Abstract—Simulations of the dynamic response of high voltage, large-scale electric grids often generate large amounts of data. This can make it difficult for engineers to understand the overall system behavior. This paper presents various techniques to help with gaining situational awareness for electric grid simulations in the time frame of milliseconds to minutes. These techniques include the use of time-domain graphs, geographic data views in which geographic information embedded in the electric grid model is leveraged to create visualizations, contouring, animation loops, machine learning and modal analysis. Results are demonstrated on a 10,000 bus synthetic grid model and an actual electric grid model with 110,000 buses.

I. INTRODUCTION

The design and operation of large-scale electric grid require a variety of different engineering studies and simulations. Some of these are static, such as power flow, contingency analysis and security constrained optimal power flow. And some are dynamic, usually involving time-domain simulations to determine the behavior of the electric grid following some disturbance (contingency). In all of these it is important that the person doing the study or simulation understand what is going on. A term that can be used to convey this concept is situational awareness (SA). While defined informally as “knowing what’s going on,” a more formal definition is “the perception of the elements in the environment within a volume of time and space, the comprehension of their meaning and the projection of their status in the near future” [1], [2]. The term is now widely used in electric grid operations and increasing with engineering studies [3], [4], [5], [6]. The focus on this paper is on techniques to help with time-domain simulations of large-scale electric grids.

Electric grid time-domain simulations can be divided based upon the time scale of the underlying dynamics with [7] presenting four groups starting with wave phenomena (with a time scale of less than a microsecond) and going out to thermodynamic (ranging up to many hours). The time-domain simulations considered here will be in the middle of this range, a scale in which the electric grid is modeled using a phasor representation. As noted in [8] and [7], this considers aspects of rotor angle stability, voltage stability, frequency stability, and to some extent converter driven stability. A typical integration step size would be ¼ or ½ electrical cycle, though the use of multirate methods [9], [10] allows for accurate modeling of the much faster models associated with devices such as exciters, loads and some renewable generators.

While historically such studies were known as transient stability simulations [11], here we’ll use the generic term “simulations.” Usually these simulations are initialized from a power flow solution, then a contingency scenario is applied to the grid and the goal is to determine the time-domain response. The simulations considered here are assumed to have a fixed duration ranging from seconds to minutes. Such simulations are extremely common throughout the electric power industry.

The SA challenges with these simulations depend upon the electric grid size, the complexity of its models, the simulation contingency scenario complexity, and the desired application. For example in many educational and some research simulations the grid size, model complexity, scenario complexity and desired application are similar to the 96-bus angular stability study presented in [11]; SA can usually be adequately maintained just using a graph or two (e.g., Figure 8 of [11] showing the rotor angles for the 20 generators). Similarly even with a large system with complex models and scenarios, if the goal is just to insure that the results for potentially thousands of different contingencies (perhaps run in parallel) meet some criterion (such as for voltage recovery as given in [12]) then likewise the SA needs would be modest.

In contrast the focus here is on improving SA associated with simulations in which there is a desire is to obtain a rather detailed understanding of the total system response. Example applications include doing simulations to insure all of the system models perform adequately, designing remedial action schemes (RASs) [13], considering the impact of unusual events on the grid [14] (with one example a high altitude electromagnetic pulse [15]), or a recent study by the authors considering an ac interconnection of the North American Eastern and Western grids in which the associated grid had 110,000 (110K) buses, there were 245 different types of device models and more than 46,000 model instances [16], [17].

Leveraging the authors’ extensive experience in doing such simulations and in developing the associated software, the paper presents a number of techniques specifically focused on improving SA for such studies. The paper has two main sections. The first focuses on SA during the initial simulation
setup including the power flow, while the second is on SA during and after the simulation. Results are demonstrated on both a 10,000-bus (10K) synthetic grid [18], [19] and on the previously mentioned 110,000-bus model. While the presented techniques are generic, they are specifically demonstrated using PowerWorld Simulator version 22.

II. SA DURING THE SIMULATION SETUP

The simulations are initialized by setting up the differential equations using the models and their individual instance parameters, and then back-solving starting with the power flow solution to determine the initial states (with classic coverage of this topic presented in [20]). Hence overall SA begins with understanding the power flow solution, particularly the aspects most likely to affect the simulation such as the generator reactive power outputs.

For the small electric grids often encountered when power flow is taught most of the values of interest can be clearly shown on a oneline diagram, with the 42-bus system provided with [21] being an example. However, for the much larger systems used in industry, which usually have large numbers of automatic controls (e.g., switched shunts, LTC and phase shifting transformers, area interchange control), the situation becomes more difficult. This difficulty is compounded if the studied system is modeling an unusual operating condition.

Over the years a number of techniques have been developed to help with SA including the use of onelines (often at the substation level), tabular displays, intelligent alarming and color contouring; some background papers in this area include [22], [23], [24]. Our experience is that all of these techniques can be quite useful, with the most important design aspect being the ability to easily get more information on anything that seems important. Or to quote [25], [26], “Overview first, zoom and filter, then details on demand.”

The focus here on wide-area visualization has been helped recently with more widespread availability of electric grid geographic information. This geographic information can be leveraged using geographic data views (GDVs) [27], [6] in which geographic information embedded in the electric grid model is used to draw symbols on a display in which the symbol’s appearance can be dynamically modified to show model object values. GDVs can be quite useful for providing the “details on demand” mentioned earlier and will be utilized throughout the paper.

The validity of different visualization techniques depends upon the desired task. As an example, Figure 1 shows the oneline for a 10,000-bus (10K) synthetic grid (available at [28]) that covers a geographic footprint of the Western US. In the oneline different colors are used to indicate the different nominal voltage levels (with green for 765 kV, orange for 500 kV, red for 345 kV, purple for 230 kV and black for lower voltages). To give some indication of the real power flow, dark green arrows are superimposed on the transmission lines with the size of the arrow proportional to the real power flow [29]. If desired, a color contour could be used to show the per unit voltage magnitudes, illustrated in Figure 2 [30] with alpha-blending used to deemphasize the transmission grid [31].

As an alternative Figure 3 shows a GDV summary visualization [6] in which the substation and line information have been aggregated based on a geographic grid. Here the size of each rectangle (a GDV summary object) is proportional to the net real power injection for the buses in the rectangle, the color and field value give the minimum per unit voltage magnitude, and the size of the black flow arrows is proportional to the real power flow between the different regions. Figure 4 shows an example substation GDV for the 110K system in which the size of the ovals corresponds to the substation’s generation, and its color is shaded based upon the generator’s percentage reactive power output relative to its Mvar limits (with red corresponding to heavy Mvar output, and blue heavy Mvar absorption).

Having good SA on the initial power flow solution is crucial to correctly initializing the simulations. However, before running the simulations it is important to address model instance parameter errors, a challenge common in larger-sized
grid models. While errors on the individual instances do gradually get corrected, this is offset by the addition of new model types and associated instances. With the current rapid change in many electric grids worldwide with the addition of more renewable generation and storage this trend shows no signs of abating.

Figure 3: 10K Summary Overview Display

Figure 4: 110K System Substation Generation Overview Display

One technique that has proven to be quite helpful in determining potential parameter errors is the use of single machine, infinite bus (SMIB) eigenvalue analysis. First presented in [32], SMIB analysis determines the eigenvalues for each generator separately as though it were connected to an infinite bus. Positive or sometimes extremely negative eigenvalues can indicate potential model errors. The SMIB analysis requires the calculation of the driving point impedance for each generator, a value determined based on the diagonal values of the inverse of the bus admittance matrix; these values can be computed quite quickly using sparse vector methods [33]. As examples all the SMIB eigenvalues for the 2485 generators in the 10K synthetic system can be computed in less than three seconds (on a Windows PC with an i7-5820K, 3.3 GHz Processor), and in about 25 seconds for the 13,800 generators in the 110K system.

Associated with SMIB analysis a useful technique is to also use the driving point impedance to calculate a two-bus equivalent for the generators with suspect eigenvalues. The stability of the two-us system can then be quickly assessed by applying a variety of events, such as a self-clearing fault at the generator’s terminal or dynamically varying the infinite bus voltage magnitude and/or frequency. Various model types, such as the generator’s stabilizer, can also be quickly disabled if needed. The eigenvalue participation factor matrix [34] can also be used to further isolate suspect parameters.

As an example, based on a previously encountered error with an actual grid case, Figure 5 shows SMIB results for one generator in the 10K synthetic grid in which the generator exciter’s line drop compensation impedance [35], [36] is erroneously set to a value much higher than the impedance of its step-up transformer. In the simulation this caused a sustained change in the generator’s reactive power output following a self-clearing fault. Figure 5 show this generator’s SMIB eigenvalues, sorted from high to low, whereas the last few columns show some of the participation factors, normalized to unity. The near unity participation value in the column associated with the exciter’s voltage feedback (VF) allowed the error to be quickly located and corrected.

Figure 5: Example SMIB Eigenvalues with Participation Factors

Another useful technique for locating parameter errors is to examine a log of any initial dynamic simulation limit violations coupled with the ability to visualize all of the parameters associated with a particular model type. As has been noted, the dynamic simulation is initialized with the state variables set by backsolving the dynamic model differential equations starting from the initializing power flow solution. During this process, some of the dynamic model limits may be violated. While the reasons for these limit violations vary, they can certainly indicate either model parameter or limit value errors.

As an example with the 110K system, which has 46,800 model instances and 202,400 states, there were about 120 initial limit violations (which on a percentage basis is quite small). While ideally all would be checked, most correspond to minor issues. However, some are potentially more significant depending upon the application. An example is one associated with the WT3E1 wind turbine model [37] in which several of the parameters from the more than 300 instances of this model are shown in Figure 6. For this model the interpretation of the Xlqmin and Xlqmax limits (shown in the eighth and ninth columns) depends upon the vltflg parameter (shown in the second column). The likely erroneous entry is given in the third line in which the Xlqmin and Xlqmax limits do not match what would be expected with vltflg set to 1, causing an initial limit violation. The use of a limit violations
simulation generated events (SGEs) that model the grid's behavior and are documented in a log. The SGEs can be divided into two categories: 1) SGEs that the simulation completed normally, with such events usually occurring when the simulation is to know what events occurred and whether the simulation completed without error. 2) SGEs that indicate a simulation failure, such as forgetting to set the time step, something that can be readily determined from the log. Other times it can be more difficult to determine what occurred, requiring a much broader consideration of the results.

The remainder of this section presents various techniques for better gaining simulation SA. Since each has its strengths and weaknesses, often an approach combining multiple techniques is best. A helpful interpretation approach is to consider the results as a set of matrices in which each matrix corresponds to a different measurement type (e.g., bus frequencies, bus voltage magnitudes, generator governor outputs, etc.). The rows in all the matrices correspond to the time samples and the columns the various measurement values at the different system locations (e.g., buses, etc.). The columns then have associated metadata (e.g., bus number or names, geographic locations, electric characteristics, etc.) that define them and relate them to the columns. The different visualization techniques are then associated with showing some or all of these matrices, or metrics derived from their elements.

With this interpretation the techniques presented here are broadly divided into three groups. First, traditional time-varying graphs in which time is the x-axis parameter and the time-variation in the values of interest (i.e., the signals) is shown. Hence they show a portion of the results matrices but without much of the metadata (beyond perhaps a label). Second, ones in which the visualizations show the grid at a particular time point and animation loops are used to show the simulation response. This approach provides an opportunity for more fully showing the metadata, such as the geographic or electric location. Third, techniques that use algorithms to aggregate the overall system response with the machine learning approach of [38] and modal analysis examples.

In the first group a common technique for gaining some simulation SA is to setup a time plot of a small sample of results (signals), usually chosen from across the grid, to get a feel for the overall grid response. As an example consider the 10K grid in which the contingency is the simultaneous loss of three large generators at the same substation after one second. Figure 7 signals show the voltage frequency response at ten buses selected from across the system (and if the goal is to rapidly identify specific locations then results from [26] indicate the number of colors should be rather modest, no more than about ten). Advantages of such figures include they are quick to draw, a key can be used to provide a label for each signal (e.g., mapping the signal to a particular bus), and they can show the complete time-variation for the signals. Disadvantages include 1) there is a lack of any spatial relationships between the signals, 2) since it is just a sample of the signals important results could be missed, and 3) it can be difficult include different types of signals such as voltage magnitudes with frequency (this could be done with multiple

<table>
<thead>
<tr>
<th>Bus</th>
<th>Voltage</th>
<th>Frequency</th>
<th>Power</th>
<th>Protection</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.95</td>
<td>0.94</td>
<td>0.95</td>
<td>0.94</td>
</tr>
<tr>
<td>2</td>
<td>0.96</td>
<td>0.95</td>
<td>0.96</td>
<td>0.95</td>
</tr>
<tr>
<td>3</td>
<td>0.97</td>
<td>0.96</td>
<td>0.97</td>
<td>0.96</td>
</tr>
<tr>
<td>4</td>
<td>0.98</td>
<td>0.97</td>
<td>0.98</td>
<td>0.97</td>
</tr>
<tr>
<td>5</td>
<td>0.99</td>
<td>0.98</td>
<td>0.99</td>
<td>0.98</td>
</tr>
</tbody>
</table>

Figure 6: Example WT3E1 Model Parameters

III. SA DURING AND AFTER THE SIMULATION

The potential SA challenge during and/or after a simulation is with interpreting the large amount of data that could be generated. For example with the 10K system following a 30-second simulation with a ¼ cycle time step (i.e., 240 time points per second) if just two values are stored per bus (say frequency and voltage magnitude) then there will be 144 million values. With the 110K system the number of values is increased by a factor of ten. While not large from a big data perspective and many of the values are related, understanding what just occurred can still be a challenge. This section presents some techniques for helping to achieve this understanding.

How much interpretation is needed depends on the application. Luckily for many simulations the SA challenge can be extremely modest. In a true transient stability study it might just be determining whether any generators lost synchronism within a few seconds following a fault at a specified time. A question that can be answered automatically by monitoring the generator rotor angle separation. Or the question could be to sequentially run thousands of parallel contingency simulations to determine the maximum amount of power that can be transferred between two utilities before a frequency deviation or voltage recovery criterion [7], [12] is no longer met. Such situations are not this paper’s focus since they are already well handled. The paper’s focus is also not particularly on understanding the response of smaller systems since existing techniques are usually adequate.

Rather its focus is on really knowing what is going on in a large grid simulation when something quite unexpected could be occurring. Examples could include debugging a new electric grid model to look for parameter errors, simulating more unusual situations, or even some of the more routine studies mentioned earlier when things don’t go as planned. Hence there is a need for more sophisticated techniques.

The starting point for understanding a fixed duration simulation is to know what events occurred and whether the simulation completed normally, with such events usually documented in a log. These events can be divided into the pre-defined scenario contingency events that take place at user set times (e.g., apply a fault, open a generator), and the simulation generated events (SGEs) that model the grid’s protection system including any RASs that may have been implemented. A first step to gaining SA is knowing whether the simulation failed (and if so, when) along with noting the number and times of occurrences of the SGEs. Sometimes the cause of a simulation failure is simple, such as forgetting to set a fault, something that can be readily determined from a log. Other times it can be much more difficult to determine what occurred, requiring a much broader consideration of the results.
y-axes though with a similar limitation on the number of signals). Such graphs have a long history (e.g., [11]) and they certainly play an important role in gaining SA.

An alternative to plotting a small signal subset is to plot all of them. Given the current speed of plotting algorithms tens of thousands of signals can be quickly rendered. As an example Figure 8 shows a plot of all 10,000 bus voltage frequencies. The advantage is that now no signals of the specified type are missed (and for this example it is clear that there is a sustained oscillation that will be considered later in the paper). Disadvantages include the loss of being able to see the individual signals, potentially longer rendering times, and because of overlap many of the signals are actually covered. An alternative for showing all of the signals is to plot the envelope of their response (i.e., the minimum and maximum at each time point). This is shown in Figure 9 in which the oscillation is also readily apparent. Hence a metric derived from the results metric is being shown, that is the maximum and minimum values from each row.

Of course similar graphs could be created for a wide variety of different values, either for a subset or for all the signals in a class. As an example Figure 10 shows all the voltage magnitudes for the 10K system. As with the frequencies there are simply too many signals to adequately show each one, but the display does show that the response is bounded within a fairly narrow range. For parameters with a range of initial values it often beneficial to show the deviation from the initial values, with Figure 11 showing the values for this same contingency. Now the sustained voltage oscillations are more apparent.

The second general group of techniques is to show data at a particular point in time, and then utilize animation loops [39], [40] to show how the system changes with time. One commonly used technique is a geographic oneline diagram, often with an associated contour [30]. An example of this is shown for the 10K system in Figure 12 in which the contour is
showing the bus voltage magnitude deviation at a simulation time of two seconds and alpha-blended is used to deemphasize the transmission grid. Hence it is showing a row of the voltage magnitude simulation results matrix with all the values shifted by the values in the first row (i.e., when simulation time is zero).

Figure 12: 10K Grid Voltage Deviation Contour at 2.0 Seconds

The contours can be combined to show other results data as well. An example is shown Figure 13, also for a simulation time of two seconds, in which the contour shows the spatial variation in the bus frequency, the yellow rectangles use the GDV summary visualization approach used in Figure 3, now showing the largest voltage deviation in the different geographic portions of the grid, and the black arrows show the change in the real power flow. While by themselves such visualizations can help with SA, they are even more effective when used in animations. One animation approach that has been particularly effective in showing generator outage scenarios is develop the animations using a variable playback speed approach. For example, creating the animation to the first 10 seconds at ½ real-time, show the next 20 seconds are real-time, and show any subsequent values (i.e., the slower automatic generation control [AGC] response) at twice or more real-time.

Figure 13: 10K Grid Voltage Deviation, Flows and Frequency at 2.0 Seconds

The third general group of techniques for gaining SA is to utilize various algorithms to aggregate and summarize the overall system response. Many different techniques exist, with an example of machine learning applied to cluster the results given in [38]. Broadly useful approach is signal-based modal analysis which was initially introduced into the power community in [41]. For the larger system applications here, the iterative matrix pencil (IMP) approach is particularly effective in determining the modes for large numbers of signals [42]. To finish the 10K example, it is clear from Figure 11 that there is an oscillation. Modal analysis with the IMP utilizing all 10,000 voltage magnitude signals as inputs can be used to quickly determine frequency (0.31 Hz) and the algorithm from [43] can determine its source and visualize the results. This is shown in Figure 14 in which the large magenta rectangle shows the source of the modal power flows in Northeast Montana, the yellow rectangles show the absorption locations (primarily in the southwest part of the grid) and the arrows show the modal power flows. The oscillation can be corrected by either disabling the dynamic model for the 37 MVA generator or correcting the model parameters (this error was actually deliberately induced to shown an oscillation, and could have been found with the SMIB since the generator had a positive eigenvalue).

Figure 14: 10K Grid Visualizing the Source of the 0.31 Hz Oscillation

To finish, Figures 15 to 18 provide example results from the 110K grid for a generator loss contingency that includes modeling some of the AGC response. Figure 15 shows the frequency variation at a dozen buses chosen across the system for the first minute of the simulation. While these signals give an indication of the overall system response, they do not show the spatial variation in the frequency response. This could be best shown with an animation. Several images from such an animation are shown in Figures 16 to 18 that include a color contour to show the frequency variation (with used to indicate less than 60 Hz), yellow or magenta ovals to show the net
change in generation in different locations in the grid (with yellow indicating increased generation), and black arrows are used to visualize the change in line flows. In the first few seconds after the event (Figure 16) the frequency is declining with the replacement power being supplied by nearby generators. At 12 seconds the event has affected the entire grid, with now the replacement power dependent upon the generators’ governor response. Finally by 150 seconds AGC is assumed to be taking place, with emergency power transfers in place. Key from an SA perspective, these visualizations are conveying a large amount of information about several different result values (bus frequency, change in generator real power output, and line flows). Of course now single display or animation can convey everything, but in combination the techniques presented here can help to provide good SA even associated with quite large electric grids.

**Figure 15: 110K All Substation Average Frequencies**

**Figure 16: 110K Grid Generation, Bus Frequency and Line Flow Variation at 2.0 Seconds**

**IV. SUMMARY AND FUTURE DIRECTIONS**

This paper has provided some coverage of techniques for improving large-scale electric grid simulation situational awareness. This is an extremely expansive topic, and clearly no single paper can cover even a small portion of the topic. Still, the paper has provided some insights in how SA can be improved. Many of these techniques can be expanded and extended for future work, particularly those associated with the visualization of transmission grid power flows and the use of various techniques for data analytics including more machine learning.

**V. ACKNOWLEDGEMENTS**

This work was partially supported through funding provided by the U.S. National Science Foundation in Award 1916142, and the Power Systems Engineering Research Center (PSERC) through projects S-91 and S-92G.

**REFERENCES**


