IEEE Access
Multidisciplinary : Rapid Review : Open Access Journal

Received 25 April 2023, accepted 14 May 2023, date of publication 18 May 2023, date of current version 31 May 2023.

Digital Object Identifier 10.1109/ACCESS.2023.3277616

RESEARCH ARTICLE

Unscented Trainable Kalman Filter Based on Deep Learning Method Considering Incomplete Information

YANJIE YU[1], QIANG LI[1], (Member, IEEE), AND HOUYI ZHANG[2]

[1]State Key Laboratory of Power Transmission Equipment and System Security and New Technology, School of Electrical Engineering, Chongqing University, Chongqing 400030, China

[2]Electric Power Research Institute, Guizhou Power Grid Company Ltd., Guiyang 550005, China

Corresponding author: Qiang Li (qiangli@cqu.edu.cn)

This work was supported by the China Southern Power Grid through the research project 'Research on reliable operation of rural microgrids based on multiple data sources' under Award 035300KK52210006.

ABSTRACT Rapid changes of states and occurrence of data missing in power systems cause accurate state estimation very hard. In this paper, an unscented trainable Kalman filter (UTKF) with a deep learning prediction model is proposed to provide accurate state estimation under incomplete information. First, the CNN-LSTM architecture, a typical deep learning model, is applied to form a trainable prediction model (TPM), which offers more accurate prediction of states. However, sometimes states are incomplete due to data losses in transmissions. To deal with incomplete information, historical time-series states are employed and fed to the TPM in order to develop a missing data filling method. In this way, the prediction errors can be lower through the online training and parameter adjustment of the TPM. Combining with the TPM and the missing data filling method, an unscented trainable Kalman filter (UTKF) is proposed to improve the state estimation of power systems when incomplete information is involved. Finally, three cases are designed, and the simulation results show that for the prediction of states, the root mean square error (RMSE), an indicator of accuracies, is reduced by about 3 multiples, if our missing data filling method is added. Furthermore, the accuracy of state estimation is improved about 5 multiples by the proposed UTKF method, even if incomplete information is involved.

INDEX TERMS Dynamic state estimation, deep learning, prediction model, missing data filling, unscented Kalman filter.

I. INTRODUCTION

Power system states, sampled by measurement units, provide the most effective method to analyze the operating condition of power systems directly, which helps the detection of latent faults and prevents the grid from sudden breakdown [1]. However, the pseudo-measurement is influenced by sensor sampling errors and transportation disturbance, resulting in cumulative errors during the operation [2]. For this reason, state estimation is proposed to eliminate the residual between the actual value and pseudo-measurement to increase the accuracy of data [3], and it plays a critical role in distribution networks [4].

Usually, the state estimation is divided into static and dynamic algorithms based on the application scenarios [5]. Widely used static state estimation (SSE) concentrates on the numerical analysis in the constant time to search the sources of errors and reduce the stable errors through the weight least square (WLS) method but it is weak in estimating the changes state due to the complex calculating process [6]. Opposite to SSE, the dynamic state estimation (DSE) is proposed to track the actual state by combining the prior probability of prediction and posterior probability of measurement to obtain the closest state to true value after regarding the noises as the probability distribution [7], [8]. Furthermore, DSE is

The associate editor coordinating the review of this manuscript and approving it for publication was Halil Ersin Soken.

commonly divided into the prediction process for predicting the next-time state and the measurement process for transporting the errors for estimation. Rather than a steady system, the states in power systems are affected by tiny disturbances easily and fluctuates in a large magnitude frequently [9]. Thus, the DSE is more suitable for distribution networks to reduce the errors in pseudo-measurement.

As an important part of the DSE, Kalman filter (KF) proposed in 1960 is an efficient method to solve linear filter problems and prediction issues, and this algorithm could be incorporated in most scenarios without any modification [10]. Besides, the estimation is regarded as two steps, the prediction process and measurement process, in KF, and the key purpose of KF is to obtain the differential equations derived from the variance matrix of the optimal estimation errors [11]. However, KF is more accurate only in the linear system because it requires the prediction process and measurement to be described in linear equations, which brings difficulty to the application in complex nonlinear systems. For this reason, the extended Kalman filter (EKF) is proposed to expand the application of Kalman filter by processing the nonlinear function [12]. The Jacobian matrix of prediction and measurement process derived from the differential equations only obtains the first-order part in the Taylor series expansion [13]. That makes it quick in the calculation but the higher-order data is abandoned, which results in the low accuracy and weak convergence rate of EKF. Combining the unscented transform (UT) and KF, the unscented Kalman filter (UKF) presents a great performance in dynamic state estimation recently [14], [15]. As an efficient transform method, UT maps the state from the prediction process into the measurement process according the linear or nonlinear function with no loss [16], [17]. That indicates all higher-order information is reserved, thus the estimation of UKF is more accurate than that of EKF.

Many researchers study dynamic state estimation in power systems. Several researchers consider the variances as a constant matrix and lack of the evaluation and the analysis of errors [18], [19]. Indeed, the errors in the prediction process are changing during each state transition. Only the errors in the measurement process could be regarded as constant due to the sensors residual [20]. It is difficult for those estimators to filter the errors derived from the forecasting process. Furthermore, some researchers consider the power systems as the simple model to predict the state such like the linear model that neglect the practical operation condition of power system [21], [22], [23]. The states of the power systems are fluctuating frequently, it is essential to forecast the next-time state in an accurate method to accelerate the convergence rate and increase the accuracy. Other works of estimation adopt algorithms that are not suitable for estimating the states of the power systems [24]. To specific, EKF is quicker than UKF in the calculation speed, but UKF is more accurate than EKF. For some scenarios, the error accumulation effect from EKF results in a larger derivation in the controller or dispatch center. Besides, most estimations do not consider the effectiveness and integrality of data under harsh operating conditions such as power line disconnection, different measurement units breakdown and the data sequences accepted by the state estimator could miss sometimes [25], [26]. The data missing could lead to the error accumulation and economic losses.

To get a more accurate prediction state, the realistic power systems model is essential to be considered. With the development of deep learning (DL), artificial neural network (ANN) becomes a critical component of the prediction algorithm over the last few decades [27], [28]. The deeper hidden architecture of ANN has begun to surpass the classical method in many scenarios. Based on that convolution neural network (CNN) is the most popular to extract key information from the specified object [29], [30], [31]. Some researchers concentrate on the excellent information processing ability of CNN, and the results demonstrate that CNN could extract the state labeled with time in power systems [32]. Considering the state is fluctuating with operating time, the CNN is suitable for the power systems in the prediction process. Nevertheless, the hidden layers of CNN limited the performance in the time of calculating for state estimation in power systems, and predicting the speed of CNN is influenced by its multilayers architecture, which makes it hard to forecast state in power systems [33]. As one of the most powerful dynamic predictors, long short-term memory (LSTM) is successfully used in a different field to forecast time-series data in the low time cost, which could also be used in state estimation in the prediction process [34]. Thus, the combining of CNN and LSTM could compensate for the shortcomings of each other in power systems for state estimation in the prediction process.

As mentioned above, the existing prediction models of power systems are weak to provide accurate state prediction during operation, and traditional DSE algorithms could not fill the missing data during operation. In this paper, an unscented trainable Kalman filter (UTKF) with the CNN-LSTM architecture is proposed, which improves the accuracy of state estimation significantly, even if data losses occur. The simulation results demonstrate that the proposed TPM decrease the RMSE in prediction and missing data filling, and the UTKF achieves better performance in accuracy of estimation in power systems involved incomplete information. Therefore, the main contributions are listed as follows

1) To improve the prediction accuracy, the deep learning (DL) model, CNN-LSTM, is introduced to dynamic state estimation (DSE), which leads to a trainable prediction model (TPM) of power systems. Through online training in the CNN-LSTM and parameter adjustment, the forecasting errors are reduced, which makes the prediction more accurate.

2) Data losses always exist in measurement and transmission. To deal with incomplete information in the DSE, historical time-series states are employed to design a missing data filling method, which offers a way to avoid error accumulation.

3) An unscented trainable Kalman filter (UTKF), combined with deep learning (DL) and Kalman filter (KF) is proposed to tackle the problem of state estimation when incomplete information is involved. The covariance matrix of states is analyzed and a proposition is proved, which indicates that the estimated errors are reduced effectively.

The paper is organized as follows. Section II describes the problem formulation in power systems. Section III details the unscented trainable Kalman filter with missing data filling. Section IV discusses the effectiveness and accuracy of the proposed algorithms by simulation. Section V draws the conclusion.

II. PROBLEM FORMULATION

In this section, the problem formulation of power systems is introduced. The DSE is revised into a nonlinear formulation according to the actual operation condition of power systems.

In this paper, power systems installed with phasor measurement units (PMU) and supervisory control and data acquisitions (SCADA) is considered. The nodes in power systems are described as $i \subset \mathcal{N} = \{1, 2, \ldots, n\}$. $\mathcal{N}$ represents the set of the bus number. If node i and node j is connected directly, they are called neighbors, and all neighbors of node i are $j \subset \mathcal{N}_i$.

The state transition of the network with n nodes could be described as

$$x(t+1) = Fx(t) + w(t) \tag{1}$$

where $x(t)$ is the state at time t; F is the transition matrix; $w(t)$ is the transition noise. The state of power systems is dynamic and influenced by any changes in the grid, the equation (1) could decrease the response time effectively of estimation by predicting the next-time state. The state vectors are

$$x(t) = [u(t), v(t)]^T \tag{2}$$

where $u(t)$ and $v(t)$ are the magnitude and phasor of node voltage, respectively. Also, the deviation exists during measurement. The measurement could be described as

$$z(t+1) = Hx(t+1) + q(t+1) \tag{3}$$

where $z(t+1)$ is the measurement at time $t+1$, H is the measurement matrix of ith node; and $q(t+1)$ is the measurement noise. The noise of measurement occurs when the sensors sample or transport the state data inevitably. The measurement vectors are

$$z(t) = \left[P(t), Q(t), U(t), V(t), P_{i,j}(t), Q_{i,j}(t)\right]^T \tag{4}$$

where $P(t)$ and $Q(t)$ are the node positive power and negative power, respectively; $U(t)$ and $V(t)$ are the measurement of magnitude and phasor of node voltage; $P_{i,j}(t)$ and $Q_{i,j}(t)$ are the branch power from i to j nodes, $j \subset \mathcal{N}_i$.

The equations (1) and (3) formulate the problem for a linear system. However, the power systems is characterized by strong nonlinearity and frequent fluctuation, which leads to the nonlinear function of F and H. Consequently, the problem could be described as

$$x(t+1) = f(x(t)) + w(t) \tag{5}$$

$$z(t+1) = h(x(t+1)) + q(t+1) \tag{6}$$

where $f(x)$ and $h(x)$ are the nonlinear transition and measurement function, respectively. The equations (5) and (6) describe the model in state transition and measurement process of multi-agent distribution networks.

In this paper, the algorithm is proposed to estimate the next-time state to cope with the dynamic and nonlinear power systems. The main aim of this paper is to filter the noise which exists in the prediction and measurement process and fill the missing data based on the historical time-series data to increase the accuracy of measurement in power systems. Power systems discussed in this paper is regarded as absolutely observable, the data missing only occurs in several buses at the same time. The data error in the prediction and measurement process is distributed as Gaussian noise.

III. UNSCENTED TRAINABLE KALMAN FILTER WITH MISSING DATA FILLING

In this section, an UTKF algorithm is proposed to improve the accuracy of the state in power systems involved incomplete information. The UTKF is divided into the prediction process and measurement process. To increase the estimation accuracy and reduce the complexity of calculation, the prediction process provides a prior mathematical expectation for the measurement process. Then, the more accurate state is estimated based on the predictions. Considering the data losses in power systems, traditional UKF is limited on solving the missing data problem in power systems, thus, the UTKF algorithm is proposed to increase the effectiveness of estimation.

A. TRAINABLE PREDICTION MODEL OF POWER SYSTEM

Although the bus voltage is dynamically changing during operation, the trend of the state is correlated with the residential and industrial consumption habits, climate, and renewable generation. Thus, it is feasible to predict the state based on DL. The trainable prediction model (TPM) is a new approach to a model complex system. As an important part of the neural network, the convolution neural network (CNN) is implemented into the information extraction. In fact, power systems could be regarded as a system with the input of renewable resource generation, injection power, and consumer power of several nodes, and the states in power systems are not only interrelated with themselves but correlated with the input mentioned above. Therefore, the input information with time-series is evaluated by CNN to extract the correlation.

The CNN is implemented to extract the information for the prediction process, and it could be divided into four layers:

1) Embedded layer. The state data is input through the embedded layer and is cut into pieces. The time-series state

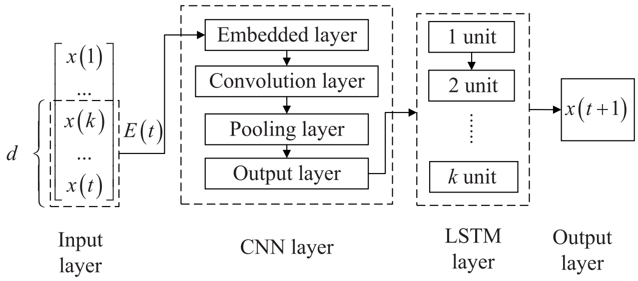


FIGURE 1. CNN-LSTM architecture.

vector is $X = [x(1), \ldots, x(t)]^T$ at time t, and the output of embedded layer of ith node is $E(t) = [x(t-d), \ldots, x(t)]^T$. The length of the output is depending on the size of convolution kernel d.

2) Convolution layer. The data of ith node is obtained by convolution layers, which could be described as

$$x^i = g(M^i E) + b^i \tag{7}$$

where x^i is the output of convolution layer; M^i is the weighed matrix and b^i is the constant term.

3) Pooling layer. The pooling layer could simplify the calculation time during prediction by the pooling algorithm, usually the max-pooling method. Similarly important with the convolution layer, the data aggregated could increase the effectiveness and accuracy of prediction.

4) Dense layer. The architecture of CNN is established for extracting the features of data, for this reason, the work of the dense layer is to output the reorganized data for the next layer.

Although CNN could identify the data effectively, the prediction of CNN is limited by its convergence time, which is unable to output prediction on time. The long short-term memory units (LSTM) present the superior prediction of time-series data, recently. After the pre-processing of data by CNN, the data features are extracted, then those would be transported to prediction units. The jth unit in LSTM architecture with k units is described as

$$\begin{aligned} c^{(j)} &= e^{(j)} \odot \phi(M^c x^i + R_i^c y^{(j-1)} + b^c) \\ &\quad + p^{(j)} \odot c^{(j-1)} &(8)\\ e^{(j)} &= \sigma(M^e x^i + R_i^e y^{(j-1)} + b^e) &(9)\\ p^{(j)} &= \sigma(M^p x^i + R_i^p y^{(j-1)} + b^p) &(10)\\ o^{(j)} &= \sigma(M^o x^i + R_i^o y^{(l-1)} + b^o) &(11)\\ y^{(j)} &= o^{(j)} \odot \rho(c^{(j)}) &(12)\end{aligned}$$

where j is the number of LSTM units; $e^{(j)}$, $p^{(j)}$ and $o^{(j)}$ are the input, forget, and output gate, respectively. $c^{(j)}$ is the state vector of jth unit; $y^{(j)}$ is the output of LSTM; $M = [M^e, M^p, M^o]$, $R = [R^e, R_i^p, R^o]$ and $b_i = [b^e, b^p, b^o]$ are the state parameters of input, forget and output gate, respectively. σ is the point-wise sigmoid function; ϕ and ρ are the hyperbolic tangent function; $\odot$ denotes the corresponding elements multiplication of matrix. The state parameters of LSTM are obtained by training to get a more accurate prediction state, and every node of power systems is installed with CNN-LSTM architecture. The architecture of CNN-LSTM is shown in Fig.1, and the prediction process is

$$\widehat{x}(t+1) = f(E(t)) \tag{13}$$

where $\widehat{x}(t+1)$ is the prediction state. Besides, the TPM of CNN-LSTM architecture could correct the parameters in neural network by learning and calculate the residual in operating. For this reason, the CNN-LSTM architecture could be implemented without an error estimator, and the residual function is

$$\epsilon(t+1) = s(x(t+1), \widehat{x}(t+1)) \tag{14}$$

where $\epsilon(t+1)$ is the residual at time $t+1$; $\widehat{x}_i(t+1)$ is the prediction state vector at time t; $s(\cdot)$ is the residual function.

B. MISSING DATA FILLING METHOD

When meeting the incomplete information in measurement, the prediction need to detect and fill the missing data. If the networks are only trained in offline data, the missing data would decrease the prediction accuracy and error accumulation. Supposed that the dispatch center could detect data missing in measurement, if the measurement data couldn't be read by estimator, it would be regarded as missing data, and the estimator need to fill it to avoid accumulation of errors. Thus, the online training method with missing data filling method is proposed.

Considering the neural network is also a project to minimize the residual between prediction and true value, the online training method could be incorporated into the DSE. The parameter information matrix of ith agent is $H_{tr}^{(j)}(t) = [M^{c,(j)}, M^{e,(j)}, M^{p,(j)}, M^{o,(j)}, R^{c,(j)}, R^{e,(j)}, R^{p,(j)}, R^{o,(j)}, b^{c,(j)}, b^{e,(j)}, b^{p,(j)}, b^{o,(j)}]^T$, and it could be adjusted to output more accuracy prediction during system operating. The state transition process according to (8) and (12) is

$$\begin{aligned} \widehat{c}(t+1) &= \Lambda(\widehat{c}(t), E_i(t), \widehat{y}(t)) + \epsilon_c(t) &(15)\\ \widehat{y}(t+1) &= \Upsilon(\widehat{c}(t), E_i(t), \widehat{y}(t)) + \epsilon_y(t) &(16)\\ \widehat{H}_{tr}(t+1) &= \widehat{H}_{tr}(t) + \epsilon_h(t) &(17)\end{aligned}$$

where $R(t) = [\epsilon_c(t), \epsilon_y(t), \epsilon_h(t)]$ are the errors should be filtered; The function $\Lambda(\cdot)$ and $\Upsilon(\cdot)$ is the nonlinear transition function of LSTM unit. Furthermore, the Jacobian matrix is adopted to linearize the transition function. The transition equations could be written as

$$\begin{aligned} \widehat{c}(t+1) &= \Lambda(\widehat{c}(t), E_i(t), \widehat{y}(t)) &(18)\\ \widehat{y}(t+1) &= \Upsilon(\widehat{c}(t), E_i(t), \widehat{y}(t)) &(19)\\ \widehat{H}_{tr}(t+1) &= \widehat{H}_{tr}(t) &(20)\\ \widehat{\Sigma}_a(t+1) &= F_{tr}(t+1)\widehat{\Sigma}_a(t)F_{tr}(t+1)^T \\ &\quad + R(t+1) &(21)\end{aligned}$$

Algorithm 1 Online Training Based on TPM

Initialization: $\widehat{\Sigma}_a(t) \leftarrow \widehat{\Sigma}_a(0)$

Information matrix calculation:

1) Construct information matrix.

$\xi(t+1) \leftarrow [\widehat{c}(t+1), \widehat{y}(t+1), \widehat{H}_{tr}(t+1)]^T$

$\psi(t+1) \leftarrow \widehat{\Sigma}_a(t)$

2) Get next-time parameters by (18), (19),(20), and (21).

3) For power systems with N nodes, calculate the information matrix.

$\eta(t+1) \leftarrow \psi(t+1) + R(t+1)$

$\xi(t+1) \leftarrow \xi(t+1) + \psi(t+1)\eta(t+1)^{-1} \cdot (x(t+1) - \widehat{x}(t+1))$

$\psi(t+1) \leftarrow \psi(t+1) - \psi(t+1)\eta(t+1)^{-1}\psi(t+1)^T$

Update parameters:

$[\overline{c}(t+1), \overline{y}(t+1), \overline{H}_{tr}(t+1)] \leftarrow \xi(t+1)/N$

$\overline{\Sigma}_a(t+1) \leftarrow \psi(t+1)$

where $\widehat{\Sigma}_{a_i}(t+1)$ is the covariance matrix for jth unit at $t+1$ time, and the Jacobian matrix $F_{i,tr}(t+1)$ is

$$F_{i,tr}(t+1) = \begin{bmatrix} \frac{\partial\Lambda(\widehat{c},E,\widehat{y})}{\partial\widehat{c}} & \frac{\partial(\Lambda\widehat{c},E,\widehat{y})}{\partial\widehat{y}} & \frac{\partial(\Lambda\widehat{c},E,\widehat{y})}{\partial\widehat{H}} \\ \frac{\partial\Upsilon(\widehat{c},E,\widehat{y})}{\partial\widehat{c}} & \frac{\partial(\Upsilon\widehat{c},E,\widehat{y})}{\partial\widehat{y}} & \frac{\partial(\Upsilon\widehat{c},E,\widehat{y})}{\partial\widehat{H}} \\ 0 & 0 & I \end{bmatrix} \tag{22}$$

where $\widehat{c} = \widehat{c}(t+1)$; $\widehat{H} = \widehat{H}_{tr}(t+1)$; $\widehat{y} = \widehat{y}(t+1)$. Consequently, for agent i at time $k+1$ with neighbor $j \in \mathcal{N}_i$ the iteration [2] of distributed CNN-LSTM is

$$\eta(t+1) = \psi(t+1) + R(t+1) \tag{23}$$

$$\xi(t+1) = \xi(t+1) + \psi(t+1)\eta(t+1)^{-1} \cdot (x(t+1) - \widehat{x}(t+1)) \tag{24}$$

$$\psi(t+1) = \psi(t+1) - \psi(t+1)\eta(t+1)^{-1}\psi(t+1)^T \tag{25}$$

where $\eta(t+1)$, $\xi(t+1)$ and $\psi(t+1)$ is the information matrix. The training result is calculated as

$$\begin{bmatrix} \overline{c}(t+1) \\ \overline{y}(t+1) \\ \overline{H}_{tr}(t+1) \end{bmatrix} = \frac{1}{N}\xi(t+1) \tag{26}$$

where N is the number of nodes in power systems. The method of online training based on trainable prediction model (TPM) is Algorithm 1.

The incomplete information leads that the secondary devices being unable to respond on time, which would influence the power supply and stability of the grid. Besides, several missing data results in the wrong calculation of state estimation, and the cumulative error could also lead to economic losses.

The missing data matrix is described as

$$x(t) = [x_1(t), \text{NaN}, \ldots, x_{n-1}(t), x_n(t)]^T \tag{27}$$

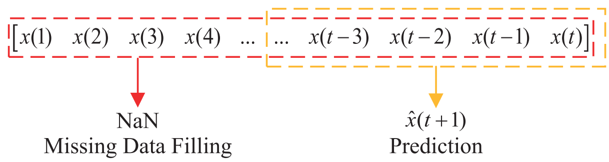


FIGURE 2. The comparison of the input matrix between the prediction and missing data filling.

where NaN represents the missing data in state matrix; $x(t)$ is the state matrix with missing data.

If the missing data appears in the state matrix, the estimation would not operate as normal because the estimator couldn't calculate the deviation between the prediction state and the actual state. Fig.2 shows the comparison of the input matrix between the prediction and missing data filling. The whole states from initial operation time are as the input matrix for missing data filling, and the several latest states are as the input matrix for prediction. For the lth missing data in $x(t)$, the filling state is

$$\widehat{x}(t+1) = f(E_{1:t}(t)) \tag{28}$$

where $E_{1:t}(t)$ is the state time-series matrix from initial time to time t, and the filled state matrix is

$$\widehat{x}(t) = [x_1(t), \widehat{x}_2(t), \widehat{x}_3(t), \ldots, x_{n-1}(t), x_n(t)]^T \tag{29}$$

where $\widehat{x}_2(t)$ and $\widehat{x}_3(t)$ is the filling state predicted by the CNN-LSTM architecture.

C. UNSCENTED TRAINABLE KALMAN FILTER

To solve the nonlinear function and decrease the error during estimation, the unscented transition (UT), which could transport the variables and not change the raw distribution of the data, is adopted to eliminate the model nonlinearization. Sigma points are adopted after the prediction process as follows

$$\chi^{(1)}(t+1) = \widehat{x}(t+1) \tag{30}$$

$$\chi^{(k)}(t+1) = \widehat{x}(t+1) + [\sqrt{(k+\lambda)\widehat{\Sigma}_x(t+1)}]_k \tag{31}$$

$$\chi^{(k)}(t+1) = \widehat{x}(t+1) - [\sqrt{(k+\lambda)\widehat{\Sigma}_x(t+1)}]_k \tag{32}$$

where $\chi^{(k)}(t+1)$ is the kth sigma point of prediction $x(t+1)$; $\widehat{\Sigma}_x(t+1)$ is the covariance matrix of predicting state; k are from 2 to $n+1$ and $n+2$ to $2n+1$ in equations (31) and (32), respectively; $(\cdot)_k$ is the k column of the matrix and $\sqrt{\cdot}$ is operation of Cholesky factorization, respectively. After the generation of sigma points, the measurement process could be described as

$$\gamma^{(k)}(t+1) = h(\chi^{(k)}(t+1)), k = 1, \ldots, 2n+1 \tag{33}$$

$$\widehat{z}(t+1) = \sum_{k=1}^{2n+1} \theta^{(k)}\gamma^{(k)}(t+1) \tag{34}$$

$$\widehat{\Sigma}_z(t+1) = \sum_{k=1}^{2n+1} \theta^{(k)}(\gamma^{(k)}(t+1) - \widehat{z}(t+1)) \cdot (\gamma^{(k)}(t+1) - \widehat{z}(t+1))^T + O(t+1) \quad (35)$$

where the weights $\theta^{(k)} = 1/2N, k = 1, 2, \ldots, 2N+1$; $\gamma^{(k)}(t+1)$ is the sigma point of measurement; $\widehat{\Sigma}_z(t+1)$ is the measurement matrix based prediction; $\widehat{\Sigma}_z(t+1)$ is the variance matrix of measurement; $O(t+1)$ is the noise matrix of measurement process. The UTKF algorithm is summarized in Algorithm 2.

Algorithm 2 Unscented Trainable Kalman Filter (UTKF)

Off-line training iteration:

1) Time-series matrix $E(t) \leftarrow [x(1), \ldots, x(t)]^T$

2) Calculate the output by (8), (9), (10), (11) and (12).

3) Calculate the residual ϵ by (14).

4) Adjust parameters according to ϵ.

If $\epsilon < \beta$

5) Calculate the initial variance matrix $\widehat{\Sigma}_x(0)$

where $E(t)$ consists of the operating state from time 1 to t; β is the tolerance error.

Initialization:

$x(t) \leftarrow x(0)$ $\widehat{\Sigma}_x(t) \leftarrow \widehat{\Sigma}_x(0)$

Prediction process:

1) Construct input matrix $E(t) = [x(t-d), \ldots x(t)]$.

2) Output the prediction state matrix $\widehat{x}(t+1)$ by (8), (9), (10), (11) and (12).

3) Operate the Algorithm 1.

4) Get residual $\epsilon(t+1)$ by (14), calculate the $\widehat{\Sigma}_x(t+1)$ and update the parameter in CNN-LSTM architecture.

5) $\widehat{\Sigma}_x(t+1) \leftarrow \widehat{\Sigma}_x(t)$

If find *NaN* **in** $x(t)$**:**

6) Fill the missing data by (28).

Measurement process:

1) Generate sigma points by (30), (31) and (32).

2) Calculate measurement based on prediction by (33), (34) and (35).

Estimation process:

$K = \widehat{\Sigma}_{xz}(t+1)[\widehat{\Sigma}_z(t+1)]^{-1}$

$\overline{x}(t+1) \leftarrow \widehat{x}(t+1) + K(z(t+1) - \widehat{z}(t+1))$

$\overline{z}(t+1) = h(\overline{x}(t+1))$

where $\overline{z}(t+1)$ and $\overline{x}(t+1)$ are the estimation of measurement and state vector which is consist whole nodes of power systems, respectively; K is the Kalman gain; $\widehat{\Sigma}_{xz}(t+1)$ is the covariance matrix between $\widehat{x}(t+1)$ and $\widehat{z}(t+1)$; $z(t+1)$ is the pseudo-measurement at time $t+1$.

Update:

$E(t+1) \leftarrow [x(2), \ldots, \overline{x}(t+1)]$, $\widehat{\Sigma}_z(t) \leftarrow \widehat{\Sigma}_z(t+1)$.

Algorithm 2 demonstrates the process of UTKF, the estimation of state is obtained according to the Proposition 1.

Proposition 1: If the $\widehat{x}(t+1)$, $\widehat{z}(t+1)$ and $\widehat{\Sigma}_z(t+1)$ are obtained by (13), (34) and (35), respectively, the optimal estimation state is

$$\overline{x}(t+1) \leftarrow \widehat{x}(t+1) + K(z(t+1) - \widehat{z}(t+1)) \quad (36)$$

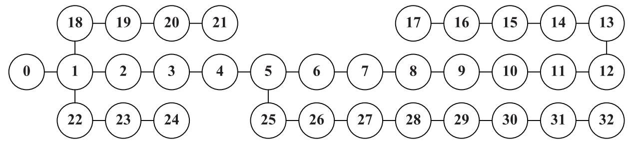


FIGURE 3. IEEE 33-bus System.

and the optimal Kalman gain is

$$K = \widehat{\Sigma}_{xz}(t+1)[\widehat{\Sigma}_z(t+1)]^{-1} \quad (37)$$

Proof 1: The main purpose of state estimation based on Kalman filter is to minimize the variance of state, the error between true state and estimated state is

$$r(t+1) = x(t+1) - \overline{x}_i(t+1) \quad (38)$$

Substituting the (14) and estimated state into (38), and the equation is

$$r_i(t+1) = \epsilon(t+1) - K[z_i(t+1) - \widehat{z}_i(t+1)] \quad (39)$$

Considering the variance of the real and predicted state is the important parameters to evaluate the error, the covariance matrix is defined as

$$\begin{aligned}\overline{\Sigma}_x(t+1) &= E[r(t+1)r(t+1)^T] \\ &= \widehat{\Sigma}_x(t+1) + K\widehat{\Sigma}_z(t+1)K_i^T - 2K\widehat{\Sigma}_{xz}(t+1)\end{aligned}$$

where $E(\cdot)$ is the calculation of expectation; $\widehat{\Sigma}_{xz}(t+1) = [x(t+1) - \widehat{x}(t+1)][z(t+1) - \widehat{z}(t+1)]^T$. To eliminate the residual in prediction and measurement, the variance between true state and estimated state need to be minimized, and the partial derivative of with $\overline{\Sigma}_x(t+1)$ respect to K is

$$\frac{\partial[tr(\overline{\Sigma}_x(t+1))]}{\partial K} = 2K\widehat{\Sigma}_z(t+1) - 2\widehat{\Sigma}_{xz}(t+1) \quad (40)$$

In order to get the optimal K, setting the $\frac{\partial[tr(\overline{\Sigma}_x(t+1))]}{\partial K} = 0$ as

$$2K\widehat{\Sigma}_z(t+1) - 2\widehat{\Sigma}_{xz}(t+1) = 0 \quad (41)$$

So the optimal Kalman gain $K = \widehat{\Sigma}_{xz}(t+1)[\widehat{\Sigma}_z(t+1)]^{-1}$.

IV. CASE STUDY

In this section, the simulation demonstrates the efficiency and accuracy of the proposed UTKF algorithm by comparing it with other popular algorithms in power systems involved incomplete information. The parameters of the simulation are set at first. In case 1, the accuracy of proposed TPM is verified by comparing the prediction with traditional model. The accuracy of prediction under incomplete information is verified in case 2. At last, the case 3 demonstrates the accuracy of proposed UTKF with the missing data filling for state estimation.

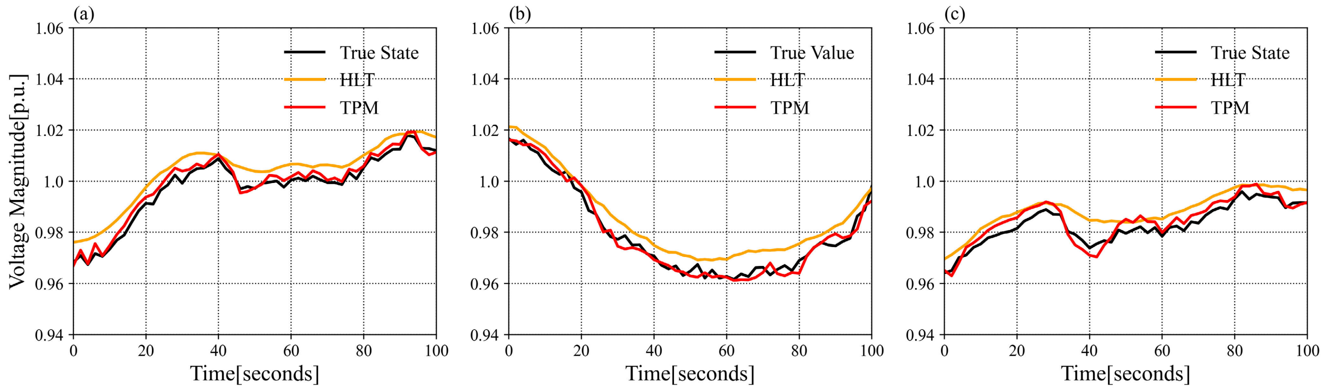


FIGURE 4. The voltage magnitude prediction of different prediction algorithms in several nodes of power systems in IEEE 33-bus system. (a) Node 1. (b) Node 11. (c) Node 25.

TABLE 1. Architecture of CNN-LSTM.

Layer	Size	Activation function
Input layer	9×1,64 cells	Linear
Convolution-1D	8×64	ReLU
LSTM	8×64, 2×16 cells, 0.2 Dropout	Sigmoid
Output layer	16×1	Linaer

A. SIMULATION SETTINGS

Simulation is conducted in IEEE 33-bus power system. Fig.3 shows the connection of nodes, and each node is installed with measurement units of PMUs. The covariance matrix of Gaussian noise $W(t) \subset \mathbb{R}^i$ and $O(t) \subset \mathbb{R}^i$ are set as $0.02\mathbf{I}$ and $0.002\mathbf{I}$, respectively. The time interval of the sampling rate is 2 seconds. The window size d is set as 3, and the architecture of CNN-LSTM is shown in Table 1.

In order to evaluate the performance of algorithms, some performance indices are selected. The root mean square error (RMSE) is

$$\tau_{\text{RMSE}} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}[\epsilon_i(t)\epsilon_i(t)^T]} \tag{42}$$

B. CASE 1: COMPARISONS OF ACCURACIES OF PREDICTION MODELS

An accurate prediction model is the most important way to pre-judge the system change and increase the accuracy of estimation. A widely used 2-parameter Holts linear trend (HLT) [35] prediction model is introduced to forecast the state of operations. It also could predict the missing data based on the historical time-series data. In this case, the data missing follows the uniform distribution based on operating time.

Fig.4 compares the prediction of different algorithms in different nodes in the system It shows that the TPM predictions follow the true state closer than those of HLT, the prediction of TPM could track most time. Consequently, the proposed CNN-LSTM architecture is more suitable for state prediction for DSE.

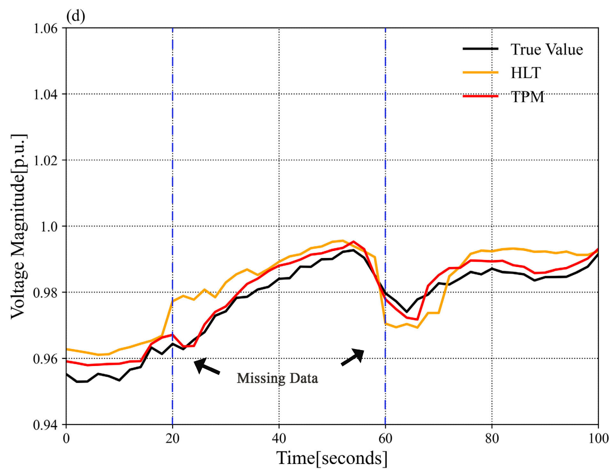


FIGURE 5. The performance of prediction involved the missing data of nodes 17 in IEEE 33-bus system.

TABLE 2. Online training cost in several nodes.

Node Number	Node 3	Node 15	Node23	Node31	Average
Time Cost(s)	0.29	0.21	0.17	0.32	0.24

C. CASE 2: COMPARISONS OF ACCURACIES OF PREDICTION MODELS UNDER INCOMPLETE INFORMATION

To guarantee the performance of TPM, the time cost of online training is considered. Table.2 shows the time costs of online training in several nodes, and the average time cost is 0.24 seconds, which means the online training could operate for about 8 iterations during the interval between two sampling. Consequently, the online training method of TPM could be used in practical implementation to fill the missing data.

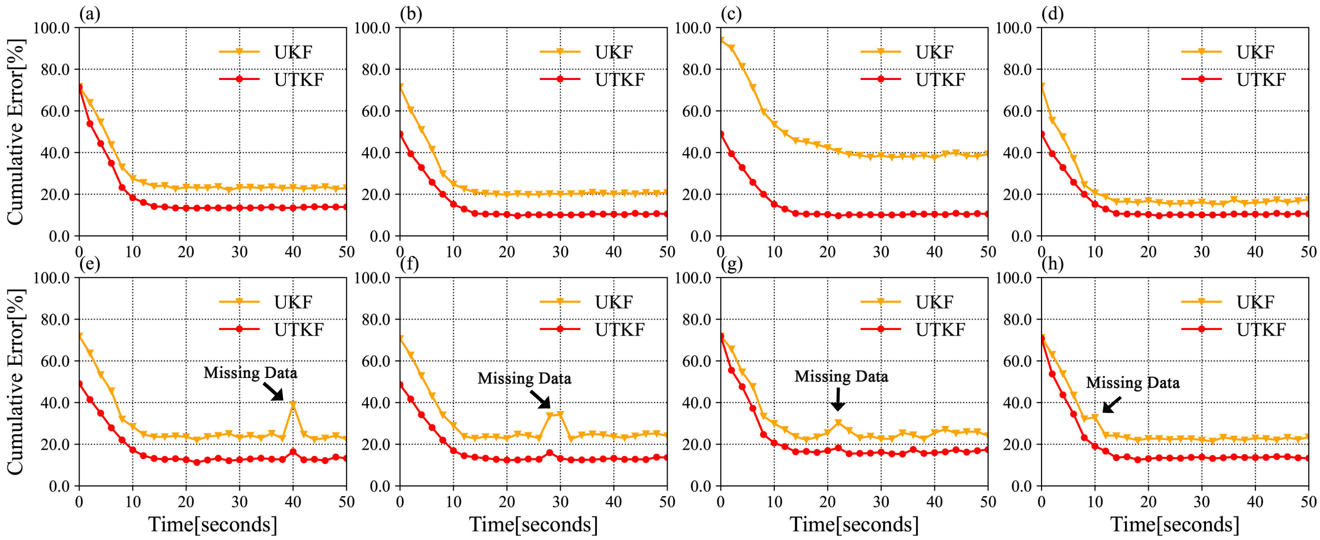


FIGURE 6. The cumulative errors of UTKF and UKF in different nodes in IEEE 33-bus system. (a) Node 1. (b) Node 5. (c) Node 11. (d) Node 17. (e) Node 21. (f) Node 25. (g) Node 29. (h) Node 31.

Considering the incomplete information, Fig.5 shows the compensation value for missing data by different algorithms in node 17. The TPM also performs better in missing data filling in the accuracy of prediction. The prediction track of TPM is influenced slightly by the missing data but the prediction track of HLT is influenced in the next fewer predictions. For traditional trend prediction model like HLT, it only considers the changing trend in the long time distance. However, the CNN-LSTM based on DL could extract the features from the historical state and promote the performances through online and offline training.

To analyze the error in detail, the RMSE of different algorithms in prediction and missing data filling is shown in Table 3. In the overview of the performance, the RMSE of TPM not only forecasts the state in lower error but also fills the missing data with lower deviation compared with the HLT. The RMSE of TPM decreases more than about 3.07 multiples in prediction and 2.96 multiples in missing data filing on average. In conclusion, the prediction model could forecast the state in power systems.

D. CASE 3: ACCURACY OF UNSCENTED TRAINABLE KALMAN FILTER WITH THE MISSING DATA FILLING METHOD FOR STATE ESTIMATION

In order to verify the accuracy of the proposed algorithm, the cumulative errors of estimation from UTKF with CNN-LSTM architecture and traditional UKF with HLT prediction model are presented in Fig.6, and the simulation is conducted at the initial time after the off-line training of the CNN-LSTM architecture. In the simulation, the 21, 25, 29, and 31 nodes are injected with the missing data which is distributed as uniform distribution, other nodes are operating steadily. Fig.6 shows that the proposed UTKF estimates the state in lower percent of cumulative error, and the convergence rate is in the same speed almost in the stable operation situation. Considering the missing data, the errors of UTKF are affected slightly because of the filling method based on the prediction of CNN-LSTM architecture. Due to the lack of predicting method based on the historical state, the cumulative errors of UKF in several nodes are influenced apparently. Thus, the proposed UTKF estimates the state in a more accurate and effective method.

After the power systems reach steady operation condition, the RMSEs are obtained for every node in Table 4. in one week, and the missing data is in a uniform distribution in all nodes. Overall speaking, the UTKF performs better than HLT in RMSE, the RMSE from UTKF is smaller than UKF in every node, and the average RMSE of UTKF is decreased by 5.99 multiples on average comparing with the UKF. The most important reason is that UTKF keeps online training during operation rather than the fixed parameters in UKF with the HLT prediction model. UTKF increases the accuracy of prediction based on the residual between the predicted state

TABLE 3. RMSE of algorithms in several nodes.

Node	$\tau_{\text{RMSE}} \times 10^{-3}$			
	TPM		HLT	
	Prediction	Filling	Prediction	Filling
Node 1	1.69	3.42	7.53	10.12
Node 5	1.95	4.14	5.45	12.82
Node 7	2.61	4.98	6.01	9.45
Node 11	2.34	5.21	6.64	11.22
Node 15	2.05	3.32	8.11	9.44
Node 17	2.78	3.94	7.43	14.26
Node 25	2.11	4.61	6.57	11.19
Node 29	2.54	4.42	8.07	15.72
Node 32	2.66	3.58	7.91	17.23

TABLE 4. The convergence time of UTKF and UKF in the simulation of different scale systems.

Node	$\tau_{\text{RMSE}} \times 10^{-3}$		Node	$\tau_{\text{RMSE}} \times 10^{-3}$		Node	$\tau_{\text{RMSE}} \times 10^{-3}$	
	UTKF	UKF		UTKF	UKF		UTKF	UKF
Node 1	0.78	1.45	Node 12	0.44	1.78	Node 23	0.47	1.24
Node 2	0.93	2.98	Node 13	0.49	3.84	Node 24	0.32	3.66
Node 3	0.21	2.67	Node 14	0.92	7.93	Node 25	0.64	1.01
Node 4	0.64	2.12	Node 15	0.23	3.17	Node 26	0.17	1.20
Node 5	0.77	1.14	Node 16	1.32	5.67	Node 27	0.91	2.37
Node 6	0.67	3.94	Node 17	0.33	4.61	Node 28	0.82	3.89
Node 7	0.26	4.12	Node 18	0.94	4.01	Node 29	0.36	2.09
Node 8	1.02	4.91	Node 19	0.59	2.37	Node 30	0.72	4.44
Node 9	0.78	3.74	Node 20	0.31	4.95	Node 31	0.33	3.15
Node 10	0.54	4.54	Node 21	0.12	1.62	Node 32	0.69	3.61
Node 11	0.48	5.39	Node 22	0.17	3.44	Node 33	0.28	4.77

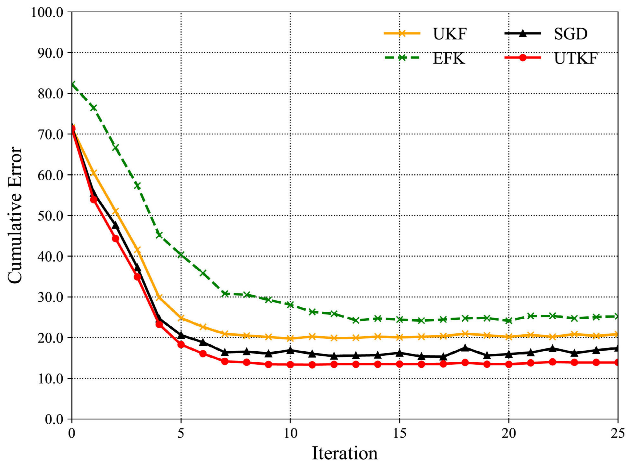


FIGURE 7. The average cumulative errors of different algorithms with the iterations of all nodes in IEEE 33-bus system.

and the estimated state, at the same time, it fills the missing data according to the historical time-series data effectively. Consequently, the proposed algorithm performs better than traditional UKF in estimation and missing data filling.

To investigate the performance of different algorithms, the stochastic gradient descent (SGD) and extended Kalman filter (EKF), performs well in convergence rate in solving the optimal problem, is introduced to verify the convergence rate of proposed algorithm. Fig.7 shows th average cumulative errors of estimation with iteration. It demonstrates that the proposed UTKF has the most accurate estimation, and the EKF has the larger cumulative errors in estimation. Due to the neglect of the higher-order in Taylor expansion, the EKF is the most inaccurate estimation when facing the missing data in power systems. Convergence rates of UKF, SGD, and UTKF are the closely the same, thus the UTKF has fast rate of convergence in operating. Consequently, the proposed algorithm estimates the state more accurately in power systems involved incomplete information.

V. CONCLUSION

As one effective way to eliminate the errors in measurement, state estimation plays an important role in the operation of power systems. However, rapid changes of states and incomplete information bring difficulty for state estimation. In this paper, an unscented trainable Kalman filter (UTKF) with a deep learning prediction model is proposed to provide accurate state estimation under incomplete information. First, based on the DL, a trainable prediction model (TPM) under the CNN-LSTM architecture is introduced to predict states in power systems in a more accurate way. Considering the data losses in estimation, the TPM is designed to fill the missing data through online training to adjust parameters in network layers of CNN-LSTM architecture. Combined with the TPM and missing data filling method, an UTKF algorithms is proposed to increase the accuracy of estimation of power systems involved the incomplete information.

Finally, three simulations are conducted in IEEE 33-bus network to verify the effectiveness and accuracy of proposed model and algorithm. In case 1, the prediction accuracies of TPM and widely used prediction model are compared by forecasting the operating state in power systems, the proposed TPM could track the real-time state in much more approximate way. Then, the case 2 verifies the accuracies of prediction models under incomplete information, the proposed method could reduce the RMSE by about 3.07 multiples in prediction and 2.96 multiples in missing data filling on average. Considering the estimation in power systems involved incomplete information, the proposed UTKF could decrease the RMSE of estimation by about 5.99 multiples on average in all nodes in long period of operation, and all proposed methods increase the accuracy and decrease the errors of the state estimation significantly.

REFERENCES

[1] Y. Liu, A. K. Singh, J. Zhao, A. P. Meliopoulos, B. Pal, M. A. B. M. Ariff, T. Van Cutsem, M. Glavic, Z. Huang, I. Kamwa, L. Mili, A. S. Mir, A. Taha, V. Terzija, and S. Yu, "Dynamic state estimation for power system control and protection IEEE task force on power system dynamic state and parameter estimation," *IEEE Trans. Power Syst.*, vol. 36, no. 6, pp. 5909–5921, Jan. 2021.

[2] T. Ergen and S. S. Kozat, "Brief papers online training of LSTM networks in distributed systems," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 29, no. 10, pp. 5159–5165, Dec. 2018.

[3] F. C. Schweppe and J. Wildes, "Power system static-state estimation, Part I: Exact model," *IEEE Trans. Power App. Syst.*, vol. 89, no. 1, pp. 120–125, Jan. 1970.

[4] M. S. Guney and Y. Tepe, "Classification and assessment of energy storage systems," *Renew. Sustain. Energy Rev.*, vol. 75, pp. 1187–1197, Oct. 2017.

[5] F. Shabani, M. Seyedyazdi, M. Vaziri, M. Zarghami, and S. Vadhva, "State estimation of a distribution system using WLS and EKF techniques," in *Proc. IEEE Int. Conf. Inf. Reuse Integr.*, Aug. 2015, pp. 609–613.

[6] J. Wan and K. N. Miu, "Weighted least squares methods for load estimation in distribution networks," *IEEE Trans. Power Syst.*, vol. 18, no. 4, pp. 1338–1345, Nov. 2003.

[7] C. Mishra, L. Vanfretti, and K. D. Jones, "Synchrophasor phase angle data unwrapping using an unscented Kalman filter," *IEEE Trans. Power Syst.*, vol. 36, no. 5, pp. 4868–4871, Sep. 2021.

[8] J. Ban, J. Im, Y.-J. Kim, and J. Zhao, "Decentralization of phasor-aided state estimation using local state vector extension," *IEEE Trans. Power Syst.*, vol. 36, no. 5, pp. 4645–4659, Sep. 2021.

[9] J. Zhao, M. Netto, and L. Mili, "A robust iterated extended Kalman filter for power system dynamic state estimation," *IEEE Trans. Power Syst.*, vol. 32, no. 4, pp. 3205–3216, Jul. 2017.

[10] R. E. Kalman, "A new approach to linear filtering and prediction problems," *J. Basic Eng.*, vol. 81, no. 1, pp. 35–45, Mar. 1960.

[11] R. Diversi, R. Guidorzi, and U. Soverini, "Kalman filtering in extended noise environments," *IEEE Trans. Autom. Control*, vol. 50, no. 9, pp. 1396–1402, Sep. 2005.

[12] R. P. Wishner, J. A. Tabaczynski, and M. Athans, "A comparison of three non-linear filters," *Automatica*, vol. 5, no. 4, pp. 487–496, Jul. 1969.

[13] T. Moller, R. Machiraju, K. Mueller, and R. Yagel, "Evaluation and design of filters using a Taylor series expansion," *IEEE Trans. Vis. Comput. Graphics*, vol. 3, no. 2, pp. 184–199, 1997.

[14] E. A. Wan and R. Van Der Merwe, "The unscented Kalman filter," *Kalman Filtering and Neural Networks*. 2001, pp. 221–280.

[15] Y. Yu, Q. Li, C. Chen, X. Zheng, and Y. Tan, "Improved dynamic state estimation of power system using unscented Kalman filter with more accurate prediction model," *Energy Rep.*, vol. 8, pp. 364–376, 2022.

[16] S. J. Julier and J. K. Uhlmann, "Unscented filtering and nonlinear estimation," *Proc. IEEE*, vol. 92, no. 3, pp. 401–422, Mar. 2004.

[17] S. J. Julier, "The spherical simplex unscented transformation," in *Proc. Amer. Control Conf.*, vol. 3, Jun. 2003, pp. 2430–2434.

[18] J. Zhao, L. Mili, and A. Gómez-Expósito, "Constrained robust unscented Kalman filter for generalized dynamic state estimation," *IEEE Trans. Power Syst.*, vol. 34, no. 5, pp. 3637–3646, Sep. 2019.

[19] P. S. Madhukar and L. B. Prasad, "State estimation using extended Kalman filter and unscented Kalman filter," in *Proc. Int. Conf. Emerg. Trends Commun., Control Comput. (ICONC3)*, Feb. 2020, pp. 1–4.

[20] M. Rostami and S. Lotfifard, "Distributed dynamic state estimation of power systems," *IEEE Trans. Ind. Informat.*, vol. 14, no. 8, pp. 3395–3404, Aug. 2018.

[21] B. Ramesh, R. Madhusudanarao, V. H. Vamsi, and S. V. R. L. Kumari, "Power system dynamic state estimation using exponential smoothing methods," in *Proc. 3rd IEEE Int. Conf. Recent Trends Electron., Inf. Commun. Technol. (RTEICT)*, May 2018, pp. 1202–1205.

[22] J. Yang, W. Zhang, and F. Guo, "Dynamic state estimation for power networks by distributed unscented information filter," *IEEE Trans. Smart Grid*, vol. 11, no. 3, pp. 2162–2171, May 2020.

[23] A. K. Singh and B. C. Pal, "Decentralized robust dynamic state estimation in power systems using instrument transformers," *IEEE Trans. Signal Process.*, vol. 66, no. 6, pp. 1541–1550, Mar. 2018.

[24] X. Wang, J. W. Bialek, and K. Turitsyn, "PMU-based estimation of dynamic state Jacobian matrix and dynamic system state matrix in ambient conditions," *IEEE Trans. Power Syst.*, vol. 33, no. 1, pp. 681–690, Jan. 2018.

[25] C. Lin, W. Wu, and Y. Guo, "Decentralized robust state estimation of active distribution grids incorporating microgrids based on PMU measurements," *IEEE Trans. Smart Grid*, vol. 11, no. 1, pp. 810–820, Jan. 2020.

[26] Y. Chakhchoukh, H. Lei, and B. K. Johnson, "Diagnosis of outliers and cyber attacks in dynamic PMU-based power state estimation," *IEEE Trans. Power Syst.*, vol. 35, no. 2, pp. 1188–1197, Mar. 2020.

[27] M. Schilling, A. Melnik, F. W. Ohl, H. J. Ritter, and B. Hammer, "Decentralized control and local information for robust and adaptive decentralized deep reinforcement learning," *Neural Netw.*, vol. 144, pp. 699–725, Dec. 2021, doi: 10.1016/j.neunet.2021.09.017.

[28] S. Dong, P. Wang, and K. Abbas, "A survey on deep learning and its applications," *Comput. Sci. Rev.*, vol. 40, Jun. 2021, Art. no. 100379.

[29] S. Albawi, T. A. Mohammed, and S. Al-Zawi, "Understanding of a convolutional neural network," in *Proc. Int. Conf. Eng. Technol. (ICET)*, Aug. 2017, pp. 1–6.

[30] J. Gawlikowski, C. R. N. Tassi, M. Ali, J. Lee, M. Humt, J. Feng, A. Kruspe, R. Triebel, P. Jung, R. Roscher, M. Shahzad, W. Yang, R. Bamler, and X. X. Zhu, "A survey of uncertainty in deep neural networks," 2021, *arXiv:2107.03342*.

[31] K. R. Mestav, J. Luengo-Rozas, and L. Tong, "Bayesian state estimation for unobservable distribution systems via deep learning," *IEEE Trans. Power Syst.*, vol. 34, no. 6, pp. 4910–4920, Nov. 2019.

[32] R. Chauhan, K. K. Ghanshala, and R. C. Joshi, "Convolutional neural network (CNN) for image detection and recognition," in *Proc. 1st Int. Conf. Secure Cyber Comput. Commun. (ICSCCC)*, Dec. 2018, pp. 278–282.

[33] T. Kattenborn, J. Leitloff, F. Schiefer, and S. Hinz, "Review on convolutional neural networks (CNN) in vegetation remote sensing," *ISPRS J. Photogramm. Remote Sens.*, vol. 173, pp. 24–49, Mar. 2021.

[34] R. C. Staudemeyer and E. Rothstein Morris, "Understanding LSTM—A tutorial into long short-term memory recurrent neural networks," 2019, *arXiv:1909.09586*.

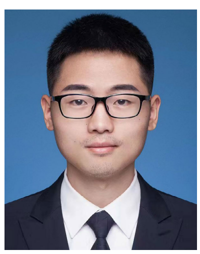
[35] G. Yapar, S. Capar, H. T. Selamlar, and I. Yavuz, "ModisCfied Holt s linear trend method," *Hacettepe J. Math. Statist.*, vol. 48, no. 2, pp. 1394–1403, Jul. 2017.

YANJIE YU received the B.S. degree in electrical engineering from the Harbin University of Science and Technology, Harbin, China, in 2020. He is currently pursuing the master's degree with Chongqing University, Chongqing, China.

His research interests include microgrid dispatch and distributed optimization.

QIANG LI (Member, IEEE) received the Ph.D. degree in electrical engineering from Zhejiang University, Hangzhou, China, in 2009.

He was a Postdoctoral Fellow with Chongqing University, from 2009 to 2012, and a Visiting Postdoctoral Scholar with The University of Adelaide, Adelaide, SA, Australia, from 2011 to 2012. He is currently an Associate Professor with Chongqing University. His current research interests include networked control systems, optimization of microgrids, and evolutionary dynamics.



HOUYI ZHANG received the master's degree from Chongqing University, Chongqing, China.

He is currently a Researcher with the Electric Power Research Institute, Guizhou Power Grid Company Ltd., Guiyang, China. His research interests include optimization and dispatching of microgrid, model and simulation of new energy grid, power quality analysis, and governance.