Network Reduction Based on Propagation of Large Disturbances for Load Model Validation

Siming Guo, *Student Member, IEEE*, Komal S. Shetye, *Member, IEEE*, and Thomas J. Overbye, *Fellow, IEEE*

***Abstract*—Wide-area load model validation using disturbances is difficult because the large number of load buses makes the problem intractable. One method to mitigate this problem is to reduce the number of buses on which dynamic models are implemented. In this paper, we define the region of influence (ROI) as those buses which are most impacted by the disturbance, and we propose using the ROI to determine the internal and external buses of the analysis. The analysis shows that a large percentage of the system is minimally impacted by the disturbance when only the voltage is considered, unlike frequency, which is a global system phenomenon. Such a voltage effect based equivalent is assumed to be adequate for load model validation due to the close coupling between load and voltage dynamics. By setting a threshold and only placing dynamic load models on the ROI region within, the number of differential equations can be reduced significantly. Finally, we investigate how the parameter estimation is affected by using this reduced load model.**

I. Introduction

Transient stability model validation involves ensuring that simulations of power system dynamic models adequately represent real-system response. Following certain outages and oscillations, for example in the Western Electricity Coordinating Council (WECC), post-event analyses showed incorrect model behavior and predictions; subsequently, significant work was done in model validation [1], [2]. Key improvements were done in generator model validation using point of interconnection PMU measurements, coupled with single machine infinite bus simulations. Apart from validating models with data, work has also been done to cross-validate models across different software packages and implementations [3]. Recently, the focus has been on load models with newer, comprehensive dynamic models being proposed [4]. Load model validation is also commonly conducted at the device and aggregate level, for transmission system studies [5]. Several papers describe parameter estimation techniques for load models [6], [7]. Unlike generator dynamic parameters that do not change drastically in a short period, load dynamics are much more transient in nature, and hence need more frequent validation or parameter calibration.

Currently, the vast majority of measurement-based parameter estimation research is performed with small systems or at a single bus [7]–[11]. However, owing to the deployment of wide-area measurement systems there is a wealth of high-fidelity, reasonably accurate data available now which can be used for this application. This data forms the cornerstone of system-level validation [2], which aims to match global system response, rather than that of individual devices. This is a challenging problem. A dynamic model for a load can have anywhere from 8 to 135 parameters [12], [4]. In [7]–[11], the authors do not need to address this issue because even with a complicated load model, the small number of buses means that the computational costs are insignificant. However, in a typical interconnect-level system of tens of thousands of buses, this could mean having to deal with an order of 10^6 parameters for estimation.

To address this, this paper investigates the concept of network reduction to perform parameter estimation for load models. It divides a large, interconnect-level system into different regions based on the propagation of voltage disturbance. These layered regions can be used to define "internal" and "external" parts of the network, such that load modeling and estimation is performed only in the internal part. This is aimed at reducing the computational requirements. The effects of choosing different sizes of the "internal" region on the validation results is also studied. The network is "reduced" in terms of the number of dynamic loads and their parameters considered in the estimation problem.

The methodology and analysis are demonstrated using only transmission system-level PMU measurements. This poses an additional challenge to load modeling, which is usually performed at lower voltages. We utilize actual PMU measurements from the high voltage transmission network of a North American electric utility. A system-wide disturbance event and its associated data are used in this paper to demonstrate the algorithms and results. A key goal of this reduction is to enable the automation of the validation process, yet keep it computationally tractable.

Section II introduces the system under study and the three-phase induction motor model used in this paper. Section III explains the concept and calculation of the "region of influence" which aims to narrow the scope of load modeling, for faster calculations. Section IV investigates the performance of parameter estimation using the reduced load model defined by the region of influence. Finally, Section V summarizes the key findings of the paper and proposes the next steps in this research.

II. Background

A. Wide Area System Model

The system considered here consists of around 13,000 buses and 3000 generators, and is henceforth called the "System". To recreate the system state as it was just before the disturbance occurred, a state estimator (SE) snapshot of the system was obtained, rather than using an offline (i.e. planning) case. Dynamic models for generators were mapped to this SE case from the offline case. Due to the changing nature of loads,

 This paper was presented at the IEEE PowerTech, Manchester, UK, June 2017.

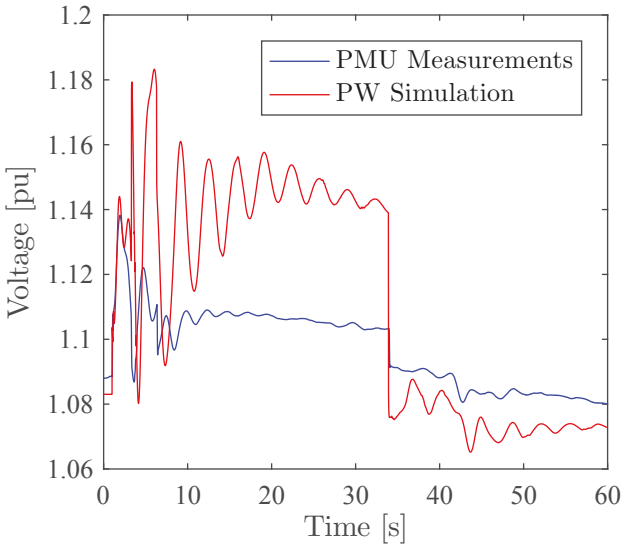

Fig. 1. PMU recorded data compared to simulations at a bus.

the mappings for loads are currently a work in progress in the industry. Hence a 20%, 15-parameter induction motor model assumption was made at each load, based on past industry assumptions [12].

The disturbance started with a single-phase fault at one circuit of a 500 kV transmission line, hereafter referred to as Line AB, in the system, while the other circuit was out of service. This is simulated at $t = 1$ second using PowerWorld (PW), and shown in Fig. 1. This fault was cleared by opening of the line. This led to the remedial action scheme in the system to operate, dropping several generators totaling 2500 MW, followed by the insertion of a 1400 MW breaking resistor, all within 0.6 seconds of the opening of the line. Between $t = 3$ and 6.3 seconds, there were more generator drops and shunt insertions. Following that, there were no more events until $t = 33.9$ seconds, when more shunts were inserted and one more generation drop event occurred at $t = 42$ seconds. The simulation was run until $t = 60$ seconds, and the results were compared to the PMU data with the corresponding event time stamps.

The said PMU data was received from a particular entity within the Interconnect System, hereafter referred to as the "Utility". Bus voltage and frequency measurements at 45 high-voltage locations that lie entirely within the Utility were received, and no measurements in other parts of the System were available for this analysis.

B. Three-Phase Induction Motor Model

The three-phase induction motor (TPIM) load model used in this paper is commonly used in industry for dynamic studies. For instance, in the WECC, power system studies were done with 20% of the load modeled as TPIM, and the remaining as static load, up until 2014 [13]. This model represents an aggregation of several motors dispersed through a load represented at a high voltage bus [14], thus making it appropriate for validation with the transmission-level PMU data considered in this paper. It is not intended for modeling the characteristics of individual motors. This model typically has six electrical parameters: synchronous reactance (L_s), transient reactance (L'), stator resistance (R_a), transient time constant (T_0'), sub-transient reactance (L''), and sub-transient time constant (T_0''). It can represent either a small or a large motor, based on certain values assigned to its parameters, as shown in Table I derived from [13].

TABLE I
TYPICAL ELECTRICAL DATA FOR THE TPIM MODEL

PARAMETER	SMALL MOTOR	LARGE MOTOR
R_a	$0.03 - 0.04$	< 0.01
L_s	$1.8 - 3.0$	$2.5 - 5.5$
L'	$0.15 - 0.18$	$0.18 - 0.20$
L''	$0.12 - 0.15$	$0.15 - 0.18$
T_0'	$0.12 - 0.20$	$0.8 - 1.8$
T_0''	$0.0024 - 0.003$	$0.003 - 0.005$

III. REGION OF INFLUENCE

The goal of this paper is to provide a means to find a subset of the system where we need to perform load model validation. This will result in a faster transient stability calculation, which means a shorter computation time will be required for each iteration of the parameter optimization. This is found by quantifying how various regions in the footprint of the system are "excited" in different degrees by the disturbance. Coupling this with geographic visualization, it is possible to divide the large interconnected power system into several "internal" and "external" regions. This is somewhat similar to the concept of partitioning the network in creating equivalents. An internal region in this case indicates an area in which the load buses are assigned the TPIM model for performing parameter estimation. The loads in the rest of the system, i.e. the "external" part, are represented using a constant impedance model. The idea is to show how the sizes of these internal regions, i.e. the load buses considered for estimation in that zone, impact the validation results. The region also proves to be an effective visualization tool for disturbance propagation.

A. Approach

The ROI seeks to capture how far the disturbance propagates on the System. Since the 45 PMU measurements are all within the Utility, we cannot determine the ROI for the interconnect from the measurements. Thus, we use the simulation to determine the ROI, since simulation results are available for the entire System. To measure the impact of the disturbance, we perform the transient stability simulation with an end time of 60 s. This time span is sufficient to capture all the events during the fault and during the recovery. Then, for each bus, we find the difference between its transient stability result and the average transient stability result across all buses at each time step, and use this to calculate the root mean square error (RMSE). The reason we use the average of

the system instead of simply the pre-fault values is because the goal is to find those buses which are most important to retain in the internal system. A bus with an average transient result is simply swinging with the rest of the system and not contributing significant dynamics of its own, and can thus be equivalenced. The signals studied are the same as the PMU measurements available, i.e. bus voltage magnitude and frequency. Mathematically, the RMSE for bus i that we use for ROI is defined as follows:

$$\underset{\text{ROI}}{\text{RMSE}}(V_i) = \sqrt{\frac{1}{T}\sum_{\tau=1}^{T}\left(V_i[\tau] - \frac{1}{I}\sum_{i=1}^{I} V_i[\tau]\right)^2} \quad (1)$$

$$\underset{\text{ROI}}{\text{RMSE}}(F_i) = \sqrt{\frac{1}{T}\sum_{\tau=1}^{T}\left(F_i[\tau] - \frac{1}{I}\sum_{i=1}^{I} F_i[\tau]\right)^2} \quad (2)$$

where V_i and F_i are the voltage and frequency simulation results for bus i, τ is the time step (note that $\tau = 1$ is equivalent to $t = 0$), T is the total number of time steps, and I is the total number of buses.

To measure the size of the ROI, a distance measure must be defined. In this paper, we use geographical distance. Ideally, the geographical distance could be easily calculated from the geographical coordinates (latitude/longitude) of the buses. However, there are two obstacles. The first is that geographical coordinates for many of the buses are not available as this data is not a standard part of either offline or SE cases in industry. Of the 12853 buses in service, we do however have the coordinates for 3025 buses. Thus, we will have to perform the analysis with approximately a quarter of the total number of buses. However, 3025 buses spread evenly throughout the System should still provide more than sufficient resolution for finding the ROI, and the results confirm this assumption.

The second obstacle is that the coordinates for some of the buses are incorrect. The reason for this is that the geography data was mapped to the SE case from the planning case, as geography data is not a part of the SE case. However, the two cases typically use starkly different bus numbers, and somewhat different bus names, to refer to the same physical bus in the system. As a result, the coordinates for some similarly named buses may be mapped incorrectly. To overcome this, an algorithm was used to detect suspicious geography data. The premise of this algorithm is that a bus with coordinates that are very distant from its neighbor buses most likely had incorrect coordinates. Thus, buses with geographically distance neighbor buses were deemed suspicious, and their coordinates were discarded.

B. Results

Figs. 2 and 3 show contours of the log of the RMSE values, based on geographical coordinates. The black cross is the location of the faulted line. We can see that the ROI for voltage is close to the fault, as expected. From this figure, it would be possible to define an ROI by simply selecting

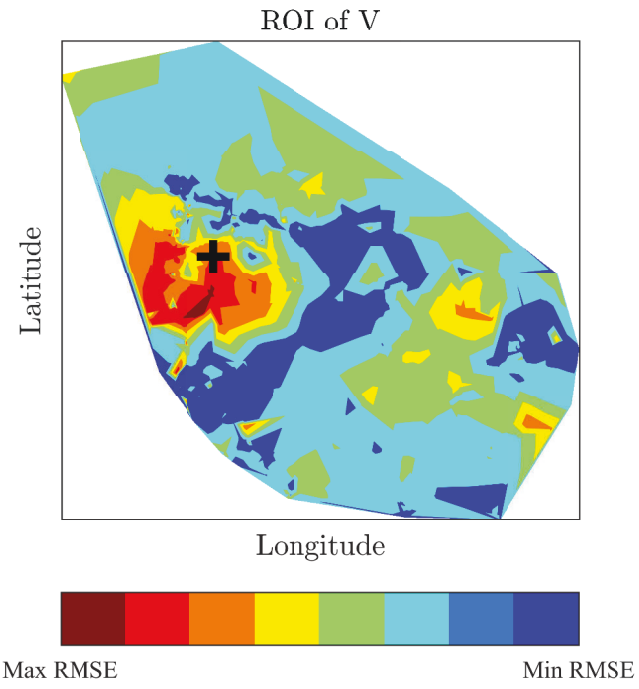

Fig. 2. The geographical region of influence for voltage. Red indicates higher RMSE.

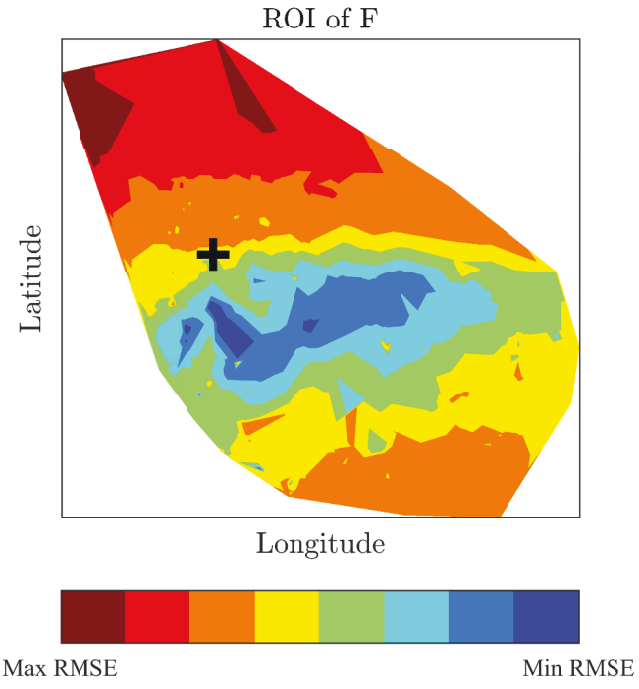

Fig. 3. The geographical region of influence for frequency. Red indicates higher RMSE.

a color and including the buses within that region. The ROI for frequency, on the other hand, provides much less insight. Because frequency is a system-wide phenomenon, all buses are affected regardless of where the fault occurs. Fig. 3 shows that the largest deviations of frequency occur at the extremities of the system, while the center shows relatively little deviation, essentially because the extremities are farther away from the sources of inertia on the system.

C. Reducing the Load Model Based on ROI

Based on the results above, we conclude that the geographical distance-based voltage ROI is the best candidate for defining the internal and external systems of the equivalent. This is an expected result, given that voltage effects tend to be localized. From a load model validation perspective, this is desirable since the newer models such as the WECC composite load model CMPLDW that are in need of validation are sensitive to voltage, and were actually introduced to better represent voltage swings in the system.

This voltage ROI is analogous to voltage control areas [15], in that they are a group of buses in a geographically compact area having similar voltage changes for disturbances [16]. VCAs can be found by methods such as Jacobian sensitivities [15], [17] but they are not suitable for topology changes, and certainly not for large disturbances. Reference [18] also proposes a method to partition the system based purely on graph theory, without accounting for any sensitivities, to eventually provide secondary voltage control and prevent disturbance propagation. In [19] dynamic VCAs are found for transient contingencies. Each VCA is composed of contingency clusters, and the most influential buses for dynamic var injections. The first step in this process is to find buses that are most impacted by the contingencies, i.e. choosing a voltage divergence criteria, followed by clustering. Our paper does not aim to find such voltage control areas, but it does start from the same point of finding the most impacted buses to eventually find these areas, or what we refer to here as "regions".

In this section, we partition the buses into internal and external systems based on the ROI. However, as we noted in Section III-A, we only have the geographical coordinates of approximately one quarter of the buses. For a bus which does not have coordinates, we must estimate its location to determine if it lies within the ROI. To accomplish this, we make use of force-directed graph drawing techniques. Force-directed graph drawing attempts to lay out the nodes and edges of a graph in an aesthetically pleasing way. It typically accomplishes this by modeling the edges of the graph as attractive springs, and the nodes of the graph as repulsive electric charges [20]. At equilibrium of the N-body simulation, the forces balance, resulting in a graph where connected nodes are close together, but all nodes are fairly evenly spaced. The motivation for using this technique comes from [21], which essentially found that in the real power system, nodes are typically connected in a lattice structure as opposed to a radial structure. The technique has also been used frequently for power system visualization [22], [23].

IV. Parameter Estimation Using the Reduced Load Model

In Section III-B, we calculated the ROI contour based on the geographical coordinates of the buses we had coordinates of. Then, in Section III-C, we estimated the coordinates of the other buses so that we could determine the internal and external systems given the selection of a contour level. Finally, in this section, we look at how the performance of parameter estimation is affected by the selection of the contour level for load modeling. In Fig. 4, we plot the cumulative distribution of the RMSE of the buses. This graph can be interpreted as follows: if we select the brown region, we retain 0.3% of all buses; if we select the red region, we retain 6.6% of all buses, including those in the brown region above it; if we select dark blue, we retain all buses in all colors. We can see that as we choose larger contours, many more buses will be included. The question is whether we can achieve relatively good parameter estimation results with a small subset of the buses. For example: Can we place load models only on buses in the orange contour and still achieve a good match between the simulated results and the measurements? If so, this will reduce the number of differential equations in the simulation, thus reducing the computation time required for the simulation in each iteration.

In this analysis, we focused on using parameter estimation to find the best fit between measurements and simulation at one specific bus out of the 45 total buses with measurements, and for only the first 3 seconds. The reason we make this simplification is that, through empirical testing, the 14 parameters in TPIM are only enough to fit this limited set of measurements. In Section V, we will elaborate on how this paper sets the

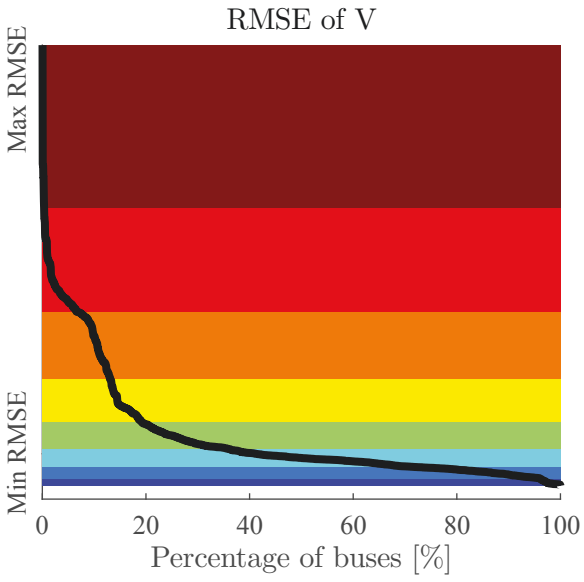

Fig. 4. The sorted RMSE values, overlaid on the contour region colors.

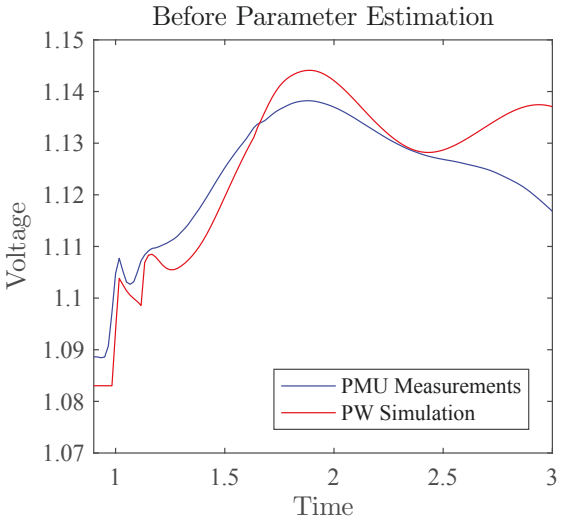

Fig. 5. Measurements and simulated results before parameter estimation.

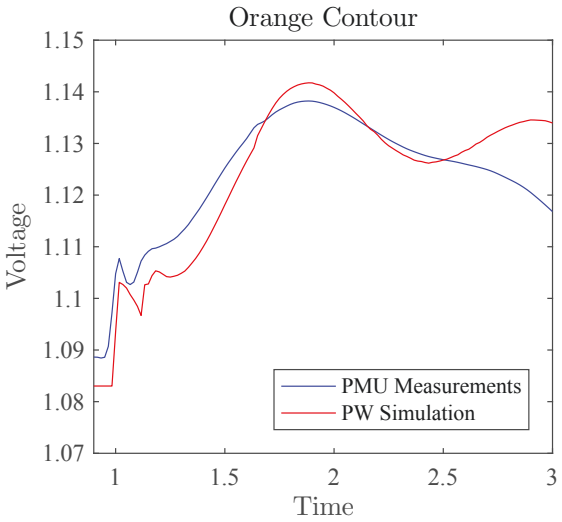

Fig. 6. Measurements and simulated results after parameter estimation of TPIM models placed on load buses within the orange contour.

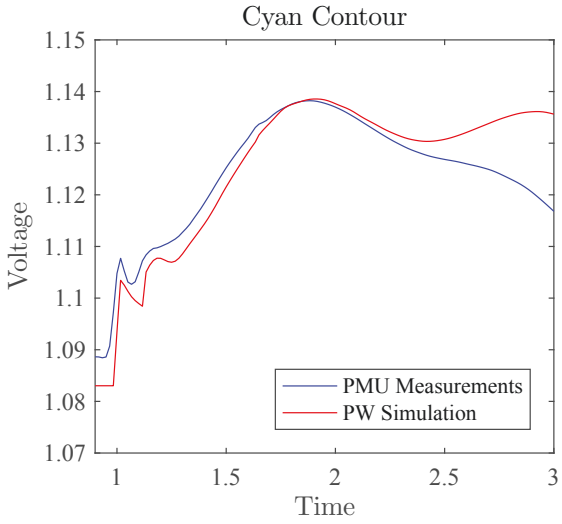

Fig. 7. Measurements and simulated results after parameter estimation of TPIM models placed on load buses within the cyan contour.

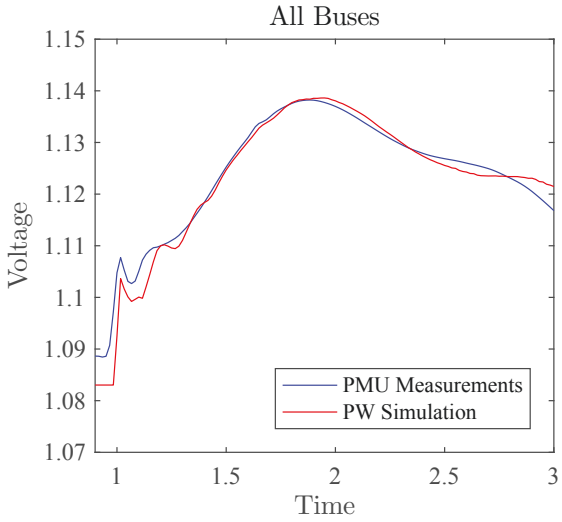

Fig. 8. Measurements and simulated results after parameter estimation of TPIM models placed on all load buses.

framework to allow us to fit longer windows, and at multiple buses.

In Fig. 5, we compare the measurements and the simulation results for the default parameters of the TPIM. In Figs. 6 and 7, we perform load modeling on the buses in the small orange contour area, and the medium cyan contour area, respectively. Finally, in Fig. 8, the TPIM is placed on all load buses. Table II also shows the parameter values before and after parameter estimation using all buses. As expected, the parameter estimation results improve as more buses contain load models. From Fig. 5 (before parameter estimation) to Fig. 6 (orange contour), we can see that the peak at $t = 2$ s is closer to the measurements. However, in general, the simulation result after parameter estimation is only marginally better than before parameter estimation. From Fig. 6 to Fig. 7, the improvement is much more pronounced. The simulation and measurements are quite close until $t = 2.5$ s. However, the last 0.5 s of simulation still swings differently than the measurements. Finally, from Fig. 7 to Fig. 8, the match is even better for $t < 2.5$ s, and more importantly, the swing after 2.5 s also matches much better. What we can take away from this analysis is that the shorter the window we wish to match, the fewer buses we need at which to perform parameter estimation.

V. Conclusion and Future Work

In this work, we first used simulations to investigate how a voltage disturbance propagates in a large network. Based on the magnitude of the voltage swings at each bus, a contour was produced which revealed possible regions of interest. A few buses near the disturbance had large deviations, while

TABLE II
MODEL BEFORE AND AFTER PARAMETER ESTIMATION

PARAMETER		BEFORE	AFTER
P_{ul}	Fraction of total load	0.2	0.8933
L_s	Synchronous reactance	3.6	2.1506
L'	Transient reactance	0.17	0.0388
R_a	Stator resistance	0.0068	0.0074
T_0'	Transient rotor time const.	0.53	0.9214
H	Inertia constant	0.5	0.2027
D	Damping factor	2	0.3730
V_T	Voltage trip threshold	0.6	0.8438
T_V	Trip pickup time	30	26.168
T_{bkr}	Breaker operation time	0.03333	0.0575
Acc	Acceleration factor	0.6	0.2563
L''	Subtransient reactance	0.17	0.3561
T_0''	Subtrans. rotor time const.	0	0
n_Δ	Time step subdivision	10	10.546
ω_Δ	Subdiv. speed threshold	0.8	0.7688

each larger contour level contained increasingly more buses. We then hypothesized that placing dynamic induction motor models at the buses in a larger contour would result in a better fit, at the expense of greater computational costs. This was indeed the case. However, the improvement was not linear. When we increased the number of load models from a small to a moderate number, the first 2.5 s matched well, while the last 0.5 s did not improve. Only by adding load models to the remaining buses were we able to achieve a close match in the last 0.5 s.

The methods discussed in this paper set the framework for continued research. Most importantly, we have confirmed our hypothesis that we would require more load models in order to achieve a match in a longer window. Eventually, we would like to match all 60 seconds of the disturbance. In order to do so, we require more degrees of freedom. In this paper, we used the same parameters at all induction motor models simultaneously; in the future, we will need to relax this constraint. The ROI contour will play a much larger role here. For the small contours near the disturbance—those that showed the largest deviation—we will perform load modeling at finer resolution. For the large contours that encompass much of the entire system, a coarser resolution can be used. For example, in the dark red region, we may use a different load model for every bus. In the orange region, we may partition the buses into several clusters, and use a different model for each cluster. For the rest of the system, we can use even larger clusters. This approach will allow us to increase the dimensionality of the optimization in a controlled fashion, with computational resources allocated first to the most important buses.

ACKNOWLEDGMENTS

The authors would like to thank the Bonneville Power Administration (project TIP 357) for supporting this work.

REFERENCES

[1] D. Kosterev, C. Taylor, and W. Mittelstadt, "Model validation for the August 10, 1996 WSCC system outage," *IEEE Transactions on Power Systems*, vol. 14, no. 3, pp. 967–979, 1999.

[2] E. Allen, D. Kosterev, and P. Pourbeik, "Validation of power system models," *IEEE PES General Meeting, PES 2010*, pp. 1–7, 2010.

[3] K. S. Shetye, T. J. Overbye, S. Mohapatra, R. Xu, J. F. Gronquist, and T. L. Doern, "Systematic determination of discrepancies across transient stability software packages," *IEEE Transactions on Power Systems*, vol. 31, no. 1, pp. 432–441, 2016.

[4] R. D. Quint, "A look into load modeling: The composite load model," NERC, Tech. Rep., 2015.

[5] D. Kosterev, A. Meklin, J. Undrill, B. Lesieutre, W. Price, D. Chassin, R. Bravo, and S. Yang, "Load modeling in power system studies: WECC progress update," *IEEE Power and Energy Society 2008 General Meeting: Conversion and Delivery of Electrical Energy in the 21st Century, PES*, pp. 1–8, 2008.

[6] S. Guo and T. J. Overbye, "Suitability of a dynamic load model to measurement-based parameter estimation," in *17th International Conference on Intelligent Systems Application to Power Systems*, 2013.

[7] H. Renmu, M. Jin, and D. Hill, "Composite load modeling via measurement approach," *IEEE Transactions on Power Systems*, vol. 21, no. 2, pp. 663–672, 2006.

[8] S. Guo and T. J. Overbye, "A maximum a-posteriori based algorithm for dynamic load model parameter estimation," in *IEEE International Conference on Smart Grid Communications*, 2015.

[9] K. Zhang, H. Zhu, and S. Guo, "Dependency analysis and improved parameter estimation for dynamic composite load modeling," *IEEE Transactions on Power Systems*, 2017.

[10] I. F. Visconti, D. A. Lima, J. M. C. S. Costa, and N. R. B. C. Sobrinho, "Measurement-based load modeling using transfer functions for dynamic simulations," *IEEE Transactions on Power Systems*, vol. 29, no. 1, pp. 111–120, 2014.

[11] H. Bai, "Novel measurement based load modeling and demand side control methods for fault induced delayed voltage recovery mitigation," Dissertation, Iowa State University, 2010.

[12] L. Pereira and D. Kosterev, "An interim dynamic induction motor model for stability studies in the WSCC," *IEEE Transactions on Power Systems*, vol. 17, no. 4, pp. 1108–1115, 2002.

[13] NERC, "Technical reference document - Dynamic load modeling," North American Electric Reliability Corporation, Tech. Rep. Dec., 2016.

[14] K. S. Shetye, T. J. Overbye, and T. L. Doern, "Assessment of discrepancies in load models across transient stability software packages," *IEEE Power and Energy Society General Meeting*, 2015.

[15] R. A. Schlueter, I. Hu, M. W. Chang, J. C. Lo, and A. Costi, "Methods for determining proximity to voltage collapse," *IEEE Transactions on Power Systems*, vol. 6, no. 1, pp. 285–292, 1991.

[16] D. Juneja, M. Prasad, and M. K. Verma, "An approach for formation of voltage control areas based on voltage stability criterion," in *16th National Power Systems Conference*, 2010, pp. 636–640.

[17] T. Lie, R. A. Schlueter, P. A. Rusche, and R. Rhoades, "Method of identifying weak transmission network stability boundaries," *IEEE Transactions on Power Systems*, vol. 8, no. 1, pp. 293–301, 1993.

[18] H. Mehrjerdi, S. Lefebvre, M. Saad, and D. Asber, "A decentralized control of partitioned power networks for voltage regulation and prevention against disturbance propagation," *IEEE Transactions on Power Systems*, vol. 28, no. 2, pp. 1461–1469, 2013.

[19] M. Paramasivam, S. Dasgupta, V. Ajjarapu, and U. Vaidya, "Contingency analysis and identification of dynamic voltage control areas," *IEEE Transactions on Power Systems*, vol. 30, no. 6, pp. 2974–2983, 2015.

[20] T. M. J. Fruchterman and E. M. Reingold, "Graph drawing by force-directed placement," *Software-Practice and Experience*, vol. 21, no. 11, pp. 1129–1164, 1991.

[21] A. B. Birchfield, K. M. Gegner, T. Xu, K. S. Shetye, and T. J. Overbye, "Statistical considerations in the creation of realistic synthetic power grids for geomagnetic disturbance studies," *IEEE Transactions on Power Systems*, vol. 8950, no. c, pp. 1–1, 2016.

[22] P. Cuffe and A. Keane, "Visualizing the electrical structure of power systems," *IEEE Systems Journal*, pp. 1–12, 2015.