A Step-by-Step Data Processing Guideline for Load Model Development Based on Field Measurements

Document Version
Accepted author manuscript

Citation for published version (APA):
Hasan, K. N., Milanovic, J., Turner, P., & Turnham, V. (2015). A Step-by-Step Data Processing Guideline for Load Model Development Based on Field Measurements. In *host publication*

Published in:
host publication

A Step-by-Step Data Processing Guideline for Load Model Development Based on Field Measurements

Kazi N Hasan, Jovica V Milanović
School of Electrical and Electronic Engineering
The University of Manchester
Manchester, UK
kazi.hasan@manchester.ac.uk,
jovica.milanovic@manchester.ac.uk

Paul Turner, Victoria Turnham
Electricity North West Ltd.
Manchester, UK
Paul.Turner@enwl.co.uk,
Victoria.Turnham@enwl.co.uk

Abstract—**Load modelling attracts renewed interests these days in maintaining peak load conditions, supplying new types of loads and accommodating more renewable generation into electricity networks. This work describes real measurement data acquisition and step-by-step signal processing for developing aggregate load models at 11kV and 6.6kV level. Challenges in analyzing real measurement data are highlighted and issues to improve measurement-based load modelling are discussed. Load models at 15 substations from a UK distribution network are presented with subsequent model parameters. These load models will provide an insight to the operational flexibility, network resilience and management requirements of the measurement sites and related up-and-downstream substations.**

Index Terms-- **Distribution system, load-voltage relationship, measurement-based load model, operational flexibility.**

I. INTRODUCTION

Distribution network requires more operational flexibility these days than ever before for accommodating distributed energy resources and new (power electronic interfaced) types and proportion of loads. During high supply period of intermittent renewable energy integration, grid requires curtailed generation or further investment in the network. On the other hand, during peak load period and excessive capacity requirement, network may be inclined to perform load curtailment/management or to consider further investment in the system infrastructure. A solution to this problem could be having an estimation of the flexibility of loads following a 'voltage control operation' in the network which could respond to the increased generation and/or peak load situation. An assessment of the actual load response (including load model parameter values) may facilitate the implementation of an intended 'voltage control mechanism'. The load modelling studies reported in this paper are part of the research that aims to establish the feasibility of increasing the network capacity by controlling the distribution level loads (i.e. load voltage).

A good number of recent studies revolve around some application-oriented network-specific studies. The capability assessment of a generation-constrained (naval) power system

has been assessed through probabilistic representation of power system loads [1]. Conservation voltage reduction has been discussed by using a time-varying stochastic version of the exponential (static) load model [2]. Example of some other application-oriented load modeling studies are in the area of electric-vehicle (and other new-type of load) connection [3, 4], and accommodating high renewable energy integration [5, 6] into an electricity network.

Most of the existing load modelling studies concentrate on mathematical optimization/ parameter estimation on deterministic/ stochastic [1, 7, 8] or transfer function [9] or AI (artificial intelligence) based [10] approach. A stochastic representation of electrical loads has been presented though *pdf* (probability density function) of probabilistic uniform, normal and exponential distributions [1]. A deterministic parameter identification method has been discussed in [7] which reduces the model variance error in developing composite load model. Another composite load model has been developed in [8], with a special attention to the dynamic IM (Induction Motor) model. A time-dependent load model based on 2nd order transfer function has been developed in [9] by utilizing genetic algorithm as the optimization process. Some of the works highlight the load model selection aspects [7, 11, 12]. Static load models have been developed in [11] considering the polynomial, linear and exponential representation. 1st order dynamic model along with model parameters has been presented in [12]. Considering the contribution of both static and dynamic characteristics in aggregate (substation level) load, composite model has been discussed in [7]. This composite model considers combined ZIP (a combination of constant impedance (Z), constant current (I) and constant power (P) load model) and IM loads in the model.

Very few papers however, have reported data processing, filtering and smoothing techniques while processing filed measurements [13]. These issues may significantly affect the developed load models and subsequent parameters. Issues of handing data quality in terms of the required resolution, amount of noise, occurrences of spikes, preserving sharpness and smoothness etc. have been partly addressed in [14, 15].

This paper discusses data processing issues related to field recording to develop measurement-based load model. Actual load responses have been collected during the summer season at 15 (5 industrial, 5 domestic and 5 mixed loads) substations. In total there are 6×15=90 initiated voltage disturbances so far, (and more data will be analyzed from this ongoing work). There are significant variations in the quality of recorded data which have been collected from different substations at different time periods. They differ in the amount of noise, spikes, variability etc. These bring different types of challenges for pre-processing and filter design. The amount of noise is considerably different in different signals, i.e., in voltage, real and reactive power etc. Also, the amount of noise varies widely (even in the same signal, i.e., real power of different substations) while taking measurements from different substations. A wide variety of filtering parameters need to be examined while eliminating noise from all recorded measurements. Furthermore, spikes from the measurement signals also need to be eliminated through filtering. The occurrences and magnitude of spikes also differ in different signals. In order to handle the abovementioned issues appropriately, this paper presents a step-by-step process to handle filed-measurement recordings for developing static and dynamic load models highlighting difficulties in data processing, filtering and smoothing. The proposed methodology has been verified through actual load data analysis for diverse load compositions present in a UK distribution system as a part of CLASS (customer load active system services) project [16].

II. STEP-BY-STEP LOAD MODELLING METHODOLOGY

A step-by-step measurement based load modelling approach, presented in Fig. 1 and adopted in this study involves 4 steps. The steps involved and performed studies within each step are described below.

- *Data collection and processing* – Actual measurement data is collected from the DNO (distribution network operator) server. The sampling rate of the data is 1second. These files are imported to the MATLAB where filtering and further processing of the data are carried out.

- *Data filtering* – A range of filters has been tested to filter recorded data as they vary over a wide range in terms of the amount of noise, occurrences of spikes, preserving sharpness and smoothness etc. Performances of different filters have been compared. Design parameters of different filters have been decided to obtain optimum filter performance.

- *Load model selection* – Extensive literature review has identified traditional and advanced load modelling techniques which have been used in measurement-based load modelling. The applicability of appropriate load models for recorded data has been investigated.

- *Parameter estimation and validation* – For a selected model and/or a set of models, corresponding parameters have been estimated. Parameter values have been modified through optimization techniques by comparing with the actually measured data. Parameters have been tested against the historical values for validation.
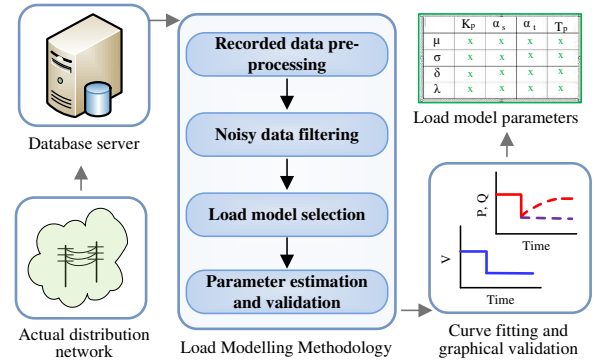

Figure 1. Step-by-step workflow for data acquisition, load modelling and parameter estimation.
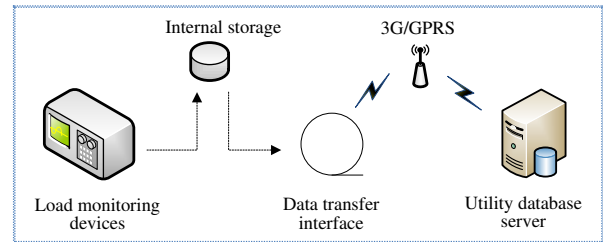

Figure 2. A schematic architecture of load-data acquisition system [14, 15].

III. DATA ACQUISITION, PRE-PROCESSING AND FILTERING

Accuracy of a measurement-based load model is heavily dependent on the quality of data, its pre-processing and filtering. This Section discusses some challenges associated with the measured data, data processing and filtering process.

A. Data Acquisition System

Fig. 2 shows a schematic representation of the data acquisition system. The load measuring devices are installed at the primary substations. Each measuring unit records the load parameters (i.e. time, voltage, real and reactive power and frequency) of each transformer at 1s resolution. Each unit has 4GB of internal storage. It transfers data to 'interface' device for further processing and updating to the database server. The data transfer communication interface could be either 3G mobile network or GPRS (General Packet Radio Service). Interface devices capture/sample data at a rate of 5 seconds. (Sampling rates though might go to 1 second.) The updates to server are every hour. As shown in Fig. 2, actual measured (sampled) data has been processed at least at three stages before it has been passed for load modelling. In this series cascaded scheme, the final data resolution could be limited by the slowest device connected to the data acquisition system.

B. Extracting Required Information

Extracting usable voltage/power signals from the recorded data stream is an important aspect of load modelling. For instance, Fig. 3 identifies 6 voltage/power windows from a snapshot of actually measured data. Window 1 & 6 are not usable as these are not capturing a real power response appropriately following a voltage disturbance. The same is true for windows 2 & 5. Windows 3 & 4 can be used for modelling steady-state and/or dynamic responses, as these windows capture consistent and realistic load behaviors.
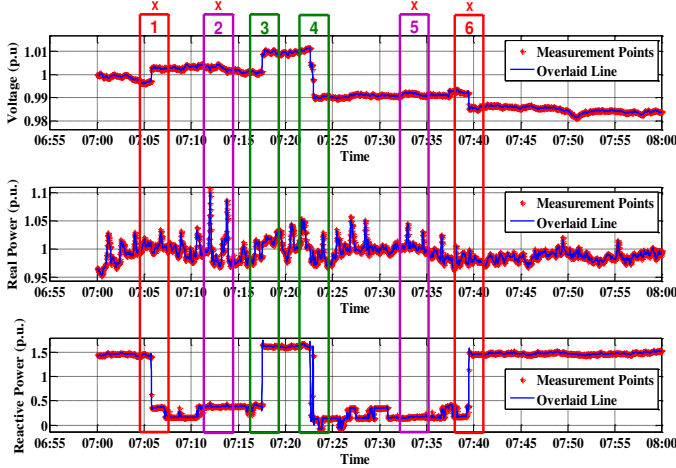
Figure 3.   Extracting required (initiated/unintended load/system) response and static/dynamic information from real measurement data.
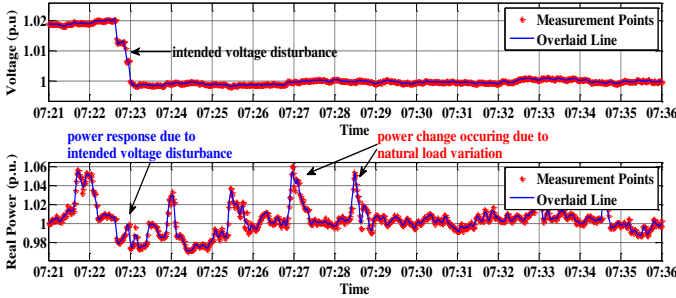

Figure 4.   Identification of power response due to voltage change or natural load variation.
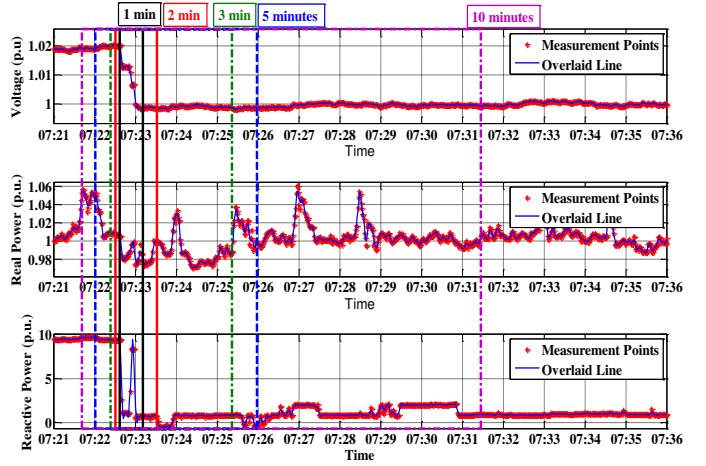

Figure 5.   Influences of the window length on capturing load response following an intended voltage disturbance.

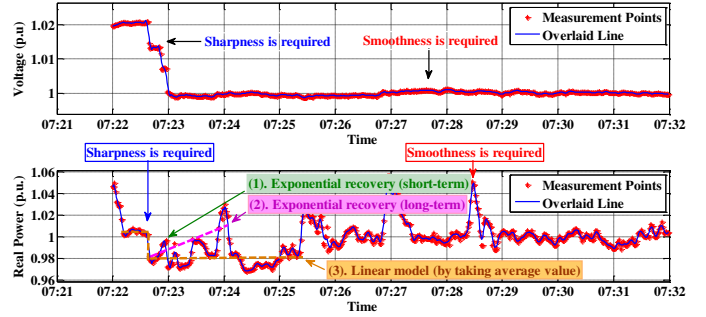Figure 6.   Significance of load recovery characteristics and preserving required smoothness and sharpness.

C. Spontaneous Load Changes

It can be observed from Fig. 4, that there are some abrupt changes in power when the voltage is almost stationary. This is the effect of some loads being connected /disconnected from the system. This may also happen at the instant of initiated voltage disturbance [12]. In some cases the power change is even higher than that which is caused by a voltage change. These unintended power changes need to be eliminated to enhance the accuracy of load model parameters.

D. Influence of Data Window Size

Fig. 5 shows different sizes of windows used for data capture. Load model parameters could be different for a short-or-long-term observation. It typically takes about 30 minutes of data to determine the steady-state characteristics while it takes anywhere up to 30 seconds to determine transient characteristics [11]. This time can be significantly shorter and of the order of a few seconds in case of single type of load connected at a certain bus [10, 17].

E. Significance of Load Recovery Characteristrics

The nature and required time for a load recovery contributes highly to the appropriateness of a developed load model. For example, induction motor exhibits a short-term recovery which is only few seconds. Tap-changers and voltage controllers show a mid-term recovery which is in the range of tens of seconds to few minutes. Resistive load shows a long-term recovery in the range of few/some minutes [10, 17].

Fig. 6 presents 3 modelling scenarios after a step-change in power. As shown in Fig. 6, a quick recovery (in the range of 10 second) is observed after the power drop, point (1); if the window size is taken longer there is another recovery pattern which is in the range of 100 second, point (2); by ignoring the abrupt power changes, the power recovery can be linearized after the step-change, point (3). The ultimate challenge here is to distinguish the natural load change from the intended load change and estimating/identifying appropriate window/load behavior.

F. Impact of Filtering Techniques

Several filtering techniques have been investigated in this research, such as the moving average (MA), Savitzky-Golay (SG), and robust local regression (RLR) techniques. A comparison of the performances among these filters has been investigated and presented later in the paper.

Moving average filtering: The moving average (MA) filtering method uses the procedure of replacing an individual sample with the average of the neighboring data points over a specified data span. The mathematical representation of the MA filter can be expressed as below:

$$y_s(i) = \frac{1}{2N+1}\left(y(i+N) + y(i+N-1) + \dots + y(i-N)\right) \quad (1)$$

where, $y_s(i)$ is the averaged value of the i-th data point, N is the data points on either side of the $y_s(i)$ and $2N+1$ is the

data span over which the value has been averaged. Data span is the design parameter of this filter.

Savitzky-Golay filtering: The Savitzky-Golay filter is designed based on the least-square polynomial approximation. This filter is also known as digital smoothing polynomial filter or least-square smoothing filter. This filter can achieve high level of smoothing without significantly distorting the data features. The polynomial order and frame size are the design criteria of this filter.

Robust local regression: This technique removes the outliers and contains the trend of the data points. Outliers cannot distort the original shape of the signal. This procedure calculates the regression weights for each data points in the selected window, as given by the following formula:

$$w_i = \left(1 - \left|\frac{x - x_i}{d(x)}\right|^3\right)^3 \qquad (2)$$

where, x is the predictor value linked with the data point to be smoothed, x_i is the nearest neighbors of x, and $d(x)$ is the distance between x to the most distant predictor value. The window size is the design parameter of the robust local regression (RLR) filter.

G. Comparison of Filtering Techniques

It has been observed from the discussions of Section III.F that 'filtering order' and 'span/window size' are the design criteria for different filters. A high order filter (and small window size) retains originality of the signal and preserve sharpness. On the other hand, a low order filter (and large window size) flattens out the natural changes in the signal. Three filtering techniques are implemented for the purpose of comparison. Fig. 7 shows a comparative snapshot of three filtering techniques. The design parameters of the filters are as follows: A 3rd order SG is with a frame size of 119. The MA is with an averaging data span of 25. The RLR is with an averaging window size of 35.

All three methods can filter curves with similar major trends but with different accuracies (Fig. 7). The MA method captures sharp peaks and the width and height of peaks can be better preserved. While optimizing for required smoothness, sharpness of the signal has been lost in some cases. Signals filtered through the S-G technique are smooth enough. However, the sharpness cannot be retained accurately. The RLR filter captures sharpness of the signal accurately, which is an expected criterion for retaining step-changes. RLR can optimize the signal for preserving both sharpness and smoothness, compared to two other methods.

H. Smoothing Signals to Eliminate Spontaneous Load Changes

As shown in Fig. 8, there are very high variations in the real power data after filtering. The spontaneous load changes cannot be eliminated through filters. These natural variations have been smoothed by taking an average value of the data points. The smoothed curves are now ready for load model parameter estimation.
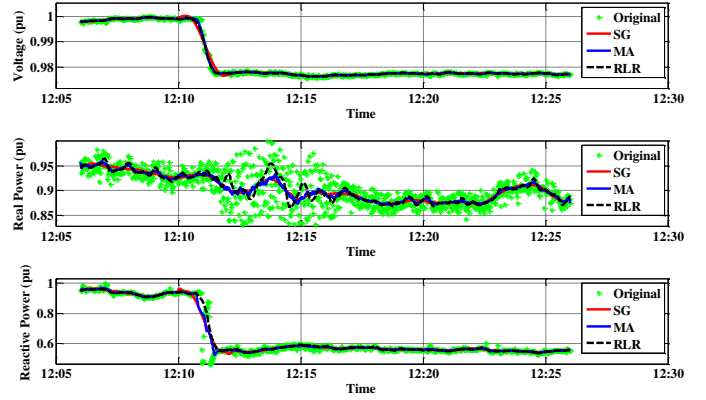

Figure 7.   Comparison of Savitzky-Golay, moving average and robust local regression filter.
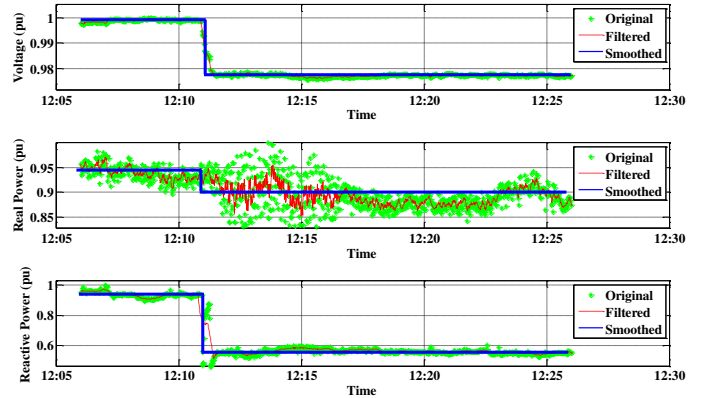

Figure 8.   Smoothening of filtered signal.

IV.   LOAD MODEL SELECTION AND ESTIMATION OF MODEL PARAMETERS

Load modelling and parameter identification process have been adopted from the literature [11, 12, 18, 19]. The process involves finding a relationship between the real (and reactive) power with the voltage. Among different load models, candidate models have been selected based on the study purpose and according to the suitability of the available recorded measurements.

A. Static Exponential Load Model

One of the most frequently used load models, which is the exponential model can be represented as follows,

$$P = P_0\left(\frac{V}{V_0}\right)^{K_p}, \quad Q = Q_0\left(\frac{V}{V_0}\right)^{K_q} \qquad (3)$$

The K_p and K_q represents the voltage exponents of real and reactive power, respectively, for a static exponential load model [11]. P, V and P_o, V_o are power and voltage after and before the disturbance, respectively.

B. Dynamic Exponential Load Model

A dynamic load model with exponential recovery, as shown in Fig. 9, is presented by the following equation [17],

$$P = P_0 \times \left[V^{\alpha_s} - V^{\alpha_t} \right] \left(e^{-\frac{t}{T_p}} \right), Q = Q_0 \times \left[V^{\beta_s} - V^{\beta_t} \right] \left(e^{-\frac{t}{T_q}} \right) \quad (4)$$

The parameters are for real and reactive power, respectively, where the specification is as follows: the steady-state voltage exponents (α_s, β_s), transient voltage exponents (α_t, β_t) and load recovery time constants (T_p, T_q).

C. Parameter Estimation – Curve Fitting

Considering the general applicability and simplicity of the model, a least square optimization procedure has been chosen for parameter estimation as shown in Fig. 10. Detail discussion in this aspect can be found in [14, 18]. The optimization procedure relates the measurement data with a mathematical model. The mathematical model represents a standard curve, where measurement points are fitted to align with the curve. This approach is also known as the curve fitting technique. Curve fitting is the process of finding a mathematical form of relationship that most closely indicates relationship between dependent and independent variables. The "curve" defined by the mathematical equation is said to "fit" the observed data. The process of curve fitting is also called "regression". Once the regression equation is obtained, it can be used to predict the output variable.

D. Goodness-of-Fit

The mean absolute percentage error (MAPE) given by (5) is a measure of accuracy of a fitted model,

$$MAPE = \frac{100\%}{N} \sum_{n=1}^{N} \left| \frac{V_n - F_n}{V_n} \right| \quad (5)$$

where, V_n is the actual value, F_n is the fitted value and N is the number of fitted points.

E. Outliers Rejection

The robustness of load model parameters have been assessed through an analytical method as presented in Fig. 11. Estimated values have been compared with historical values from the literature to validate the accuracy and precision of parameters [11, 12, 19]. In the validation procedure, this study implements the approach of 'high precision' at 'high accuracy', the red box in Fig. 11. Accuracy of the parameters is determined by considering the differences of obtained value from the historical value. However, as more trials might be conducted in the future, a more reliable statistical distribution for load model parameters can be obtained.

F. Ranges of Load Model Parameters

Table I presents the static exponential load model parameters where voltage exponents of real power varies between 0.79 and 1.62 and for reactive power varies between 2.69 and 5.73. The exponent values resemble the step-change in power due to a step-change in voltage signal. This power-voltage relationship for different substations at different times could be used as a benchmark for "voltage-controlled demand management" mechanism of the CLASS project [16].
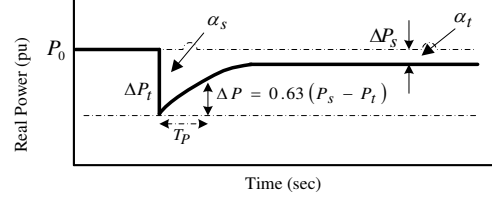

Figure 9. Load recovery following a voltage step-change, which represents the dynamic response of a load.
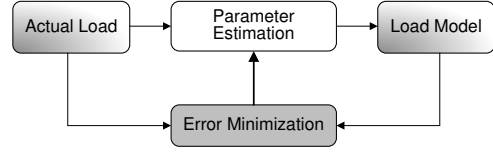

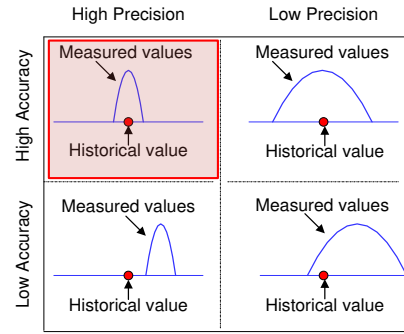Figure 10. Flowchart of the parameter estimation technique [18].

Figure 11. An analytical procedure to remove outliers [14].

TABLE I.   STATIC EXPONENTIAL LOAD MODEL PARAMETERS

Customer types / Substations	K_p		Kq	
	Avg., μ	Std. Dev., σ	Avg., μ	Std. Dev., σ
5 Largely industrial and commercial	1.62	1.34	4.87	2.50
5 Largely Domestic	0.79	0.56	2.69	1.99
5 Mixed	1.08	0.91	5.73	2.36
15 Substations	1.16	0.94	4.43	2.28

Table II shows the dynamic exponential load model parameters. The (real power) steady-state voltage exponent (α_s) varies between 0.66 and 1.10, transient voltage exponent (α_t) varies between 1.80 and 2.99, and load recovery time constant (T_p) varies between 78s and 186s. The reactive power steady-state voltage exponent (β_s) varies between 5.62 and 13.19, transient voltage exponent (β_t) varies between 14.07 and 24.36, and load recovery time constant (T_q) varies between 92s and 184s. Physical significance of dynamic load model parameters can be understood from Fig. 9, where α_s characterizes the load recovery up to a steady state value, α_t relates the transient nature of power change following an abrupt voltage change and the recovery time constant, T_p exhibits the pattern and speed of load recovery following a disturbance.

TABLE II. DYNAMIC EXPONENTIAL LOAD MODEL PARAMETERS

		Ind.-Com.	Domestic	Mixed	All
α_s	μ	1.10	0.90	0.66	0.89
	σ	0.85	0.52	0.62	0.66
α_t	μ	2.99	2.62	1.80	2.47
	σ	2.22	1.60	1.68	1.83
β_s	μ	13.19	6.41	5.62	8.41
	σ	17.75	3.95	3.87	8.53
β_t	μ	24.36	17.35	14.07	18.59
	σ	25.44	10.68	9.00	15.04
T_p (s)	μ	77.55	169.67	186.01	144.41
	σ	54.65	52.78	68.96	58.80
T_q (s)	μ	92.00	184.23	165.78	146.50
	σ	91.28	79.24	76.55	82.36

V. CONCLUSIONS

This paper discusses inherent challenges and possible solutions of processing filed measurement data for developing aggregate load models at distribution network buses. The analysis highlights the underlying issues of processing recorded data for load modelling and modification of the evaluation procedure at different stages. Different types of filters, design parameter(s) of a filter, optimum window selection, smoothing techniques, optimization, and validation procedures are the necessary steps of measurement-based load modelling. A step-by-step data processing guideline is provided for accurate development of load models. Accuracy of a measurement-based load model heavily depends on the quality of the data. This paper highlights the data quality issues such as extracting usable information, selecting required window, avoiding spontaneous load changes, identifying load recovery characteristics etc. Discussions on the filter design parameters are also presented. An analytical procedure (to reject the outliers from the estimated parameter sets) implemented in this study is based on "accuracy and precision" as it is applicable for cases where there is a lack of data for obtaining a statistical distribution.

Developed methodologies have been applied in a realistic, modern distribution network. Obtained load model parameters give an indication of the change in real and reactive power with respect to a change in network voltage. This power-voltage relationship could be used as a benchmark for "voltage-controlled demand management" mechanism in a modern distribution network.

REFERENCES

[1] U. Orji, B. Sievenpiper, K. Gerhard, S. B. Leeb, N. Doerry, J. L. Kirtley, and T. McCoy, "Load Modeling For Power System Requirement and Capability Assessment," *IEEE Transactions on Power Systems,*, vol. PP, pp. 1-9, 2014.

[2] W. Zhaoyu and W. Jianhui, "Time-Varying Stochastic Assessment of Conservation Voltage Reduction Based on Load Modeling," *IEEE Transactions on Power Systems,* vol. 29, pp. 2321-2328, 2014.

[3] S. Yang, M. Wu, X. Yao, and J. Jiang, "Load Modeling and Identification Based on Ant Colony Algorithms for EV Charging Stations," *IEEE Transactions on Power Systems,* vol. PP, pp. 1-7, 2014.

[4] D. Qian, C. Tao, D. Shanxu, and Z. Feng, "Stochastic Modeling and Forecasting of Load Demand for Electric Bus Battery-Swap Station," *IEEE Transactions on Power Delivery,* vol. 29, pp. 1909-1917, 2014.

[5] S. Wogrin, P. Duenas, A. Delgadillo, and J. Reneses, "A New Approach to Model Load Levels in Electric Power Systems With High Renewable Penetration," *Power Systems, IEEE Transactions on,* vol. 29, pp. 2210-2218, 2014.

[6] L. Yingliang and P. J. Wolfs, "A Hybrid Model for Residential Loads in a Distribution System With High PV Penetration," *IEEE Transactions on Power Systems,,* vol. 28, pp. 3372-3379, 2013.

[7] H. Renmu, M. Jin, and D. J. Hill, "Composite load modeling via measurement approach," *IEEE Trans on Power Systems,* vol. 21, pp. 663-672, 2006.

[8] L. Pereira, D. N. Kosterev, P. Mackin, D. Davies, J. Undrill, and Z. Wenchun, "An interim dynamic induction motor model for stability studies in the WSCC," *IEEE Transactions on Power Systems,* vol. 17, pp. 1108-1115, 2002.

[9] I. F. Visconti, D. A. Lima, J. M. C. de Sousa Costa, and N. Rabello de B.C.Sobrinho, "Measurement-Based Load Modeling Using Transfer Functions for Dynamic Simulations," *IEEE Trans on Power Systems,* vol. 29, pp. 111-120, 2014.

[10] I. R. Navarro, "Dynamic Load Models for Power Systems," PhD Thesis, Lund University, Sweden, 2002.

[11] L. M. Korunovic, D. P. Stojanovic, and J. V. Milanovic, "Identification of static load characteristics based on measurements in medium-voltage distribution network," *IET Gen, Trans & Distr,* vol. 2, pp. 227-234, 2008.

[12] D. P. Stojanović, L. M. Korunović, and J. V. Milanović, "Dynamic load modelling based on measurements in medium voltage distribution network," *Electric Power Systems Research,* vol. 78, pp. 228-238, 2008.

[13] CIGRE_WorkingGroup_C4.605, "Chapter 4 Recommended Methodologies for Load Model Development. in *Modelling and Aggregation of Loads in Flexible Power Networks*," 2013.

[14] K. N. Hasan and J. V. Milanovic, "Interim Profile Modelling Study " The University of Manchester & Electricity North West Limited, UK, 2015, [available online]: http://www.enwl.co.uk/docs/default-source/class-documents/university-of-manchester-interim-report-wp1.pdf?sfvrsn=4.

[15] K. N. Hasan and J. V. Milanović, "Practical aspects of developing load models at distribution network buses based on field measurements," in *CIRED International Conference and Exhibition on Eelctricity Distribution, (submitted),* Lyon, France, 2015.

[16] Electricity_North_West_Ltd_UK. (2015). *CLASS (Customer Load Active System Services).* Available: http://www.enwl.co.uk/class#

[17] D. J. Hill, "Nonlinear dynamic load models with recovery for voltage stability studies," *IEEE Trans on Power Systems,* vol. 8, pp. 166-176, 1993.

[18] EPRI, "Measurement-based load modeling. [available online]: http://www.epri.com/abstracts/Pages/ProductAbstract.aspx?ProductId=000000000001014402," 2006.

[19] CIGRE_WG_C4.605, "Modelling and aggregation of loads in flexible power networks, Jovica V. Milanović, (Convenor), Julija Matevosiyan, Anish Gaikwad, Alberto Borghetti, Saša Ž. Djokić, Zhao Yang Dong, Andrew Halley, Lidija M. Korunović, Sergio Martinez Villanueva, Jin Ma, Pouyan Pourbeik, Fernanda Resende, Stefan Sterpu, Fortunato Villella, Koji Yamashita, Odin Auer, Karim Karoui, Dimitry Kosterev, Shu Kwan Leung, Dumisani Mtolo, Samila Mat Zali, Adam Collin, Yizheng Xu, ISBN: 978-2-85873-261-6," 2014.