

Cellular Computational Networks Based Hierarchical Data-driven Dynamic State Estimation Method Considering Uncertainties

Lili Wu, Yi Wang, *Member, IEEE*, Yaoqiang Wang, *Senior Member, IEEE*, and Jikai Si, *Member, IEEE*

Abstract—Accurate generator information is crucial for the efficient control and operation of a power system. This study proposes a hierarchical data-driven approach for dynamic state estimation (DSE) of generators using cellular computational networks (CCNs) structure. The proposed method initially divides the problem of dynamic state estimation into multiple layers through hierarchical architecture. In the prediction layer, CCNs are employed to reduce the system scale by considering only relevant generators. In the correction layer, a novel adaptive filter is utilized to increase data abundance. Simulation results demonstrate that the proposed hierarchical data-driven method can accurately estimate states using PMU data alone while maintaining high computational efficiency. Additionally, it offers easy scalability and strong robustness against uncertainties. The proposed method has potential applications in online dynamic state estimation and real-time security monitoring.

Index Terms—Cellular computational networks, data driven, dynamic state estimation, hierarchical, model uncertainty.

I. INTRODUCTION

With the widespread adoption of distributed energy resources, extensive research has been conducted worldwide on dynamic state estimation in renewable power systems. Most of these studies have primarily focused on enhancing the Bayesian filter framework under stringent assumptions, such as a known system model or Gaussian noise distribution.

However, practical applications often present challenges in meeting these strict assumptions. Consequently, numerous advanced methods have been proposed to enhance the robustness of traditional ap-

proaches against various types of uncertainties including model uncertainty, non-Gaussian noise, disturbances, outliers, etc.

Bayesian-based adaptive Kalman filter (KF) methods are proposed in [1], [2] to mitigate estimation errors caused by uncertain noise. Several improved solutions based on the extended Kalman filter (EKF) are adopted in [3]–[5] to enhance robustness against disturbances. The techniques described in [6] utilize a robust cubature Kalman filter (CKF) to handle unknown measurement noise and outliers. A significant amount of work is conducted using enhanced unscented Kalman filter (UKF) [8]–[14] for addressing uncertainty problems. Reference [8] proposes a dual-stage UKF method for handling measurement errors and noise. The approach presented in [9] employs a decentralized H-infinity UKF dynamic state estimation technique to solve model unknown problems. In [10], a novel UKF algorithm is adopted to address both noise and model uncertainty issues. The work described in [11] introduces a generalized maximum-likelihood-type UKF (GM-UKF) method aimed at mitigating the negative effects of measurement noise and data anomalies. Reference [12] designs a derivative-free robust UKF estimator to maintain its robustness under outliers and non-Gaussian noise conditions. Additionally, reference [13] proposes a square-root unscented Kalman filter (SR-UKF) that enhances the numerical stability of the UKF algorithm, while reference [14] presents a robust UKF method specifically developed for mitigating outliers and non-Gaussian noise.

Research findings demonstrate that these novel strategies can enhance the robustness of DSE method against unknown noise, disturbance or outliers. However, it is impossible to eliminate the negative impact caused by model uncertainties as Bayesian-based methods require accurate models.

To address the issue of model dependence mentioned above, researchers worldwide have proposed artificial intelligence (AI)-based data-driven dynamic state estimation methods [15]–[21]. These AI-based approaches eliminate the reliance on actual power system models and instead utilize RTU or PMU measurement

Received: April 8, 2024

Accepted: July 26, 2024

Published Online: March 1, 2025

Lili Wu, Yi Wang (corresponding author), Yaoqiang Wang, and Jikai Si are with School of Electrical and Information Engineering, Zhengzhou University, Zhengzhou 450001, China (wuli1988me@gmail.com; wangyi1414599008@163.com; wangyqEE@163.com; sijikai@zzu.edu.cn).

DOI: 10.23919/PCMP.2023.000307

values to accurately map the nonlinear power system model. For instance, reference [15] introduces a power system state estimation method based on conditional generation adversarial network. In [16], a data-driven power system state estimation method is employed using deep ensemble learning algorithm to account for measurement data loss. Furthermore, a hybrid-learning online DSE approach is presented in [17], which incorporates feedforward neural networks (FFNN) considering random noise, PMU data fault, and model uncertainty. Additionally, a fully data-driven approach based on generative adversarial network is adopted in [18]. Lastly, the work described in [19] proposes a distribution system state estimation method that combines deep recurrent neural network with physics-aware linear proximal scheme.

As mentioned above, the AI-based data-driven methods possess the advantage of being model-independent; however, they still suffer from the drawbacks of high computational complexity and limited scalability. The number of parameters in AI-based DSE methods increases significantly with the size of the power system, rendering them impractical for online usage in large-scale power networks. To address these limitations, several hierarchical data-driven approaches have been proposed. For instance, a robust distributed power system state estimation approach based on a t-distribution noise model and maximum likelihood estimation is presented in [20]. Additionally, reference [21] introduces a multi-level false data identification scheme that utilizes long-short term memory neural networks for temporal identification and convolution neural networks for spatial identification.

The proposed paper presents a hierarchical dynamic state method based on CCNs, which is a distributed scalable neural network architecture consisting of computing elements (neural networks or others) in each cell. This architecture is suitable for modeling complex nonlinear network systems with unknown actual models and learning their dynamic characteristics in time and space [22]. The main characteristics of the proposed CCNs-based method are as follows:

1) The proposed dynamic state method based on CCNs is completely data-driven, offering the advantage of being model-independent. Unlike traditional dynamic state estimation methods that require exact generator models, this method can estimate the generator state using only PMU data.

2) The proposed dynamic state method exhibits high scalability by utilizing CCNs to decompose large power networks into smaller ones. Increasing the scale of the power network will only increase the number of CCNs cells without introducing computational complexity.

3) The hierarchical design reduces problem difficulty and enhances system robustness. In cases where measurement data loss or communication delay occurs, ab-

normal values can be replaced with predicted values, thereby increasing data redundancy and improving robustness against data loss.

The remainder of this paper is organized as follows. Section II introduces the basic formulations of DSE. Section III illustrates the proposed data-driven method based on CCNs. Section IV focuses on implementing the proposed CCNs-based method. Section V provides discussions and results regarding generator dynamic state estimation. Section VI concludes the paper by suggesting potential future work in this field.

II. PROBLEM FORMULATION

The objective of power system dynamic state estimation (DSE) is to accurately estimate the dynamic states of power angle δ and rotor speed ω for synchronous generators, utilizing both the state transition model and measurement model. In this study, we employ the second-order classical generator model along with the measurement model.

A. Second Order State Transition Model

The classic generator model is represented by a second-order differential equation in the local dq reference frame. The state vector \mathbf{x}_k is chosen as $[\delta, \omega]^T$ given as:

$$\begin{cases} \dot{\delta} = (\omega - \omega_0) \\ \dot{\omega} = \frac{\omega_0}{2H} \left[T_m - T_e - \frac{D}{\omega_0} (\omega - \omega_0) \right] \end{cases} \quad (1)$$

where the rated value of angular frequency is represented by $\omega_0 = 2\pi f_0$; H and D correspond to the inertia constant and damping factor respectively; and the parameters T_m and T_e represent the mechanical torque and electric air-gap torque, respectively.

B. Measurement Model

The advancement of PMU technology enables the direct selection of power angle δ and rotor speed ω as measurement vectors. Therefore, in this study, the measurement vector \mathbf{z}_k is chosen as follows:

$$\mathbf{z}_k = [\delta, \omega]^T \quad (2)$$

And the measurement function is:

$$\mathbf{z}_k = \mathbf{H}_{\text{diag}} \mathbf{x}_k + \mathbf{v}_k \quad (3)$$

$$\mathbf{H}_{\text{diag}} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \quad (4)$$

where \mathbf{H}_{diag} is the output matrix; \mathbf{x}_k is state vector; and \mathbf{v}_k is the measurement noise.

C. Nonlinear Discrete Dynamical System Model

Equations (1)–(4) can be formulated as a discrete-time model, consisting of a state difference equation and an observation equation, as:

$$\begin{aligned}\hat{x}_k &= f(\mathbf{x}_k, \mathbf{u}_{k-1}) + w_k \\ \mathbf{z}_k &= h(\mathbf{x}_k) + v_k\end{aligned}\quad (5)$$

where \mathbf{u}_{k-1} denotes previous input vector. The function $f(\cdot)$ captures the nonlinearity between states and inputs, illustrating state transitions, while $h(\cdot)$ represents the observation function that establishes the relationship between states and measurements. The process noise w_k and measurement noise v_k are assumed to be independent with Gaussian stationary distributions.

D. Bayesian-derived DSE Methods

Traditionally, the dynamic state problems of nonlinear models are addressed in a recursive manner using Bayesian methods, involving two steps: 1) prediction step; and 2) correction step.

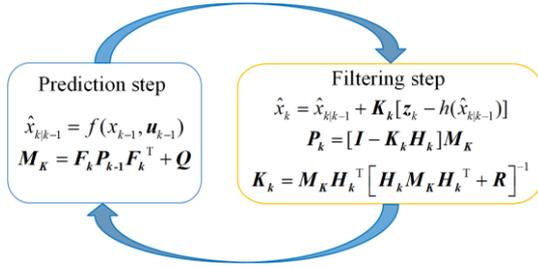


Fig. 1. Bayesian-derived method.

Figure 1 shows the Bayesian-derived method, where x_{k-1} donates previous state; $\hat{x}_{k|k-1}$ is state prediction based on previous state; \mathbf{P}_{k-1} refers to posteriori error covariance matrix at step $k-1$; \mathbf{F}_k is the partial derivative matrices of $f(\cdot)$; \mathbf{K}_k represents gain coefficient matrices; \mathbf{M}_k donates priori error covariance matrix; \mathbf{H}_k represents partial derivative matrices of $h(\cdot)$; \mathbf{P}_k corresponds to posteriori error covariance at step k ; \mathbf{I} is identity matrix; while \mathbf{Q} and \mathbf{R} denote the covariance matrices of w_k and v_k , respectively.

III. THE PROPOSED METHODOLOGY

Motivated by the Bayesian-derived method, this paper proposes a hierarchical data-driven approach based on CCNs, incorporating both prediction step and correction step. In the prediction step, a CCNs predictor is employed to generate prior state estimation. In the correction step, an adaptive filter is applied to obtain posterior state estimation. Further details are illustrated in Fig. 2, where k represents sample time and N denotes the number of ahead prediction steps. \mathbf{u}_k is input vector; $\hat{x}_{k+N|k}$ refers to priori state estimation; $\hat{x}_{k|k-N}$ is previous priori state estimation; while \hat{x}_{k+N} corresponds to posteriori state estimation.

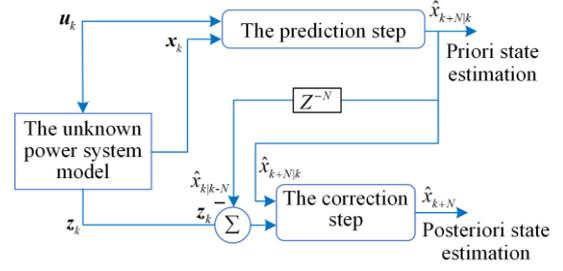


Fig. 2. The prediction-correction design of proposed CCNs based method.

A. The Prediction Step

The generator state prediction in each cell of the CCNs is performed individually using a small-scale neural network, as illustrated in Fig. 3.

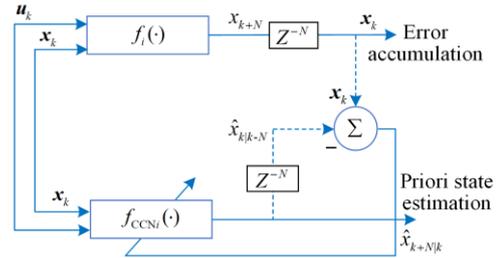


Fig. 3. Each single CCNs cell of prediction step.

The priori state estimation can be given by:

$$\hat{x}_{k+N|k} = f_{CCNi}(\mathbf{x}_k, \mathbf{u}_k) \quad (6)$$

where the generator G_i (nonlinear function $f_i(\cdot)$) is represented by CCNs cell with function $f_{CCNi}(\cdot)$. The distributed CCNs cells are interconnected based on the physical connection of real generator buses. The weights of each $f_{CCNi}(\cdot)$ are optimized through minimizing the fitness function, as illustrated in Fig. 4, where $i = 1, 2, 3, \dots, N_{\text{cell}}$, and N_{cell} is the total cell number.

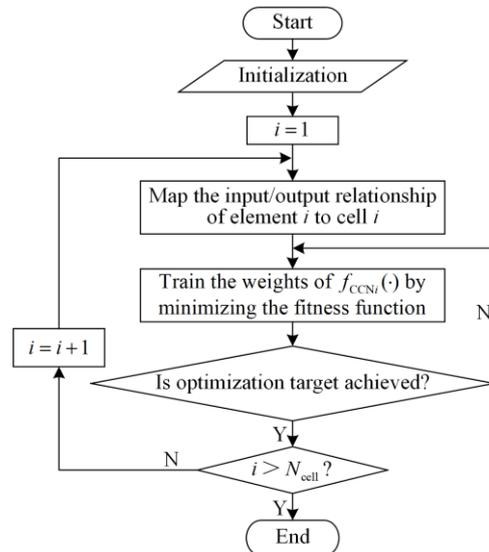


Fig. 4. The design of prediction step.

B. The Correction Step

In the correction step, we propose an adaptive filter with a time-varying gain coefficient $\mathbf{K}_k^{\text{CCN}}$ that can change based on the innovation vector:

$$\mathbf{K}_k^{\text{CCN}} = \sqrt{\frac{\sum_{t_1=k}^{t_2=k-l} |z_t - \hat{x}_{t|t-N}|^2}{l+1}} \quad (7)$$

where z_t is the measurement value at time t ; $\hat{x}_{t|t-N}$ is the priori state estimation at time t ; $[t_1, t_2]$ is a time window to calculate the priori error; and l is length of time window.

The posteriori state estimation \hat{x}_{k+N} is modified based on both priori state estimation $\hat{x}_{k+N|k}$ and gain coefficient $\mathbf{K}_k^{\text{CCN}}$ as:

$$\hat{x}_{k+N} = \hat{x}_{k+N|k} + \mathbf{K}_k^{\text{CCN}} [z_k - h(\hat{x}_{k|k-N})] \quad (8)$$

Besides, the proposed adaptive filter takes into account potential measurement data loss. In case of any such loss, the missing data will be substituted with predicted values to enhance data observability. The specific details are illustrated in Fig. 5.

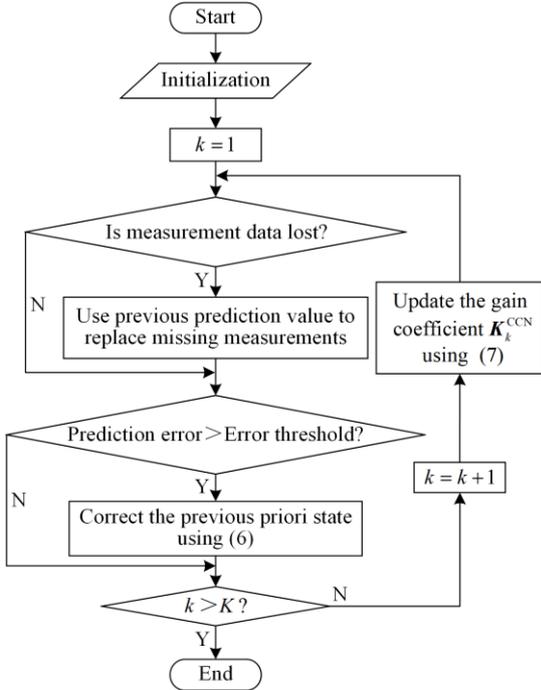


Fig. 5. The design process of correction step.

IV. THE IMPLEMENTATION OF PROPOSED METHOD

In this section, we illustrate the implementation process of the proposed method using a 12-bus infinite power system [23] as an example.

A. Hierarchical Design of Proposed Method

As mentioned above, this paper adopts power angle δ and rotor speed ω as the state vectors. Initially, these two variables are trained together as a multi-objective problem. However, due to training difficulties, it is challenging for both δ and ω to reach their optimal values simultaneously. To address this issue, we split the multi-objective optimization problem into two single-objective optimization problems: one for power angle δ and another for rotor speed ω . By incorporating the two-step design illustrated in Fig. 2, we decompose the generator state estimation problem into four parts using the proposed data-driven method. The corresponding hierarchical structure is depicted in Fig. 6.

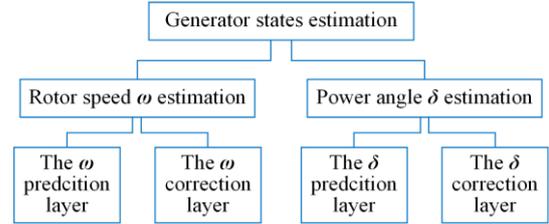


Fig. 6. The hierarchical design of proposed data driven method.

B. Geographical Connection of 12-bus Power System

The power system is a sparsely connected two-dimensional network, where the injection of bus from any generator only affects its immediate geographical neighbors. Concentric relaxation begins by initially treating the system or a portion of it as electrically rigid, and subsequently gradually relaxing the outcomes tier by tier to account for the actual flexibility of power transmission [24].

In Fig. 7, the 12-bus power system comprises three generators, namely G_2 , G_3 , and G_4 , excluding the infinite source G_1 . Assuming a fault occurs at bus 1, the system is categorized into multiple tiers based on concentric relaxation theory. These tiers include the first tier (encircled by red line), second tier (encircled by blue line), and third tier (encircled by orange line). In Fig. 7, generators G_2 and G_4 fall within the third tier, and thus, they are considered geographically adjacent. Similarly, generators G_3 and G_4 are also considered to be geographically adjacent. The geographical connections of all generators are summarized in Table I.

Generator i	Adjacent generator j
G_2	G_4
G_3	G_4
G_4	G_2, G_3

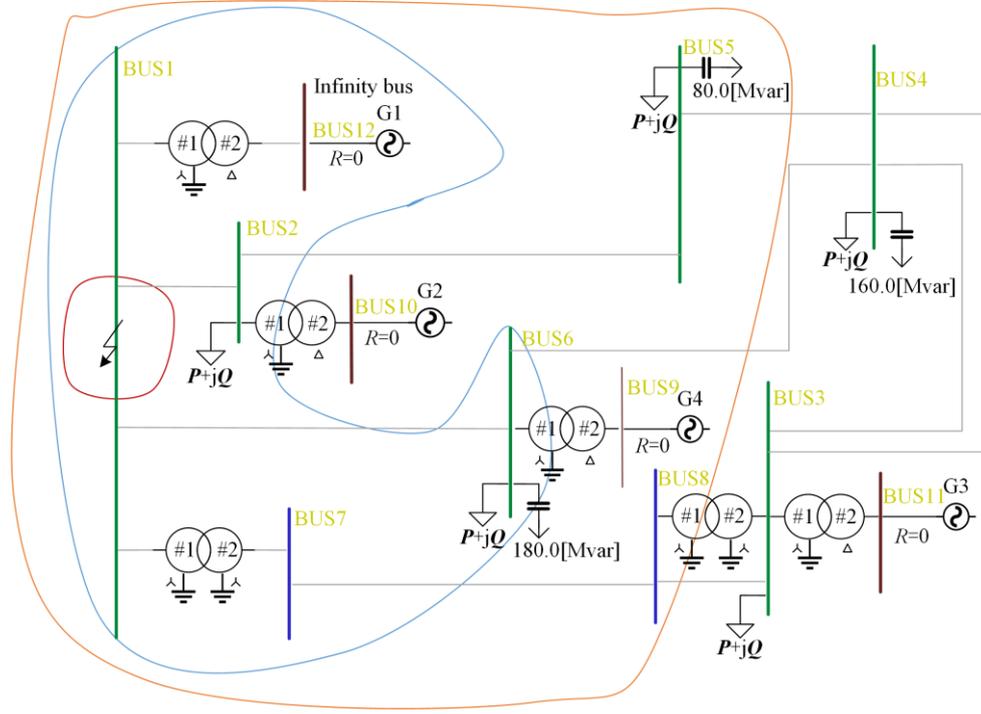


Fig. 7. Tier relaxation of 12-bus power system.

C. Prediction Layers Design

The δ and ω prediction models are built for the 12-bus power system utilizing the structure of CCNs in this part.

1) Power Angle δ Prediction Model

The power angle prediction of generator G_i at time t , denoted as $\hat{\delta}_{G_i}(t+1)$ ($i = 2, 3, 4$ corresponding to generators G_2, G_3 , and G_4), is simulated using a sequence of CCNs cell. Each CCNs cell employs a multilayer perceptron (MLP) neural network with the function defined as $f_{CCNs_oG_i}$ ($i = 2, 3, 4$):

$$\hat{\delta}_{G_i}(t+1) = f_{CCNs_oG_i} \{ \mathbf{x}_t, \mathbf{u}_t \} = f_{CCNs_oG_i} \{ \delta_{G_i}(t), \hat{\delta}_{G_j}(t), \Delta V_{G_i}(t) \} \quad (9)$$

where t is sample time, and $t = \Delta T \times k$; ΔT is sample interval; and k is sample step. The state vector $\mathbf{x}_t = [\delta_{G_i}(t), \hat{\delta}_{G_j}(t)]$ represents power angle of generator G_i at time t and its neighbor generator G_j 's previous predicted power angle. The input vector \mathbf{u}_t is $\Delta V_{G_i}(t)$, which represents voltage deviation of generator G_i . Generator number i and j defined as shown in Table I, where if i corresponds to G_2 ; then j corresponds to G_4 indicating their geographical adjacent. Further details can be found in Fig. 8.

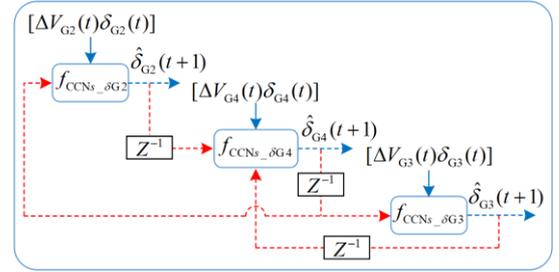


Fig. 8. The power angle δ prediction layer.

2) Rotor Speed ω Prediction Model

Similarly, the rotor speed ω prediction layer employs three cells to forecast rotor speed of three generators G_2, G_3 , and G_4 , as illustrated in Fig. 9.

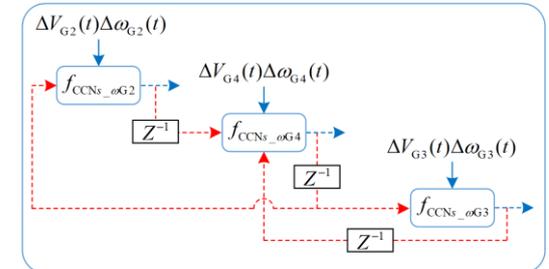


Fig. 9. The rotor speed ω prediction layer.

In each cell, the MLP neural network is adopted with function $f_{CCNs_oG_i}$ ($i = 2, 3, 4$):

$$\Delta \hat{\omega}_{G_i}(t+1) = f_{CCNs_oG_i} \{ \mathbf{x}_t, \mathbf{u}_t \} = f_{CCNs_oG_i} \{ \Delta \omega_{G_i}(t), \Delta \hat{\omega}_{G_j}(t), \Delta V_{G_i}(t) \} \quad (10)$$

where $\omega_{G_i}(t)$ is the rotor speed at time t ; ω_0 is the synchronous rotor speed, and $\Delta\omega_{G_i}(t) = \omega_{G_i}(t) - \omega_0$; the state vector $\mathbf{x}_i = [\Delta\omega_{G_i}(t), \Delta\hat{\omega}_{G_j}(t)]$; $\Delta\omega_{G_i}(t)$ denotes the speed deviation of generator G_i ; $\Delta\hat{\omega}_{G_j}(t)$ is previous predicted speed deviation of neighbor generator G_j . The input vector \mathbf{u}_i is the voltage deviation of generator G_i , donated as $\Delta V_{G_i}(t)$, while the output vector is given by $\Delta\hat{\omega}_{G_i}(t+1)$, which signifies the prediction for speed deviation of generator G_i at time t .

The proposed prediction model operates in two modes: 1) training mode; and 2) prediction mode. In the prediction mode, it utilizes measurements at time t to forecast the next state and subsequently updates the database. If the prediction error surpasses a predefined threshold, the training mode is activated to refine the weight using the most recently updated database.

3) CCNs Model Training

Taking the power angle δ as an illustrative example, the training process is depicted in Fig. 10. As seen, $[W(t)V(t)]$ are initial MLP weights and $[W(t+1)V(t+1)]$ are optimized MLP weights; f_{com} is a novel combination fitness function which can be found in (13). Various optimization methods including BP learning, PSO and differential evolution (DE) are explored during the training process to identify the most effective algorithm. After comparison, a dynamic multi-swarm particle swarm optimization algorithm (DMPSO) [25] is employed. The results demonstrate that this DMPSO approach with local search exhibits excellent performance in rapidly determining optimal MLP weights.

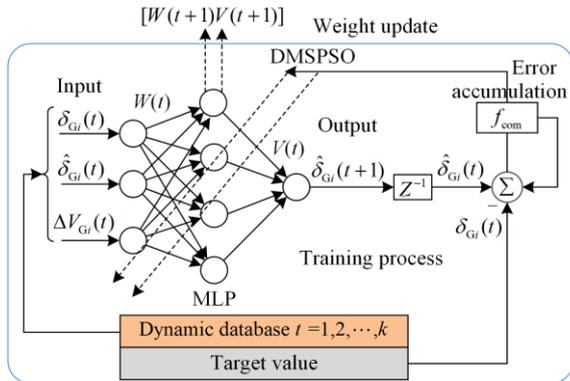


Fig. 10. The MLP training process in each cell.

Mean square error (MSE) is the average square of absolute error (AE), while mean absolute percent error (MAPE) is the mean of the relative error multiplied by 100%. The MSE and MAPE are employed as fitness functions during the learning process, corresponding fitness function are shown as:

$$f_{\text{MSE}} = \frac{1}{k} \sum_{i=1}^k |x_i - \hat{x}_i|^2 \quad (11)$$

$$f_{\text{MAPE}} = \frac{100\%}{k} \sum_{i=1}^k \left| \frac{x_i - \hat{x}_i}{x_i} \right| \quad (12)$$

where x_i represents the ground true value at sample i ; \hat{x}_i signifies the estimate of sample i ; and K denotes the size of sample dataset.

The experimental results demonstrate that the MAPE objective function performs well in capturing peak values, while the MSE objective function exhibits superior performance in training with large samples. Therefore, we propose a composite objective function f_{com} that combines the advantages of MSE and MAPE as follows:

$$f_{\text{com}} = \omega_1 f_{\text{MSE}} + \omega_2 f_{\text{MAPE}} \quad (13)$$

where ω_1 and ω_2 are the weights for f_{MSE} and f_{MAPE} , respectively.

D. Correction Layers Design

The correction layers for power angle δ and rotor speed ω are designed as follows.

1) Power Angle δ Correction Layer Design

The δ correction layer is designed as Fig. 11 following the steps below:

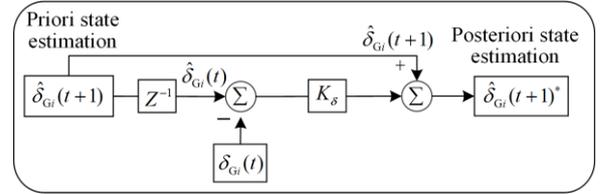


Fig. 11. The power angle δ correction layer.

Step 1: Verify the integrity of the measurement data by checking if $\delta_{G_i}(t)$ is missing. In case of any missing data, replace it with the previous priori state estimation $\hat{\delta}_{G_i}(t)$. Otherwise, predict $\hat{\delta}_{G_i}(t+1)$ uses (9) with measurement data.

Step 2: Examine the previous prediction error, if the previous priori state estimation error $|\delta_{G_i}(t) - \hat{\delta}_{G_i}(t)|$ exceeds the pre-established error threshold, go to **Step 3**. Otherwise, skip to **Step 4**.

Step 3: Generate the posteriori estimation of power angle $\hat{\delta}_{G_i}(t+1)^*$ at time t :

$$\hat{\delta}_{G_i}(t+1)^* = \hat{\delta}_{G_i}(t+1) + K_\delta [\delta_{G_i}(t) - \hat{\delta}_{G_i}(t)] \quad (14)$$

where $\hat{\delta}_{G_i}(t+1)$ and $\hat{\delta}_{G_i}(t)$ are the respective priori state estimations at time t and $t-1$ from (9); and K_δ is the gain coefficient of power angle, given as:

$$K_\delta = \sqrt{\frac{\sum_{t=k}^{t=k-l} |\delta_{G_i}(t) - \hat{\delta}_{G_i}(t)|^2}{l+1}} \quad (15)$$

where l is length of selected time window.

Step 4: The end.

2) Rotor Speed ω Correction Layer Design

The ω correction layer is designed as shown in Fig. 12 following the steps below:

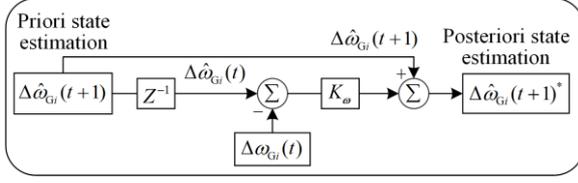


Fig. 12. The rotor speed ω correction layer.

Step 1: If there is a data loss of rotor speed $\Delta\omega_{Gi}(t)$ at time t , use previous priori state estimation $\Delta\hat{\omega}_{Gi}(t)$ to replace the missing data. Otherwise, use the measured value $\Delta\omega_{Gi}(t)$ to predict $\Delta\hat{\omega}_{Gi}(t+1)$ using (10).

Step 2: Check previous prediction error, and if the previous priori state estimation error $\Delta\omega_{Gi}(t) - \Delta\hat{\omega}_{Gi}(t)$ exceeds the pre-established error threshold, go to **Step 3**. Otherwise, jump to **Step 4**.

Step 3: Generate the posteriori estimation of speed deviation $\Delta\hat{\omega}_{Gi}(t+1)^*$ at time t :

$$\Delta\hat{\omega}_{Gi}(t+1)^* = \Delta\hat{\omega}_{Gi}(t+1) + K_{\omega}[\Delta\omega_{Gi}(t) - \Delta\hat{\omega}_{Gi}(t)] \quad (16)$$

where $\Delta\hat{\omega}_{Gi}(t+1)$ and $\Delta\hat{\omega}_{Gi}(t)$ are the respective priori state estimations at time t and $t-1$ from (10); and K_{ω} is the gain coefficient of rotor speed, given as:

$$K_{\omega} = \sqrt{\frac{\sum_{t=k}^{t=k-l} |\Delta\omega_{Gi}(t) - \Delta\hat{\omega}_{Gi}(t)|^2}{l+1}} \quad (17)$$

Step 4: The end.

V. RESULTS AND DISCUSSION

This section presents the generator dynamic state estimation results under uncertainties for a 12-bus power system and an NPCC 48-machine 140-bus system.

A. Simulation Results on a 12-bus Power System

The proposed method is evaluated on the 12-bus power system by subjecting it to various types of uncertainties, including non-Gaussian noise, line contingency, outliers, and data loss, to assess its performance and robustness.

1) Robust Against Non-Gaussian Noise

In this section, two types of non-Gaussian noise are applied to the 12-bus system to evaluate the noise robustness of the proposed CCNs based method:

Type 1: Non-Gaussian noise on generators G_2 , G_3 , G_4 with covariance of 10^{-5} .

Type 2: PRBS noise (0.5 Hz, 1 Hz and 2 Hz mixed) on generators G_2 , G_3 , G_4 with zero mean, covariance of 10^{-3} and variance of 0.2.

The rotor speed performance of generator G_2 under different noise conditions is depicted in Fig. 13. The estimation results demonstrate the robustness of the proposed method against various types of noise.

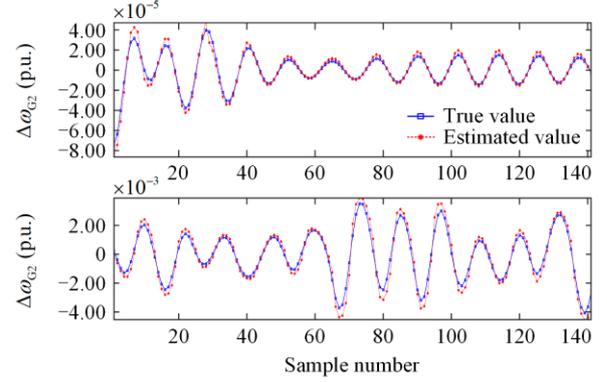


Fig. 13. Speed deviation estimation under different disturbances: above panel (small disturbance), bottom panel (large disturbance).

The simulation results under (**Type 2**) PRBS noise at different time steps are shown on Fig. 14, while the corresponding RMSE errors are presented in Table II.

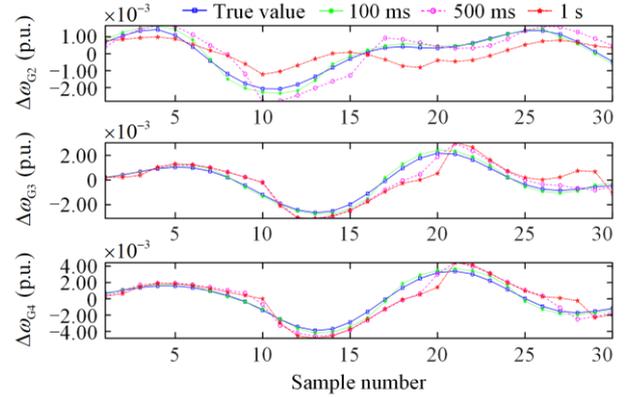


Fig. 14. Rotor speed estimation at different prediction time steps under PRBS noise.

In Fig. 14, the estimation errors increase with the prediction time step, the largest prediction time step (red star line at 1 s prediction) leads to the highest fluctuation. This is consistent with the RMSE (square root of MSE) errors at Table II.

TABLE II
RMSE UNDER PRBS NOISE AT DIFFERENT TIME STEPS

State	Gen. No.	100 ms	500 ms	1 s	1 s ⁽¹⁾
ω	G_2	1.939×10^{-4}	5.717×10^{-4}	7.635×10^{-4}	1.839×10^{-3}
	G_3	1.979×10^{-4}	5.420×10^{-4}	8.129×10^{-4}	1.917×10^{-3}
	G_4	3.249×10^{-4}	8.548×10^{-4}	9.998×10^{-4}	2.143×10^{-3}
δ	G_2	5.047×10^{-3}	1.307×10^{-2}	2.121×10^{-2}	5.444×10^{-2}
	G_3	9.194×10^{-3}	1.858×10^{-2}	1.987×10^{-2}	6.314×10^{-2}
	G_4	1.951×10^{-2}	2.835×10^{-2}	2.944×10^{-2}	9.063×10^{-2}

In Table II, the values at the third column represents the posteriori estimation errors, while the values at the fourth column represents the priori state estimation errors. The priori state estimation (fourth column) has maximum RMSE errors as the increase of prediction time step introduces stronger nonlinearity to the prediction model. However, after correction, the posteriori estimation errors can be reduced to as low as 500 ms time step. This reduction occurs because when the prediction error exceeds a pre-established limit, an adaptive filter is activated and corrects the priori state estimation value.

In conclusion, regardless of noise types or different time steps for rotor speed or power angle measurements, the proposed method demonstrates excellent performance in estimating generator states.

2) Robust Against Contingency, Outliers, and Data Loss

In this section, the performance of the proposed CCNs method under data uncertainties are tested. The experimental cases are listed in Table III. As seen, Case 1 includes a line trip contingency, Case 2 tests abnormal outliers, while Case 3 simulates missing data situation. The comparison results of these three cases can be observed in Fig. 15 and Table IV.

TABLE III
EXPERIMENTAL SCHEME

Abnormal cases	Disturbance type	Description
1	Line contingency	Line _{4,6} is tripped from sample 50
2	Line contingency and outliers	Line _{4,6} tripped, 30% error on PMU measurements of G ₂ , G ₃ , G ₄ from sample 201 to 210
3	Line contingency and data loss	Line _{4,6} tripped, data loss on PMU measurements of G ₂ , G ₃ , G ₄ from sample 201 to 210

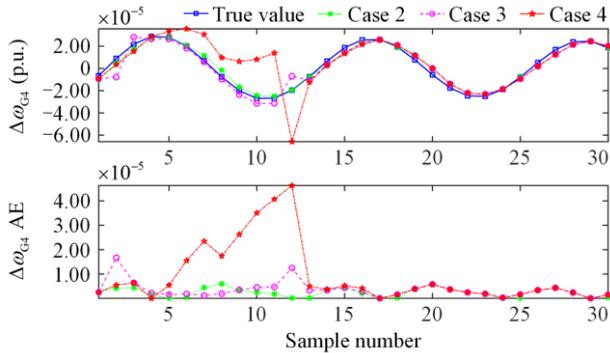


Fig. 15. Rotor speed estimation of generator G₄ under different cases: above panel (state estimation results), bottom panel (state estimation error).

In Fig. 15, the estimation lines start to exhibit more fluctuations (from first sample) as the abnormal cases applied, particularly for the worst case (Case 3).

TABLE IV
ROTOR SPEED RMSE OF GENERATOR G₄ UNDER VARIOUS CASES

State	Case No.	Maximum AE	Minimum AE	RMSE
Rotor speed of	2	5.961×10^{-6}	6.660×10^{-8}	2.983×10^{-6}
	3	1.669×10^{-5}	7.015×10^{-8}	4.836×10^{-6}
G ₄	4	4.629×10^{-5}	4.457×10^{-8}	1.537×10^{-5}
	normal	3.274×10^{-5}	1.188×10^{-8}	9.763×10^{-6}

Correspondingly, In Table IV, Case 3 has larger values for estimation AE and RMSE compared with other abnormal cases. However, regardless of the cases, the estimation results can restore to normal state soon after the abnormal situation ended, this is also verified by the similar minimum AE in Table IV. These findings validate the effectiveness of the proposed state estimation method under various abnormal scenarios.

The performance of the worst case (Case 3) is fully depicted in Fig. 16, showcasing the estimation results at 1 s time step throughout the entire transition process (pre-contingency, during-contingency, and post-contingency).

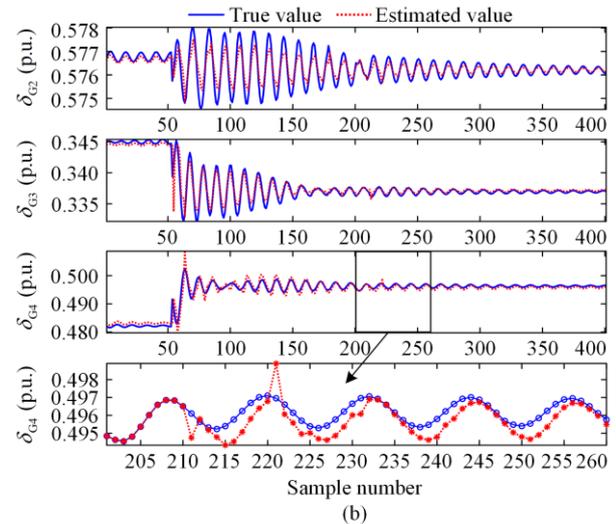
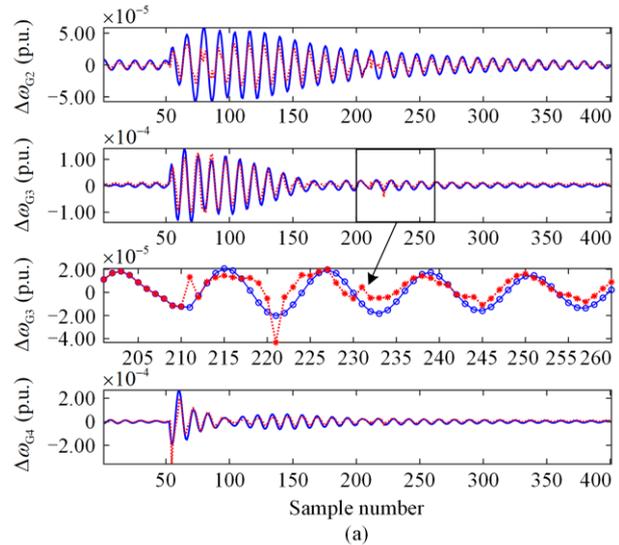


Fig. 16. The 1 s prediction time step generator state estimation under line_{4,6} trip and data loss. (a) Rotor speed. (b) Power angle.

In Fig. 16, oscillations in rotor speed and power angle are observed starting from sample 50 when line₄₋₆ is tripped, significantly affecting generator G₄ due to its closeness to line₄₋₆. In Fig. 16(b), the power angle of generator G₄ rises from 0.47 p.u. to 0.51 p.u. to redeem the shortage of load power. Besides, the error during-contingency is larger than post-contingency, which indicates that the line trip brings difficulty to generator state estimation, making it harder for the estimated value (red dot curve) to track the true value (blue circle curve).

The sampling interval is 100 ms, so 1 s prediction time step indicates 10 samples ahead. Suppose a data missing between samples 201 to 210, the predicted values of samples 201 to 210 are utilized for state prediction of samples 211 to 220. These predictions for samples 201 to 210 can propagate further and impact the prediction of samples 211 to 220, and even further to samples 221 to 230.

To see clearly, generator G₃ is zoomed at the third panel of Fig. 16(a) between samples 201 to 260 (after data missing happened). The data loss between samples 201 to 210 leads to fluctuations at samples 211 to 220, spreading further to samples 221 to 230. Similarly, in the fourth panel of Fig. 16(b), the data loss causes fluctuations on the power angle of generators G₄. But after several prediction cycles, the state estimations of the rotor speed and power angle all return to normal. This verifies the strong robustness of the proposed method to line contingency and data loss.

B. Simulation Results of NPCC 48-machine System

The scalability of the proposed method is evaluated by adopting a larger NPCC power system consisting of 48 machines, 140 buses, and 233 lines in this section, which is extracted from Power System Toolbox [26].

A sample frequency of 120 Hz PMU (sample interval 8.3 ms) is selected to compare with referenced methods [13], [17]. Each scenario involves a 10 s simulation comprising 1200 PMU data points. The proposed method is trained using the first 80 percent of the PMU data (8 s) and tested using the remaining 20 percent (2 s). All the methods are implemented in MATLAB on a computer running Microsoft Windows 10 environment equipped with Intel Core i7-8850H CPU@2.59-GHz Processor and 32-GB RAM. The proposed method is compared with other methods considering two aspects: 1) computational efficiency and accuracy; and 2) robustness against uncertainties.

1) Computational Efficiency and Accuracy

The computational efficiency and accuracy of the proposed method are tested under two scenarios: 1) normal operation state; and 2) base fault case.

a) Normal Operation State

The estimation results of EKF, method [13], method [17] and the proposed method under normal operation state are presented in Fig. 17. The RMSE error of each generator is depicted in Fig. 18. The average RMSE of δ and ω for all generators are displayed in Table V.

In Fig. 17, it can be observed that the EKF estimation exhibits a noticeable steady-state error, while method [17] successfully captures the true value trend. Method [13] and the proposed approach demonstrate commendable performance. Similarly, in Fig. 18, both the proposed method and method [13] exhibit significantly smaller RMSE errors compared to other methods. The quantified difference in error between the proposed approach and method [13] is presented in Table V.

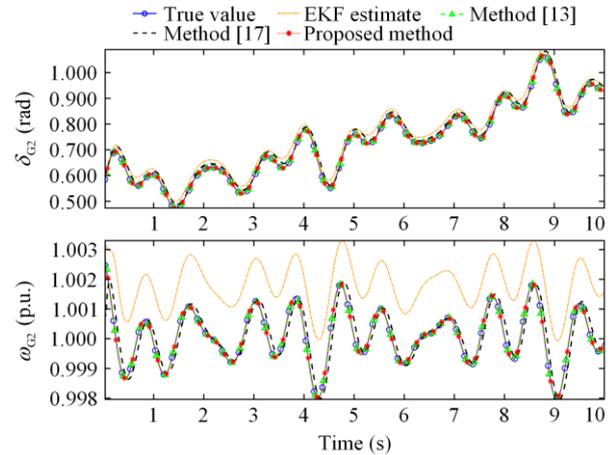


Fig. 17. State estimation of generator G₂ under normal operation state.

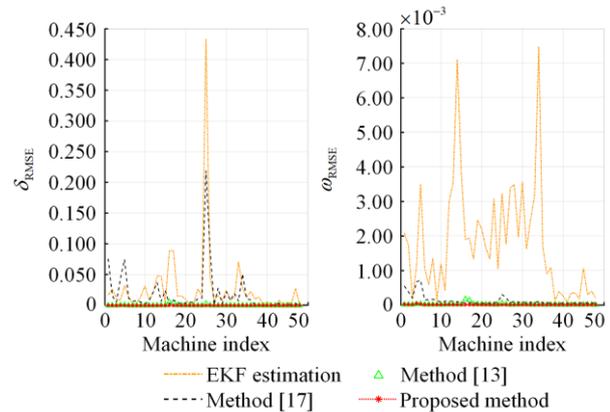


Fig. 18. RMSE of all generators under normal state.

Error type	Average ω_{RMSE}	Average δ_{RMSE}	Computation Time (s)	Time per iteration (ms)
Proposed method	4.185×10^{-6}	1.988×10^{-4}	1.173	0.977
EKF	1.812×10^{-3}	2.802×10^{-2}	4.142	3.451
Method in [13]	1.449×10^{-5}	4.334×10^{-4}	54.966	45.804
Method in [17]	1.243×10^{-4}	1.871×10^{-2}	0.675	0.563

In Table V, the proposed data-driven approach achieves the smallest average RMSE values (highlighted) for both rotor speed and power angle estimations while also being one of the fastest computation methods available. These findings indicate that under normal operating conditions, the proposed data-driven approach not only matches Bayesian-derived DSE methods in terms of accuracy but also rivals other AI-based approaches in terms of efficiency.

b) Base Fault Case

The RMSE errors of all methods during a three-phase fault are presented in Fig. 19 and Table VI, where the three-phase fault is applied on bus 40 of line₄₀₋₄₁. The results demonstrate that even under line contingency, the proposed method outperforms other methods with the smallest average RMSE as well as the second shortest computation time, which are 1.662×10^{-5} , 4.265×10^{-4} and 0.901 ms/iteration respectively. This indicates the effectiveness and accuracy of the proposed method under contingency conditions.

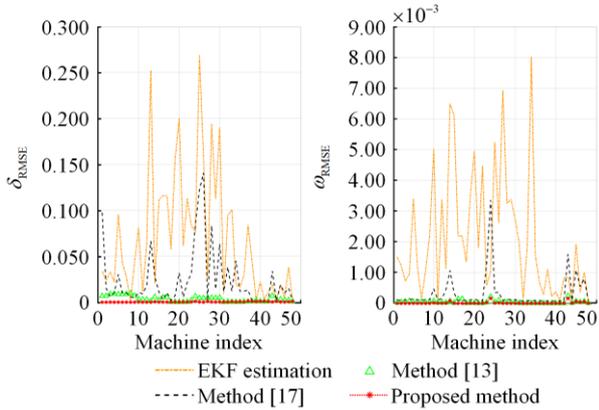


Fig. 19. RMSE of all generators under line fault.

TABLE VI
AVERAGE RMSE UNDER LINE FAULT

Error type	Average ω_{RMSE}	Average δ_{RMSE}	Computation Time (s)	Time per iteration (ms)
Proposed method	1.662×10^{-5}	4.265×10^{-4}	1.082	0.901
EKF	2.132×10^{-3}	6.905×10^{-2}	4.188	3.490
Method in [13]	4.601×10^{-5}	3.577×10^{-3}	53.187	44.322
Method in [17]	3.216×10^{-4}	2.478×10^{-2}	0.645	0.538

2) Robustness Against Uncertainties

The proposed method is compared with SR-UKF method [13], to assess its robustness under non-Gaussian noise and model parameter ambiguity using a larger estimation time size. Simulation data are generated as follows:

1) The sample frequency is 120 Hz PMU (sample interval 8.33 ms), and the estimation time size is set to be 8 times of sample interval, which equals 66.67 ms.

2) To generate dynamic response, a three-phase fault

is applied at one bus of the branches with the highest line flows and is cleared at both near and remote ends after 0.05 s and 0.1 s respectively. The testing window covers a 10 s postfault period.

3) In the system model, independent Gaussian noise is selected for processing noise and measurement noise, in accordance to [13].

The simulations related to the SR-UKF are as follows:

1) For SR-UKF method, it is assumed that both the noise type and the power system parameters are accurately known.

2) In the case of SR-UKF-non-Gaussian, it is assumed that neither the process noise nor measurement noise strictly adhere to a Gaussian distribution. Instead, they follow a Gaussian mixture model with 50% $Q=10^{-5}I, R=10^{-5}I$ and 50% $Q=10^{-6}I, R=10^{-6}I$.

3) In the scenario of SR-UKF-model ambiguity, the generator mechanical power parameter (PM) is considered ambiguous. The uncertainty range for PM is set at 5%.

The estimation results and errors of the proposed method and SR-UKF under uncertainties are depicted in Fig. 20 and Fig. 21.

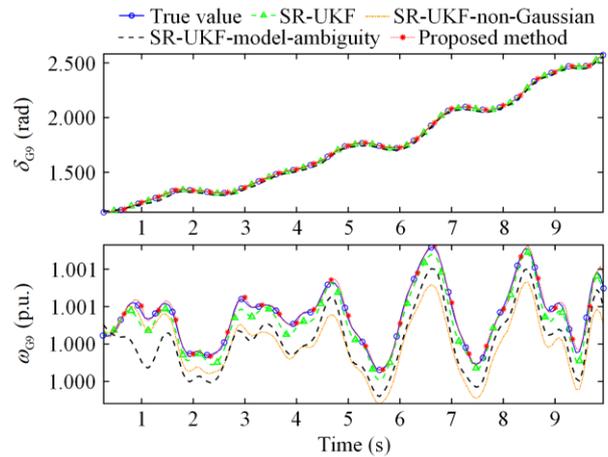


Fig. 20. State estimation of generator G_9 under uncertainties.

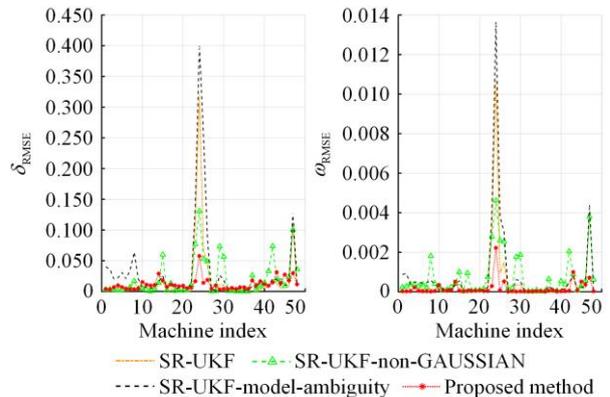


Fig. 21. RMSE of all generators under uncertainties.

In Fig. 20, the proposed method remains unaffected by the uncertainties due to its model independence. However, for SR-UKF method, noticeable errors can be observed in rotor speed when subjected to non-Gaussian noise or model parameter ambiguity cases. In addition, from Fig. 21, it is seen that the proposed method consistently exhibits the smallest RMSE for both rotor speed and power angle throughout most of the time. These findings demonstrate that the proposed data driven method possesses a high robustness against uncertainties.

VI. CONCLUSION

To tackle the challenges of model dependency and low robustness in generator state estimation, this paper proposes a hierarchical data-driven dynamic state estimation method based on CCNs. In the prediction layer, the CCNs-based model accurately simulates nonlinear state transitions and calculates priori state estimations. To mitigate the adverse effect of uncertain situations, an adaptive filter is introduced in the correction step. By integrating CCNs with a prediction-correction framework, this approach reduces complexity and enhances estimation accuracy by providing initial predictions followed by error correction using time-varying gain coefficients. Simulation results demonstrate that the proposed data-driven method achieves high accuracy and computational efficiency for both small and large power networks. Furthermore, it exhibits strong robustness against uncertainties such as abnormal measurements, non-Gaussian noise types, and ambiguous model parameters.

ACKNOWLEDGMENT

Lili Wu is truly grateful for the guidance provided by prof. G. K. Venayagamoorthy during my student visiting time at Clemson University, as his supervision greatly influenced the design of this paper. She also extends her appreciation to prof. Jinfen Gao for his invaluable support during my postdoctoral research period, as he continuously encouraged me to persevere through difficulties and never give up on my research endeavors.

AUTHORS' CONTRIBUTIONS

Lili Wu: simulation and manuscript writing. Yi Wang: revision. Yaoqiang Wang and Jikai Si: supervision. All authors read and approved the final manuscript.

FUNDING

This work is supported in part by the National Natural Science Foundation of China (No. 62203395), in part by the China Postdoctoral Science Foundation (No. 2023TQ0306), in part by the Natural Science Foundation of Henan Province (No. 242300421167), and in part by the Postdoctoral Research Project of Henan Province (No. 202101011).

AVAILABILITY OF DATA AND MATERIALS

Please contact the corresponding author for data material request.

DECLARATIONS

Competing interest: The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this article.

AUTHORS' INFORMATION

Lili Wu received her M.S. degree and Ph.D. degree in the Electrical and Computer Engineering Department of Zhengzhou University, Zhengzhou, China at 2013 and 2021 respectively. She was a visiting student of Georgia Tech at Atlanta, Georgia from October 2015 to December 2017. She is currently a postdoctoral at the Electrical and Computer Engineering Department of Zhengzhou University, Zhengzhou, China. Her research interest is artificial intelligence-based dynamic state estimation and situation awareness regarding power system stability and security.

Yi Wang received his B.S. degree from Luoyang Institute of Science and Technology, Luoyang, China, in 2014; and received his Ph.D. degree from Hohai University, Nanjing, China, in 2020. He was a visiting scholar at the University of Alberta between 2018 and 2019. He is currently a lecturer at Zhengzhou University. His current research interests include theoretical and algorithmic studies in power system estimation, parameters identification, power system dynamics, signal processing, and cyber security. He is also an active reviewer for many international journals.

Yaoqiang Wang received his B.S. degree from Hangzhou Dianzi University, Hangzhou, China, in 2006; and his M.S. and Ph.D. degrees from the Harbin Institute of Technology, Harbin, China, in 2008 and 2013, respectively. He is currently a professor with the School of Electrical and Information Engineering, Zhengzhou University, Zhengzhou, China. He is also serving as the director of the Henan Provincial Engineering Research Center of Power Electronics and Energy Systems (HERC-PEES). He has authored more than 100 technical papers including over 80 journal papers and is the holder of more than 30 patents. His current research interests include power conversion and control, new energy power system, power system analysis and planning, integrated energy operation and optimization.

Jikai Si received the B.S. degree in electrical engineering and automation from the Jiaozuo Institute of Technology, Jiaozuo, China, in 1998; the M.S. degree in electrical engineering from Henan Polytechnic Univer-

sity, Jiaozuo, China, in 2005; and the Ph.D. degree in 2008 from the School of Information and Electrical Engineering, China University of Mining and Technology, Xuzhou, China, in 2008. He is currently a distinguished professor at Zhengzhou University. His main research interests include the theory, application, and control of special motor. He has authored and co-authored over 160 technical papers in these areas. Prof. Si is a member of the Green Motor System Professional Committee, China.

REFERENCES

- [1] P. Jiao, D. Yang, and G. Cai, "Dynamic state estimation for a synchronous generator based on the Koopman operator and Kalman filter," *Power System Protection and Control*, vol. 52, no. 9, pp. 27-35, May 2024. (in Chinese)
- [2] Y. Chen, Y. Yao, and Y. Lin *et al.*, "Dynamic state estimation for integrated electricity-gas systems based on Kalman filter," *CSEE Journal of Power and Energy Systems*, vol. 8, no. 1, pp. 293-303, Jan. 2022.
- [3] R. Xiao, G. Wang, and X. Hao *et al.*, "Dynamic state estimation of medium-voltage DC integrated power system with pulse load," *Journal of Modern Power Systems and Clean Energy*, vol. 8, no. 4, pp. 689-698, Jul. 2020.
- [4] R. Xiao, G. Wang, and L. Fu *et al.*, "An adaptive hybrid dynamic state estimation method of the medium-voltage DC integrated power system with pulse load," *International Journal of Electrical Power & Energy Systems*, vol. 134, no. 107441, Jan. 2022.
- [5] X. Wang, "Power systems dynamic state estimation with the two-step fault tolerant extended Kalman filtering," *IEEE Access*, vol. 9, pp. 137211-137223, 2021.
- [6] Y. Li, J. Li, and J. Qi *et al.*, "Robust cubature Kalman filter for dynamic state estimation of synchronous machines under unknown measurement noise statistics," *IEEE Access*, vol. 7, pp. 29139-29148, 2019.
- [7] Y. Wang, Y. Sun, and V. Dinavahi *et al.*, "Adaptive robust cubature Kalman filter for power system dynamic state estimation against outliers," *IEEE Access*, vol. 7, pp. 105872-105881, Oct. 2019.
- [8] A. K. Singh and B. C. Pal, "Decentralized robust dynamic state estimation in power systems using instrument transformers," *IEEE Transactions on Signal Processing*, vol. 66, no. 6, pp. 1541-1550, Mar. 2018.
- [9] J. Zhao and L. Mili, "A decentralized H-infinity unscented Kalman filter for dynamic state estimation against uncertainties," *IEEE Transactions on Smart Grid*, vol. 10, no. 5, pp. 4870-4880, Sept. 2019.
- [10] D. Hou, Y. Sun, and J. Wang *et al.*, "Dynamic state estimation of power systems with uncertainties based on robust adaptive unscented Kalman filter," *Journal of Modern Power Systems and Clean Energy*, vol. 11, no. 4, pp. 1065-1074, Jul. 2023.
- [11] J. Zhao and L. Mili, "A robust generalized-maximum likelihood unscented Kalman filter for power system dynamic state estimation," *IEEE Journal of Selected Topics in Signal Processing*, vol. 12 no. 4, pp. 578-592, 2018.
- [12] J. Zhao, L. Mili, and A. Gómez-Expósito, "Constrained robust unscented Kalman filter for generalized dynamic state estimation," *IEEE Transactions on Power Systems*, vol. 34, no. 5, pp. 3637-3646, Nov. 2019.
- [13] J. Qi, K. Sun, and J. Wang *et al.*, "Dynamic state estimation for multi-machine power system by unscented Kalman filter with enhanced numerical stability," *IEEE Transactions on Smart Grid*, vol. 9, no. 2, pp. 1184-1196, Mar. 2018.
- [14] T. Chen, H. Ren, and P. Li *et al.*, "A robust dynamic state estimation method for power systems using exponential absolute value-based estimator," *IEEE Transactions on Instrumentation and Measurement*, vol. 71, pp. 1-10, 2022.
- [15] Y. He, S. Chai, and Z. Xu *et al.*, "Power system state estimation using conditional generative adversarial network," *IET Generation, Transmission & Distribution*, vol. 14, no. 24, pp. 5823-5833, Dec. 2020.
- [16] N. Bhusal, R. M. Shukla, and M. Gautam *et al.*, "Deep ensemble learning-based approach to real-time power system state estimation," *International Journal of Electrical Power & Energy Systems*, vol. 129, 2021.
- [17] G. Tian, Q. Zhou, and R. Birari *et al.*, "A hybrid-learning algorithm for online dynamic state estimation in multimachine power systems," *IEEE Trans. on Neural Networks and Learning Systems*, vol. 31, no. 12, pp. 5497-5508, Dec. 2020.
- [18] C. Ren and Y. Xu, "A fully data-driven method based on generative adversarial networks for power system dynamic security assessment with missing data," *IEEE Transactions on Power Systems*, vol. 34, no. 6, pp. 5044-5052, Nov. 2019.
- [19] L. Zhang, G. Wang, and G. B. Giannakis, "Real-time power system state estimation and forecasting via deep unrolled neural networks," *IEEE Transaction on Signal Processing*, vol. 67, no. 15, pp. 4069-4077, Aug. 2019.
- [20] T. Chen, Y. Cao, and L. Sun *et al.*, "A distributed robust power system state estimation approach using *t*-distribution noise model," *IEEE Systems Journal*, vol. 15, no. 1, pp. 1066-1076, Mar. 2021.
- [21] Z. Gao, S. Hu, and H. Sun *et al.*, "Dynamic state estimation of new energy power systems considering multi-level false data identification based on LSTM-CNN," *IEEE Access*, vol. 9, pp. 142411-142424, Oct. 2021.
- [22] B. Luitel and G. K. Venayagamoorthy, "Cellular computational networks-a scalable architecture for learning the dynamics of large networked systems," *Neural Networks*, vol. 50, pp.120-123, 2014.
- [23] S. Jiang, U. D. Annakkage, and A. M. Gole, "A platform for validation of FACTS models," *IEEE Transactions on Power Delivery*, vol. 21, no. 1, pp. 484-491, Jan. 2006.
- [24] J. Zaborszky, K. Whang, and K. Prasad, "Fast contingency evaluation using concentric relaxation," *IEEE Transactions on Power Apparatus and Systems*, vol. PAS-99, no. 1, pp. 28-36, Jan. 1980.
- [25] J. Liang and P. N. Suganthan, "Dynamic multi-swarm particle swarm optimizer with local search," in 2005 *IEEE Congress on Evolutionary Computation*, Edinburgh, UK, vol. 1, pp. 522-528, 2005.
- [26] J. Chow and G. Rogers. User manual for power system toolbox, version 3.0.,1991-2008.