

Regression Model Forecasting for Time-Skew Problems in Power System State Estimation

Gavin Trevorrow

Electrical and Computer Engineering Department
State University of New York at Binghamton
Vestal NY, USA
E-mail: gtrevor1@binghamton.edu

Ning Zhou

Electrical and Computer Engineering Department
State University of New York at Binghamton
Vestal NY, USA
E-mail: ningzhou@binghamton.edu

Abstract—The negative impact of measurement time skew on the static state estimation of the power grid has been exacerbated by increasing variation of system operating conditions. To mitigate the time skew problem, this paper proposes a regression model forecasting (RMF) method to forecast the time-skewed measurements, along with a confidence interval estimation (CIE) method to determine the weights associated with the forecasted measurements. The proposed RMF-CIE method is compared against several benchmark methods through Monte-Carlo simulation on the IEEE 16-machine, 68-bus model. It was observed that the proposed RMF-CIE consistently achieved more accurate state estimation on average. In addition, it was found that its estimation accuracy increases with the decrease of the skew time and variation levels.

Index Terms—Forecasting, Regression analysis, State estimation, Time skew.

I. INTRODUCTION

State estimation plays a critical role in the efficient and reliable operation of the power grid [1]. To make a well-informed decision, a power grid operator needs to obtain the complete and accurate operating status of a power grid. Yet, the measurements from a supervisory control and data acquisition (SCADA) system are often limited in their number and accuracy. To increase the monitoring accuracy and scope, static state estimation (SSE) has been widely adopted in the control center of utilities to monitor the operating conditions of the power grid by integrating SCADA measurement data together with power flow models [2].

While the SSE may function well when operating conditions change very slowly, they often fail to converge when the grid experiences rapid changes because of time-skew problems [3]. SSE assumes that all the measurements are taken at the same time instant, but SCADA measurements are not synchronized and are taken at different instants [4]. The differences among the sampling time of the SCADA measurements are known as time-skew problems, which may lead to large estimation errors and even SSE divergence. With the accelerating penetration of renewable generation, the variation of the power grid will increase, which exacerbates the time-skew problem in the state estimation. As such, there is a

need to address the time-skew problem in SSE when the power grid experiences rapid changes.

To address these challenges, some initial studies have been carried out. In [4], Holt's linear (HL) approach is used to forecast all SCADA measurements in the same time instant. In [5], the Winter's Multiplicative Seasonal Model and Autoregressive Moving Average (ARMA) techniques are used to mitigate the power mismatch of buses at the boundary of different network areas, some of which have a time delay on the order of several minutes. Reference [3] uses neighboring PMU buses to create pseudo-measurements for SCADA buses affected by time skew, then weighs the measurement's covariance matrix entries according to the innovation analysis. In [6], an extended Kalman filter is developed to provide optimal estimates for systems with random measurement delay of either 0- or 1-time sample. A wealth of literature on the leveraging of hybrid state estimation schemes (SCADA/PMU) to mitigate the time-skew issue has also been presented [7]- [8]. Additionally, convolutional neural networks (CNNs) with long-short term memory (LSTM) layers have been shown to excel at extracting spatial and temporal correlations respectively, and have been used to forecast power generation and consumption time series [9] [10]. These studies have laid a solid foundation for solving the time-skew problem.

In the literature, an essential step in mitigating the negative impact of the time-skew problem is to create pseudo-measurements at the estimation time. Yet, there lacks a systematic study in building an accurate forecasting model for pseudo-measurements. To bridge the gap, a regression model forecasting (RMF) method is proposed in this paper, which leverages both temporal and spatial correlation in the SCADA measurements to forecast the measurements into an aligned time instant. Leveraging the stepwise regression method [11], a two-step procedure is developed to establish the regression forecasting models to create the pseudo-measurements. The covariance entries are then updated based on the confidence interval estimate (CIE) of the measurement forecast. As a result, the estimation accuracy of SSE is improved. The proposed RMF-CIE method is compared with the persistence method [12], the RMF with Lin/Pan (RMF-LP) method [13], and the long-short term memory convolutional neural network (LSTM-

CNN) forecast method [9] [10], and is shown to be more accurate.

The rest of the paper is organized as follows. In Section II, conventional SSE is reviewed, and the time-skew problem is defined. The proposed RMF-CIE method is discussed in Section III. The simulation and case studies are presented in Section V. Conclusions and future works are discussed in Section VI.

II. CONVENTIONAL SSE AND TIME-SKEW PROBLEM

To lay the ground for discussion, the conventional SSE is reviewed under the condition of a time series of measurements, and its associated time-skew problem is formulated.

A. Conventional State Estimation Goals and Methods

At the time instant of state estimation (t_{se}), the objective function of the conventional SSE problem for a power grid with M measurements and N buses is often formulated into a weighted least squared (WLS) problem (1).

$$J(\mathbf{x}_{t_{se}}) = \sum_{m=1}^M \frac{(z_{m,t_{se}} - h_{m,t_{se}}(\mathbf{x}_{t_{se}}))^2}{\sigma_{mm}^2} \quad (1.a)$$

$$\hat{\mathbf{x}}_{t_{se}} = \arg \min_{\mathbf{x}_{t_{se}}} \{J(\mathbf{x}_{t_{se}})\} \quad (1.b)$$

Here, $z_{m,t_{se}}$ denotes the m^{th} measurement at time t_{se} . Symbols $\mathbf{x}_{t_{se}} = [\theta_{2,t_{se}} \cdots \theta_{N,t_{se}} \ V_{1,t_{se}} \cdots V_{N,t_{se}}]^T$ are the states to be estimated, in which $\theta_{n,t_{se}}$ and $V_{n,t_{se}}$ are the voltage angles and magnitudes of bus n . Symbol $h_{m,t_{se}}(*)$ is the measurement function, which relates the system states and measurements according to the power flow model. Symbol σ_{mm}^2 is the covariance of the measurement noise of $z_{m,t_{se}}$.

Because $h_{m,t_{se}}(*)$ is often nonlinear, the solution to (1.b), $\hat{\mathbf{x}}_{t_{se}} = [\hat{\theta}_{2,t_{se}} \cdots \hat{\theta}_{N,t_{se}} \ \hat{V}_{1,t_{se}} \cdots \hat{V}_{N,t_{se}}]^T$, is often found through iteratively solving its normal equations [1]. To assess the estimation accuracy of the SSE, the total vector error (TVE) defined by (2) is used.

$$TVE_{t_{se}} = \frac{\sum_{n=1}^N |\hat{V}_{n,t_{se}} \angle \hat{\theta}_{n,t_{se}} - V_{n,t_{se}} \angle \theta_{n,t_{se}}|}{\sum_{n=1}^N |V_{n,t_{se}} \angle \theta_{n,t_{se}}|} \quad (2)$$

Here, $V_{n,t_{se}} \angle \theta_{n,t_{se}}$ is the true value of the bus voltage phasor at bus n , and $\hat{V}_{n,t_{se}} \angle \hat{\theta}_{n,t_{se}}$ is its estimated value. A smaller TVE indicates a more accurate estimate of the states.

B. Time-Skew Problem

Traditional SSE assumes that all the measurements (z_m) are snapshots simultaneously taken at the same time instant t_{se} . In reality, however, due to issues of synchrony and latency in the communication network, the most recent data point from each measurement channel is not from the same time instant [4]. This means that the traditional SSE will be performed using measurements of the system at different times which can be detrimental to estimation accuracy, especially in the presence of severe dynamics.

To quantify the time-skew problem, let t_i and t_j be the latest sampling time of the i^{th} and j^{th} measurement channel [14] respectively. Then, skew time Δt can be defined as (3) and illustrated in Fig. 1.

$$\Delta t = \max_{i \in [1,M]} (t_i) - \min_{j \in [1,M]} (t_j) \quad (3)$$

The latest sampling time can be modeled as a random variable T_m , which follows the uniform distribution as in (4).

$$T_m \sim U(\min(t_m), \min(t_m) + \Delta t) \quad (4)$$

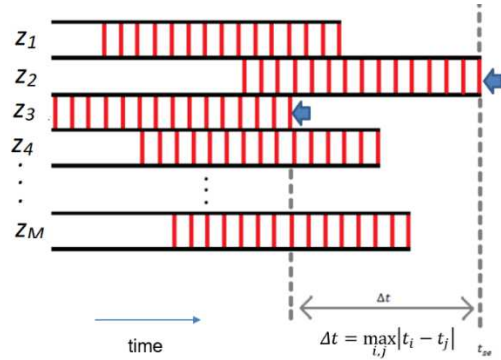


Fig. 1. Time-skew problem in measurements with z_m denoting the SCADA data from the m^{th} measurement channel and red vertical lines denoting the sampling times.

Because many measurements are not available at the moment of state estimation $t = t_{se}$ due to the time-skew problem, measurement $z_{m,t_{se}}$ needs to be estimated ($\hat{z}_{m,t_{se}}$), which is then used as a pseudo measurement in the conventional SSE as in (5). Note that the major difference between (1.a) and (5.a) is that $z_{m,t_{se}}$ is used in (1.a) whereas $\hat{z}_{m,t_{se}}$ is used in (5.a).

$$J(\mathbf{x}_{t_{se},\Delta t}) = \sum_{m=1}^M \frac{(\hat{z}_{m,t_{se}} - h_{m,t_{se}}(\mathbf{x}_{t_{se}}))^2}{\sigma_{mm}^2} \quad (5.a)$$

$$\hat{\mathbf{x}}_{t_{se},\Delta t} = \arg \min_{\mathbf{x}_{t_{se}}} \{J(\mathbf{x}_{t_{se},\Delta t})\} \quad (5.b)$$

Denoting z_{m,T_m} as the latest available measurements taken at time instant T_m , a simple estimation model that assumes no change (known as the *persistence method*) can be used by setting $\hat{z}_{m,t_{se}} = z_{m,T_m}$. When the measured variables change, the skew time Δt can incur the measurement time-skew errors of $\Delta z_{m,\Delta t} = z_{m,T_m} - z_{m,t_{se}}$, which will be super-imposed to the measurement noise. Note that $\Delta z_{m,\Delta t}$ is proportional to the skew time Δt and system changing rate. Thus, when systems significantly change during Δt , $\Delta z_{m,\Delta t}$ will be significant and can statistically degrade the estimation accuracy. Because the *persistence method* is logically simple and computationally cheap, most SSE deployed in control centers adopts this method.

III. REGRESSION MODEL FORECASTING AND CONFIDENCE INTERVAL COVARIANCE METHOD

To reduce the negative impacts of the time-skew problems on the estimation accuracy of SSE, this section proposes the RMF method to forecast $\hat{z}_{m,t_{se}}$, and a CIE method to determine its weight.

A. Forecasting Models of the RMF Method

The estimation accuracy of $\hat{z}_{m,t_{se}}$ can be improved by building a forecasting model. Leveraging the temporal and spatial correlations among different measurements in the power grid [14] [15], a general forecasting model can be constructed for $\hat{z}_{m,t_{se}}$ through (6).

$$\hat{z}_{m,t_{se}} = F_m(z_{i,t_i-k_i} | i = 1, \dots, M; k_i = 0, \dots, K_i) \quad (6)$$

Here, $F_m(\cdot)$ is the forecasting model. Symbol z_{i,t_i} stands for the latest (t_i) value at measurement channel i , while K_i is the time window of the predictor. Assuming linear models, the forecasting model of the RMF method can be built as a multiple regression model in (7).

$$\hat{z}_{m,t_{se}} = \sum_{i=1}^M \sum_{k_i=0}^{K_i} z_{i,t_i-k_i} \beta_{i,t_i-k_i} + \beta_{m,0} \quad (7)$$

Here, $\beta_{*,*}$ are the coefficients of the multiple regression model to be estimated. Because the correlation between $z_{m,t_{se}}$ and many of z_{i,t_i-k_i} can be trivial, including them as the predictors may be detrimental to the forecasting accuracy of the models. A systematic approach is thus needed to identify the irrelevant predictors and remove them by directly setting the corresponding coefficients to 0. At the same time, the coefficients of the relevant predictors shall be estimated and used in (7) to forecast $\hat{z}_{m,t_{se}}$. Note that the persistent model discussed in the previous section is a special case of (7) if one sets $\beta_{m,t_m} = 1$ and all the other coefficients 0s.

B. Variable Selection for the RMF Method

To build a reliable and accurate forecasting model of (7), a two-step approach is used to identify the important predictors from many candidate variables for the RMF method.

In the first step, a long list of candidate predictors is selected by applying the users' engineering judgment. The following criteria are proposed based on the authors' experience and used to identify the candidate predictors, which may be able to make significant contributions to the forecasting model in explaining response variables:

- (i) If $\hat{z}_{m,t_{se}}$ is for real power, only real power measurements will be selected as candidate predictors because real power and reactive power/voltage magnitudes are often decoupled. A similar rule is applied for reactive power.
- (ii) Only the 12 most up-to-date measurement channels shall be considered for the model building to ensure that the model will produce forecasts for times as close to the SE time as possible. The channel number 12 was chosen heuristically based on a tradeoff between the lower TVE and greater execution time observed as the channel number increased.

- (iii) Only variables within the time windows that have the same topology as $\hat{z}_{m,t_{se}}$ will be included as candidates because different topology may lead to different correlation models.

The first step is important because it can significantly reduce the size of candidate predictors based on the knowledge and experience accumulated by an experienced engineer. While a machine learning approach could be used in selecting predictors, the required training data size and computation expenses will grow exponentially with the number of candidate predictors, and the approach could become prohibitively complex.

The variables in the long list are further trimmed down using the stepwise regression method [16] to form a short list of the key predictors. The stepwise-regression method adopts an iterative procedure to identify the key predictors by adding and removing the predictors in the long list, then comparing the resulting coefficients of determination, R_m^2 , defined in (8).

$$R_m^2 = 1 - \frac{\sum_{t_{se} \in \text{training set}} (z_{m,t_{se}} - \hat{z}_{m,t_{se}})^2}{\sum_{t_{se} \in \text{training set}} (z_{m,t_{se}} - \bar{z}_{m,t_{se}})^2} \quad (8)$$

Here, $\hat{z}_{m,t_{se}}$ is the forecasted value of $z_{m,t_{se}}$ while $\bar{z}_{m,t_{se}}$ is the mean value of $z_{m,t_{se}}$. The R_m^2 represents how well a forecasting model explains the variation of the response variable. Thus, adding a predictor into an existing forecasting model usually increases R_m^2 . The amount of the increase is proportional to how significantly the added predictor contributes to the forecasting model. Readers are referred to [16] for more details on the stepwise regression. Its procedure is briefly reviewed as follows.

- (i) Initiate the short list with a linear term and intercept for each predictor in the long list.
- (ii) If z_{i,t_i-k_i} was not in the short list, its contribution to the forecasting model is evaluated by comparing the R_m^2 -values before and after adding it to the forecasting model. If its contribution is significant (i.e., the increase of the R_m^2 -value after adding it to the model is greater than or equal to a pre-selected threshold, e.g., $P_{enter} = 0.10$), z_{i,t_i-k_i} shall be added to the short list.
- (iii) If z_{i,t_i-k_i} was in the short list, check its contribution by comparing the R_m^2 -values before and after removing it from the forecasting model. If its contribution is not significant (i.e., its R_m^2 -value decrease is smaller than a pre-selected threshold, e.g., $P_{remove} = 0.05$), z_{i,t_i-k_i} shall be removed from the short list.

Steps (ii) and (iii) shall be looped through $i = 1, 2, \dots, M$; $k_i = 0, 1, \dots, K_i$ for the variables in the long list. Then, the procedure shall be repeated until no variable is added or removed from the short list.

With the short list of key predictors identified, the forecasting model of the RMF method can be denoted by (9) by removing the irrelevant terms in (7). In (9), M_m is the total number of key predictors $\mathbf{z}_m = [1 \ z_{m_1} \ \dots \ z_{m_{M_m}}]^T$ in the short list. Symbol $\boldsymbol{\beta}_m = [\beta_{m_0} \ \beta_{m_1} \ \dots \ \beta_{m_{M_m}}]^T$ is the coefficient vector, which is estimated through (10) by applying

the least squared method on the training set. Let T be the total number of forecasting instances that can be constructed from the training data. In (10), $\hat{\mathbf{z}}_{m,T_{se}} = [\hat{z}_{m,se1} \ \hat{z}_{m,se2} \ \dots \ \hat{z}_{m,seT}]^T$ is the response vector while $\mathbf{Z}_m = [\mathbf{z}_{m,1} \ \mathbf{z}_{m,2} \ \dots \ \mathbf{z}_{m,T}]^T$ is the measurement matrix whose rows are the regressor vector $\mathbf{z}_m$ in the training data at time instants of $1, 2, \dots, T$. Once formed, the RMF forecasting model (9) was used to output the dependent variables' predicted response to the remainder of the predictor data, up to t_{se} . The latest projected response was then used as $\hat{z}_{m,t_{se}}$ when performing SSE using (5.a).

$$\hat{z}_{m,t_{se}} = \sum_{j=1}^{M_m} z_{m,j} \beta_{m_j} + \beta_{m_0} = \mathbf{z}_m^T \boldsymbol{\beta}_m \quad (9)$$

$$\boldsymbol{\beta}_m = (\mathbf{Z}_m^T \mathbf{Z}_m)^{-1} \mathbf{Z}_m^T \hat{\mathbf{z}}_{m,T_{se}} \quad (10)$$

C. CIE Method

To perform SSE, the uncertainty of forecasted measurement $\hat{z}_{m,t_{se}}$ needs to be quantified and translated into the covariance of σ_{mm}^2 in (5). Following the CIE method in [17], the regressor vector for channel m at t_{se} can be defined as $\mathbf{z}_{m,t_{se}} = [1 \ z_{m,1,t_{se}} \ \dots \ z_{m,M_m,t_{se}}]^T$. The $100(1-\alpha)$ percent confidence interval (CI) on $\hat{z}_{m,t_{se}}$ can be computed using (11).

$$CI_{m,1-\alpha} = t_{\frac{\alpha}{2}, T-M_m-1} \sqrt{\hat{\sigma}_m^2 \mathbf{z}_{m,t_{se}}^T (\mathbf{Z}_m^T \mathbf{Z}_m)^{-1} \mathbf{z}_{m,t_{se}}} \quad (11)$$

Here, $t_{\frac{\alpha}{2}, T-M_m-1}$ is the inverse of Student's t cumulative distribution function with degree of freedom of $T - M_m - 1$ at the significant level of $\alpha/2$. $\hat{\sigma}_m^2 = \frac{(\hat{\mathbf{z}}_{m,T_{se}} - \mathbf{Z}_m \boldsymbol{\beta}_m)^T (\hat{\mathbf{z}}_{m,T_{se}} - \mathbf{Z}_m \boldsymbol{\beta}_m)}{T - M_m - 1}$ is the mean squared error (MSE) of the forecast using the training data set. The estimated 95% confidence interval, i.e., $CI_{m,1-0.05}$, is used as the covariance of σ_{mm}^2 to set up the weight for the forecasted measurement $\hat{z}_{m,t_{se}}$.

D. Summary on the proposed RMF-CIE method

In summary, the procedure of the SSE using the proposed RMF-CIE method to reduce the negative impacts of the time-skew problem can be executed as follows:

- (i) Build a training dataset using historical measurements.
- (ii) Determine the structure of an RMF model (9) using the proposed two-step procedure based on the stepwise regression method.
- (iii) Estimate the coefficients of the RMF model using (10).
- (iv) Apply the RMF model to generate pseudo-measurements $\hat{z}_{m,t_{se}}$.
- (v) Estimate the $CI_{m,1-0.05}$ of $\hat{z}_{m,t_{se}}$ using (11)
- (vi) Solve (5) using $\hat{z}_{m,t_{se}}$ as measurements and $CI_{m,1-0.05}$ as the covariance to estimate the states.

Finally, the TVE defined in (2) shall be used to evaluate the accuracy of the estimates.

IV. CASE STUDY

In this section, the proposed RMF-CIE method is applied to estimate the static states of the IEEE 16-machine, 68-bus test system shown in Fig. 2 [18]. The system's dynamic responses

to the fault were simulated using power system toolbox (PST) in MATLAB [19]. To trigger the dynamic responses, a three-phase fault was triggered on line 11 (from bus 5 to bus 8) at 5.1 s and cleared at 5.2 s by tripping off the line. An example line power profile for the time interval is shown in Fig. 3. Real and reactive power flow was measured on each of the 86 lines for forward and backward flows, as well as real and reactive power injection at each bus for a total of 480 unique measurements. The magnitude and angle of each bus voltage needs to be estimated and, subtracting a reference bus angle, leads to 135 states to be estimated. Also, a random Gaussian variable with variance $\sigma^2 = 9 \times 10^{-4}$ was added to measurements to simulate measurement noise. To build the training dataset, the measuring units sampled from the PST simulation, leveraging data only from after the fault is cleared.

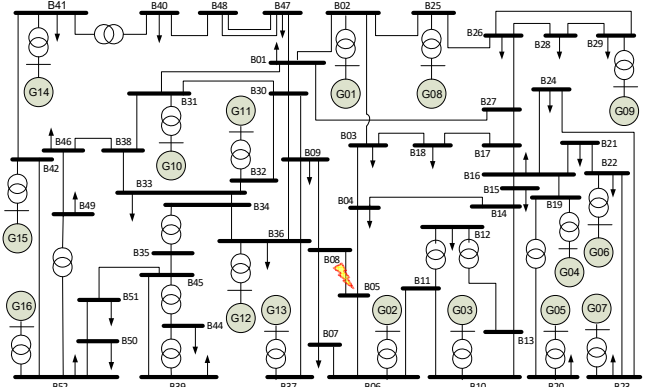


Fig. 2. IEEE 16-machine, 68-bus test system with the faulty line marked out.

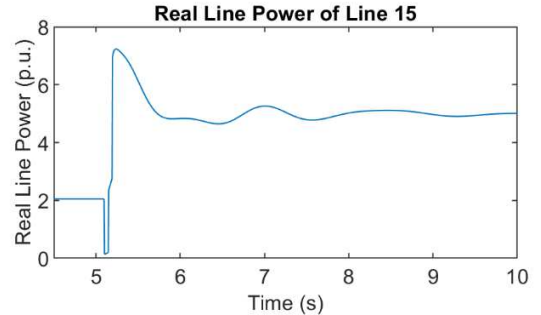


Fig. 3. Real power of Line 15. Note the transient response to the fault at 5.1 s and the relaxation to the new steady state.

To study the impacts of the time-skew problem on estimation accuracy, the initial conditions of the state estimator were set to a flat start (i.e., bus voltage magnitudes of one p.u. and angles of zero). To apply the stepwise regression method, the thresholds for adding and removing a predictor are set $P_{enter} = 0.10$ and $P_{remove} = 0.05$, respectively. Considering the randomness of the measurement noise, a Monte Carlo approach was used to evaluate the estimation accuracy statistic metrics with 100 iterations completed for each simulation configuration. To evaluate the estimation accuracy of the SSE, whisker plots are used to express the TVE data collected from 100 instances of the Monte Carlo simulation for each configuration.

A. Benchmark Methods

To assess its advantages and disadvantages, the proposed RMF-CIE method is compared with the persistence model [12], the RMF-LP method, and the LSTM-CNN method [20] under different time skew scenarios. As the persistence method has been reviewed in section II, the LSTM-CNN method and the RMF-LP method are briefly reviewed in this subsection to make this paper self-contained.

1) LSTM-CNN Method

A thorough treatment of the theory of LSTM-CNNs can be found in [9] [10]. In this study, a convolutional neural network with a long-short term memory layer was trained on the set of time series measurements that occurred after the fault at 5.1s (the same set given to the proposed RMF model). The network architecture consists of an input layer, a LSTM layer with 400 hidden units, a dropout layer with probability 0.5 (to avoid overfitting), a convolutional layer with 32 filters of size 12 and zero padding at the sequence start, another dropout layer with probability 0.2, a normalization layer, a rectified linear unit activation layer, a fully connected layer, and a regression layer for output. An adaptive moment estimation optimizer was used, with a batch size of 32 and initial learning rate of 0.01. A total of 1500 steps were allowed for the training, at which point the root MSE settled to a fairly static value. The trained network was then used to forecast each measurement channel up to the state estimation time. The last forecasted value was used in the state estimation along with the same static covariance used in the persistence forecast.

2) RMF-LP Method

Like the CIE method, the LP method is an approach of estimating the covariance of forecasted measurement $\hat{z}_{m,t_{se}}$ [13]. Because the forecasting errors often increase with the increase of forecasting horizon, a data processing approach developed by Lin and Pan in [13] was modified into (12) to weigh the covariance of the forecasted pseudo-measurement according to the time between measurement and estimation.

$$\sigma'^2_{mm} = \sigma^2_{mm} e^{\frac{7t_a}{\Delta t}} \quad (12)$$

where t_a is the age of the measurement when the state is estimated, and Δt is the skew time. Eq (10) effectively puts less “trust” in the forecasted value from older measurements. The LP method can work together with the RMF method for the SSE, which is named RMF-LP method.

B. Impact of Skew Time

To evaluate the estimation accuracy of SSE using the proposed RMF-CIE method under different skew time Δt , the TVEs of the estimated states were summarized in Fig. 4. The TVE results were obtained by setting the SE time $t_{se} = 10$ s and varying the skew time $\Delta t = 0.8:0.8:3.2$. It can be observed that the TVEs increased with the increase of skew time Δt , which indicated that Δt negatively influenced the estimation accuracy. Also shown in Fig. 4 were the TVEs of the SSE using the LSTM-CNN method, the persistence (Pers) method, RMF-LP method. It can be observed that the proposed RMF-CIE method had smaller TVEs than the other three benchmark methods.

C. Impact of Variation Levels

To evaluate the estimation accuracy of SSE using the proposed RMF-CIE under different variation levels, the SE time (t_{se}) was varied from 8 to 10 s while fixing the skew time at $\Delta t = 2.0$ s. It can be observed in Fig. 3 that the variation levels decreased as t_{se} increased because the oscillations from dynamical responses were damped out. The variation levels were quantified using the normalized standard deviation (STD) of the measurements and was plotted in Fig. 5. The normalized STD was obtained by subtracting the median from the measurements, dividing each signal by its standard deviation, then taking the standard deviation of that data set.

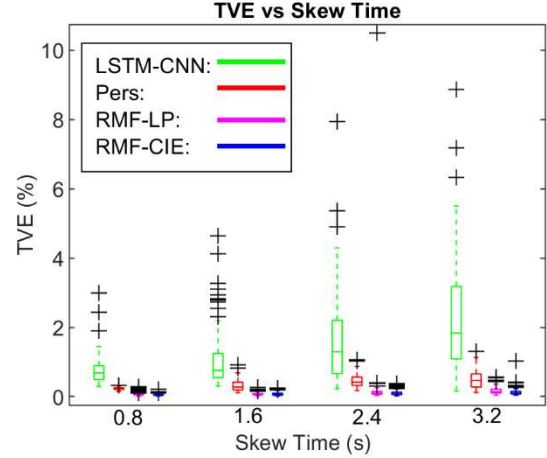


Fig. 4. TVEs of the SSE using the proposed RMF-CIE and other benchmark methods for different skew times (Δt).

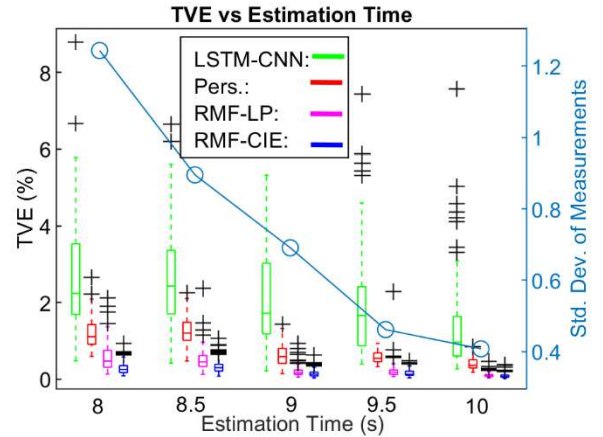


Fig. 5. TVEs of the SSE using the proposed RMF-CIE and other benchmark methods for different estimation time (t_{se}), which corresponds to different variation levels of the dynamic responses. (Note that SE close to the fault at 5.1 s was more burdened by system transients). Omitted from this figure for ease of visual comparison are two outliers at roughly 20 and 28 from the LSTM-CNN method at $t_{se} = 8$ s.

It can be observed that for the same estimation method, the TVE decreased with the decrease of variation levels, which indicated that the variation levels of measurements exacerbate the time-skew problem in the accuracy of SSE. Also observe that the SSE using the proposed RMF-CIE method had smaller TVEs than the other three benchmark methods.

V. CONCLUSIONS AND FUTURE WORK

It can be concluded from this study that the estimation errors of SSE increased with the increase of skew time during the transient responses of the system. Leveraging the temporal and spatial correlations, the proposed RMF method outperformed the persistence method and the LSTM-CNN method in constructing more accurate pseudo measurements at state estimation time through forecasting, resulting in smaller TVEs in SSE. Within the RMF method, the proposed RMF-CIE method produces smaller TVE than the RMF-LP method. In the future, the time skew that spans over a topology change will be studied, and a more systematic approach will be proposed to quantify forecasting errors under a statistical framework.

REFERENCES

- [1] A. Abur and A. G. Exposito, *Power system state estimation: theory and implementation*. New York, NY, USA: Marcel Dekker, 2004.
- [2] A. J. Wood, B. F. Wollenberg and G. B. Sheblé, *Power generation, operation, and control*, 3rd ed. New York, NY, USA: John Wiley & Sons, 2013.
- [3] X. Ma, C. Liu, J. Wu and C. Long, "On-demand state estimation with sampling time skew in power systems," in *IEEE 28th Canadian Conf. Electrical and Computer Engineering*, Halifax, NS, Canada, May 3-6, 2015, pp. 259-264.
- [4] B. A. Alcaide-Moreno, C. R. Fuerte-Esquivel, M. Glavic and T. V. Cutsem, "Electric power network state tracking from multirate measurements," *IEEE Trans. Instrumentation and Measurement*, vol. 67, no. 1, pp. 33-44, Jan. 2018.
- [5] R. C. Leou and C. N. Lu, "Adjustment of the external network's measurements and its effect on the power mismatch analysis," in *Conf. Rec. 1999 IEEE Industry Applications Conf.*, Phoenix, AZ, USA, Oct. 3-7 1999, pp. 2072-2076.
- [6] C.-L. Su and C.-N. Lu, "Interconnected network state estimation using randomly delayed measurements," *IEEE Trans. Power Systems*, vol. 16, no. 4, pp. 870-878, Nov. 2001.
- [7] M. Glavic and T. Van Cutsem, "Tracking network state from combined SCADA and synchronized phasor measurements," in *IREP Symposium-Bulk Power System Dynamics and Control - IX*, Reythmno, Greece, Aug. 25-30 2013, pp. 1-10.
- [8] A. Jovicic, M. Jereminov, L. Pileggi and G. Hug, "A linear formulation for power system state estimation including RTU and PMU measurements," in *IEEE PES Innovative Smart Grid Technologies Europe*, Bucharest, Romania, Sep. 29-Oct. 2 2019, pp. 1-5.
- [9] T.-Y. Kim and S.-B. Cho, "Predicting residential energy consumption using CNN-LSTM neural networks," *Energy*, vol. 182, no. 1, pp. 72-81, May 2019.
- [10] Q. Wu, F. Guan, C. Lv and Y. Huang, "Ultra-short-term multi-step wind power forecasting based on CNN-LSTM," *IET Renewable Power Generation*, vol. 15, no. 5, pp. 1019-1029, Jan. 2021.
- [11] R. R. Hocking, "The analysis and selection of variables in linear regression," *Biometrics*, vol. 32, no. 1, pp. 1-49, Mar. 1976.
- [12] R. G. Kavasseri and K. Seetharaman, "Day-ahead wind speed forecasting using f-ARIMA models," *Renewable Energy*, vol. 34, no. 5, pp. 1388-1393, May 2009.
- [13] J.-M. Lin and H.-Y. Pan, "A static state estimation approach including bad data detection and identification in power systems," in *IEEE Power Engineering Society General Meeting*, Tampa, FL, USA, Jun. 24-28, 2007, pp. 1-7.
- [14] D. Osipov and J. H. Chow, "PMU missing data recovery using tensor decomposition," *IEEE Trans. Power Systems*, vol. 35, no. 6, pp. 4554-4563, Nov. 2020.
- [15] P. Gao, M. Wang, S. G. Ghiocel, J. H. Chow, B. Fardanesh and G. Stofopoulos, "Missing data recovery by exploiting low-dimensionality in power system synchrophasor measurements," *IEEE Trans. Power Systems*, vol. 31, no. 2, pp. 1006-1013, Apr. 2015.
- [16] N. R. Draper and H. Smith, *Applied Regression Analysis*, Hoboken, NJ, USA: Wiley-Interscience, 1998.
- [17] D. C. Montgomery, E. A. Peck and G. G. Vining, "CI estimation of mean responses," in *Introduction to linear regression analysis*, Hoboken, NJ, USA: Wiley, 2012, p. 107.
- [18] T. Jiang, H. Jia, H. Yuan, N. Zhou and Li.Fangxing, "Projection pursuit: a general methodology of wide-area coherency detection in bulk power grid," *IEEE Trans. Power Systems*, vol. 31, no. 4, pp. 2776-2786, Sep. 2015.
- [19] J. H. Chow and K. W. Cheung, "A toolbox for power system dynamics and control engineering education and research," *IEEE Trans. Power Systems*, vol. 7, no. 4, pp. 1559-1564, Nov. 1992.
- [20] P. Kumari and D. Toshniwal, "Long short term memory-convolutional neural network based hybrid approach for solar irradiance forecasting," *Applied Energy*, vol. 295, no. 117061, May 2021.