IEEE Access
Multidisciplinary : Rapid Review : Open Access Journal

Received October 24, 2019, accepted November 5, 2019, date of publication November 8, 2019, date of current version January 3, 2020.

Digital Object Identifier 10.1109/ACCESS.2019.2952561

Validation Method for Simulation Model of Internet of Things-Aided Power System

HUICAN CHEN[1], JIECONG WANG[2], (Student Member, IEEE), JUNLEI LIU[1], BILAWAL REHMAN[2], (Member, IEEE), FENG QIAN[1], AND CHONGRU LIU[2], (Senior Member, IEEE)

[1]Electric Power Dispatching and Control Center of Guangdong Power Grid Company, Guangzhou 510000, China
[2]State Key Laboratory for Alternate Electrical Power System with Renewable Energy Sources, North China Electric Power University, Beijing 102206, China

Corresponding author: Chongru Liu (chongru.liu@ncepu.edu.cn)

This work was supported by the China Southern Power Grid Company Limited under Grant GDKJXM20162608.

ABSTRACT Validation of power system simulation models is essentially a similarity analysis problem based on multivariate time series. With the development of the internet of things (IoT) technology in the power system, the interoperability and integration of devices in the practical project are improved, and the cross interaction in the simulation process becomes more complex correspondingly. It is critical to explore the inherent correlation from the high dimensional data to evaluate the credibility and to locate the error of the simulation model. Thus, a model validation method based on factor analysis and the Prony method is proposed in this paper. Firstly, the multivariate time series of the simulation model and the practical/acknowledged system are replaced by a low number of common factors with physical meanings by factor analysis. Secondly, the modified adaptive Prony method is applied to extract the features of each common factor to ensure the best fitting of the non-stationary signal. Then the complete similarity evaluation model of the simulation system is established based on energy proportion, information entropy, and variance of the contribution rate. Finally, the error location is identified in the evaluation process based on the physical meaning of extracted features. The feasibility and effectiveness of the proposed method are verified by an application in the simulation model of a power electronics system developed in PSASP.

INDEX TERMS Validation of simulation model, power system, Internet of Things, similarity evaluation, high dimensional time series.

I. INTRODUCTION

The current energy network is being reconstructed with more renewable energy sources like wind power and solar power [1]. The power system is being more complicated with more and more power electronics, which is combined with the development of renewable energy [2]. Compared with the conventional power system, the features of the power system with high proportion electronic devices changes with the behaviors of power electronic switches and their non-stationary characteristics [3]. Internet of things is identified as one of the technologies to solve the stability and power flow reversal issues created by power electronics [4], and the modern and intelligent power system will not be possible without the IoT technology [4]-[6]. The interoperability and integration of devices in the practical system are enhanced [5], and the cross interaction in the simulation process becomes more sophisticated correspondingly, which introduces challenges to the validation of the simulation model.

Validation of simulation models is an important process before performing simulation work, which ensures that the model or modeling technique behaves in a way that we would expect. The validation of the simulation model includes model verification and model error location. Model verification plays a pivotal role in order to estimate whether the model is credible or not. Locating the errors in the model is important for those models with unsatisfactory evaluation results. One of the most basic methods of simulation model validation is to evaluate the similarity between the output of the simulation model and practical/ acknowledged results [7]. The similarity evaluation extracts the similarity elements of

The associate editor coordinating the review of this manuscript and approving it for publication was Takuro Sato.

comparing systems and gives the weights of each element, and the results could in some sense describe the degree of model authenticity.

The Feature Selective Validation (FSV) method is recommended by the IEEE Standard 1597.1/2 as an automated validation method for computational electromagnetics modeling and simulations [8]. One of the attractive advantages of FSV is the use of natural language descriptors to bridge the gap between expert opinion and the quantitative indicators [8]–[9]. However, the FSV method has its limitations: 1) Based on discrete Fourier transformation (DFT), the FSV method is only suitable for stationary signals; 2) the FSV method is only applicable to the estimation of one-dimensional data.

Compared with DFT, the Prony method has advantages in processing the non-stability signal in power system with high proportion electronic devices, due to its excellent capacity of exactly describing transient signal and directly acquiring the frequency, amplitude, and phase and decay factor of signals [10]–[12]. Thus, some researchers have proposed validation methods based on the Prony method to evaluate model credibility. The frequency, damping and amplitude credibility indexes are employed in [13] to represent the feature error of transient signals. But this method is unable to process the fault signals containing components with discontinuous or abrupt changes. An EMMD-Prony approach for dynamic validation is proposed in [14], which processes signals into stationary components by extremum field mean mode decomposition (EMMD) and then applies the Prony method to feature identification. However, the effect of the mode mixing problem of EMMD on the evaluation results is not considered.

Most of the existing methods of model validation are based on similarity assessment of one-dimensional time series. In theory, however, the similarity assessment of simulation results is said to be accurate if and only if all elements of all output variables are processed. Therefore, model validation is actually a problem of similarity analysis based on multivariate time series. Some methods to calculate the credibility of the simulation model were carried out by considering a finite set of partial variables. FSV method is extended in [15], [16] for comparison of data with multiple degrees of freedom. However, the data dimension is still very low and the correlation between variables was ignored. The analytic hierarchy process (AHP) was combined in [17] to assess the credibility of multivariate simulation results, but the assessment model is subjective and the evaluation results cannot be used for error location of simulation. According to the cross-iteration process, a factor space is built in [18] to perform the validation of the simulation model, but it can only be implemented when the internal structure of the model is fully understood. Nevertheless, with the application of IoT technology, the integration of the power system in monitoring, controlling, protection and operation has been strengthened. It is unrealistic to transfer a power system into a white-box model for verification. At the same time, considering the tremendous amount of data from the perception layer of the IoT construction, it is critical to explore the internal correlation from high-dimensional data, so that the data after dimensionality reduction can be used to verify the simulation results, and more information of model error location can be obtained to improve the accuracy of the model.

In this paper, a validation method of power system simulation model is proposed through evaluating the similarity of multiple simulation results, in response to the development trend of the power system with the IoT. Firstly, the factor analysis is used to reduce the dimensions of the multiple output results and the Prony method is used to identify the features of signals. In order to validate the power system dynamic simulation, including fault simulation, the modified adaptive Prony method proposed in [12] is adopted. Then a similarity evaluation model of multiple simulation results is established in which weights of each element are objectively obtained based on factor analysis and Prony analysis results. Finally, the error and its location in the simulation model based on the Factor analysis results and the physical meaning of these features are identified if the evaluation results are unsatisfactory.

II. FEATURE EXTRACTION METHOD FOR MULTIVARIATE TIME SERIES

This section proposes a feature extraction method for high dimensional time series. The common factors with low dimensions of the high dimensional time series are estimated by factor analysis and the features of each factor are extracted by using the modified adaptive Prony method.

A. FACTOR ANALYSIS

Factor analysis is a statistical method used to describe variability among observed and correlated variables. The basic idea of factor analysis is to find out a potentially lower number of unobserved factors to describe all or most of the observed variables by studying the internal structure of the correlation coefficient matrix or covariance matrix of variables. Let's assume a normalized time series $\mathbf{X} = [X_1, X_2, \ldots, X_v]^{\mathrm{T}}$ of dimensional v, where the length of each time series is N. The factor analysis model for $\mathbf{X}$ is defined as follows:

$$\mathbf{X} = \mathbf{AF} + \boldsymbol{\varepsilon} \tag{1}$$

where $\mathbf{F} = [F_1, F_2, \ldots, F_r]^{\mathrm{T}}$ is the common factor and $\mathbf{A} = (a_{ji})_{v\times r}$ is the factor loading matrix. The a_{ji} is the factor loading between the observed variable X_j and the common factor F_i. The $\boldsymbol{\varepsilon} = [\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_v]^{\mathrm{T}}$ is the stochastic error with zero mean and finite variance.

The common factor $\mathbf{F}$ can be estimated as the following steps:

1) Calculate $\mathbf{A}^* = (\mathrm{a}^*_{ji})_{v\times v}$ as (2) based on the method of principal component analysis [19].

$$\mathbf{A}^* = \mathrm{diag}(\sqrt{\lambda_1}, \sqrt{\lambda_2}, \ldots, \sqrt{\lambda_v})\,[\boldsymbol{u}_1, \boldsymbol{u}_2, \ldots, \boldsymbol{u}_v]^{\mathrm{T}} \tag{2}$$

where $\lambda_1, \lambda_2, \ldots, \lambda_v$ and $\boldsymbol{u}_1, \boldsymbol{u}_2, \ldots, \boldsymbol{u}_v$ are the eigenvalues and the corresponding eigenvectors of the normalized correlation coefficient matrix $\mathbf{R}$.

2) Determine the number of common factors based on the variance contribution rate g_i^2 which reflects the influence of common factor F_i on observed variables and can be calculated as (3).

$$g_i^2 = \sum_{j=1}^{v} \left(a_{ji}^*\right)^2 \quad i = 1, 2, \cdots, r \tag{3}$$

Common factors are sorted in descending order by contribution rate g_i^2. When the cumulative contribution rate of the top r common factors $\sum_{i=1}^{r} g_i^2 / \sum_{i=1}^{v} g_i^2$ is greater than 85%, it is approximated that these specific r common factors can explain the original data.

3) Factor loading matrix $\boldsymbol{A}$ can be obtained through the orthogonal rotational transformation shown as (4), which is required to ensure the variance of the squared loadings of a factor (column) on all the variables (rows) in $\mathbf{A}$ is maximized, so that each variable has a larger load on only a few common factors and a smaller load on the other common factors. In this way, explain the physical meaning of common factors can better.

$$\mathbf{A} = \mathbf{A}^*\mathbf{T} \tag{4}$$

4) The factor weight coefficient matrix $\mathbf{W}$ is defined as (5), which reflects the importance of observed variables to common factors.

$$\mathbf{W} = \mathbf{A}^{\mathrm{T}}\mathbf{R}^{-1} \tag{5}$$

5) Common factors can be estimated by (6).

$$\mathbf{F} = \mathbf{WX} \tag{6}$$

B. FEATURE EXTRACTION BASED ON PRONY METHOD

In the Prony method, a time series $[x(1), x(2), \cdots, x(n), \ldots, x(N)]$ is decomposed into a linear combination of exponential functions as (7),

$$\begin{aligned} \hat{x}(n) &= \sum_{k=1}^{p} b_k z_k^n \\ &= \sum_{k=1}^{p} A_k e^{j\theta_k} e^{(\zeta_k + j2\pi f_k)(n-1)\Delta t} \; (n = 1, 2, \mathrm{L} \cdots, N) \end{aligned} \tag{7}$$

where A_k, f_k, ζ_k, and θk are amplitude, frequency, damping and phase angle of k^{th} exponential function, respectively. The Δt is the sampling period and p is the order of the Prony model.

The Prony method is unsuitable for processing the fault signals containing components with discontinuous or abrupt changes. One solution is to apply the Prony method to a number of short contiguous time windows inside the signal. The modified adaptive Prony method was discussed in [12] based on an adaptive technique that acts with the aim of minimizing the mean square relative fitting error of signal estimation. The steps of the modified adaptive Prony method are:

1) Select an initial short time window length $L_{\min}$ and initial step size ΔL for increments in $L_{\min}$.
2) Apply the Prony method to samples in the short time window i.e. $[x(n_{\mathrm{s}}), x(n_{\mathrm{s}+1}), \cdots, x(n_{\mathrm{e}})]$ in order to obtain the model parameters (amplitude, damping, frequency, and phase of Prony exponentials), where n_{s} and n_{e} are the start and end numbers of signals in each time window, respectively.
3) Calculate the mean square relative fitting error (MSRFE) e_{f} with (8) by using the exponentials obtained in step 2.

$$e_{\mathrm{f}} = \frac{1}{n_{\mathrm{nz}}} \sum_{n=n_{\mathrm{s}}, x(n)\neq 0}^{n_{\mathrm{e}}} \frac{\left[x(n) - \hat{x}(n)\right]^2}{x(n)^2} \tag{8}$$

where n_{nz} is the number of $x(n)$ which is not equal to zero in the short time window.

4) Compare e_{f} with the threshold E_{thr} and:
   a) If $e_f \leq E_{\mathrm{thr}}$, set ΔL equal to the maximum step size $\Delta L_{\max}$ and increase the short time window length, and then jump to step 2.
   b) If $e_f > E_{\mathrm{thr}}$, set $\Delta L = \mathrm{INT}(\Delta L/2)$ and decrease the short time window length, and then jump to step 5.
5) Repeat step 2 and 3, and then compare e_{f} with the threshold E_{thr} and:
   a) If $e_f \leq E_{\mathrm{thr}}$, set $\Delta L = \mathrm{INT}(\Delta L/2)$ and increase the short time window length, repeat step 5 until $\Delta L = 0$, and then jump to step 6.
   b) If $e_f > E_{\mathrm{thr}}$, set $\Delta L = \mathrm{INT}(\Delta L/2)$ and decrease the short time window length and repeat step 5.
6) Compare n_{e} with the length N and:
   a) If $n_{\mathrm{e}} < N$, store the Prony model exponential parameters and jump to step 2 to analyze the next contiguous short window.
   b) If $n_{\mathrm{e}} \geq N$, store the Prony model exponential parameters and jump to stop.

In addition, the order of the Prony model of each time window is determined by (9) to assure the best curve fitting.

$$p_{\mathrm{tw}} = \mathrm{INT}\left(\frac{n_{\mathrm{e}} - n_{\mathrm{s}} + 1}{2}\right) - 1 \tag{9}$$

where INT() is Integral Function.

Each time window can be extracted a p_{tw}-dimension feature matrix by the modified adaptive Prony method. The p_{tw}, nevertheless, is relatively large to ensure the accuracy of the Prony model. Thus the energy proportion is taken as a criterion to reduce the dimension of the feature matrix. The exponential components are sorted in descending order and the components having an energy proportion less than 0.1% are ignored. The energy proportion is defined as,

$$\eta_k = En_k \Big/ \sum_{k=1}^{p_{\mathrm{tw}}} En_k \tag{10}$$

where the subscript k denotes the k^{th} exponential component. The En_k is the energy of an exponential component, which can be calculated as,

$$En_k = \sum_{n=0}^{N-1} \left(A_k e^{j\theta_k} e^{(\zeta_k + j2\pi f_k)n\Delta t}\right)^2 \tag{11}$$

Finally, an $l \times 4$ feature matrix **FM** is obtained from each time window, as (12).

$$\mathbf{FM} = \begin{bmatrix} A_1 & f_1 & \zeta_1 & \theta_1 \\ \vdots & \vdots & \vdots & \vdots \\ A_l & f_l & \zeta_l & \theta_l \end{bmatrix} \tag{12}$$

III. SIMULATION VALIDATION AND ERROR LOCATION

A. DEFINITION OF SYSTEM SIMILARITY

The system similarity is an overall reflection of the similarity between each similarity element of two systems. Suppose the system A consists of $n_{\mathrm{e\_A}}$ elements and the system B consists of $n_{\mathrm{e\_B}}$ elements; the n_{se} elements are similar between A and B system, hence, called similarity elements. The similarity of each similarity element is recorded as $s(e_{h\_\mathrm{A}}, e_{h\_\mathrm{B}})$, and the weight of each similar element is denoted by w_h. The similarity between A and B system can be defined as:

$$S\,(\mathrm{A,B}) = \frac{n_{\mathrm{se}}}{n_{\mathrm{e\_A}} + n_{\mathrm{e\_B}} - n_{\mathrm{se}}} \sum_{h=1}^{n_{se}} w_h \cdot s(e_{h\_A}, e_{h\_B}) \tag{13}$$

where $S \in [0, 1]$, the S =1 means that the two systems are exactly similar and the S =0 means that the two systems do not have anything similar. In between, the systems have higher similarity as the S approaches to 1. The similarity of each similarity element is defined by (14).

$$s(e_{h\_\mathrm{A}}, e_{h\_\mathrm{B}}) = 1 - \frac{\left|e_{h\_\mathrm{A}} - e_{h\_\mathrm{B}}\right|}{\left|e_{h\_A} + e_{h\_\mathrm{B}}\right|} \tag{14}$$

B. SIMILARITY ELEMENTS EXTRACTION

Let's suppose $\mathbf{X}_\mathrm{S} = \left[X_{1\_\mathrm{S}}, X_{2\_\mathrm{S}}, \cdots, X_{v\_S}\right]^\mathrm{T}$ is the normalized multivariate time series of the practical/acknowledged system and $\mathbf{X}_\mathrm{M} = \left[X_{1\_M}, X_{2\_M}, \cdots, X_{v\_M}\right]^\mathrm{T}$ is that of the simulation model.

In order to compare the two multivariate time series, the dimensions of common factors, the division of short time windows and the orders of corresponding feature matrixes from the two series must be identical. Therefore, the multivariate time series $\mathbf{X}_\mathrm{S}$ from the practical/acknowledged system is analyzed first to obtain the factor score coefficient matrix $\mathbf{W}_\mathrm{S}$. The common factor $\mathbf{F}_\mathrm{M}$ of the simulation model is determined based on $\mathbf{W}_\mathrm{S}$ as (15).

$$\begin{cases} \mathbf{F}_\mathrm{S} = \mathbf{W}_\mathrm{S}\mathbf{X}_\mathrm{S} = [F_{1\_\mathrm{S}}, F_{2\_\mathrm{S}}, \ldots, F_{r\_\mathrm{S}}]^\mathrm{T} \\ \mathbf{F}_\mathrm{M} = \mathbf{W}_\mathrm{S}\mathbf{X}_\mathrm{M} = [F_{1\_\mathrm{M}}, F_{2\_\mathrm{M}}, \ldots, F_{r\_\mathrm{M}}]^\mathrm{T} \end{cases} \tag{15}$$

Similarly, the common factor $F_{\mathrm{M}i}$ should be divided in the same way as common factor $F_{\mathrm{S}i}$. Also, the orders of the corresponding feature matrices can be unified by (16).

$$l_{ij} = \min\left(l_{ij\_\mathrm{S}}, l_{ij\_\mathrm{M}}\right) \tag{16}$$

where the subscript i denotes the common factor F_i and j denotes the j^{th} short time window.

A total of $\sum_{i=1}^{r} m_i$ feature matrices are extracted from one system, where m_i means the number of the short time windows of common factor F_i. These feature matrices will be used for similarity analysis to evaluate the similarity of models.

C. FEATURE SIMILARITY

The feature similarity of each short time window is described by the matrix $\mathbf{S}_{\mathrm{STW}\_ij}$ as follows.

$$\mathbf{S}_{\mathrm{STW}\_ij} = \left[\, s_{ij}^{A} \quad s_{ij}^{f} \quad s_{ij}^{\zeta} \quad s_{ij}^{\theta} \,\right] \tag{17}$$

where s_{ij}^{u} is calculated by (18) as,

$$s_{ij}^{u} = \sum_{k=1}^{l_{ij}} w_{k\_ij} s_{k\_ij}^{u} \;\; (u = A, f, \zeta, \theta) \tag{18}$$

where w_{k_ij} is the weight of the exponential component, which can be obtained by the following formula.

$$w_{k\_ij} = \left(\eta_{k\_ij\_\mathrm{S}} + \eta_{k\_ij\_\mathrm{M}}\right) \Big/ \sum_{k=1}^{l_{ij}} \left(\eta_{k\_ij\_S} + \eta_{k\_ij\_M}\right) \tag{19}$$

where $\eta_{k\_ij\_\mathrm{S}}$ and $\eta_{k\_ij\_\mathrm{M}}$ are obtained by (10).

D. COMMON FACTOR SIMILARITY

Information entropy weight (IEW) is widely used as an index in comprehensive evaluation [20]. IEW is computed according to the amount of information that can be transferred by the index, which describes the importance of this index in the comprehensive evaluation. The smaller the entropy of the index is, the more information it provides, and the higher its weight is. The weight of each short time window is determined by information entropy weight in this paper. The information entropy for a normalized time series $X = [x(1), x(2), \ldots, x(N)]^\mathrm{T}$ is defined as below.

$$\begin{aligned} e(X) &= -\sum_{n=1}^{N} p(x\,(n)) \log p(x\,(n)) \\ \text{s.t.} &\sum_{n=1}^{N} p(x\,(n)) = 1 \end{aligned} \tag{20}$$

The weight w_{ij} of the j^{th} short time window of the common factor F_i is derived according to the corresponding information entropy e_{ij} by (21).

$$w_{ij} = \left(1 - e_{ij}\right) \Big/ \sum_{j=1}^{m_i} \left(1 - e_{ij}\right) \tag{21}$$

The similarity of each common factor is obtained based on the IEW as below.

$$\mathbf{S}_{\mathrm{CF}} = \begin{bmatrix} s_1^A & s_1^f & s_1^\zeta & s_1^\theta \\ s_2^A & s_2^f & s_2^\zeta & s_2^\theta \\ \vdots & \vdots & \vdots & \vdots \\ s_r^A & s_r^f & s_r^\zeta & s_r^\theta \end{bmatrix} \tag{22}$$

where s_i^u is calculated by (23).

$$s_i^u = \sum_{j=1}^{m_i} w_{ij} s_{ij}^u \;\; (u = A, f, \zeta, \theta) \tag{23}$$

E. EVALUATION RESULT

The variance contribution rate of each common factor reflects the importance of that factor in factor analysis of multivariate time series. The weight of common factor F_i can be calculated as follows.

$$w_i = g_i^2 \Big/ \sum_{i=1}^{r} g_i^2 \tag{24}$$

where g_i^2 is the variance contribution rate of common factor F_i.

The similarity between the simulation model and the practical/acknowledged system, based on the weights of common factors, is calculated by (25).

$$\mathbf{S} = \begin{bmatrix} S^A & S^f & S^\zeta & S^\theta \end{bmatrix} \tag{25}$$

where s_{ij}^u can be obtained by (28) as,

$$S^u = \sum_{i=1}^{r} w_i s_i^u \quad (u = A, f, \zeta, \theta) \tag{26}$$

The similarity of these four features is independent of each other. For each of the four specific features, the simulation model satisfies the credibility requirement when the similarity of feature reaches a predefined criterion.

F. EVALUATION PROCESS AND MODEL ERROR LOCATION

Fig. 1 shows the complete process of simulation model validation. The model errors are located in which specific subsystems by analyzing the common factor similarity, and the error parameters are identified according to the features with low similarity. Once the errors are identified, the simulation model can be improved for accurate results.

IV. CASE STUDIES

Over the past few years, the MMC has become a subject of interest for power systems and industrial applications including HVDC transmission systems, FACTS, medium-voltage variable-speed drives, and medium/high voltage DC/DC converters [21]. The modeling and simulation play an important role in analyzing the operational characteristics of MMC applied in power system [22]. A 2-terminal MMC based HVDC system is shown in Fig. 2, where the AC grid is replaced by equivalent Thevenin models. The simulation is carried out in PSCAD/EMTDC (an electromagnetic transient simulation software) and PSASP (an electromechanical transient simulation software). The simulation results of PSASP are validated with reference to EMTDC simulation results.

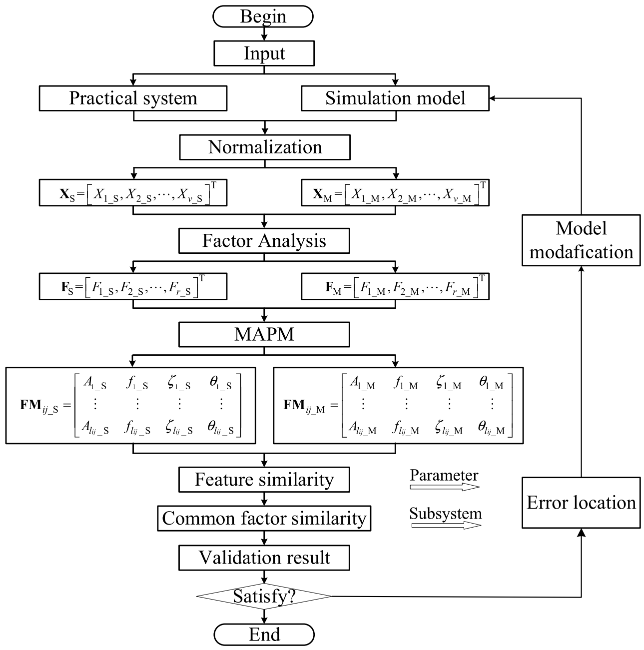


FIGURE 1. Process of credibility evaluation and error location.

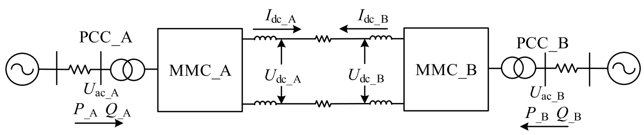


FIGURE 2. Circuit diagram of the simulation system.

A three-phase short-circuit grounding fault is applied at bus PCC_A for 0.1s to verify the results under transient conditions. The simulation results of PSCAD/EMTDC and PSASP are provided in Fig.3.

A. FEATURE EXTRACTION

The time series from PSASP and PSCAD/EMTDC simulation models are extracted by the proposed feature extraction method. The factor analysis for multivariate time series results in the variance contribution rate, the top 3 common factors are 48.87%, 22.43%, and 18.96%, respectively. The cumulative contribution rate is 90.26% which is greater than 85%; So, these 3 factors reflect the original multivariate time series. The factor loading matrix is shown in Table 1. The common factor F_1 reflects the information of AC voltage, active power, reactive power and DC current at the faulty terminal. The common factor F_2 reflects the DC voltage and the common factor F_3 reflects the information of active and reactive power. The MMC model can be divided into three modules; the AC side model, DC side model, and the control

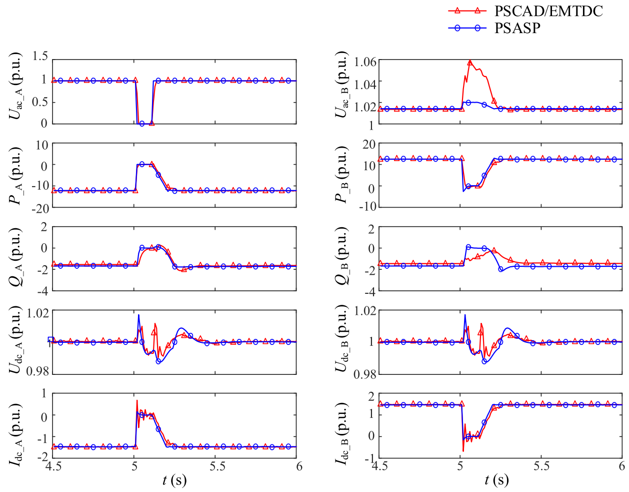


FIGURE 3. Curves of simulation results.

TABLE 1. Factor loadings matrix.

Observed variables	Common Factor F_1	F_2	F_3
U_{ac_A}	-0.8197	0.0522	-0.2759
$P_{\_A}$	0.8237	-0.1634	0.5403
$Q_{\_A}$	0.7493	-0.3508	0.5157
U_{dc_A}	-0.1475	0.9831	-0.0967
I_{dc_A}	-0.9773	0.0143	-0.2054
U_{ac_B}	0.2622	0.3313	0.2891
$P_{\_B}$	-0.8113	0.1995	-0.5467
$Q_{\_B}$	0.5423	-0.0896	0.6132
U_{dc_B}	-0.1475	0.9831	-0.0967
I_{dc_B}	0.9773	-0.0143	0.2054

system model [23]. Therefore, the three common factors can be classified as the above three models.

The modified adaptive Prony method is applied on common factors to extract their feature matrices. The common factors with short time windows are shown in Fig.4. As shown in Fig.4, piecewise points appear at the point where the

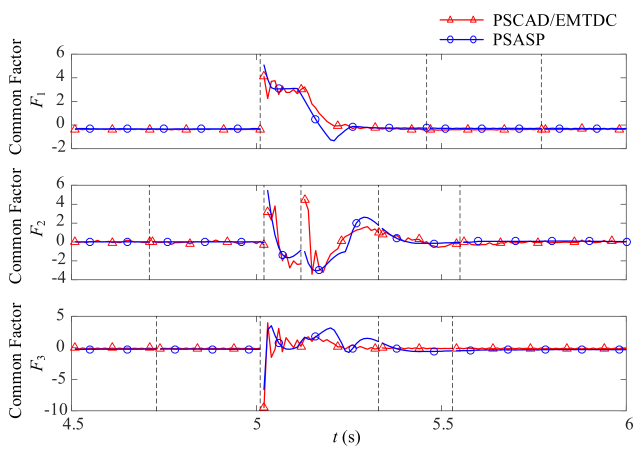


FIGURE 4. Common factors and short time windows.

signal mutated and signal in each time window is relatively stationary.

A feature matrix can be extracted from each time window. The feature matrices, extracted from 2nd short time window in the common factor F_1 of two multivariate time series, are shown in TABLE 2.

TABLE 2. Feature matrix.

No.	Amplitude (A) M^a	Amplitude (A) S^b	Frequency (f) M	Frequency (f) S	Damping (ζ) M	Damping (ζ) S	Phase (θ) M	Phase (θ) S	w_k
1	3.898	3.810	1.290	1.240	0.354	0.119	-5.75	-6.79	0.783
2	1.379	2.276	3.830	3.760	2.054	1.375	-7.10	-7.51	0.098
3	0.539	1.858	35.16	35.37	1.530	0.180	-10.4	-11.5	0.031
4	0.608	1.275	17.32	17.29	0.367	0.677	-11.0	-10.7	0.019
5	0.365	0.472	8.290	8.220	0.118	0.020	-9.15	-9.80	0.019
6	0.559	0.442	10.61	10.50	1.515	2.779	-9.31	-9.91	0.016
7	0.283	0.296	6.110	6.090	2.065	2.555	-8.82	-9.38	0.015
8	0.392	0.291	46.64	46.63	2.470	1.933	-10.9	-11.8	0.008
9	0.772	0.179	48.87	48.87	2.822	2.584	-10.9	-11.9	0.007
10	0.613	0.175	15.13	15.05	-1.52	-2.00	-10.8	-10.4	0.004

[a] PSASP
[b] PSCAD/EMTDC

B. VALIDATION AND ANALYSIS

The three parameters i.e. the weights of Prony exponential components, the weights of the short time windows and the weights of the common factors need to be determined in order to get feature similarity, common factor similarity and the simulation model similarity.

Table 3 shows the weights and the duration proportion (DP) to the total duration of each short time window. The weight of the short time window in the duration of the fault is greater than the DP of that, while the weight of the short time window in time period before the fault is smaller than the DP, which means the short time window in the duration of the fault conveys more information and its weight is higher.

TABLE 3. Weights and Duration Proportions of time windows.

Common Factor No.		Time Window No. 1	2	3	4	5	6
1	Weight (%)	27.96	39.25	20.14	12.65	-	-
	DP (%)	34	30	20.67	15.33	-	-
2	Weight (%)	12.22	22.25	12.68	12.76	13.24	30.85
	DP (%)	14	20.67	6.67	7.33	14.67	30
3	Weight (%)	13.70	17.82	22.86	15.07	30.55	-
	DP (%)	15.33	18.67	21.33	13.33	31.33	-

The common factor similarity can be described as follows by (27).

$$\mathbf{S}_{\mathrm{CF}} = \begin{bmatrix} A & f & \zeta & \theta \\ 0.8640 & 0.9951 & 0.7967 & 0.5630 \\ 0.1728 & 0.8473 & 0.7317 & 0.3226 \\ 0.3527 & 0.9719 & 0.5681 & 0.6774 \end{bmatrix} \quad (27)$$

The common factor similarity matrix $\mathbf{S}_{\mathrm{CF}}$ shows that the common factors F_1 obtained from two systems have similar features for amplitude, frequency, and damping. However, their features for the phase are different. The similarity of common factors F_2 is poor in the amplitude and phase features. The similarity features of common factors F_3 are not high for amplitude and damping. The results are in line with Fig. 4.

The weights of common factors are obtained as 0.5414, 0.2485 and 0.2101, according to the variance contribution rate, which means the AC side model represented by F_1 has the greatest impact on the accuracy of the simulation system. Hence, the similarity of simulation models is obtained.

$$\mathbf{S} = \begin{bmatrix} A & f & \zeta & \theta \\ 0.5848 & 0.9568 & 0.7325 & 0.5273 \end{bmatrix} \quad (28)$$

This means that the simulation results of PSASP and PSCAD/EMTDC are almost similar except the amplitude and phase; where the similarity is not so high. Therefore, the similarity of MMC model in PSASP should be considered for further improvement.

The errors on the DC and control side can also be identified by $\mathbf{S}_{\mathrm{CF}}$. The common factor F_2 has a poor similarity for the amplitude and phase feature. The difference between the DC side models is due to the difference in mechanism between electromechanical transient and electromagnetic transient simulation tools. The result of common factor F_3 shows the difference between the control module of two MMC simulation models. The model errors are identified based on amplitude and damping parameters that have great influence. The MMC control system in PSASP is simple than PSCAD/EMTDC, like the modulation process is simplified to a first-order lag model; which has lack of data because the time constant of the lag cannot be obtained from the field project. This parameter must also be optimized to get a higher similarity between the two models.

C. COMPARISON WITH THE FSV METHOD

The credibility of the simulation model is assessed by the FSV method [3] to verify the method proposed in this paper. The FSV overall evaluation indexes ADM, FDM and GDM of the 10 groups of simulation data are given in Table 4.

The results of the validation method proposed in this paper are consistent with those of the FSV method. The common factors F_1 of the two groups of simulation results have a high similarity as shown in (24). The variables with higher factor loading on F_1 are $U_{\mathrm{ac\_A}}$, P_A, Q_A, P_B, and $I_{\mathrm{dc\_B}}$, which have good evaluation results in FSV analysis. The evaluation results of common factor F_2 are unsatisfactory, and the FSV evaluation results of $U_{\mathrm{dc\_A}}$ and $U_{\mathrm{dc\_B}}$ with high factor loading to F_2 are poor too. The FSV evaluation results of P_A, Q_A, and P_B are good, but the results of Q_B are poor. Correspondingly, the similarity features of common factor F_3 is high for frequency but low for amplitude and damping.

TABLE 4. Overall evaluation results of FSV.

Variable	ADM	FDM	GDM
$U_{\mathrm{ac\_A}}$	0.0773 (Excellent)	0.2790 (Good)	0.3101 (Good)
$P_{\_A}$	0.0797 (Excellent)	0.1368 (Very good)	0.1748 (Very good)
$Q_{\_A}$	0.1701 (Very good)	0.3049 (Good)	0.3925 (Good)
$U_{\mathrm{dc\_A}}$	0.3012 (Good)	0.5921 (Fair)	0.7261 (Fair)
$I_{\mathrm{dc\_A}}$	0.0810 (Excellent)	0.2502(Good)	0.2867 (Good)
$U_{\mathrm{ac\_B}}$	0.7368 (Fair)	1.0733 (Poor)	1.4438 (Poor)
$P_{\_B}$	0.1243 (Very good)	0.3011 (Good)	0.3627 (Good)
$Q_{\_B}$	0.6499 (Fair)	0.6561 (Fair)	1.0359 (Poor)
$U_{\mathrm{dc\_B}}$	0.3012 (Good)	0.5920 (Fair)	0.7261 (Fair)
$I_{\mathrm{dc\_B}}$	0.0810 (Excellent)	0.2502 (Good)	0.2867 (Good)

The advantage of FSV is that its evaluation results can be described in natural language, and the corresponding relationship between expert opinions and quantitative evaluation is established. However, applying FSV in multivariable simulation, evaluation results of each variable rather than the overall evaluation results are obtained. In comparison with FSV, the proposed method can not only effectively evaluate the overall situation of the model, but also can be the evidence about model error location based on the physical meaning of factor analysis and Prony analysis in the evaluation process.

V. CONCLUSION

In this paper, a validation method for the simulation model of power system integrated with the internet of things is proposed which has the ability to process high-dimensional simulation data and provide evidence for model error location. The method consists of two main parts. First, a feature extraction method for multivariate time series is proposed based on factor analysis and modified adaptive Prony method. Second, a validation model based on the similarity evaluation is established. The validation discussed in this paper identifies the model errors and their locations; which can be used to improve the simulation model against the practical/acknowledged system. The method is verified by an application to MMC-HVDC model in PSASP.

REFERENCES

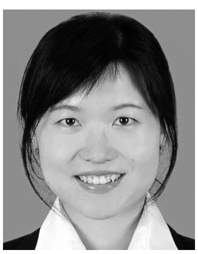
[1] B. K. Bose, "Global energy scenario and impact of power electronics in 21st century," *IEEE Trans. Ind. Electron.*, vol. 60, no. 7, pp. 2638–2651, Jul. 2013.

[2] Z. Shu, L. Kaipei, Q. Liang, L. Gang, H. Xianlai, and L. Dong, "Analysis of transient stability of power electronics dominated power system: An overview," *Proc. CSEE*, vol. 37, no. 14, pp. 3948–3962, 2017.

[3] Q. Peng, Y. Yang, H. Wang, and F. Blaabjerg, "On power electronized power systems: Challenges and solutions," in *Proc. IEEE Ind. Appl. Soc. Annu. Meeting (IAS)*, Portland, OR, USA, Sep. 2018, pp. 1–9.

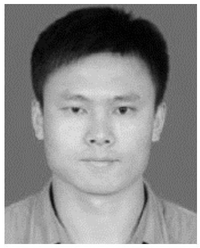
[4] Y. Saleem, N. Crespi, M. H. Rehmani, and R. Copeland, "Internet of Things-aided smart grid: Technologies, architectures, applications, prototypes, and future research directions," *IEEE Access*, vol. 7, pp. 62962–63003, 2019.

[5] J. Cárdenas and D. Menéndez, "Internet of Things: How the electrical grid can be controlled and managed in other dimensions," *J. Eng.*, vol. 2018, no. 15, pp. 918–923, Oct. 2018.
[6] Y. Wang, J. Yan, Q. Sun, J. Li, and Z. Yang, "A MobileNets convolutional neural network for GIS partial discharge pattern recognition in the ubiquitous power Internet of Things context: Optimization, comparison, and application," *IEEE Access*, vol. 7, pp. 150226–150236, 2019.
[7] S. Liu, X. Liu, and W. Zhang, "Fixed quantity evaluation to reliability of simulation system with similar degree," *J. Syst. Simul.*, vol. 14, no. 2, pp. 143–145, Feb. 2002.
[8] A. P. Duffy, A. J. Martin, A. Orlandi, G. Antonini, T. M. Benson, and M. S. Woolfson, "Feature selective validation (FSV) for validation of computational electromagnetics (CEM). Part I—The FSV method," *IEEE Trans. Electromagn. Compat.*, vol. 48, no. 3, pp. 449–459, Aug. 2006.
[9] R. Jauregui, G. Zhang, J. Rojas-Mora, O. Ventosa, F. Silva, A. P. Duffy, and H. Sasse, "Analyzing transient phenomena in the time domain using the feature selective validation (FSV) method," *IEEE Trans. Electromagn. Compat.*, vol. 56, no. 4, pp. 825–834, Aug. 2014.
[10] H. Ouyang and J. Wang, "Power system transient signal analysis based on Prony algorithm and neural network," presented at the IEEE Innov. Smart Grid Technol.-Asia (ISGT Asia), Tianjin, China, May 2012.
[11] A. Bracale, P. Caramia, and G. Carpinelli, "Adaptive Prony method for waveform distortion detection in power systems," *Elect. Power Energy Syst.*, vol. 29, no. 5, pp. 371–379, Jun. 2007.
[12] X. Sun, M. Gao, D. Liu, X. Hou, and R. Yuan, "Analysis and processing of electrical fault signals with modified adaptive Prony method," *Proc. CSEE*, vol. 30, no. 28, pp. 80–87, 2010.
[13] X.-D. Jia, G.-Y. Li, C.-Y. Zhao, X.-N. Xiao, W.-Q. Han, and Q. Guo, "Study of the credibility evaluation method for the power system simulation," *Proc. CSEE*, vol. 30, no. 19, pp. 51–57, Oct. 2010.
[14] Y. Chen, X. Wu, X. Bu, and R. Bai, "EMMD-Prony approach for dynamic validation of simulation models," *J. Syst. Eng. Electron.*, vol. 26, no. 1, pp. 172–181, Feb. 2015.
[15] G. Zhang, A. P. Duffy, A. Orlandi, D. Di Febo, L. Wang, and H. Sasse, "Comparison of data with multiple degrees of freedom utilizing the feature selective validation method," *IEEE Trans. Electromagn. Compat.*, vol. 58, no. 3, pp. 784–791, Jun. 2016.
[16] A. Duffy, G. Zhang, and H. Sasse, "Data comparison with many degrees of freedom: The FSV (feature selective validation) in multiple dimensions," presented at the 9th Int. Conf. Ultrawideband Ultrashort Impulse Signals (UWBUSIS), Odessa, Ukraine, Sep. 2018.
[17] Z. Gang, "Research on key issues of FSV method used for credibility assessment of electromagnetic simulation result," Ph.D. dissertation, Harbin Inst. Technol., Harbin, China, 2014.
[18] K. Fang, K. Zhao, and Y. Zhou, "Validation method for simulation models with cross iteration," *J. Syst. Eng. Electron.*, vol. 30, no. 3, pp. 555–563, Jun. 2019.
[19] N. Dehak, P. J. Kenny, R. Dehak, P. Dumouchel, and P. Ouellet, "Front-end factor analysis for speaker verification," *IEEE Trans. Audio, Speech Language Process.*, vol. 19, no. 4, pp. 788–798, May 2011.
[20] C. E. Shannon, "A mathematical theory of communication," *Bell Syst. Tech. J.*, vol. 27, no. 3, pp. 379–423, Jul. 1948, doi: 10.1002/j.1538-7305.1948.tb01338.x.
[21] S. Debnath, J. Qin, B. Bahrani, M. Saeedifard, and P. Barbosa, "Operation, control, and applications of the modular multilevel converter: A review," *IEEE Trans. Power Electron.*, vol. 30, no. 1, pp. 37–53, Jan. 2015.
[22] J. Wang, C. Liu, D. Xu, G. Xie, and Y. Zhu, "Simulation method of modular multilevel converter blocking state based on RTDS," *Trans. China Electrotech. Soc.*, vol. 33, no. 16, pp. 3686–3696, Aug. 2018.
[23] S. Liu, Z. Xu, W. Hua, G. Tang, and Y. Xue, "Electromechanical transient modeling of modular multilevel converter based multi-terminal HVDC systems," *IEEE Trans. Power Syst.*, vol. 29, no. 1, pp. 72–83, Jan. 2014.



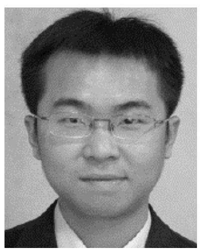
HUICAN CHEN was born in Chibi, Hubei, China. She received the B.S. and M.S. degrees from the South China University of Technology, Guangzhou, China. She is currently an Engineer with the Electric Power Dispatching and Control Center of Guangdong Power Grid Company, Guangzhou. Her research interests include the analysis, control, and operation of power systems.

JIECONG WANG (S'15) received the B.S. degree from the School of Electrical and Electronic Engineering, North China Electric Power University, Beijing, China, where she is currently pursuing the Ph.D. degree. She is also a Visiting Student Researcher with the Department of Electrical and Computer Engineering, University of Manitoba, Winnipeg, Canada. Her research interests include modeling, simulation, and analysis of power system with MMC-HVDC.



JUNLEI LIU was born in Linying, Henan, China. He received the B.S., M.S., and Ph.D. degrees from the South China University of Technology, Guangzhou, China. He is currently an Engineer with the Electric Power Dispatching and Control Center of Guangdong Power Grid Company, Guangzhou. His research interests include the analysis of safety and stability and control of reactive power and voltage for power systems.

BILAWAL REHMAN (S'12–M'19) received the B.S. and M.S. degrees in electrical engineering from the University of Management and Technology, Lahore, Pakistan, in 2012 and 2015, respectively. He is currently pursuing the Ph.D. degree in power system and automation with the School of Electrical and Electronic Engineering, North China Electric Power University, Beijing, China. His current research interests include HVDC operation, control, simulation and AC/DC power flow.



FENG QIAN received the Ph.D. degree from the China Electric Power Research Institution, Beijing, China. He is currently a Senior Engineer with the Electric Power Dispatching and Control Center of Guangdong Power Grid Company, Guangzhou, China. His research interests include the analysis, control, and operation of power systems.

CHONGRU LIU (M'10–SM'15) received the B.S., M.S., and Ph.D. degrees in electrical engineering from Tsinghua University, Beijing, China. She was a Visiting Professor with The University of Hong Kong, from 2009 to 2010, and a Visiting Professor with Washington State University, from 2010 to 2011. She is currently a Professor with the School of Electrical and Electronic Engineering, North China Electric Power University. She is a member of the National Power System Management and Information Exchange Standardization Committee of China. She is also a member of Beijing Nova and supported by the program of New Century Excellent Talents in University.