transactions on Mathematics*, 7 (6), 2008, pp. 333-338, doi:10.5555/1466915.1466916

[50] K. G. Sheela and S. N. Deepa, "Review on methods to fix number of hidden neurons in neural networks", *Mathematical problems in engineering*, 2013, *2013*

[51] A. Zilouchian and M. Jamshidi (Eds.). "*Intelligent control systems using soft computing methodologies*, CRC press, 2001

[52] S. Hochreiter and J. Schmidhuber, "Long short-term memory," *Neural Computation*, 9 (8), 1997, pp. 1735-1780, doi: 10.1162/neco.1997.9.8.1735

[53] Q. Xiaoyun, K. Xiaoning, Z. Chao, J. Shuai, and M. Xiuda, "Short-term prediction of wind power based on deep long short-term memory," in *2016 IEEE PES Asia-Pacific Power and Energy Engineering Conference (APPEEC)* Oct. 2016, pp. 1148-1152, doi: 10.1109/APPEEC.2016.7779672