Error Modeling in Distribution Network State Estimation Using RBF-Based Artificial Neural Network

A. Hassannejad Marzouni* and A. Zakariazadeh*(C.A.)

Abstract: State estimation is essential to access observable network models for online monitoring and analyzing of power systems. Due to the integration of distributed energy resources and new technologies, state estimation in distribution systems would be necessary. However, accurate input data are essential for an accurate estimation along with knowledge on the possible correlation between the real and pseudo measurements data. This study presents a new approach to model errors for the distribution system state estimation purpose. In this paper, pseudo measurements are generated using a couple of real measurements data by means of the artificial neural network method. In the proposed method, the radial basis function network with the Gaussian kernel is also implemented to decompose pseudo measurements into several components. The robustness of the proposed error modeling method is assessed on IEEE 123-bus distribution test system where the problem is optimized by the imperialist competitive algorithm. The results evidence that the proposed method causes to increase in detachment accuracy of error components which results in presenting higher quality output in the distribution state estimation.

Keywords: State Estimation, Distribution Network, Error Modeling, ANN, Radial Basis Function.

1 Introduction

DISTRIBUTION System State Estimation (DSSE) is a data processing algorithm aiming at reducing metering data errors to provide an accurate supervisory and process in distribution systems. When a failure occurs in measurement devices, the observability of the system reduces [1, 2]. In this case, DSSE can be implemented to increase the observability and, as a result, the supervisory of the system [3]. Methodologies of DSSE started by Roytelman and Shahidehpour in the 1990s when they presented a technique of DSSE by consideration of the minimum number of measurements available in the distribution system [1-4]. The conventional DSSE methods were tackled by many researchers to improve coverage speed and accuracy.

The branch current based DSSE in [4, 5] has been proposed to reduce pseudo measurement for accuracy and coverage speed enhancement. In [6], a bad data filter for measurement data has been introduced based on the weighted least square method. In [7], state estimation in distribution networks aiming at improving forecasted load data by using real-time measurements has been presented. State estimation using pseudo-measurements and artificial neural networks (ANN) is presented in [8]. In this approach, an ANN is used for power injections and voltage magnitudes to estimate directly both of them. However, meaningful and reasonable state estimation using limited measurements is challenging. To increase the observability of the system, pseudo-measurements are introduced. Additionally, pseudo measurements need to be accurately modeled to improve the quality of the estimations. State estimation of a distribution network is never perfect as it contains inaccuracies such as network parameters, topology, measurements, data correlation and operating conditions. Using a large number of pseudo measurements with uncertainties aiming at making a distribution network observable may result in
a deviation of the estimated state from the real one. Measurement devices allocation is an important application of DSSE in recent literature. In [9], optimal placement of synchronized phasor measurement in unbalance condition of distribution systems for observability and state estimation, using genetic algorithm has been proposed. In [10], a multi-objective optimization method aiming at seeking number and location of the measurement devices for accurate distribution state estimation has been proposed. In [11], a measurement device allocation method using distribution state estimation has been presented in which the uncertainty introduced by the measurement devices and the tolerance in the knowledge of line impedances have been considered. In [12], aggregation of smart metering data is introduced as a restriction in DSSE. To improve DSSE accuracy under this restriction, the correlations among pseudo loads’ errors have been considered. The robust placement of a limited number of phasor measurement units and voltage magnitude meters for state estimation has been presented in [13] where the Fisher information matrix has been preferred as a criterion for the estimation accuracy. The authors in [14] have studied the influence of possible correlations between data of measurement devices and pseudo measurement in the weighted least square estimation method. One of the most important usages of state estimation is to diagnose and reduce the error of measurement data [15]. A methodology for customers’ load allocation based on the probabilistic neural network is proposed in [16]. In [17], an approach to modeling the loads as pseudo measurements based on a Gaussian mixture model (GMM) has been presented. To compare the DSSE methods, a classification has been presented in Table 1.

To the best of our knowledge, error modeling using radial distribution function in DSSE has not been reported in the literature. In this paper, the ANN method is also implemented to model pseudo measurements of active and reactive power injections. The error between the target (power injection) and the output (load profile) is modeled using the radial basis function (RBF). This model is used to obtain the pseudo measurements’ variance. On the other hand, it has been firstly assumed that the distribution network is observable and the focus of this paper is to truthfully model errors corresponding to pseudo measurement data. The innovative contributions of this paper are highlighted as follows:

- To model pseudo measurements using multilayer perceptron artificial neural network and consider the error of the modeled neural network as pseudo measurement errors.
- To extract effective Gaussian kernel in pseudo measurement errors of the Gaussian mixture model using radial basis function neural network.
- To optimize the radial basis function neural network by the imperialist competitive algorithm and carry out self-tuning detection of pseudo measurement errors’ kernel number of the Gaussian mixture model.

The rest of the paper is organized as follows: The state estimation formulation based on the weighted least square method is described in Section 2. The pseudo measurements modeling with multilayer perceptron neural network is presented in Section 3. The measurement data encountering strategy is described in Section 4. The error modeling and effective error extraction are described in Section 5. The state estimation application and their roles are discussed in Section 6. The overall methodology and application of the error modeling and the usage of state estimation are presented in Section 7. The simulation results on the real network as a case study are aggregated in Section 8. The discussions about the methods that have been used for this application of state estimation are prepared in Section 9 and finally, the conclusion of this paper is presented in Section 10.

<table>
<thead>
<tr>
<th>Table 1 Comparison of DSSE methods.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Refs.</td>
</tr>
<tr>
<td>[17]</td>
</tr>
<tr>
<td>[18]</td>
</tr>
<tr>
<td>[19]</td>
</tr>
<tr>
<td>[20]</td>
</tr>
<tr>
<td>[21]</td>
</tr>
<tr>
<td>[22]</td>
</tr>
<tr>
<td>[23]</td>
</tr>
<tr>
<td>[24]</td>
</tr>
<tr>
<td>[25]</td>
</tr>
<tr>
<td>This Paper</td>
</tr>
</tbody>
</table>

---

Gaussian probability density [26]. The objective of the WLS algorithm aims at minimizing the following function:

\[ F(x) = \sum_{i=1}^{m} C_{ii} (s_i - j_i(x))^2 \]

\[ = \left[ s - j(x) \right]^T C^{-1} \left[ s - j(x) \right] \]  

Subject to:

\[ s_k = j_k(x) + c_k, \quad k = 1.2.3…m \]  

where \( s \) is measurement vector, \( x \) is state vector of components, \( j \) is the nonlinear function describing measurement \( k \) to the state vector \( x \) and \( c \) is the vector of measurement error. It is assumed that measurements have an independent error, i.e., \( \text{Cov}(e) = E[ee^T] = R = \text{diag}[\sigma_1^2, \sigma_2^2, ..., \sigma_m^2] \). Thus, \( \sigma_i^2 \) indicates variance of measurement \( k \) which is considered to calculate the accuracy of corresponding metering data. The state vector can be optimized by a limit number of iterations using the Newton recursive method as given in (3)–(5) [26]:

\[ x^{i+1} = x^i - \left[ G(x^i) \right]^{-1} g(x^i) \]  

\[ G(x^i) = \frac{\partial^2 g(x^i)}{\partial x^2} = J^T(x^i)C^{-1}J(x^i) \]  

\[ g(x^i) = -J^T(x^i)C^{-1}(s - j(x^i)) \]  

where \( i \) is the iteration index, \( x^i \) is the solution vector at iteration \( i \), \( G \) is gain and \( J \) is Jacobian matrices. The state vector should estimate voltage magnitude and voltage angle at the bus \( k \) where the reference bus is bus 1 with zero voltage angle.

3 Artificial Neural Network

In this paper, a two-layer feed-forward ANN is used [17]. A layer, called hidden layer, locates between input nodes and the output layer. The hidden layer activation function is a sigmoid and linear functions are used to activate the hidden and the output layers, respectively. The ANN is trained by the scaled conjugate gradient backpropagation method which is a network training function. This function updates weight and bias values according to the scaled conjugate gradient method which is adequate for large-scale problems. The two ANNs are trained in which values of power flow measurements and power injections are considered as the input and the output, respectively. The first ANN relates to real active power flow measurements and active power injections, and the second ANN corresponds to real reactive power flow measurements and reactive power injections. The ANNs compare with load profiles and generate errors; then, the ANN updates its weight and is adjusted by those. To summarize the ANN training procedure, the following steps are given:

1. Active and reactive power data are generated using the Monte Carlo simulation for the whole year.
2. Load profiles corresponding to active and reactive power data are generated using the load flow calculation.
3. The calculated active and reactive power flows and the load profiles are used as the input and the output of ANN, respectively.
4. To save the ANN errors in the previous step for the following calculations.

4 Measurement Data Encountering Strategy

The measurement data has been built by means of Monte Carlo simulation using real historical data. However, in each run of the program, a new series of data is generated with predetermined mean and variance values. This methodology has been selected in order to show the robustness of the proposed method against various input measurement data. The input data for carrying out state estimation has been provided as follows:

1. The load data of the test network has been generated using Monte Carlo simulation in a 30 min period for a year (17520 load profiles) based on the network data given in [29].
2. For all 17520 load profiles, the load flow problem is calculated and the corresponding voltage of buses and flow of lines are provided.
3. After that, errors are randomly added to the values of bus voltages and line powers using the Monte Carlo method.

The aim of this data production is to provide both actual and noisy data within the test network for analyzing the robustness of the proposed method in properly modeling the pseudo measurement data error. Like a real network, data of bus voltage and line power vary according to load variations. Then, these values may receive errors when measured by measurement devices. In this paper, in order to have correct data, power flow calculations are executed based on some randomly generated load profiles. Then, to provide measurement data with error, randomly errors are added to the correct data.

5 Error Modelling

Errors refer to a difference between load profiles and power injection outputs of ANN. In DSSE, typically, this error is assumed to have a Gaussian density distribution. To analyze the validation of the assumption, the probability density distribution of a typical ANN output error is illustrated in Fig. 1. As shown, the error between the input and output of ANN is not similar to a Gaussian density distribution. In this
paper, the radial basis function network is proposed to process the ANN error.

5.1 Radial Basis Function Neural Network

The Radial Basis Function Network (RBF) is an ANN that uses a radial basis function as the activation function. The output of the network is a radial basis function of inputs; neurons parameters are linearly summed as illustrated in Fig. 2. The RBF network commonly uses for function approximation, time series prediction, classification and system control [27]. The RBF network is designed using three layers which include input, hidden and output layers. In the hidden layer, a non-linear RBF activation function is used where the linear function is implemented in the output layer. The output of the network is a scalar function of the input vector and as expressed in (6):

$$\phi(x) = \sum_{k=1}^{m} \alpha_k \varphi(x - \gamma_k)$$

(6)

where $m$ is the number of neurons in the hidden layer, $\gamma$ is the center vector for neuron $k$, and $\alpha$ is the weight of neuron $k$ in the linear output neuron. The Mahalanobis distance and Gaussian function respectively are used to model the norm and RBF as given in (7):

$$\varphi(x - \gamma_k) = e^{-(x - \gamma_k)^2}$$

(7)

The RBF a Gaussian kernel on feature vector is defined as (8):

$$\kappa(x,x') = e^{-\frac{(x-x')^2}{2\sigma^2}}$$

(8)

5.2 Function Approximation

5.2.1 Non-Normalized Radial Basis Function

The non-normalized RBF is made out by (9):

$$\phi(x) = \sum_{k=1}^{m} \alpha_k \varphi(x - \gamma_k)$$

(9)

where

$$\varphi(x - \gamma_k) = e^{\left(\frac{(x - \gamma_k)^2}{\sigma^2}\right)}$$

(10)

5.2.2 Normalized Radial Basis Function

The normalized RBF is made out by (9):

$$\phi(x) = \sum_{k=1}^{m} \alpha_k \varphi(x - \gamma_k) \left( \sum_{k=1}^{m} \alpha_k \varphi(x - \gamma_k) \right)^{-1}$$

(11)

and also

$$\varphi(x - \gamma_k) = e^\left(\frac{(x - \gamma_k)^2}{\sigma^2}\right) - e^\left(\frac{(x - \gamma_k)^2}{\alpha^2}\right)$$

(13)

The objective function of error modeling problem can be defined as (14) which is obtained from (13):

$$\text{cost function} = \sum_{k=1}^{m} \alpha_k \left( \frac{1}{\alpha^2} \right) e^\left(-\frac{(x - \gamma_k)^2}{\alpha^2}\right)$$

(14)

The variables of this problem are $\sigma$, $\alpha$ and $\gamma$ and the number of these variables is equal to the number of Gaussian kernels that should be optimized. The optimization is carried out via a meta-heuristic algorithm named the imperialist competitive algorithm (ICA) [28]. In this optimization problem, the ANN output that is a non-linear function error can be approximated by Gaussian kernels. Each of these kernels has its own weight which represents the effectiveness of the corresponding kernel in the
optimization problem. The largest weight in the optimization corresponds to the error of pseudo measurement data. Actually, a Gaussian function with higher variance has more impact. Also, the optimization algorithm recognizes that how many Gaussian functions should be generated for an error function.

6 State Estimation Application

In the previous section, pseudo measurements are generated by ANN and errors are created using the RBF network. A state estimation problem includes the following steps:

**Step 1:** To produce active and reactive power injection in period $t$ by measuring real active and reactive power flows as the input of ANN and then, save them as pseudo measurements.

**Step 2:** To compare the output of ANN (active and reactive power injection) with the corresponding load profile and calculate pseudo measurements errors in period $t$.

**Step 3:** To create a model for errors in RBF network with the Gaussian kernel and optimize it by ICA algorithm in order to obtain the variance of pseudo measurement errors.

**Step 4:** To construct the covariance matrix by the matrix created in the previous step and then carry out the state estimation.

In the state estimation function, inputs are real measurements, pseudo measurements, network parameters and the network topology.

7 Methodology

The aim of this paper is to accurately model the error corresponding to pseudo measurement data in order to provide an acceptable pseudo measurement data while the real measurement data is not sufficient. The pseudo measurement data is calculated from available real measured parameters and is validated by state estimation criterion. At first, the injected active and reactive powers are estimated using ANN from load demand data and real lines current measured data. The estimated values enter into RBF network aiming at detecting error. Now, there are some pseudo measurement data with known errors that can be used in state estimation. In other words, it is like a situation in which there are some measurement devices and the measurement error of these devices are known to the operator. As a result, more accurate detection of pseudo measurement error results in a more accurate estimation of system state. In this case, when the error of ANN enters the RBF network, the components of errors are separated. Then, the error distribution with the most impact is considered as pseudo-measurement error.

8 Case Study

The proposed method was applied to IEEE 123-bus distribution test system given in [29]. The system comprises 126 buses, 125 branches, and 85 active and reactive loads. Half-hourly load profiles over one year have been provided. State variables in the state estimation problem are voltage magnitude and voltage angle. The first bus assumed to be reference bus which has zero voltage angle. Two scenarios represented in Table 2 are defined to evaluate the proposed method. To model pseudo measurements, two ANN are considered. In the first ANN, variables are two real power flow measurements at the substation ($P_{116-1}$, $Q_{116-1}$, $P_{17-1}$, $Q_{17-1}$ and $P_{13-1}$, $Q_{13-1}$), the variables in the second ANN are three additional sets of measurements ($P_{13}$, $Q_{13}$, $P_{7-8}$, $Q_{7-8}$ and $P_{118-52}$, $Q_{118-52}$). It is assumed that the real measurement is available at the substation and DG locations. Thus, the voltage magnitude at bus $85$ ($V_{85}$), the reactive power flows in lines $116-1$, 1-17 and 13-18 ($P_{116-1}$, $Q_{116-1}$, $P_{17-1}$, $Q_{17-1}$ and $P_{13-1}$, $Q_{13-1}$), the active and reactive power injection at bus 1, 2, 6, 7 and 10 ($P_1$, $Q_1$, $P_2$, $Q_2$, $P_6$, $Q_6$, $P_7$, $Q_7$ and $P_{10}$, $Q_{10}$), were considered as real measurements in scenario 1. The additional voltage measurement at bus 84 and 94 ($V_{84}$, $V_{94}$), the active and reactive power flows in lines 8-13, line 7-8 and line 118-52 ($P_{8-13}$, $Q_{8-13}$, $P_{7-8}$, $Q_{7-8}$ and $P_{118-52}$, $Q_{118-52}$) were used as additional measurements in scenario 2. Real measurements are obtained by running a load flow calculation in which a Gaussian uncertainty with 3% error around the mean values has been considered. Also, load flow with Gaussian uncertainty values is implemented to provide the input of ANN and the function of state estimation. The mixture component is recognized by RBF with Gaussian Kernel and self-tuning by ICA algorithm. To test ANN training performance, error modeling and state estimation are considered in 30-minute steps during a year. In scenarios 1 and 2, the voltage magnitude and the voltage

<table>
<thead>
<tr>
<th>Scenario</th>
<th>ANN Input</th>
<th>Real Measurements</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>$P_{116-1}$, $Q_{116-1}$, $P_{17-1}$, $Q_{17-1}$, $P_{13-18}$, $Q_{13-18}$</td>
<td>$V_{85}$, $P_{116-1}$, $Q_{116-1}$, $P_{17-1}$, $Q_{17-1}$, $P_{13-18}$, $Q_{13-18}$, $P_{7-8}$, $Q_{7-8}$, $P_{118-52}$, $Q_{118-52}$</td>
</tr>
<tr>
<td>2</td>
<td>$P_{116-1}$, $Q_{116-1}$, $P_{17-1}$, $Q_{17-1}$, $P_{13-18}$, $Q_{13-18}$, $P_{7-8}$, $Q_{7-8}$, $P_{118-52}$, $Q_{118-52}$</td>
<td>$V_{85}$, $V_{84}$, $V_{94}$</td>
</tr>
</tbody>
</table>
Error Modelling in Distribution Network State Estimation Using … A. Hassannejad Marzouni and A. Zakariazadeh

Fig. 3 Voltage magnitude and voltage angle at bus 83 in scenario 1.

Fig. 4 Voltage magnitude and voltage angle at bus 83 in scenario 2.

$$err_v = 100 \left( \sum_k \left( \left| \frac{v^e_k - v^t_k}{v^t_k} \right| \right) \right) n^{-1}$$  (15)

$$err_\delta = 100 \left( \sum_k \left( \left| \frac{\delta^e_k - \delta^t_k}{\delta^t_k} \right| \right) \right) n^{-1}$$  (16)

where \(v^e\) is voltage magnitude estimation value, \(v^t\) is voltage magnitude true value, \(\delta^e\) is voltage angle estimation value, \(\delta^t\) is voltage angle true value, \(k\) is sampling step and \(n\) is measurement sampling steps.

The errors of voltage angle in scenario one and two for bus 83 are shown in Figs. 5 and 6. The results showed that the average of the state estimation in scenario 2 has been improved if compared with the one in scenario 1. The increase in voltage measurement at ANN input improves state estimation accuracy. Table 3 presents the mean of comparative voltage magnitude and voltage angle errors for buses 114, 83 and 109 in two scenarios. In order to show the advantage of the proposed method, the simulation has been also carried out using the error modeling method presented in [21] where the network and input data are the same. The comparison results are shown in Table 4. It should be noted that the maximum amount of relative voltage magnitude error provided by the method in [21] was 0.07%. As a comparison, the relative voltage magnitude error provided by the proposed method is 0.03%. Also, regarding relative voltage angle errors, the maximum values provided by the methods presented in [21] and this paper are 25% and 4.05%, respectively. So, the proposed method results in lower relative error if compared to the one presented in [21]. Load profile with 25% uncertainty is used where ANN-based approach provides pseudo measurements. It is observed that voltage magnitude errors are the same in scenarios but the proposed approach generates an estimated voltage angle with a lower error. Meanwhile, for the accuracy of power flow estimations, the performance of the proposed method is
evaluated via a scenario. It is assumed that the load demand is high in a critical day of winter. The ANN-based pseudo measurement modeling approach is used to estimate the line power flow. To verify the concept, it is assumed that the mean value of power flow estimations was around 96% of the line rating. The mean and variance values used for probability densities of line power flow estimations in all lines are computed for scenario 1. The state estimated line flow probability densities for comparison purposes are computed using 25%, 35%, 50%, and 65% uncertainty in pseudo measurements of the load which derived from the load profile mean and are also shown in Fig. 7. The uncertainty in the measurement of the load increases while the variation around the mean increases.

9 Discussion

On the other hand, in the absence of real measurement data from measurement devices, the observability of the network reduces. In this case, pseudo measurement data, as an alternative, can be taken into account for measurement data. It is worth to be noted that accurate modeling of pseudo measurement data results in accurate state estimation of the system.

A new approach for error modeling of pseudo measurement data has been presented in this paper. In the related works, some approaches such as GMM-MLE [21] have been presented aiming at modeling...
The proposed method uses offline load flow and load profiles to train two ANNs. An RFB based approach including ANNs can effectively synchronize the average load profiles with the real measurements. The results showed that the main advantage of the proposed method is to provide more accuracy in terms of error modeling. Also, in the case of large-scale networks and the network with low available metering data, the proposed method has acceptable compatibility.

### References


A. Hassannejad Marzouni received the B.Sc in the Mazandaran University of Science and Technology, Babol, Iran, in 2016 and M.Sc in the University of Science and Technology of Mazandaran, Behshahr, Iran in 2018. His research work focused on state estimation application in distribution systems, distributed measurement system applied to electric distribution grids, and active and reactive power market. His current research interests include electric vehicles charging station placement with GIS considerations, electric vehicles load predictions and, estimation of electric vehicles’ owner behavior. Also, he is a freelance and full-stack programmer in Python, Java, JavaScripts, and Matlab programming languages and he is conducting and developing desktop and web-based applications.
A. Zakariazadeh received the B.Sc. (Hons.) degree in Electrical Engineering from University of Mazandaran, Babol, in 2007 and the M.Sc. (Hons.) and the Ph.D. degrees in Electrical Engineering from Iran University of Science and Technology, Tehran in 2010 and 2014, respectively. Since 2016, he has been an Assistant Professor with University of Science and Technology of Mazandaran, Behshahr, Iran. His research interests include smart grid, demand response, electric vehicles, and stochastic optimization.

© 2020 by the authors. Licensee IUST, Tehran, Iran. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution-NonCommercial 4.0 International (CC BY-NC 4.0) license (https://creativecommons.org/licenses/by-nc/4.0/).