Automatic Identification of Power System Load Models Based on Field Measurements

DOI: 10.1109/TPWRS.2017.2763752

Document Version
Accepted author manuscript

Link to publication record in Manchester Research Explorer

Citation for published version (APA):

Published in:
IEEE Transactions on Power Systems

Citing this paper
Please note that where the full-text provided on Manchester Research Explorer is the Author Accepted Manuscript or Proof version this may differ from the final Published version. If citing, it is advised that you check and use the publisher's definitive version.

General rights
Copyright and moral rights for the publications made accessible in the Research Explorer are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

Takedown policy
If you believe that this document breaches copyright please refer to the University of Manchester’s Takedown Procedures [http://man.ac.uk/04Y6Bo] or contact uml.scholarlycommunications@manchester.ac.uk providing relevant details, so we can investigate your claim.
Automatic Identification of Power System Load Models Based on Field Measurements

Yue Zhu, Student Member, IEEE, Jovica V. Milanović, Fellow, IEEE

Abstract—With an ever growing complexity, the power grids are
designed and operated with an increasingly reduced stability
margin. Under such circumstances, the adequate modeling of
existing and new power system loads, with all its challenges, is
receiving renewed attention. To simplify and to large extent
automate the task of load modelling, this paper presents a
methodology and associated software tool—Automated Load
Modelling Tool (ALMT) to automatically (without human
intervention) develop load models and derive corresponding
model parameters from recorded power system responses. The
approach facilitates automatic identification of load models and
derivation of corresponding parameters for three different types
of load models, i.e. polynomial, dynamic exponential, and
composite load model. Several case studies using real life
measurements at 11 kV distribution buses are used to test and
validate developed methodology and software tool.

Index Terms—load modelling, optimisation, data filtering, power
system dynamics.

I. INTRODUCTION

Secure power system operation and control are closely
related to application of appropriate load models in system
studies [1]. If the load during the disturbance is appropriately
captured by the load models, the results of simulation studies
would be more reflective of actual system behavior and
consequently the operators will be more likely to successfully
handle the emergency conditions and better control the power
system in general. On the other hand, using inappropriate load
models, can lead to unexpected and potentially disastrous
results [2]. A number of massive black-outs in the past, to
name a few, the Swedish blackout of 1983, the Tokyo grid
collapse in 1987 [3], and the Western North America blackout
in 1996 [4], have proven the importance of using appropriate
load models in system studies. Historically, there has been a
large body of literature devoted to load modelling and the
effects that inappropriate load modelling could have on system
operation, both static and dynamic. The most up to date
overview of past efforts, extensive bibliography on load
modelling, latest industrial practices and recommendations on
load modelling are given in [5] and [6].

However important load modelling has proven to be, it is
impractical, both in terms of required human and financial
resources, to develop load models for every bus in the system
by staging dedicated field tests and perform subsequent data
analysis and model development. Selection of only important
load buses for which load models should be developed is a
solution to this problem. However, this is not straightforward
and it is still a matter of extensive research. The other option is
to develop an automated procedure without human
intervention, and therefore with limited financial and personnel
costs, for load model development, using routinely recorded
data by monitors installed at power system buses/substations.

Based on the findings and recommendations provided in
[6], this paper employs measurement based load modelling i.e.,
a top-down method that derives load models from the power
system disturbance data recorded at bulk supply power system
buses. Those data can be recorded by data acquisition devices
either during normal operation periods or at the time when
disturbance happens. The measurement based approach has the
advantages of capturing dynamic responses of load directly
from real power systems for both small and large disturbances
under variety of operating conditions and that prior information
of load composition is not needed [6]. It only needs a
hypothesis of load model according to the recorded power
system data.

The actual recorded system responses however, always
contain some noise induced by data acquisition devices. The

LIST OF SYMBOLS

<table>
<thead>
<tr>
<th>Symbol</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>$V_s$</td>
<td>Voltage symmetrical components</td>
</tr>
<tr>
<td>$V_p$</td>
<td>Voltage phase components</td>
</tr>
<tr>
<td>$\alpha_d$</td>
<td>Steady state exponential parameter</td>
</tr>
<tr>
<td>$\alpha_t$</td>
<td>Transient state exponential parameter</td>
</tr>
<tr>
<td>$V_0$</td>
<td>Initial steady state voltage</td>
</tr>
<tr>
<td>$P_0$</td>
<td>Initial steady state active power</td>
</tr>
<tr>
<td>$Q_0$</td>
<td>Initial steady state reactive power</td>
</tr>
<tr>
<td>$E_{d}$</td>
<td>d-axis transient EMF</td>
</tr>
<tr>
<td>$E_{q}$</td>
<td>q-axis transient EMF</td>
</tr>
<tr>
<td>$X_s$</td>
<td>Stator reactance</td>
</tr>
<tr>
<td>$X_r$</td>
<td>Rotor reactance</td>
</tr>
<tr>
<td>$\omega_r$</td>
<td>Rotor speed</td>
</tr>
<tr>
<td>$E_{d}$</td>
<td>d-axis transient EMF</td>
</tr>
<tr>
<td>$E_{q}$</td>
<td>q-axis transient EMF</td>
</tr>
<tr>
<td>$T_0$</td>
<td>Open circuit time constant</td>
</tr>
<tr>
<td>$H$</td>
<td>Rotor inertia constant</td>
</tr>
<tr>
<td>$I_d$</td>
<td>d-axis stator current</td>
</tr>
<tr>
<td>$I_q$</td>
<td>q-axis stator current</td>
</tr>
<tr>
<td>$V_d$</td>
<td>d-axis bus voltage</td>
</tr>
<tr>
<td>$V_q$</td>
<td>q-axis bus voltage</td>
</tr>
<tr>
<td>$Q_{eq}$</td>
<td>Initial reactive power of the equivalent motor</td>
</tr>
<tr>
<td>$P_{eq}$</td>
<td>Initial active power of the equivalent motor</td>
</tr>
</tbody>
</table>

The authors are with the School of Electrical and Electronic Engineering, The University of Manchester, PO Box 88, Manchester, M60 1QD, UK. (email: yue.zhu-5@postgrad.manchester.ac.uk, milanovic@manchester.ac.uk).
resolution of data captured is affected by the sampling rate used by data acquisition system and the observation window affects the observed load response as natural time variation of load could be confused with load dynamic response to disturbances. All these effects have to be filtered out and suitably taken care of, prior to load model development.

The methodology and tools presented in this paper take all the above in consideration and select an optimal configuration at “signal pre-processing stage” of the methodology for identifying true load response before identifying load models and associated load model parameters. Three different filtering techniques, namely Moving Average (MA) filter, Savitzky-Golay (SG) filter, and Butterworth (BW) filter are compared in order to select the most appropriate one [7, 8]. This is followed by selecting the appropriate parameter identification technique for parameter fitting. For a linear model, its parameters can be properly fitted by most parameter identification approaches [9, 10]. However, there are many challenges when fitting parameters for non-linear models. In this paper, the optimisation based approach is used for parameter. Three different types of optimisation algorithms have been implemented and tested, which are Least-Squares method (LS) [11], Genetic Algorithm (GA) [12], and Simulated Annealing (SA) [12, 13] prior to selecting the Least-Squares method as the optimal for the task in hand.

With optimally selected signal pre-processing and parameter identification methodology, the recorded signals by data monitors, i.e., the input to the Automated Load Modelling Tool (ALMT) developed in this study, are automatically processed; the appropriate load models based on filtered responses automatically identified among three most widely used load models, i.e. polynomial, dynamic exponential and composite load model; and finally parameters of those load models fitted. The output of the ALMT is the load model with all parameters identified. The ALMT is tested and validated using large number of real life measurements at 11 kV distribution buses.

The main contribution of this paper is making the load modelling a practically fully automated process without human intervention. The paper discusses a software tool that can automatically identify load models and corresponding parameters from data routinely measured in real power system. The three processes involved in load model development, namely the data pre-processing, load model selection, and parameter fitting are integrated into a single automated tool. The developed ALMT facilitates automatic load model identification at power system buses based on on-line measurements and can be implemented as an additional feature of standard power quality monitors or advanced fault recorders. As such it greatly facilitates power system load modelling at virtually no additional expense to power system utilities and consequently facilitates more accurate power system analysis leading to overall more secure operation and control of power systems.

II. THE AUTOMATED LOAD MODELLING

The ALMT is developed in MATLAB R2013a. It contains three major parts: data processing, load model selection, and load model parameter identification. The whole process, from

A. Data Processing

After the recorded power system signals are imported into the program, they need to be processed before identifying load models. Each recorded data set contains three phase voltages, three phase currents, active power (P), reactive power (Q), and frequency. As the phase voltages tend to be unbalanced, they have to be converted into symmetrical components using (1).

\[
V_s = A^{-1}V_p
\]  

In the case of composite load model [14], however, the direct axis voltage and quadrature axis voltage components, \(V_d\) and \(V_q\) respectively are needed as input voltage values. Thus, the three phase voltage also needs to be converted into d-axis and q-axis voltages by using Park’s abc to dq transformation. In measurement – based load modelling, not all recorded data can be directly used to build load models. The key prerequisite is that recorded voltage signal must contain obvious voltage change. The voltage change due to routine transformer tap changing often falls in the range of 0.5% to 2.5% [15, 16]. And the consequential \(P\) and \(Q\) changes should be also within a reasonable range. Therefore, after data conversion, to symmetrical and \(d-q\) components, the time slots that contain significant voltage change need to be identified first. The neighbouring (adjacent) data comparison method is used to identify data points that contain large voltage changes. In addition to change in magnitude, the change should also last long enough to clearly separate it from random spikes in signal and to separate consecutive voltage changes from one another.

The next step is choosing the appropriate filters to filter the selected events. Three filters widely used for this purpose in the past, the SG, MA, and BW are used to process the same sample data initially in order to select the most appropriate one. The

Fig. 1. Flow chart of the measurement – based modelling

Then the positive sequence of the symmetrical voltage components is used as \(V\) in the subsequent analysis for development of static exponential, polynomial and dynamic exponential load model.

\[
V_s = A^{-1}V_p
\]
MA filter is easy to apply and has been widely used to filter digital signals. It has a very good performance in reducing Gaussian white noise while not losing sharp step responses at the same time [7]. The SG filter can increase the signal to noise ratio (SNR) and keep the useful signal undistorted [8]. It is implemented using local least-squares polynomial approximation. Finally, the Butterworth filter has a distinct characteristic of having maximum level of flat magnitude response in the pass-band as it removes ripples from the original signal [8]. Comparison of performance of three different filters was performed for voltage, active power, and reactive power filtering, as illustrated in Fig. 2 – 4. A step change disturbance in recorded signals, observable at the time of 20 s in these figures, was caused by transformer tap change.

![Fig. 2. Voltage filtered by three different filters.](image)

![Fig. 3. Real power filtered by three different filters.](image)

![Fig. 4. Reactive power filtered by three different filters.](image)

The MA filter, on the other hand, cannot appropriately filter the start point and the end point of a signal, which results in an obvious spike at each end. Because there is no data point at one side of these two points, no average value can be obtained for the start point and the end point. Even though the application of MA filter results in the smoothest filtered signal among the tested filters, some useful peak features of the original curve may not be captured accurately. For example, at 20 s, when there is a voltage step change, the filtered V, P, and Q curves have a slower response than when BW and SG filters were used. At that point, a sharp change in signal value is desired for better parameter fitting, so this is an important drawback of MA filter.

The SG filter is found to be the best at maintaining the dynamic feature of the original signal, especially some useful peaks. At 20 s, when there is a voltage step change, the filtered curve has a sharp response, which preserves the feature of the original signal very well. It can be seen clearly from the figures that in addition to persevering the dynamic feature of the original signal, the SG filter is also good in removing the noise from the signal.

In summary, there were three criteria used for determining which filter is the most appropriate for filtering recorded signals, namely efficiency in noise removing, preserving dynamic feature of the original signal, and the quality at the end (start and the end of disturbance) points. For noise removing, MA filter has the best performance, followed by the SG and BW filter. For preserving dynamic feature of the original signal, the performance order is SG, MA, and BW. For the quality at end points, the performance order is SG, MA, and BW. Since SG filter performs the best for two out of three criteria and comes the second for the third it has been selected for filtering the recorded data and implementation in ALMT.

After filtered responses are obtained, they are converted into per unit value to facilitate subsequent parameter identification. In conversion to per unit values the rated voltage at the bus and the initial steady state power are used as base values.

Following the filtering and normalization stage, the recorded events have to be divided into different categories based on possible generic load responses. Three possible responses of P and Q following the voltage change can be distinguished. They are static response, first order dynamic recovery response, and high(er) order dynamic oscillatory response, as shown in Table I. The tool is capable of automatically distinguishing the type of response based on its shape by comparing it to these three generic responses. It is based on the identification and comparison of the shape features of these responses. A flow chart illustrating the process of identification of (differentiation between) different responses following a voltage step is given in Fig. 5.

<table>
<thead>
<tr>
<th>TABLE I</th>
</tr>
</thead>
<tbody>
<tr>
<td>THREE TYPES OF POSSIBLE LOAD RESPONSE SHAPES</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Response Shape</th>
<th>Response Name</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Static Response</td>
</tr>
<tr>
<td></td>
<td>First Order Dynamic Recovery Response</td>
</tr>
</tbody>
</table>
The last step of data pre-processing is data modification. After filtering, the overall quality of recorded responses has been improved a lot as they are cleared from noise and only the relevant responses are stored for further processing. Nevertheless, there are still some fluctuations in filtered signals, i.e., in what is presumed to be the new steady state of the responses as well as data processing (filtering) induced distortions. Some of these fluctuations may be caused by remaining noise and some may be due to the natural fluctuation of load. The signal distortions (the filter actually reduced the quality of response) are the consequence of filter “slowing down” the response, i.e., at the point when the voltage step change happens, the filter depending on selected parameters introduces different ramp rate in a signal instead of a sharp change. This “delay” in response needs to be compensated prior to parameter identification. The Fig. 6, Fig. 7, and Fig. 8 illustrate the modifications (red line) of filtered (green line) signals introduced prior to load model parameter identification.

B. Load Model Selection

For the load model identification and parameter fitting in developed ALTM, three types of load models are considered – second order polynomial load model, exponential dynamic load model and composite load model. The second order polynomial load model, also known as ZIP model as it represents combination of constant impedance (Z), constant current (I), and constant power (P) parts, is a commonly used static load model [17, 18]. This model can be expressed by (3) and (4).

\[
P = P_0 \left( p_1 \left( \frac{V}{V_0} \right)^2 + p_2 \left( \frac{V}{V_0} \right) + p_3 \right)
\]  

(3)

\[
p_1 + p_2 + p_3 = 1
\]  

(4)

Fig. 6. Modification of a static response.

Fig. 7. Modification of first order recovery response.

Fig. 8. Modification of high order oscillatory response.
Where \( p_i \), \( i=1,2,3 \), are parameters describing the contribution of constant impedance, constant current and constant power load, respectively, in the load mix. The similar notation applies to reactive power.

The first order exponential dynamic load model is given by (5) and (6) [17]. The parameters of the model are defined in Fig. 9. It is best suited for long term voltage stability analysis and dynamic studies where there is no major participation of induction motors (IM) in load mix.

\[
T_p \frac{dP_L}{dt} + P_L = P_o(V) - P_i(V) = P_o \left( \frac{V}{V_0} \right)^{a_S} - P_i \left( \frac{V}{V_0} \right)^{a_T} \tag{5}
\]

\[
P_i = P_r + P_o \left( \frac{V}{V_0} \right)^{a_T} \tag{6}
\]

Fig. 9. Exponential dynamic model response

The reactive power dependence on voltage for the load models presented above can be described by the equations of the same form. They are not included here due to space limitation.

The composite load model is the most advanced and widely used load model in dynamic studies [6, 18], as it is valid under most power system stability analysis scenarios. The most commonly used composite load model consist of parallel connection of ZIP load model and an IM model [6, 14, 19-22], as shown in Fig. 10.

Fig. 10. A schematic representation of the ZIP-IM composite load model

The IM part can be described as (7) and (8).

\[
\begin{align*}
\frac{dE_p'}{dt} &= -\frac{1}{T_p} \left[ E_p' + (X - X')I_q \right] - (\omega_r - 1)E_q' \\
\frac{dE_q'}{dt} &= -\frac{1}{T_q} \left[ E_q' - (X - X')I_d \right] + (\omega_r - 1)E_d' \\
\frac{d\omega_r}{dt} &= -\frac{1}{2H} \left[ (A\omega_r^2 + B\omega_r + C)T_0 - (E_d'I_d + E_q'I_q) \right]
\end{align*}
\]

\[
\begin{align*}
I_d &= \frac{1}{R_d + X'^2} \left[ R_s(V_d - E_d') + X'(V_q - E_q') \right] \\
I_q &= \frac{1}{R_q + X'^2} \left[ R_s(V_q - E_q') + X'(V_d - E_d') \right]
\end{align*}
\]

Where:

\[
T' = \frac{X_r + X_m}{R_r}, \quad X = X_s + X_m
\]

\[
X' = X_s + \frac{X_mX_r}{X_m + X_r}, \quad A + B + C = 1
\]

And its static part is a ZIP load model, described by (9)

\[
\begin{align*}
P_{ZIP} &= P_0 \left( p_1 \left( \frac{V}{V_0} \right)^2 + p_2 \left( \frac{V}{V_0} \right) + p_3 \right) \\
Q_{ZIP} &= Q_0 \left( q_1 \left( \frac{V}{V_0} \right)^2 + q_2 \left( \frac{V}{V_0} \right) + q_3 \right)
\end{align*}
\]

It satisfies

\[
\begin{align*}
p_1 + p_2 + p_3 &= 1 - K_p \\
q_1 + q_2 + q_3 &= 1 - \frac{Q_{eq}}{Q_0}
\end{align*}
\]

\[A, B, \text{ and } C\] are Mechanical torque coefficients. \( K_p \) is used to define the proportion of initial active power of equivalent motor in the composite load model, given by (10)

\[
K_p = \frac{P_{eq}}{P_0}
\]

Parameter \( M_{eq} \) represents the equivalent motor capacity base. It is calculated using (11).

\[
M_{eq} = \frac{P_{eq}}{V_0/V_p}
\]

Where \( V_b \) is the chosen voltage base, and \( S_{eqb} \) is the equivalent motor capacity base.

After the responses are obtained and classified, they can now be used to build load models. For static responses, i.e., time invariant responses of real and reactive power, the software will choose ZIP load model or exponential static load model. For first order recovery response, the exponential dynamic load model will be chosen. For high order oscillatory response, the software will choose composite load model to fit parameters on. Only one load model is used to represent power recovery in a certain time slot. If different types of responses are recorded in that time slot, then the load model is chosen based on the dominant type of response. Once the load model is chosen and its parameter fitted, the model response is simulated to verify its match with recorded load response. If the simulated response does not match the original measured response, then another load model is selected and the procedure is repeated until adequate match of simulated and recorded load response is obtained.

The implementation of static load models and exponential dynamic load models is straightforward, because there is only one equation for \( P \) and \( Q \) respectively. However, the composite load model consists of several differential equations, which do not express direct relationship between \( V, P, \text{ and } Q \). Thus it cannot be implemented in the same way as other models.
instead, it is represented by a differential algebraic equations set [23].

\[
\begin{align*}
\dot{x} &= f(x, u, p) \\
y &= g(x, u, p)
\end{align*}
\]  

(12)

Where \( f \) and \( g \) are nonlinear functions of state variable \( x \), input \( u \), and the parameter vector \( p \). To express composite load model in the form of (12), the expressions for active power and reactive power of the induction motor are needed, as in (13) and (14)

\[
\begin{align*}
P_{IM} &= V_d I_d + V_q I_q \\
Q_{IM} &= V_q I_d - V_d I_q
\end{align*}
\]  

(13)

(14)

Then \( P \) and \( Q \) of composite load model are represented as the sum of \( P \) and \( Q \) of induction motor model and ZIP model, as in (15) and (16)

\[
\begin{align*}
P &= P_{IM} + P_{ZIP} \\
Q &= Q_{IM} + Q_{ZIP}
\end{align*}
\]  

(15)

(16)

Therefore, Equation (12) can be expressed in the state space form in (17):

\[
\begin{align*}
\dot{X} &= AX + BU \\
Y &= CX
\end{align*}
\]  

(17)

Where

\[
X = [E_d' \ E_q' \ \omega_e]^T \\
U = [V_d \ V_q]^T \\
Y = [P \ Q]^T
\]

C. Load Model Parameter Fitting

The performance of three parameter fitting algorithms, i.e., Least Squares Method, Genetic Algorithm, and Simulated Annealing is tested within MATLAB, using sample data to identify the most suitable algorithm for further use. The MATLAB functions used for this purpose were “lscurvefit”, “ga”, and “simulannealbnd” respectively. The parameter fitting for ZIP load model is used for comparison of fitting algorithms.

The accuracy of the identified load model and its parameters is assessed based on calculation of the mean square errors between the simulated response obtained with chosen load model and measured response. The errors, in case of real power are calculated by (18):

\[
P_{error} = \frac{1}{n} \sum_{i=1}^{n} (P_{model_i} - P_{measured_i})^2
\]  

(18)

Where \( P_{model_i} \) and \( P_{measured_i} \) are the \( i \)-th simulated and measured real power values and \( n \) is the number of recorded events considered. The errors of \( Q \) are calculated in the same way. The results of comparison of parameter fitting using different algorithms are given in Table II – Table IV for eight recorded events in the test data. The average value and the variance of each parameter are also given. The average computation time of each event is 0.84s.

From Table II – Table IV, it can be concluded that based on the average error and the variances of parameters, the least squares method performs the best among the three algorithms. Simulated annealing has smaller average of error than genetic algorithm, but it has very large variance. The genetic algorithm and simulated annealing algorithms are more advanced methods and, at least in theory, they should perform better than the least squares method. In this application though the emphasis was not on extremely high accuracy of fitted/simulated responses since various filtering and response adjustments had to be made anyway to the original response at the pre-processing stage to make the load model identification problem practical and realistic. Therefore, the maximum iteration number used for genetic algorithm and simulated annealing was not very high in order to increase the speed of identification as program may have to fit parameters for hundreds of events to obtain a final load model parameter set. If very high accuracy of performance of GA and SA algorithms were the objective, the parameter identification process would become computationally very expensive and impractical. Table V compares the performance of the three methods in terms of the accuracy and computational time requirements. It can be seen that to achieve similar parameter variance as with the least squares method, the computational time for GA and SA should be more than 10 times longer. Therefore, the least squares method is implemented in the software as computationally more efficient.

### Table II

<table>
<thead>
<tr>
<th>Event</th>
<th>p1</th>
<th>p2</th>
<th>p3</th>
<th>P error</th>
<th>q1</th>
<th>q2</th>
<th>q3</th>
<th>Q error</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.510</td>
<td>0.401</td>
<td>0.090</td>
<td>5.65 × 10⁻⁵</td>
<td>2.676</td>
<td>0.406</td>
<td>-2.082</td>
<td>1.75 × 10⁻⁵</td>
</tr>
<tr>
<td>2</td>
<td>0.593</td>
<td>0.399</td>
<td>0.007</td>
<td>2.03 × 10⁻⁵</td>
<td>2.128</td>
<td>0.396</td>
<td>-1.523</td>
<td>2.15 × 10⁻⁵</td>
</tr>
<tr>
<td>3</td>
<td>1.061</td>
<td>0.402</td>
<td>-0.462</td>
<td>2.37 × 10⁻⁵</td>
<td>2.788</td>
<td>0.406</td>
<td>-2.194</td>
<td>2.76 × 10⁻⁵</td>
</tr>
<tr>
<td>4</td>
<td>0.849</td>
<td>0.401</td>
<td>-0.251</td>
<td>6.36 × 10⁻⁵</td>
<td>2.453</td>
<td>0.405</td>
<td>-1.858</td>
<td>3.62 × 10⁻⁵</td>
</tr>
<tr>
<td>5</td>
<td>0.938</td>
<td>0.401</td>
<td>-0.339</td>
<td>1.19 × 10⁻⁵</td>
<td>2.741</td>
<td>0.405</td>
<td>-2.146</td>
<td>4.87 × 10⁻⁵</td>
</tr>
<tr>
<td>6</td>
<td>0.199</td>
<td>0.400</td>
<td>0.401</td>
<td>7.65 × 10⁻⁵</td>
<td>2.195</td>
<td>0.405</td>
<td>-1.600</td>
<td>4.05 × 10⁻⁵</td>
</tr>
<tr>
<td>7</td>
<td>0.361</td>
<td>0.400</td>
<td>0.239</td>
<td>5.27 × 10⁻⁵</td>
<td>2.313</td>
<td>0.406</td>
<td>-1.719</td>
<td>5.03 × 10⁻⁵</td>
</tr>
<tr>
<td>8</td>
<td>1.155</td>
<td>0.402</td>
<td>-0.556</td>
<td>5.61 × 10⁻⁵</td>
<td>3.280</td>
<td>0.406</td>
<td>-2.685</td>
<td>3.64 × 10⁻⁵</td>
</tr>
<tr>
<td>Average</td>
<td>0.708</td>
<td>0.401</td>
<td>-0.109</td>
<td>4.52 × 10⁻⁵</td>
<td>2.572</td>
<td>0.404</td>
<td>-1.976</td>
<td>3.48 × 10⁻⁵</td>
</tr>
<tr>
<td>Variance</td>
<td>0.118</td>
<td>0.000</td>
<td>0.119</td>
<td>-</td>
<td>0.144</td>
<td>0.000</td>
<td>0.145</td>
<td>-</td>
</tr>
</tbody>
</table>

### Table III

**ZIP Parameters Obtained by Least Squares Method**
III. Case Studies and Results

A. Case Studies

Two case studies are presented in the paper to illustrate the performance of developed ALMT. In the first case the load response is recorded using 1 Hz sampling rate. Responses recorded at this sampling rate can be used to identify parameters of static load model and first order exponential dynamic load model. In the second case the load response is recorded using 50 Hz sampling rate and as such a composite load model can be used for parameter fitting since higher order load dynamics (2nd or higher order oscillatory responses) can be identified and modeled. (Note: The higher order load dynamics is characterized by oscillatory response of real and/or reactive power, as shown in Fig. 8. These oscillatory responses can only be adequately captured and subsequently reproduced if a sufficiently high sampling rate is used, typically between 1 ms and 16.67/20 ms for 60/50 Hz system.)

The first case study involves 3 GB of recorded load responses (with 1 Hz sampling rate, i.e., 86400 samples per day) at 15 11kV distribution substations over a period of 5 months. The data includes three phase voltages, three phase currents, active power, and reactive power for each substation. The example of the recorded data is shown in Table VI.

The load composition at a substation is different at different times. Therefore different load models will be identified for different time periods. During the week, there will be different load models for weekdays and weekends, and during weekdays, there will be different load models for daytime and night. In the available data set there were between 5 and 8 events during each time period that could be used for parameter fitting.

The second case study involves data recorded with 50 Hz sampling rate at a single substation. There are only several events that could be used for model identification. The data includes three phase voltage, active power and reactive power recorded with Phasor Measurement Unit (PMU) at a substation.

<table>
<thead>
<tr>
<th>Event</th>
<th>p1</th>
<th>p2</th>
<th>p3</th>
<th>P error</th>
<th>q1</th>
<th>q2</th>
<th>q3</th>
<th>Q error</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.770</td>
<td>0.054</td>
<td>0.171</td>
<td>8.94 x 10^-5</td>
<td>0.345</td>
<td>0.151</td>
<td>0.496</td>
<td>4.35 x 10^-4</td>
</tr>
<tr>
<td>2</td>
<td>0.337</td>
<td>-0.109</td>
<td>0.777</td>
<td>3.42 x 10^-5</td>
<td>0.273</td>
<td>0.451</td>
<td>0.295</td>
<td>2.22 x 10^-4</td>
</tr>
<tr>
<td>3</td>
<td>0.473</td>
<td>0.328</td>
<td>0.196</td>
<td>4.29 x 10^-5</td>
<td>0.813</td>
<td>0.057</td>
<td>0.114</td>
<td>2.58 x 10^-4</td>
</tr>
<tr>
<td>4</td>
<td>-0.562</td>
<td>1.868</td>
<td>-0.309</td>
<td>3.03 x 10^-5</td>
<td>3.259</td>
<td>-0.287</td>
<td>-1.972</td>
<td>5.41 x 10^-4</td>
</tr>
<tr>
<td>5</td>
<td>0.131</td>
<td>0.985</td>
<td>-0.110</td>
<td>1.07 x 10^-5</td>
<td>0.356</td>
<td>0.505</td>
<td>0.315</td>
<td>3.94 x 10^-4</td>
</tr>
<tr>
<td>6</td>
<td>0.954</td>
<td>0.298</td>
<td>-0.245</td>
<td>3.61 x 10^-5</td>
<td>0.430</td>
<td>1.315</td>
<td>-0.752</td>
<td>1.45 x 10^-4</td>
</tr>
<tr>
<td>7</td>
<td>0.608</td>
<td>0.275</td>
<td>0.115</td>
<td>2.93 x 10^-5</td>
<td>0.991</td>
<td>0.391</td>
<td>-0.377</td>
<td>4.59 x 10^-4</td>
</tr>
<tr>
<td>8</td>
<td>0.385</td>
<td>0.045</td>
<td>0.565</td>
<td>3.32 x 10^-5</td>
<td>0.660</td>
<td>0.867</td>
<td>-0.543</td>
<td>2.07 x 10^-4</td>
</tr>
</tbody>
</table>

Average: 0.387 0.468 0.145 3.83 x 10^-5 0.891 0.431 -0.215 2.72 x 10^-4

Variance: 0.213 0.428 0.143 - 0.980 0.245 0.757 -

<table>
<thead>
<tr>
<th>Event</th>
<th>p1</th>
<th>p2</th>
<th>p3</th>
<th>P error</th>
<th>q1</th>
<th>q2</th>
<th>q3</th>
<th>Q error</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>-2.341</td>
<td>5.682</td>
<td>-2.342</td>
<td>5.83 x 10^-5</td>
<td>8.169</td>
<td>-10.84</td>
<td>3.675</td>
<td>1.80 x 10^-5</td>
</tr>
<tr>
<td>2</td>
<td>-1.091</td>
<td>3.656</td>
<td>-1.565</td>
<td>2.03 x 10^-5</td>
<td>4.825</td>
<td>-5.339</td>
<td>1.515</td>
<td>2.23 x 10^-5</td>
</tr>
<tr>
<td>3</td>
<td>-0.173</td>
<td>2.148</td>
<td>-0.977</td>
<td>2.94 x 10^-5</td>
<td>9.020</td>
<td>-10.99</td>
<td>2.974</td>
<td>3.92 x 10^-5</td>
</tr>
<tr>
<td>4</td>
<td>1.230</td>
<td>0.669</td>
<td>-0.895</td>
<td>2.02 x 10^-5</td>
<td>2.986</td>
<td>-0.352</td>
<td>-1.633</td>
<td>3.86 x 10^-5</td>
</tr>
<tr>
<td>5</td>
<td>-0.930</td>
<td>4.177</td>
<td>-2.246</td>
<td>1.21 x 10^-5</td>
<td>9.507</td>
<td>-10.02</td>
<td>1.520</td>
<td>1.46 x 10^-5</td>
</tr>
<tr>
<td>6</td>
<td>-2.671</td>
<td>5.822</td>
<td>-2.152</td>
<td>8.95 x 10^-5</td>
<td>14.34</td>
<td>-20.00</td>
<td>6.672</td>
<td>2.55 x 10^-5</td>
</tr>
<tr>
<td>7</td>
<td>-1.209</td>
<td>2.529</td>
<td>-0.32</td>
<td>4.18 x 10^-5</td>
<td>3.104</td>
<td>-4.345</td>
<td>2.227</td>
<td>2.43 x 10^-5</td>
</tr>
<tr>
<td>8</td>
<td>0.181</td>
<td>3.455</td>
<td>-2.633</td>
<td>1.47 x 10^-5</td>
<td>6.709</td>
<td>-5.115</td>
<td>-0.590</td>
<td>5.18 x 10^-5</td>
</tr>
</tbody>
</table>

Average: -0.875 3.517 -1.641 2.57 x 10^-5 7.332 -8.375 2.045 1.02 x 10^-4

Variance: 1.654 3.060 0.693 - 14.41 35.89 6.570 -

<table>
<thead>
<tr>
<th>Computation</th>
<th>GA</th>
<th>SA</th>
<th>LS</th>
</tr>
</thead>
<tbody>
<tr>
<td>time [s]</td>
<td>p1</td>
<td>p2</td>
<td>p3</td>
</tr>
<tr>
<td>1</td>
<td>0.204</td>
<td>0.416</td>
<td>0.138</td>
</tr>
<tr>
<td>10</td>
<td>0.122</td>
<td>0.143</td>
<td>0.121</td>
</tr>
<tr>
<td>100</td>
<td>0.107</td>
<td>0.000</td>
<td>0.115</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Table VI</th>
<th>CASE I A SAMPLE OF RECORDED DATA</th>
</tr>
</thead>
<tbody>
<tr>
<td>Timestamp</td>
<td>T11 V1</td>
</tr>
<tr>
<td>7:00:00</td>
<td>6574</td>
</tr>
<tr>
<td>7:00:01</td>
<td>6574</td>
</tr>
<tr>
<td>7:00:02</td>
<td>6573</td>
</tr>
<tr>
<td>7:00:03</td>
<td>6573</td>
</tr>
<tr>
<td>7:00:04</td>
<td>6573</td>
</tr>
</tbody>
</table>
in a real power system. The example of the recorded data is shown in Table VII.

### TABLE VII

<table>
<thead>
<tr>
<th>Timestamp UTC</th>
<th>P [MW]</th>
<th>Q [Mvar]</th>
<th>VA [V]</th>
<th>VB [V]</th>
<th>VC [V]</th>
</tr>
</thead>
<tbody>
<tr>
<td>16:44:32.100</td>
<td>3.464</td>
<td>0.101</td>
<td>6564.0</td>
<td>6593.7</td>
<td>6580.3</td>
</tr>
<tr>
<td>16:44:32.120</td>
<td>3.461</td>
<td>0.100</td>
<td>6562.7</td>
<td>6592.4</td>
<td>6579.1</td>
</tr>
<tr>
<td>16:44:32.140</td>
<td>3.461</td>
<td>0.098</td>
<td>6562.1</td>
<td>6531.5</td>
<td>6578.5</td>
</tr>
<tr>
<td>16:44:32.160</td>
<td>3.462</td>
<td>0.098</td>
<td>6561.6</td>
<td>6591.6</td>
<td>6577.9</td>
</tr>
<tr>
<td>16:44:32.180</td>
<td>3.462</td>
<td>0.098</td>
<td>6560.3</td>
<td>6591.9</td>
<td>6577.5</td>
</tr>
</tbody>
</table>

### TABLE VIII

**CASE STUDY RESULTS OF SUBSTATION 6, CASE I**

<table>
<thead>
<tr>
<th>Parameters</th>
<th>( p_1 )</th>
<th>( p_2 )</th>
<th>( p_3 )</th>
<th>( q_1 )</th>
<th>( q_2 )</th>
<th>( q_3 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Weekday</td>
<td>1.284</td>
<td>0.398</td>
<td>-0.694</td>
<td>3.046</td>
<td>0.399</td>
<td>-2.470</td>
</tr>
<tr>
<td>daytime</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>1.253</td>
<td>0.382</td>
<td>-0.655</td>
<td>3.122</td>
<td>0.412</td>
<td>-2.463</td>
</tr>
<tr>
<td></td>
<td>1.203</td>
<td>0.374</td>
<td>-0.744</td>
<td>2.917</td>
<td>0.369</td>
<td>-2.558</td>
</tr>
<tr>
<td></td>
<td>1.382</td>
<td>0.421</td>
<td>-0.635</td>
<td>3.223</td>
<td>0.435</td>
<td>-2.359</td>
</tr>
<tr>
<td></td>
<td>0.759</td>
<td>0.400</td>
<td>-0.159</td>
<td>2.799</td>
<td>0.401</td>
<td>-2.200</td>
</tr>
<tr>
<td>Weekday</td>
<td>0.738</td>
<td>0.411</td>
<td>-0.145</td>
<td>2.807</td>
<td>0.398</td>
<td>-2.188</td>
</tr>
<tr>
<td>night</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.717</td>
<td>0.382</td>
<td>-0.223</td>
<td>2.674</td>
<td>0.377</td>
<td>-2.316</td>
</tr>
<tr>
<td></td>
<td>0.807</td>
<td>0.411</td>
<td>-0.099</td>
<td>2.955</td>
<td>0.424</td>
<td>-2.084</td>
</tr>
<tr>
<td>weekends</td>
<td>0.732</td>
<td>0.400</td>
<td>-0.132</td>
<td>2.814</td>
<td>0.402</td>
<td>-2.216</td>
</tr>
<tr>
<td></td>
<td>0.749</td>
<td>0.388</td>
<td>-0.147</td>
<td>2.806</td>
<td>0.400</td>
<td>-2.211</td>
</tr>
<tr>
<td></td>
<td>0.703</td>
<td>0.385</td>
<td>-0.166</td>
<td>2.754</td>
<td>0.374</td>
<td>-2.338</td>
</tr>
<tr>
<td></td>
<td>0.751</td>
<td>0.412</td>
<td>-0.115</td>
<td>2.889</td>
<td>0.418</td>
<td>-2.097</td>
</tr>
</tbody>
</table>

### TABLE IX

**CASE STUDY RESULTS OF SUBSTATION 1, CASE I**

<table>
<thead>
<tr>
<th>Parameters</th>
<th>( a_x )</th>
<th>( a_z )</th>
<th>( T_p )</th>
<th>( \beta_1 )</th>
<th>( \beta_2 )</th>
<th>( T_q )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Weekday</td>
<td>2.533</td>
<td>0.643</td>
<td>15.3</td>
<td>4.012</td>
<td>1.233</td>
<td>20.8</td>
</tr>
<tr>
<td>daytime</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>2.586</td>
<td>0.632</td>
<td>16.5</td>
<td>4.143</td>
<td>1.216</td>
<td>22.4</td>
</tr>
<tr>
<td></td>
<td>2.305</td>
<td>0.582</td>
<td>13.5</td>
<td>3.753</td>
<td>0.999</td>
<td>16.7</td>
</tr>
<tr>
<td></td>
<td>2.778</td>
<td>0.812</td>
<td>17.8</td>
<td>4.287</td>
<td>1.473</td>
<td>24.4</td>
</tr>
<tr>
<td>Weekday</td>
<td>2.467</td>
<td>0.655</td>
<td>16.6</td>
<td>3.745</td>
<td>0.987</td>
<td>24.8</td>
</tr>
<tr>
<td>night</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>2.425</td>
<td>0.617</td>
<td>16.9</td>
<td>3.784</td>
<td>0.956</td>
<td>25.1</td>
</tr>
<tr>
<td></td>
<td>2.165</td>
<td>0.536</td>
<td>14.1</td>
<td>3.522</td>
<td>0.748</td>
<td>19.5</td>
</tr>
<tr>
<td></td>
<td>2.651</td>
<td>0.823</td>
<td>18.4</td>
<td>3.986</td>
<td>1.221</td>
<td>29.2</td>
</tr>
<tr>
<td>Weekends</td>
<td>2.356</td>
<td>0.594</td>
<td>15.2</td>
<td>3.464</td>
<td>1.168</td>
<td>18.7</td>
</tr>
<tr>
<td></td>
<td>2.388</td>
<td>0.583</td>
<td>14.8</td>
<td>3.422</td>
<td>1.175</td>
<td>18.7</td>
</tr>
<tr>
<td></td>
<td>2.076</td>
<td>0.437</td>
<td>12.4</td>
<td>3.193</td>
<td>0.974</td>
<td>15.9</td>
</tr>
<tr>
<td></td>
<td>2.648</td>
<td>0.746</td>
<td>16.5</td>
<td>3.605</td>
<td>1.381</td>
<td>21.3</td>
</tr>
</tbody>
</table>

B. Results and Verification

The tables below show the results of identified load models for some of the buses.

### TABLE VIII

**CASE STUDY RESULTS OF SUBSTATION 6, CASE II**

### TABLE X

The load model derived from data recorded at substation 6 (Table VIII), is static polynomial load model, and that is because this substation is located at residential areas. In terms of accuracy, for example, for weekdays night load model, the mean square error between the simulated (average parameter values) and measured active power is \( 5.27 \times 10^{-6} \), and it is \( 5.03 \times 10^{-5} \) for reactive power. The mean square errors for most probable parameter values are \( 6.73 \times 10^{-6} \) and \( 9.64 \times 10^{-5} \) respectively. The time responses obtained in simulations with identified load model are plotted in Fig. 11 together with the measured responses. Fig. 12 is zoomed plots of active and reactive power around the voltage step change.

In Fig. 11 and Fig. 12, black solid line represent the original recorded responses, and the black dashed line the modified response after pre-processing of data as discussed in Section II. The green line with plus signs represents the fitted/simulated response for this event. The grey area represents the distribution of fitted/simulated responses with individual parameters obtained from each recorded event. Blue dotted and red dash-dot lines represent the response obtained with most probable and average values of those parameters. As it is shown in the figures, there is a very good match between the measured response and simulated response. The fitted/simulated responses with average parameter values are close to those with the most probable parameter values, and they have a good fit with the measured response.
The load model obtained from substation 1, (Table IX) is first order exponential dynamic load model. That is due to that it is close to the factory areas. The mean square error between the simulated (average parameter values) and measured active power is \(9.86 \times 10^{-6}\), and for reactive power it is \(6.14 \times 10^{-5}\). The graphical comparison of the measured response and fitted/simulated response are illustrated in Fig. 13, and Fig. 14. In the figure, the lines of different types and colors have the same meaning as in Fig. 11. It can be seen that both responses, using either average parameter values or the most probable parameter values, match very well the originally recorded load response.

The load model obtained from Case II (Table X) is composite load model. The mean square error between the simulated (average parameter values) and measured active power is \(4.13 \times 10^{-5}\), and for reactive power it is \(8.74 \times 10^{-5}\). Unlike previous two substations, where there are hundreds of events, for Case II, there are only several events. Therefore, only the comparison of measured response and fitted/simulated response of one event is shown in Fig. 15, and zoomed plots are shown in Fig. 16. In the figure, the blue solid lines, and green dashed lines, represent the measured responses and modified response respectively. The red dotted lines represent the fitted simulated response. The fitted simulated response still matches very well with the originally recorded response. The period and amplitude of the oscillation are basically the same. The times reaching the lowest value and steady state are accurate.

![Graphs showing comparison of responses](image)

**Fig. 13.** Comparison of estimated response and measured response for first order dynamic load model.

**Fig. 14.** The relevant parts of active power (left) and reactive power (right) responses from Figure 12 zoomed between 20s and 35s

**Fig. 15.** Comparison of estimated response and measured response for composite load model.

**Fig. 16.** Zooming of active power (left) and reactive power (right) responses after voltage step change.

### IV. CONCLUSIONS

Accurate load models are essential for reliable power system analysis and control. One of most frequently used approaches to load modelling and probably the most reliable is the measurement based method. It however, often involves significant time and resources and is strictly valid for load mix at monitored bus at the time of measurement. Thus, an automated load modelling application tool can significantly improve the efficiency of load modelling and consequently contribute to more reliable assessment of system dynamic performance.

The developed ALMT tool achieves full automation, from the data input to development of load models with corresponding parameters. The first stage of the tool focuses on data processing which involves identification of disturbances, data conditioning and classification. The second stage is load model selection, i.e., the identification of the most suitable load model for a recorded disturbance based on the shape of the dominant response. The third stage is iterative load model parameter fitting until required accuracy is achieved (measured by the similarity of recorded and simulated load responses).

The illustrative results presented in the paper demonstrate the ability of the developed ALMT to identify with appropriate accuracy and without any human intervention, load models for a range of routinely recorded disturbances in real power systems. The ALTM can be added as standard software application to existing power quality monitors or fault
recorders and such greatly enhance their functionality and contribute to effortless load modelling for power system steady state and dynamic studies.

REFERENCES


Yue Zhu (S’14) received a BEng degree in Electrical and Electronic Engineering from the University of Manchester, Manchester, U.K. and North China Electric Power University, Baoding, China, in 2015. He is currently working towards a Ph.D. degree at the University of Manchester.

Jovica V. Milanović (M’95–SM’98–F’10) received Dipl.Eng. and M.Sc. degrees from the University of Belgrade, Belgrade, Yugoslavia, a Ph.D. degree from the University of Newcastle, Newcastle, Australia, and a D.Sc. degree from the University of Manchester, Manchester, U.K., all in Electrical Engineering. He is currently a Professor of Electrical Power Engineering, Deputy Head of the School and Director of External Affairs in the School of Electrical and Electronic Engineering at the University of Manchester, Manchester, U.K., a Visiting Professor at the University of Novi Sad, Novi Sad, Serbia and University of Belgrade, Belgrade, Serbia and a Conjoint Professor at the University of Newcastle, Newcastle, Australia.