NASPI WHITE PAPER

Advanced Model Validation and Calibration using Synchrophasors

Prepared by NASPI Engineering Analysis Task Team (EATT)

2024

Disclaimer

Acknowledgments

Lead:
Honggang Wang

Contributing Authors (Alphabetical Order):
Urmila Agrawal, Graham Dudgeon, Pavel Etingov, Evangelos Farantatos, Renke Huang, Kaveri Mahapatra, Slava Maslennikov, Neeraj Nayak, Junjian Qi, Matthew Rhodes, Mani Venkatasubramanian, Junbo Zhao, Gang Zheng

Acronyms

AGM	Angle based Grid Management
BA	Balancing Authority
DOE	Department of Energy
EMS	Energy Management System
GO	Generator Owner
GPV	EPG's Generator Parameter Validation Tool
MVC	Model Validation and Calibration
MRE	Mean Response Error
MPE	Mean Parameter Error
MWEX	Minnesota-Wisconsin Exchange
NASPI	North American Synchrophasor Initiative
NERC	North American Electric Reliability Corporation
IEEE	Institute of Electrical and Electronics Engineers
ISO	Independent System Operator
PC	Planning Coordinator
PMU	Phasor Measurement Unit
POI	Point of Interconnection
PPMV	PNNL's Power Plant Model Validation Tool
PPPD	EPRI's Power Plant Parameter Derivation Tool
PSAT	PowerTech Lab's Powerflow & Short-circuit Analysis Tool
PSLF	GE's Positive Sequence Load Flow Software
PSS/E	Siemens' Power System Simulator for Engineering
RC	Reliability Coordinator
RP	Resource Planner
SCADA	Supervisory Control and Data Acquisition
SVD	Singular Value Decomposition
TMRE	Total Mean Response Error
TO	Transmission Owner
TP	Transmission Planner
TSAT	PowerTech Lab's Transient Security Assessment Tool
UKF	Unscented Kalman Filter
WAMS	Wide Area Monitoring System
WECC	Western Electricity Coordinating Council

Contents

1 Introduction

Power system models could be categorized into different types using different criteria [1]. Based on the scope of system to be modeled, there are device-level models and system-level models. The former describes the behavior of a single asset or device in the power system, such as governor or hydro-generator, while the latter describes interactive behavior of the electric network formed by multiple electrical assets including generators, transmission, and loads. Based on the time scale that the models characterize, they can be divided into static or steady-state models and dynamic or transient models. The former describes the slower or steady state behavior of the system such as power flow and generator economic dispatch. The latter describes the faster, transient system behavior such as frequency regulation, voltage excursion, line trip and Remedial Action Scheme (RAS) activation.

The power grid is like a living creature as it keeps adapting to the extreme weather, human activities and its own health degradation [2]. The interconnected and fast transient nature of the power grid requires that its planning, operation, and maintenance rely heavily on mathematical models, which should reflect the significant characteristics of the real power asset and system. Even though power grid algorithms and toolsets have become more and more sophisticated and smart, their effectiveness can only be as good as the mathematical model of the power system and power asset. Ensuring a verified and up-to-date power asset model for generator, excitation system, governor and other control system is essential for reliable power system planning and operation studies.

One popular way to verify and validate the power system models is to use synchrophasor data from Phasor Measurement Units (PMUs). Compared to traditional Supervisory Control and Data Acquisition (SCADA) system, the PMU measurement provides unprecedented observability to the power grid with 100 times faster reporting rate and unique synchronized view of the power system dynamic state including phase angles. These unique features of PMU measurement together with the available grid disturbance data enable a low-cost solution to validate and calibrate the power system models. Although there are many other ways to conduct model validation and calibration (MVC), this white paper will focus on power system MVC using disturbance data captured by PMU or high frequency disturbance recorders.

This white paper, as a continual effort of prior art [1], intends to capture recent development and the latest lessons learned as the industry moves forward. Although there are various models, including static and dynamic models for generators, transmission lines, transformers, HVDCs, short circuit models, etc., this white paper primarily focus on the generator dynamic models, due to well-established mandates from reliability standards relevant to many power grid entities. However, the challenges and developed technical framework may be transferable to other power system assets and system validation applications.

1.1 Motivation for Model Validation & Calibration

Models are the foundation of virtually all power system studies. Calculation of operating limits, planning studies for assessment of new generation and load growth, performance assessments of system integrity protection schemes (SIPS) – all these studies depend on an approximate mathematical representation of the transmission, generation, and load [3].

Although many prior arts have elaborated the importance of validating and calibrating power system models [1], the values of MVC can be best demonstrated by facilitating compliance with the four MOD Standards released and mandated by the North American Electric Reliability Corporation (NERC) for more reliable power grid planning. Table 1.1 summarizes each NERC MOD standard including its model validation focus, suggested model validation method, related functional entities, and mandated model validation interval. It is evident that MOD-026-1 [5] and MOD-027-1 [6] focus on generator model validation while MOD-032-1 [7] in conjunction with MOD-033-1 [8] validate the interconnected transmission system model.

For both MOD-026-1 and MOD-027-1, The Generator Owner (GO) of a generation unit larger than 100 MVA (East), 75 MVA (West), or 50 MVA (ERCOT) must provide "a verified generator model, including documentation and data," to its Transmission Planner every 10 years, and update these models if there are any changes to the plant controls that change the plant's response characteristics. The GO owns the model validation effort, and the Transmission Planner (TP) shall decide whether the provided model is "usable" or not. Therefore, both GO and TP may need to conduct model validation process. The GO may primarily leverage staged tests which take the generator offline, and the TP may leverage grid disturbances to conduct the model validation. The difference between MOD-026-1 and MOD-027-1 lies in the model validation focus. MOD-026-1 focuses on validating the generator voltage and reactive power response, while MOD-027-1 focuses on validating the generator frequency and active power response.

MOD-032-1 in conjunction with MOD-033-1 intends to validate the interconnected transmission system model, with a two-step approach. The first step, as described in MOD-032-1, develops the power system modeling data requirement by the Planning Coordinator (PC) and its Transmission Planner (TP). Accordingly, those modeling data shall be provided every 13 months by Balancing Authority (BA), GO, Load Serving Entity (LSE), Resource Planner (RP), Transmission Owner (TO), and Transmission Service Provider (TSP). As the second step, the PC shall validate the power system model using the disturbance or "dynamic local event" data provided by the Reliability Coordinator (RC) and Transmission Operators (TO), at least once every 24 months. Those who are interested in the role of the different entities mentioned above could refer to the reliability functional model defined by NERC [9].

Besides power grid planning, an accurate dynamic power system model with better stability margin would enable higher asset utilization. For stability constrained corridors, more accurate calculation of system operating limits can unlock latent capacity across transmission corridors. A study in the British grid identified a transient stability limit in the corridor between Scotland and England, with 1.5GW of wind generation connected between the centers of inertia. The study indicates that a definition of the limit by angle difference can enable 10–12% uplift in the corridor transfer during high wind conditions when the capacity is most valuable [11].

Table 1.1 NERC MOD Standards Relevant to Power Grid MVC.

NERC Standards	Validation Focus	Validation Method	Entities	Interval
MOD-026-1	Validate generator voltage and reactive power response	Staged test (for GO) and POI disturbance-based model validation (for TP)	TP, GO	Every 10 year or significant changes to the plant that modify its response capability
MOD-027-1	Validate generator frequency and active power response	Staged test (for GO) and POI disturbance-based model validation (for TP)	TP, GO	Every 10 year or significant changes to the plant that modify its response capability
MOD-032-1	Interconnected transmission system model	NA	PC, TP, BA, GO, LSE, RP, TO, TSP	Every 13 calendar months
MOD-033-1	Interconnected transmission system model	disturbance based model validation (for PC)	PC, RC, TO	Every 24 calendar months

1.2 Power System Model Validation Overview

Over the past years, several methods have been developed to facilitate power asset model validation and calibration. They can be classified into two major categories: intrusive based approach and non-intrusive based approach.

Generator model validation using staged tests is a well-established intrusive based approach as the generator needs to be taken offline from normal operation for a series of predefined tests. Typical stage tests include a generator test (to determine machine reactance and time constants), an exciter test (to determine exciter gains, time constants, limiter function, etc.), a governor test (to determine time constants, speed droop settings, etc.), a PSS test, and/or a reactive power capability test [12]. After all field tests are complete, model validation is conducted based on the field measurements using mathematical techniques. Staged tests are the most direct way to extract the desired model parameters and are very simple and time efficient. The major disadvantage of this method is its high cost, e.g., $15,000-$35,000 per generator per test in the United State [13].

A complimentary approach to the staged tests is to leverage the various types of disturbances across locations in the power system along with large installed base of PMUs. These disturbance-based model validation approaches are non-intrusive compared to the staged tests. As highlighted in [1], PMU data recorded during grid disturbances captures the underlying relationships among various grid assets more accurately than stand-alone stage testing of individual assets where it is not feasible to replicate the same grid disturbance. Therefore, a grid disturbance driven model validation process can result in more accurate representation of the asset of interest during real system events. This makes it possible to monitor and validate the dynamic models of the generators frequently at different operating conditions so that they can stay updated, even after the model validation with the staged tests [14].

Fig. 1-1 shows the general overview of the disturbance-based model validation using PMU data, which is adopted by most available state-of-the-art methods and toolsets. The high-level process can be described in the following steps [15]:

1. Data Preparation: Obtain the grid disturbance data, typically PMU data from the point of interconnection or a nearby location close to the generator. The PMU data may comprise of voltage magnitude, frequency or phase angle, active power and reactive power. Also, the dynamic model files for the asset and the network model for subsystem creation is collected.
2. Model Verification: The PMU measurements such as voltage magnitude and frequency (or phase angle), are injected into the network while running a dynamic simulation engine in an 'event playback' mode. The simulation outputs, including active and reactive power from the power plant, will be compared against the corresponding PMU measurements. If the difference between simulated response and measured response is sufficiently small, the model accuracy is good, and parameter calibration is not needed. Otherwise, the model calibration process is invoked.
3. Model Calibration: Based on event data mining, the most sensitive parameters will be down selected and tuned so that the difference between simulated response and measured response gets closer. Since this searching process is inherently nonlinear, it may take multiple iterations before a good solution (model parameter values) is found that makes the simulation output match the measurements. A robust model will have a good match against multiple events with the same asset parameters and will ultimately deliver better predictions of the plant's response over a wide range of grid events.

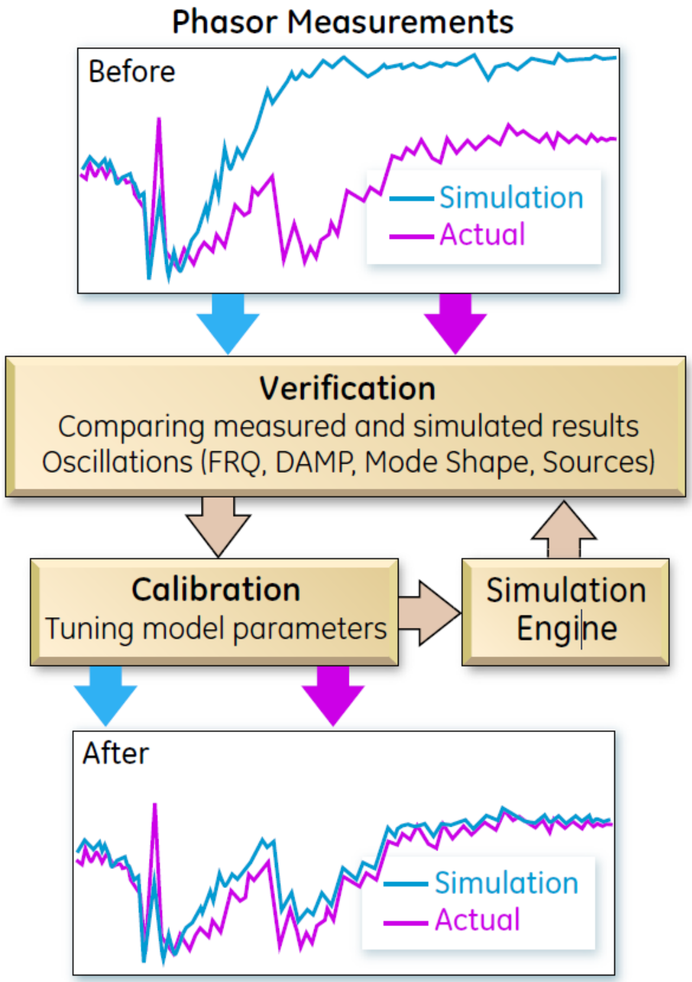


Figure 1-1: General Process of Model Validation and Calibration

1.3 State-of-the-Art Toolsets

Most dynamic simulation engines such as GE PSLF, Siemens PTI PSS/E, PowerTech's TSAT and PowerWorld Simulator are equipped with standard IEEE model library and playback simulation capability for model validation. Once significant model discrepancy is observed, there is a need to identify and calibrate the incorrect parameters. Production grade model calibration software have been developed in recent years, including Power Plant Parameter Derivation (PPPD) by EPRI [17][18][19], Power Plant Model Validation (PPMV) by BPA/ PNNL[20][21][22][23][24], Phasor Grid Dynamics Analyzer (PGDA) by EPG [25] [26][27][28][29][30], model calibration toolbox by MathWorks [25] and GE's Model Validation & Calibration Module in PhasorAnalytics [31][32]. A summary of these toolsets is found in Table 1.2.

Table 1.2 State-of-the-Art Model Validation and Calibration Toolset.

Toolset Name	Vendor	Simulation Engine	Model Validation	Model Calibration	Users
Power Plant Parameter Derivation Tool (PPPD)	Electric Power Research Institute (EPRI)	Simulink model	Playback to perform visual inspections and comparisons between the model and measured data.	User selects initial models and parameter bounds.	MISO, NYISO, PJM, and 20+ more users
Power Plant Model Validation Tool (PPMV)	BPA/PNNL	PSLF, PSSE	Same as above	Automatically identify the sensitive parameters for user review, and then perform Kalman filtering based calibration to fine tune the user selected model parameters.	BPA, PG&E
Model-Based Calibration Toolbox	MathWorks	Simulink model	Same as above	Support multiple events-based model calibration.	ERCOT, PG&E
Phasor Grid Dynamics Analyzer (PGDA)	Electric Power Group	PSSE	Playback to perform visual inspections and comparisons between the model and measured data. Automatic report generation.	Automatically identify the sensitive parameters for user review, and then use Simultaneous Perturbation Stochastic Approximation Particle Swarm Optimization algorithm to fine tune the user selected model parameters.	MISO
PhasorAnalytics (PA)	General Electric	PSLF, TSAT	Playback to perform visual inspections and comparisons between the model and measured data for multiple events	Automatically identify the sensitive parameters for user review, and then perform optimization-based calibration to fine tune the user selected model parameters. Support multiple events-based model calibration.	FPL, NG, ISONE, PG&E

1.3.1 Power Plant Parameter Derivation (PPPD)

The power plant parameter derivation (PPPD) tool is a MATLAB® based tool developed by EPRI [16]. The PPPD tool is widely used by EPRI members to validate and tune parameters of positive sequence models of synchronous generation plants, wind and photovoltaic (PV) plants, and static var systems (SVS). Specifically, the tool can validate:

- The synchronous generator and its inertia constant
- The excitation system including the power system stabilizer model
- The turbine-governor
- Type 3 (double fed induction machine) and Type 4 (permanent magnet synchronous generator) wind plants
- PV power plants, and
- Static var devices like SVCs and Statcoms.

PPPD can be used for validating models both based on staged field-testing data or online recorded event measurement data. It is however recommended that if an event data is used, a good baseline parameter set for the model being verified is already available.

For validating or tuning the parameters of a power system equipment, the relevant terminal quantities are obtained from a digital fault recorder (DFR) or a PMU. DFR measurements have a higher resolution as compared to PMUs but are only available if an event triggers the DFR. PMU measurements on the other hand are continuously available and hence these measurements can be conveniently leveraged when electrically distant disturbances are not sufficient to trigger a local DFR. PMU data however is heavily filtered, but this data can still be used to validate certain aspects of positive sequence models. It should however be noted that the measurements from DFR or PMU should be obtained from a device local to equipment being validated and that the measured quantities are the terminal quantities. Remote measurement is typically not suited for unit validation as these will be contaminated by responses from other electrically close equipment. An example of terminal quantities for a synchronous machine model validation are terminal voltage, active and reactive power, speed, and field voltages and current. Many DFRs can already provide positive sequence values of voltages as well as active and reactive power measurements. However, if the positive sequence values are not available, the raw sinusoidal measurements should be converted to the relevant positive sequence quantities by using standard signal processing techniques.

An example of a generator model validation is shown next. In the validation process of the generator model the problem may be formulated by using a standard sixth order synchronous generator model like the GENROU or the GENTPJ model. The inputs for the generator validation process are the field voltage (E_{fd}), the stator currents (I_d and I_q) and the shaft speed (ω). The output of the model, the stator voltage (V_t) and the field current (I_{fd}) are then compared to the corresponding measured values to derive the parameter set. To start the process the model is first initialized using the vector

$$y_0 = [P_0, Q_0, V_0, \omega_0] \tag{1}$$

Once initialized the input quantities given by the vector

$$u = [E_{fd}, I_d, I_q, \omega] \tag{2}$$

are used to solve the differential-algebraic equations given by

$$\begin{aligned} \dot{\varphi} &= g(\varphi, x, u) \\ z &= f(\varphi, x, u) \end{aligned} \tag{3}$$

where the vector,

$$x = [X_d, X'_d, X''_d, X_q, X'_q, X_l, T'_{d0}, T''_{d0}, T''_{q0}, S_{10,} S_{12}\ ] \tag{4}$$

is the set of parameters to be tuned (the notations in the vector x have usual meaning) and the state variables are given by the vector,

$$\varphi = [\Psi''_d, \Psi''_{dq}, \Psi_{fd}, \Psi_{kd}\ , \Psi_{fq}, \Psi_{kq}, \delta, I_{fd}] \tag{5}$$

where the notations have the usual meanings.

As described earlier, the model output

$$z = [V_t, I_{fd}] \tag{6}$$

is then compared to the measurement and an error vector given by

$$\Delta z = [\Delta V_t, \Delta I_{fd}]$$

is computed. The parameter tuning is then framed as a nonlinear least squares optimization problem

$$arg\ min_x\ \Delta z^t \Delta z \tag{7}$$

given, $x_{lb} \leq x \leq x_{ub}$ where x_{lb} and x_{ub} are the upper and lower limits on the parameters. Note that engineering judgement should be exercised to select the upper and lower limits on the parameters such that the optimization problem in (7) is tractable. Several standard methods are conveniently available in commercial as well as open-source packages to solve (7). The FMINCON routine in Matlab® is leveraged in PPPD to solve (7) and attain the tuned parameter set x.

As mentioned earlier PPPD has been used extensively by EPRI members for model validation and parameter tuning. A list of users is given at Table 1.3. A few examples of synchronous machine plant model validation can be found in [17][18][19].

Table 1.3 PPPD users

Ameren	Arkansas Electric Cooperative (AECC)	Associated Electric Cooperative (AECI)
Tri-State	Saudi Electric Company	LGE & KU
BC Hydro	CenterPoint Energy	Dominion Energy
Duke Energy	ERCOT	Entergy
FirstEnergy	Great River Energy	ISO New England
Korea Electric Power	New York Power Authority	NRG Energy
Arizona Public Service (Pinnacle West Capital Cooperation)	Southern Company	Southwest Power Pool (SPP)
Tennessee Valley Authority (TVA)	WEC Energy Group	Xcel Energy
Bonneville Power Administration (BPA)	American Electric Power (AEP)	

1.3.2 Power Plant Model Validation Tool (PPMV)

The PPMV tool is an open-source standalone Windows application. It automates the power plant model validation process based on disturbance recordings. The tool was developed by Pacific Northwest National Laboratory (PNNL) in collaboration with Bonneville Power Administration (BPA). The tool development was funded by the US Department of Energy (DOE) through the Grid Modernization Laboratory Consortium (GMLC) program and by BPA through the Technology Innovation program.

The overall PPMV tool structure is shown in Fig 1-2. The tool supports different data sources commonly used by electrical companies to store PMU measurements including OSIsoft PI database, BPA PDAT, COMTRADE, and CSV formats. The PPMV tool stores historical disturbance information in the events database. The tool also includes the xml-based database of power plants (mapping power plants with corresponding PMU and SCADA measurement signals) and the database of model validation results. To perform the model validation, the PPMV tool interacts with an external Play-In module. The current version of PPMV tool supports GE PSLF and Siemens/PTI PSS®E Play-In functions. Interaction between the PPMV application and PSLF is performed through Engineering Process Control Language (EPCL) scripts and with PSS®E through Python scripts.

The tool also has built-in advanced visualization and automatic reporting capabilities. A screenshot of the PPMV tool's main graphical user interface (GUI) is shown is Fig. 1-3. Using the GUI, the user can view existing model validation studies or can create a new model validation project. After the user selects required events and plants, the PPMV tool will interact with PSLF or PSS®E through scripting language to perform the model validation [22][23].

The validation process consists of three major stages: (1) Mini state estimation to match the initial power flow conditions; (2) Model Validation run using Play-In function; and (3) Information extraction from the PSLF/ PSS®E channel files. The PPMV tool also has capabilities to perform sensitivity studies and interact with external model calibration models. The tool was integrated with a PNNL-developed external model calibration solver based on Kalman filter approach.

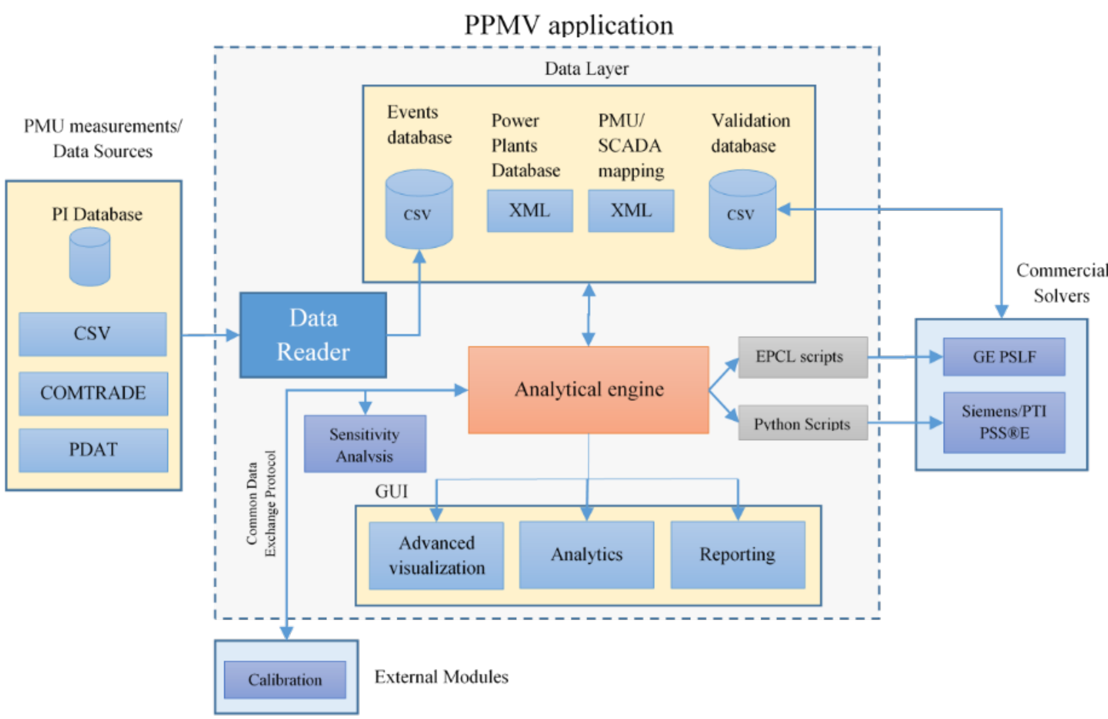


Figure 1-2. PPMV tool overall framework.

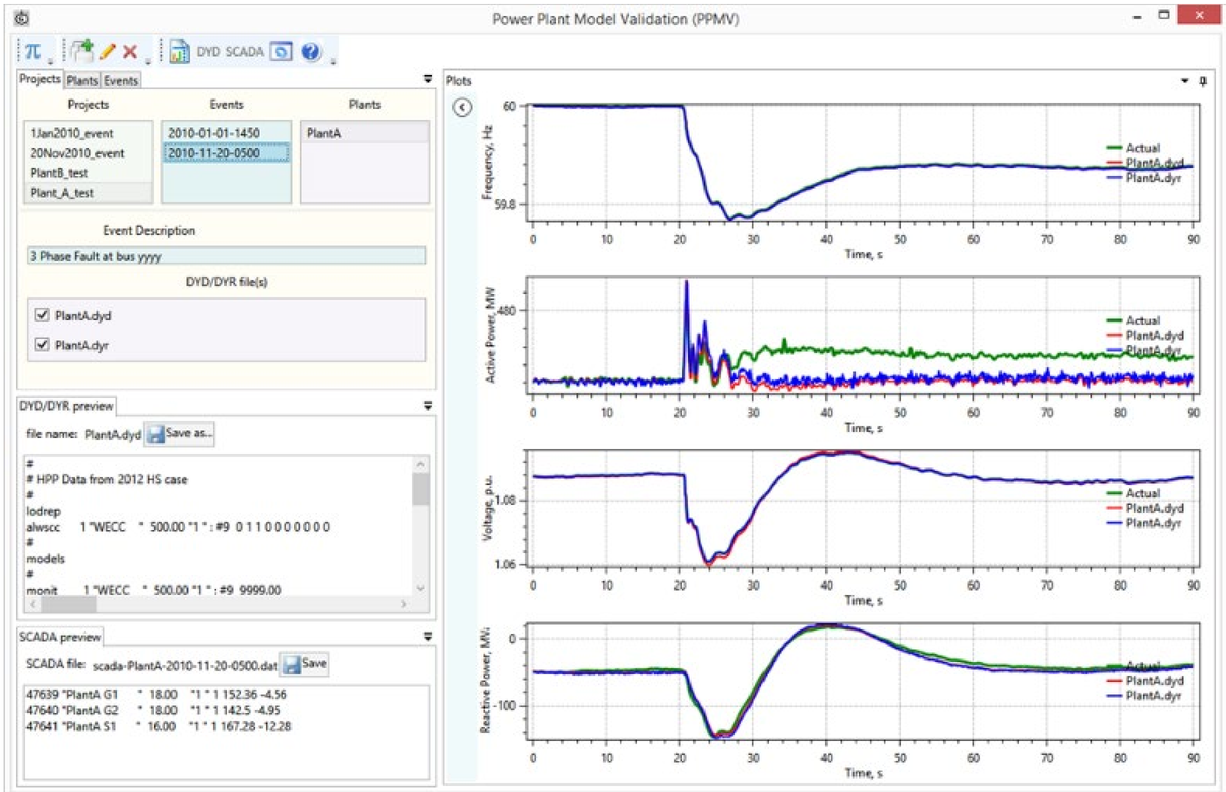


Figure 1-3. PPMV tool main GUI.

1.3.3 Power Plant Model Validation Simscape Design Solution

The Power Plant Model Validation Simscape Design Solution is a tool provided by MathWorks and is available for download from MathWorks File Exchange at the following link [25].

The Power Plant Model Validation Simscape Design Solution demonstrates power plant model validation as applied to online performance monitoring of grid events using phasor measurement unit (PMU) data, through a workflow that includes both manual adjustments and automated techniques. Both gas plant and steam plant reference examples are included.

The following MathWorks products are required to run the tool:

- MATLAB®
- Simulink®
- Control System Toolbox™
- Optimization Toolbox™
- Simscape™
- Simscape Electrical™
- Simulink Control Design™
- Simulink Design Optimization™

The Power Plant Model Validation Simscape Design Solution provides a versatile computational platform that includes:

- Flexible data import
- Data pre-processing
- Power plant simulation with support for online-disturbance data replay (provided as reference examples) and staged-event (with model modification)
- Support for assessing multiple events simultaneously
- Support for different data replay paradigms – i.e. VF, PQ
- Support for manual parameter adjustment
- Support for automated parameter adjustment

A getting started slide deck is included with the download of the Power Plant Model Validation Simscape Design Solution that provides step-by-step information on using the provided reference examples.

1.3.4 Generator Model Validation (GMV)

GMV (Generator Model Validation) uses synchrophasor data to validate and calibrate generator, excitation system and turbine control system models. PMU data along with the power flow and dynamic data for the generator is then used to validate the model and generate a report that compares simulated data with PMU data. GMV provides an automated report that identifies whether the model accurately represents the response of the generator and control system to different events and disturbances. Furthermore, sensitivity analysis can be performed to identify key parameters that should be considered for tuning when the model response does not match the actual response. The sensitivity analysis results quantify the change in the generator response for change in each parameter. This helps in identifying parameters in the model that have the most impact on the generator response and narrows down from several parameters to a few for fine tuning and calibration. For each event, multiple generating units can be validated if PMU data is available from the individual generating units. GMV can also use DFR point on wave data as input and convert it to phasor data for model validation. This application can be installed locally at the generating station or in a central location such as control center which can collect data from multiple generators at different locations. Fig. 1-4. shows overview of data flow and methodology for GMV.

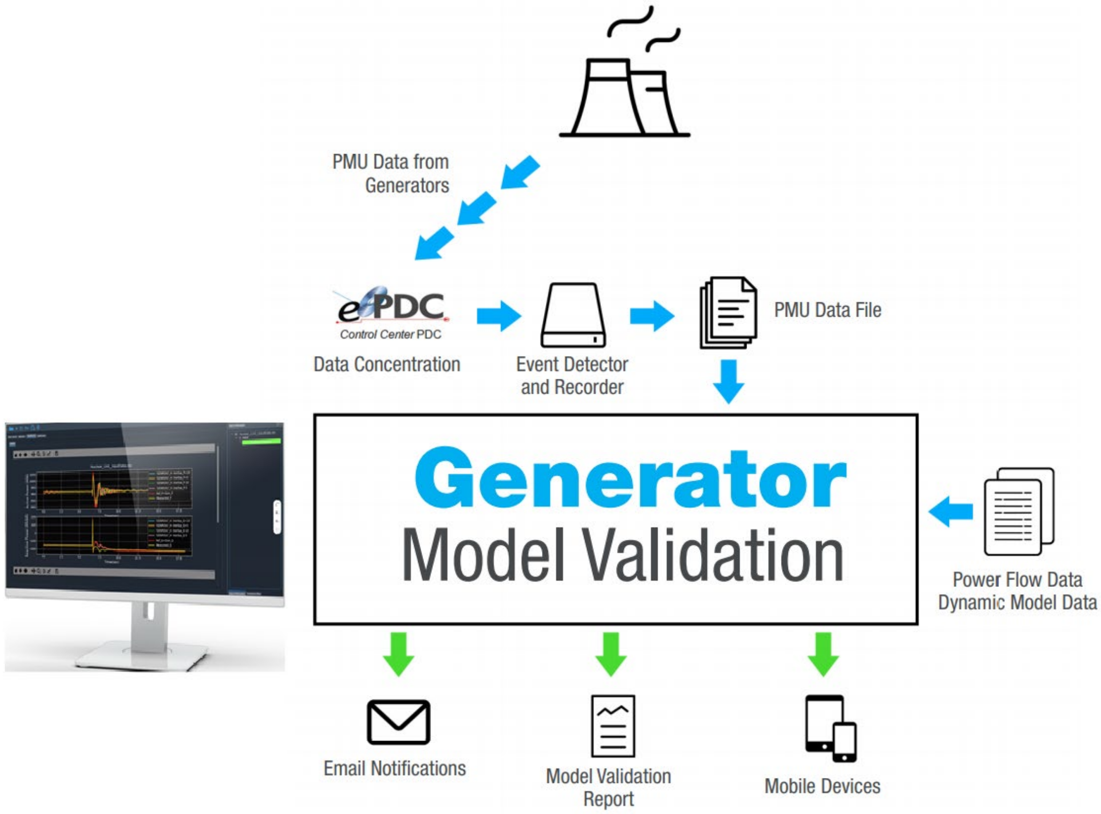


Figure 1-4. Automated Process for Generator Model Validation (GMV)

Key Features include:

- Validate generator models for multiple events
- Validate all types of conventional generating units - Hydro, Nuclear, Combined Cycle, Other Steam & Gas Turbine Generators
- Validate renewable energy resources - Solar, Wind etc. (Planned)
- Validate multiple generators simultaneously
- Automatically Quantify mismatch and identify good vs questionable models
- Perform detailed analysis offline
- Sensitivity analysis to help identify key parameters
- Perform calibration and tuning
- Automated report generation and email notification
- Help meet NERC MOD-26, MOD-27 Compliance

GMV was first developed for use at ERCOT as part of CCET project in 2014. EPG has used GMV to perform model validation and calibration studies for TVA. GMV has also been procured by ComEd and planned for future use. EPG has performed demonstration using GMV for PJM and currently in process for installation and commissioning at PJM.

EPG's Phasor Grid Dynamics Analyzer (PGDA) application is designed for offline analysis using data from various sources such as PMUs, DFRs, simulations etc. Analysis includes oscillation analysis, event analysis, fault analysis, statistical baselining analysis, frequency response analysis, system model validation etc. PGDA integrates with GMV and can perform further analysis of events and model validation results. PGDA is used by several ISOs and Utilities including ERCOT, PJM, NYISO, SPP, ComEd etc.

More information on GMV can be obtained from references [26].

1.3.5 PhasorAnalytics Dynamic Model Validation & Calibration

GE's Model Validation & Calibration improves models to meet emerging NERC requirements. WAMS data and DFR data measure equipment's response to system event or disturbance [31]. Event data are played back to selected dynamic models and the models' response will be benchmarked against actual equipment's response. Validated and calibrated dynamic models provide confidence to system planners and operators about the quality of the models and outcome of security simulations. Furthermore, the Event Library allows users exam historical events and perform multi-event model calibration. Model Validation & Calibration is compatible with industry standard models (PSLF, TSAT and PSS/E).

Figure 1-5. GE's MVC Module in PhasorAnalytics

The Key features of GE's MVC comprise of:

- Compliance oriented study.
- Support multiple data sources.
- User friendly event library.
- Interactive user-interface design.

- Generic compatibility across multiple time-domain simulation models and engines.
- PMU PDC and DFR agnostic.
- Calibration of models in one-click.
- Optimization across multiple events.

The Key Outcome of GE's MVC module comprise of:

- Reliability standards Compliance.
- Sanity check dynamic models.
- Reduce model deficiency.
- Dynamic models tuning.
- Increase renewables penetration by assuring models quality.
- Realization of Digital Twin of physical equipment.

1.4 Present Limitations

Even though significant efforts have been made to streamline the model validation and calibration process, some gaps are identified for its wide adoption in the industry:

1. Systematic performance metrics for model validation and calibration: The current performance metrics primarily focus on "curve fitting" to the measured active power and reactive power time series data. There is a need to enrich the performance metrics to determine to what extent the model represents the actual process.

2. Automatic grid event selection algorithm and data processing: Engineers have disclosed the challenge to manage large amounts of grid events and how to down select them with more automatic approach. Even though some preliminary guideline on selection of the frequency excursion events has been provided by NERC, there is a gap to standardize the event selection process so that it could be automated. Large portion of event data received tend to be unnecessarily long. There is significant amount of data with little or no dynamic information, leading to long computation times and suboptimal results. Manual crafting is possible but takes time and experience. Therefore, there is a need to preprocess the event data or extract relevant features from the event data before conducting model validation and calibration.

3. Improvement for reasonableness of the tuned parameters: Though the end user might not be sure about which parameters need to be tuned, they prefer to achieve a good response match with as few tuned parameters as possible. There is a need to improve the model parameter selection process so that a sparse parameter subset could be identified to achieve the fittings on the measurement space. There is also a need to provide software user interface to interact with the end user to capture the engineering judgement and domain knowledge and capture that reusable domain knowledge into the software database for continuous learning.

4. MVC framework to leverage multi-events: There is a need for a systematic approach to leverage recorded data from multiple qualitatively different disturbances to validate the power system model and estimate model parameters.

5. Trustable model calibration algorithms: There is a need for a robust algorithm to generate interpretable and deterministic solution with reasonable speed, to overcome the challenge from data quality and non-linear property of the dynamic models.

6. Generic production grade software: There is a need for a generic production grade software which could utilize multiple vendors' simulation engines (TSAT, PowerWorld, PSSE and PSLF packages) for validation and calibration. There is a further need for open-source model validation and calibration software package for power system asset model validation and calibration.

The objective of advanced model validation and calibration is to address the challenges listed above. The latest industry developments in this area are summarized in the following sections.

2 Advanced Model Validation

Most existing model validation modules only provide playback simulation and response comparison for user to determine whether the model is "acceptable" or not. To enable a wide adoption for GO and TP, there is a need to enhance the automation level and robustness of the model validation process [33]. It is desirable to explicitly include adequate domain knowledge or include a means to verify if the model or parameter is valid in the model validation software. It is further desirable to offer more comprehensive and interpretable performance metrics instead of the current "curve fitting" metric to justify when the model is "acceptable".

2.1 Enhanced Model Validation Procedure

An effective model validation process should provide comprehensive validation on the model type, model configuration and model parameters before looking at the simulation response curve [33].

- **Model Type Check**. The currently used dynamic model shall be automatically checked against the NERC Approved Dynamic Model List (or NERC Model Notification) [35]. There are some recommended models to be used and prohibited model lists. Machine model shall be updated to GENTPJ based on NERC Notification. Entities using the GENSAL or GENROU model are advised to consider using the GENTPJ model for new generators and where generator data is to be newly (re)verified. Prohibited machine model may include GENSAL, GENSAE, GENCLS, CGEN1, GENTRA, FRECHG.
- **Parameter Validity Check**. The bounds for the key model parameter value and their relative logic relationship (such as inequality relationship) based on physics shall be automatically evaluated. Some preliminary metric has been provided in the NERC Case Quality Dynamic Metrics [37]. The other parameter values not covered in the NERC Case Quality Dynamic Metrics shall be automatically evaluated against the parameter bounds from a historical database, such as the dynamic model files (dyd or dyr file) from WECC and Eastern Interconnection. The readers can refer to [33] for more details.
- **Model Configuration Check**. It has been found to be a common case that the models of governor and Power System Stabilizer (PSS) provided by the dynamic files are not consistent with the reality. Their different combination should be simulated if the simulation response does not match the measurement data.
- **Simulation Response Check**. The simulated response shall be compared with the measurement response for multiple typical scenarios using engineering acceptable and applicable metrics, including but not limited to time domain and frequency domain metrics such as phase shift, amplitude, and damping ratio.

Fig. 2-1 shows an exemplified flowchart on how the Model Type Check, Parameter Validity Check, Model Configuration Check and Simulation Response Check can be integrated into the Model Validation module. Instead of immediately starting the playback simulation to evaluate the curve fitting performance, the new scheme will first check the validity of model types based on NERC List of Validate Models. The user will be notified if any prohibited model or missing excitation model in the dynamic model file has been identified. Based on this information, the user can further correct the dynamic model file if there is human error, or to use the model conversion module to convert any prohibited model to a valid one before evaluating the curve fitting performance.

Once the model passes the validity check, the playback simulation response using the given dynamic model parameters will be compared with the real measurement. The response matching result can be either acceptable or not, depending on the performance criteria and engineer judgement. The model will further go through the Parameter Validity Check even though the curve matching is acceptable, wherein the bounds and inequality constraints will be evaluated for relevant parameters. This additional Check ensures that the generator model is "acceptable" at both response space and parameter space. If the Parameter Validity Check fails, the model needs to go through the Model Calibration step.

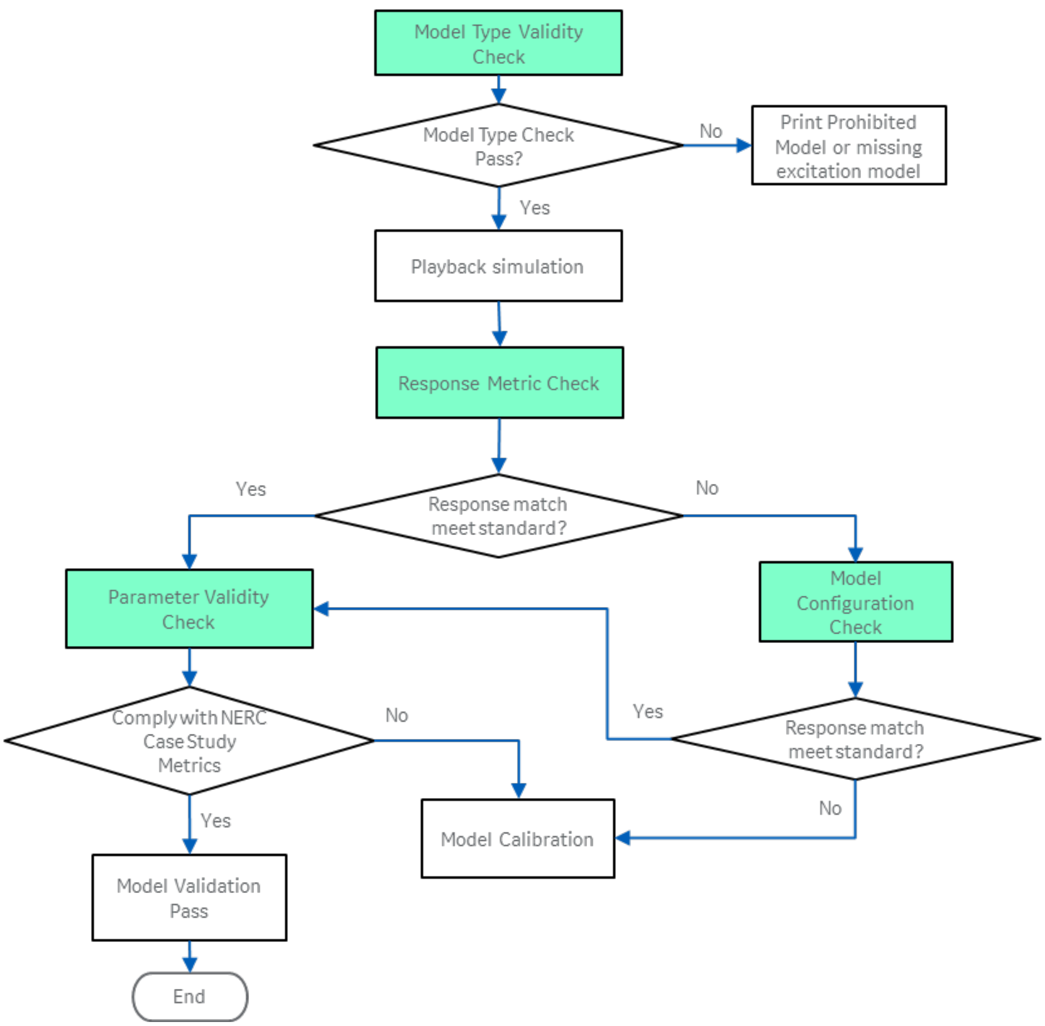


Figure 2-1: Enhanced Model Validation Process from GE [33]

If the Response Matching result is “unacceptable”, the model will go through some Model Configuration Check before preceding to Model Calibration. To avoid unnecessary model calibration on the already accurate enough model, the model validation will check the simulation response at different model configuration such as governor mode and PSS mode. If any mode provides a more reasonable response, then that mode may be the true governor modes. In that case, the unnecessary model calibration could be eliminated, and the robustness of the model validation is improved.

Fig. 2-2 shows a user interface that reports the result of the model validation. It does not only provide the response match results, but also provides the validity of the model types and model parameters, as well as the compliance results to the related NERC published guidelines.

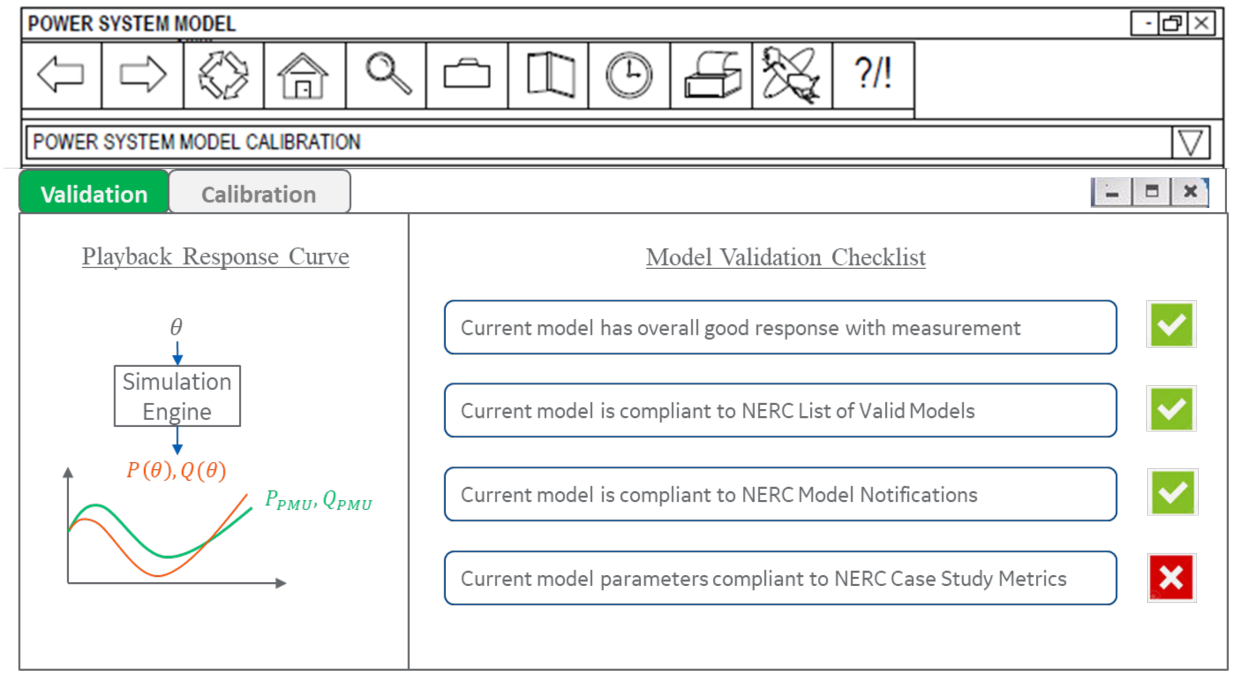


Figure 2-2: Enhanced Model Validation User Interface [33]

Fig. 2-3 shows the enhanced model validation process in the commercial software Generator Parameter Validation (GPV). The model type check is incorporated. Also, the interaction with the end user (GO in this case) is highlighted at different decision points.

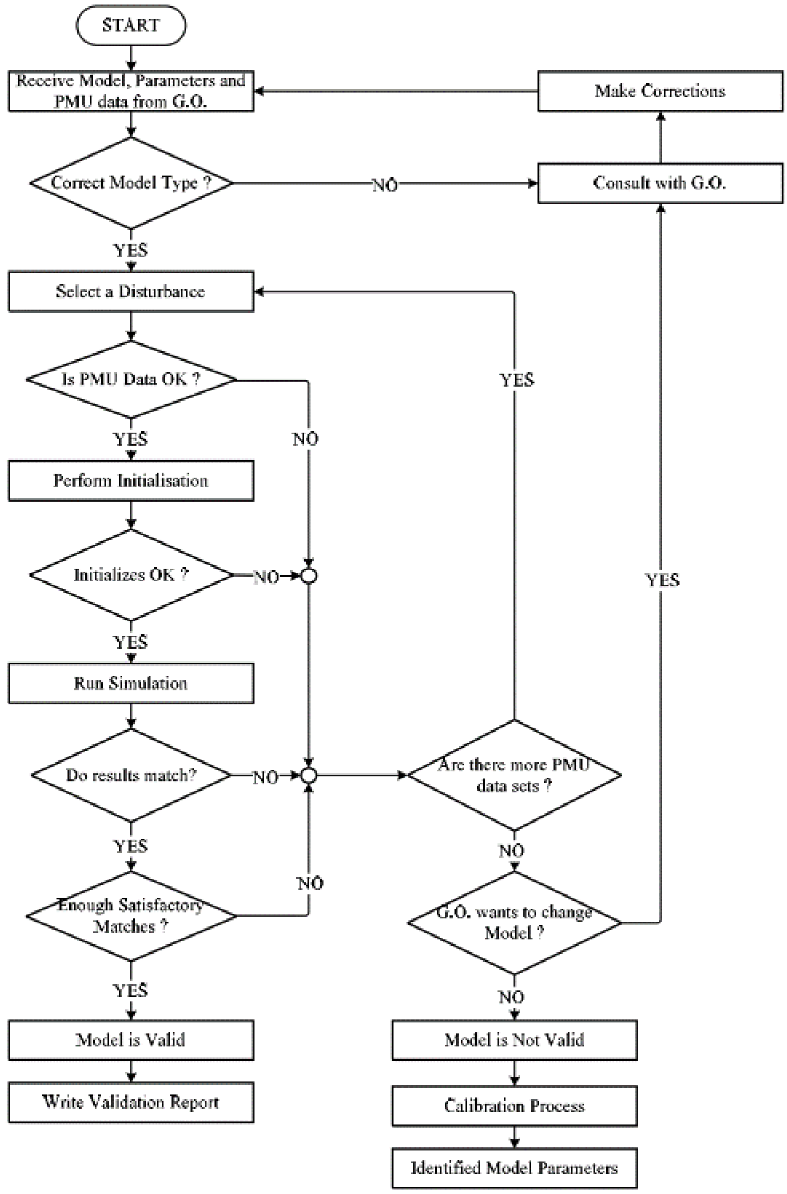


Figure 2-3: Enhanced Model Validation Process from EPG [38]

2.2 Performance Metrics

For generator model validation, an important step is to identify whether the generator model is good or questionable. This involves verifying the performance of the generator model for a given event and determining if model calibration is required. Typically, a visual comparison between the model response and measured response from PMUs for grid events is performed to judge the validity of the model.

2.2.1 Time and Frequency Domain Mismatch Indices

In [30], the EPG team proposes new indices to automatically assess the validity of a generator model. The indices can quantify the mismatch accurately between the model and measured responses from both time and frequency domains. The normalized Root Mean Square Deviation (NRMSD) is proposed to quantify the deviation in terms of magnitude for the model and measured responses in the time domain. The Comprehensive Similarity (CS) is considered by averaging three indices: i.e., Correlation, Frequency Magnitude Similarity, and Frequency Phase Similarity. Among them, correlation can quantify the linear similarity of the model and measured responses in terms of curve pattern. Frequency Magnitude Similarity and Frequency Phase Similarity can quantity the mismatch of the model and measured responses in the frequency domain. The active power and reactive power from the model response are compared with the measured response using the indices of NRMSD and CS to automatically identify whether the generator model is good or questionable. If the calculated NRMSD and CS fall within the corresponding pre-defined thresholds both for active power and reactive power, the validated generator model is good; otherwise, the validated generator model is questionable. Case studies verify the effectiveness of the proposed indices with 25 actual PMU measured events provided by different utilities and grid operators.

2.2.2 Governor and Oscillation Mismatch Indices

Reference [24] proposes a new set of performance metrics to analyze the model validation results by quantifying the mismatch between the actual and model-based response in a comprehensive, accurate and automated manner, which can then facilitate the automation of the model validation process. For this, the new set of metrics developed in [24] takes into consideration mismatches corresponding to several aspects of the generator response, such as phase and magnitude of oscillatory modes, and characteristics of the governor response. For this, first, the slow governor response and comparatively faster oscillatory response are separated, and then a separate set of performance metrics is calculated for each of these two components [23]. These proposed metrics quantify the mismatch between the actual and model-based response in a comprehensive manner without missing any information enabling automation of the process. Furthermore, in [24], it is also proposed that the sensitivity analysis for model calibration be performed with respect to the proposed metrics for the systematic identification of key parameters. A detailed description of the methodology to calculate these performance metrics, shown in Fig. 2-4, is discussed next.

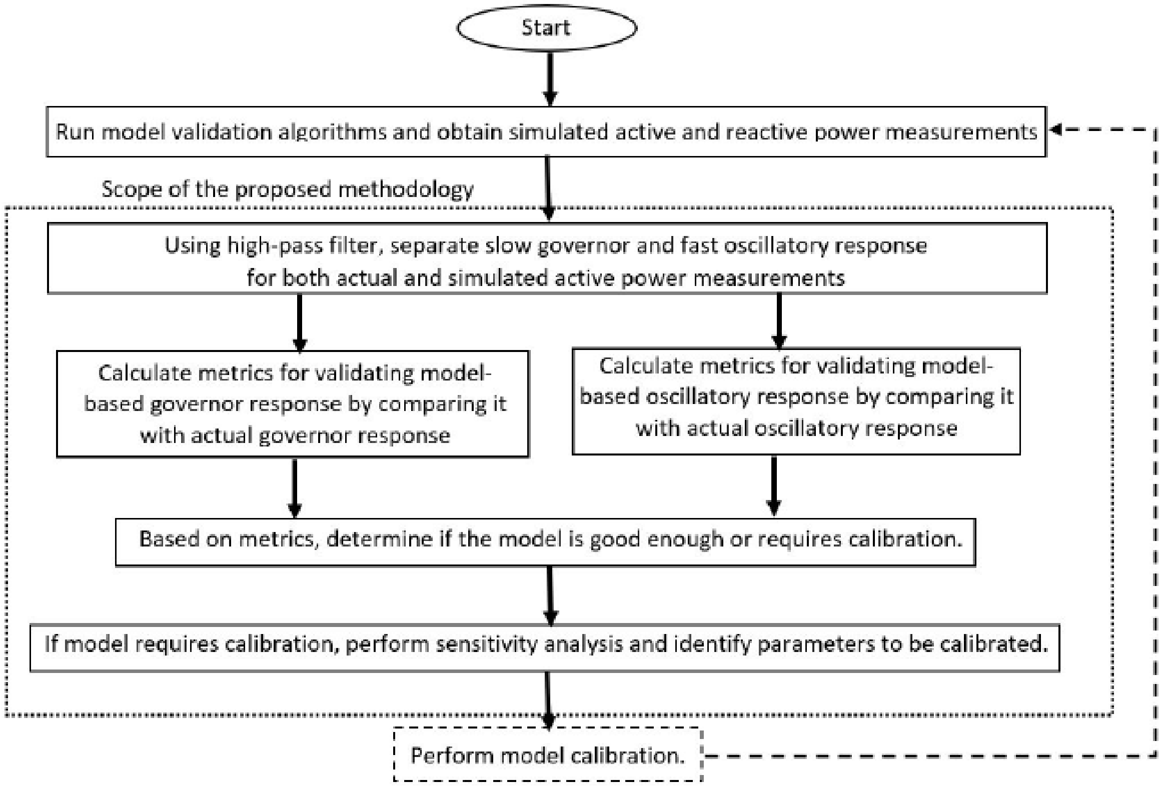


Figure 2-4. Flowchart of the methodology for calculating advanced performance metrics described in [24].

Methodology: The calculation of the set of metrics proposed in [24] consist of two main steps:

Step-1: Separating governor and oscillatory response

During system faults, generator dynamic response can be broken down into two components, one is the slow governor response and the other fast oscillatory response. The generator oscillatory response is determined by system modes and for that reason the frequency range of this response lies between 0.1 and 2.0 Hz. Therefore, the slow governor response and the oscillatory response can be separated by passing the generator response through a high-pass filter having a cut-off frequency of less than 0.1 Hz as illustrated in Fig. 2-5. The governor response is then obtained by taking the difference of the generator response and the oscillatory response and passing the resultant signal through median filter to smooth out any oscillatory components present in the signal. This is the first and the important step in calculating proposed metrics and performing sensitivity analysis.

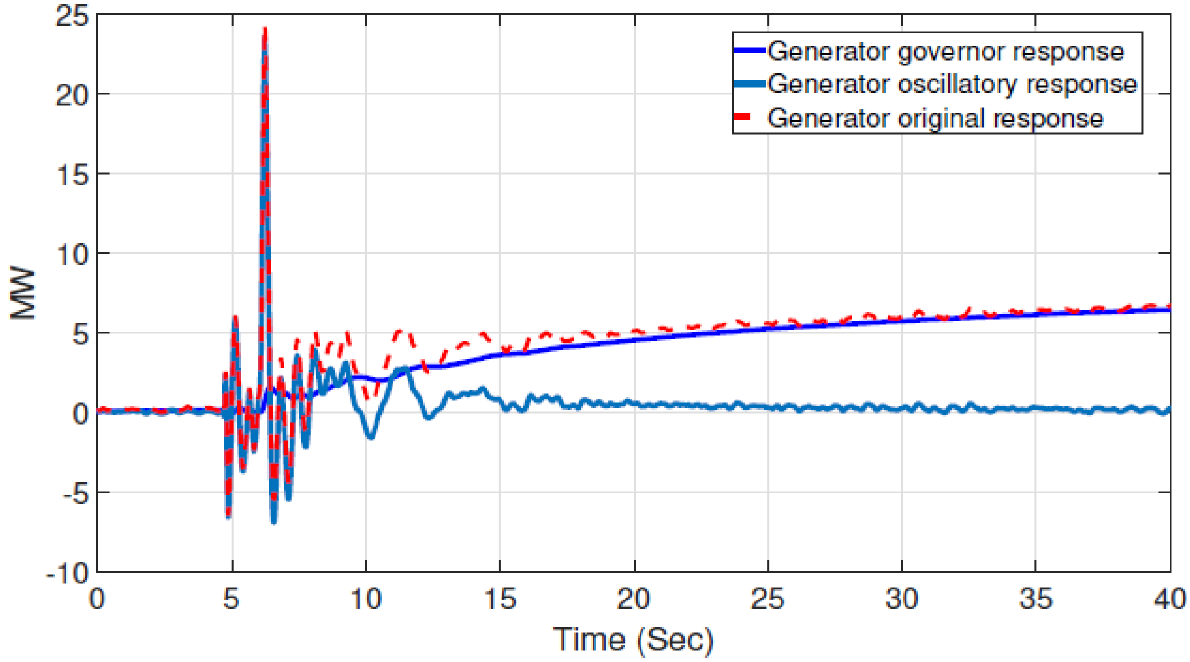


Figure 2-5. Generator response and the corresponding decoupled oscillatory and governor response[24].

Step-2: Calculation of performance metrics

In the second step, metrics are calculated for the separated governor and oscillatory response of the active power.

Metrics for oscillatory response: The metric for validating generator oscillatory response is calculated based on the properties of the oscillatory modes observed in the PMU and simulated measurements. Two metrics are proposed in [23][24] for validating generator oscillatory response, one quantifying magnitude similarity and the other quantifying phase similarity of oscillatory modes. The metric for magnitude, Osc_m, incorporates any discrepancy associated with initial amplitude, damping-ratio, and frequency of system modes between the model-based response and actual response. The metric for phase, Osc_{ph}, calculates any phase difference between the two signals. In the proposed methodology, first the system modes and their mode shapes are estimated separately for both actual and simulated measurements. In [23][24], even though Prony's method was used for estimating system modes, any other method can also be used. Once the modes are estimated, for modes common to both of these signals that have close frequencies, error in estimated magnitude (refers to the estimated contribution of a mode to the signal calculated using that mode's frequency, damping ratio and initial amplitude) and initial phase is calculated for each mode. In the next step, a single magnitude error metric and a single-phase error metric are calculated by taking a weighted average of the corresponding errors between each mode using the energy of the mode as a weight factor. For modes that are observed in actual measurements, but not in the simulated measurements, a maximum error of 1 is assigned to that mode.

Metrics for governor response: Based on the step-response characteristics of a system, several metrics are defined to validate the model-based governor response by comparing it with the actual governor response. Each metric looks into a specific aspect of the governor response, which are as follows:

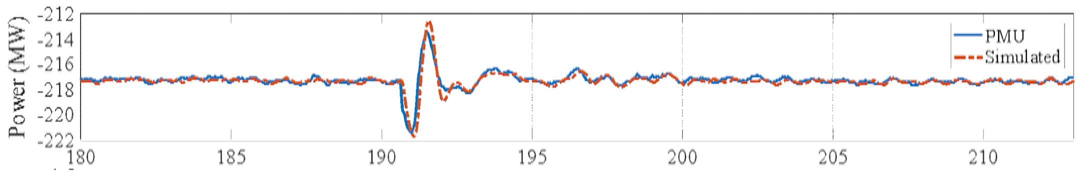
1. Delay (G_d): Obtained by taking the difference of the time taken by the model-based and actual governor response to reach 10% of their respective peak value with respect to a common time reference.
2. Peak value (G_P): Obtained by taking the difference of the peak value of the model-based and actual governor response.
3. Peak time (G_{PT}): Obtained by taking the difference of the time taken by the model-based and actual governor response to reach peak-value
4. Steady-state error (G_{SS}): Obtained by taking the difference of the final value of the model-based and actual governor response
5. Rise time (G_{RT}): Obtained by taking the difference of the time taken by the model-based and actual governor response to change from 20% to 90% of their respective peak-value.

Once these error metrics are calculated for given actual and simulated measurements, performing sensitivity analysis with respect to these metrics can also help with the identification of key model parameters that need to be calibrated in a systematic manner.

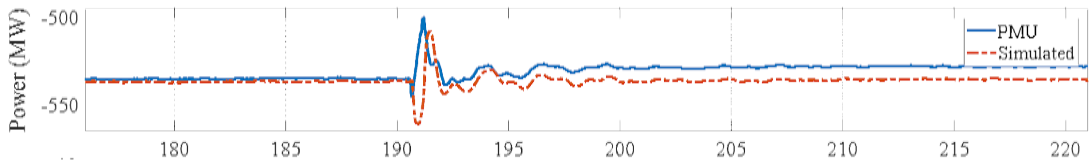
More detail on the calculation of these metrics and performing sensitivity analysis using these metrics can be found in [24]. Results for both simulated measurement- and real-world measurement-based use-cases are also included in [24] that show the effectiveness of the proposed metrics in analyzing the model validation results in a comprehensive and accurate manner, and in identification of key model parameters that require calibration.

2.2.3 Correlation and Frequency Domain Metrics

Quantitative metrics have been recently proposed in [35] for comparing the simulated system responses from the model with the PMU archived system responses. These metrics are useful for assessing whether the model simulations are reasonable and whether they match well with the actual system behavior. Three different metrics are proposed for the comparison: 1) correlation, 2) magnitude similarity in frequency domain, and 3) phase similarity in frequency domain, and an average of the three metrics could be useful for providing the user with a “score” for the match or for the fitness. The frequency domain scores in different frequency ranges can be used for assessing whether the error is contributed by specific system components such as speed governors versus exciters. Fig. 2-6 (a) and (b) show two comparisons between simulated and archived MW responses of power plants with scores of 91% and 50% respectively for the match between the simulated and archived responses.



(a) Match score of 91%



(b) Match score of 50%

Figure 2-6. Example of scores for comparison of archived and simulated responses for two different generators [35]

3 Advanced Model Calibration

3.1 Advanced Parameter Selection

The standard power system component model library involves a large number of interpretable parameters; on the other hand, the measured grid disturbance event may not carry rich enough dynamic information to excite all the parameters in the various system component models. The combined effect of many parameters with low richness information in the measurement may lead to an ill-conditioned parameter estimation problem [39].

Successful model parameter identification using measurements depends on the nature of influence of model parameters on measured quantities. If a parameter has a very weak effect on the measured output, successful estimation of such a parameter is unlikely because its effect may not be accurately quantified. If the effects of certain parameters on measured output are nearly linearly dependent, successful estimation of such parameters is unlikely because the individual parameter effects may not be distinguishable.

The presence of parameters with weak and/or nearly linear dependency is manifested by non-unique solutions to the estimation problem for different initial parameter values. This is observed in the optimization problem of minimizing the difference between measured quantity and model output. It has been verified that selecting the "well-conditioned" model parameter subset for tuning, while fixing the other "ill-conditioned" model parameters to prior estimates, have greatly enhanced the quality and reliability of the estimation results [39]. Furthermore, it may be beneficial to identify sets of parameters with strong and linearly independent effects across qualitatively different disturbances, and to identify the "best" disturbances to use for model tuning [40].

3.1.1 Trajectory Sensitivity Approach

The trajectory sensitivity approach for power plant model parameter identification can be traced back to the 1990s [42][43]. The trajectory sensitivity matrix can be calculated by performing playback simulation, perturbing the value of each parameter, and taking the difference between the two responses over the size of the parameter perturbation [44]. The readers can refer to the paper [45] for one implementation example. Due to its simplicity for shortlisting the key parameters to be tuned and potential use in the optimization iteration step, this approach has been widely used in most model calibration studies.

3.1.2 Approach Informed by Governor and Oscillation Indices

Based on the calculated metrics described in Section 2.2.2, if it is determined that the model needs calibration, the next step is to perform the sensitivity analysis to identify those parameters that need to be tuned. In [24], the sensitivity analysis was carried out with respect to the proposed metrics that quantified the impact of each model parameter, especially on error metrics that showed mismatch between actual and model-based responses. For example, if it is identified that the generator model response has mismatch with the actual governor's response time, given by delay metric, then the sensitivity analysis needs to be performed with respect to the delay metric to identify parameters that have significant impact on delay metric. In this approach, instead of selecting any parameters that can affect the generator response, selection of model parameters is limited to the ones that affect only those error metrics for which significant mismatch was observed between the actual and simulated measurements.

3.1.3 Global Sensitivity Approach

As illustrated in trajectory sensitivity approach, the sensitivity matrix can be calculated by performing playback simulation where each parameter is perturbed. However, to construct a comprehensive and precise matrix, this process can be time-consuming due to massive simulations in the presence of high-dimension parameters. Also, the identification of suspicious parameters can be challenging if the parameters are strongly dependent on each other.

To efficiently build the parameter sensitivity matrix, the generalized polynomial chaos-based analysis of variance (ANOVA) method, a global sensitivity analysis (GSA) approach is advocated. It can be formulated as:

$$y = a_0 + \sum_{i=1}^{N} a_i \phi\left(\xi_i\right) + \sum_{i=1}^{N}\sum_{j=1}^{i} a_{i,j}\phi\left(\xi_i, \xi_j\right) + \sum_{i=1}^{N}\sum_{j=1}^{i}\sum_{r=1}^{j} a_{i,j,r}\phi\left(\xi_i, \xi_j, \xi_r\right) + \dots \tag{1}$$

where, y is the measurement/system response and $\xi = \left[\xi_1, \xi_2, \dots \xi_N\right]$ is random variable vector containing parameters and state variables; $\phi_i\left(\xi_1, \xi_2, \dots \xi_N\right)$ is polynomial chaos basis and a_i is the *i*-th polynomial chaos coefficient. Then, the sensitivity for each parameter can be calculated:

$$T_{j_1} = \frac{\sigma^2\left(g_{j_1}\right)}{\sigma^2(g)} = \frac{a_i^2 E\left[\phi\left(\xi_i\right)^2\right] + a_{i,j}^2 E\left[\phi\left(\xi_i, \xi_j\right)^2\right]_{i=j}}{\sigma^2} \tag{2}$$

where σ^2 means the variance. The dependence between each parameter can also be obtained:

$$T_{j_1,j_2} = \frac{\sigma^2\left(g_{j_1,j_2}\right)}{\sigma^2(g)} = \frac{a_{i,j}^2 E\left[\phi\left(\xi_i,\xi_j\right)^2\right]_{i\neq j}}{\sigma^2} \tag{3}$$

By using the above two equations, the sensitivity of each parameter to the response y can be analytically determined. [46] uses ANOVA to detect suspicious parameters and investigate the parameter dependence. Combined with unscented Kalman filter (UKF), sensitivity of each parameter is utilized as the weight for parameters' calculation, leading to good robustness against non-Gaussian noise. Regarding the dependence between each parameter, several parameters that are identified to have strong dependence will be divided into two different groups, and updated group by group, avoiding the local optimality of parameter estimation due to the dependence.

Table 3-1 Comparison results of parameter estimation in pu with and without adaptive weights, where the values shown in the brackets represent their standard deviations.

True parameters	Wrong values	Estimated without adaptiveness	Estimated with adaptiveness
X_d=0.1218	0.2	μ=0.1938 (σ=0.036)	μ=0.127 (σ=0.004)
M=0.7568	0.9	μ=0.7628 (σ=0.015)	μ=0.7617 (σ=0.012)
$T'd$=29.7	32	μ=32.5 (σ=0.14)	μ=32.0 (σ=0.16)
KF=0.063	0.1	μ=0.0622 (σ=0.001)	μ=0.0623 (σ=0.002)
KA=60	40	μ=70.20 (σ=10.8)	μ=60.48 (σ=0.32)
RD=1	0.3	μ=1 (σ=0.001)	μ=1 (σ=0.001)

Case studies are carried out on the IEEE 39-bus system to evaluate the effectiveness and robustness of the proposed method. In Table 3-1, X_d, T_d' and M are generator parameters; KF and KA are exciter parameters; RD is the governor parameter. μ and σ are respectively the mean value and standard deviation of each parameter. Only the results for Adaptive generalized maximum-likelihood (GM-UKF) with and without the adaptive weights are derived from the sensitivity analysis. It can be found that due to the strong dependency among parameters KF , KA and X_d, the direct augmenting 6 parameters with system states for joint estimation will yield local optimal solutions of KA and X_d. By contrast, with the derived adaptive weights and the strategy of breaking down the parameter dependence, the proposed adaptive GM-UKF can achieve much higher accuracy of calibrating the suspicious parameters. By further looking at the standard deviation for each estimated parameter, it is observed due to the lack of capability in dealing with correlated parameters, the method without adaptive weights has much larger variances. This is expected as the searching space for it is quite random and as a result, a good solution may be obtained sometimes. However, the solutions are not satisfactory most of the time. By contrast, with the adaptive weights, the correlations among erroneous parameters can be addressed and the search space is around the optimal values, yielding small standard deviations of the estimated parameters.

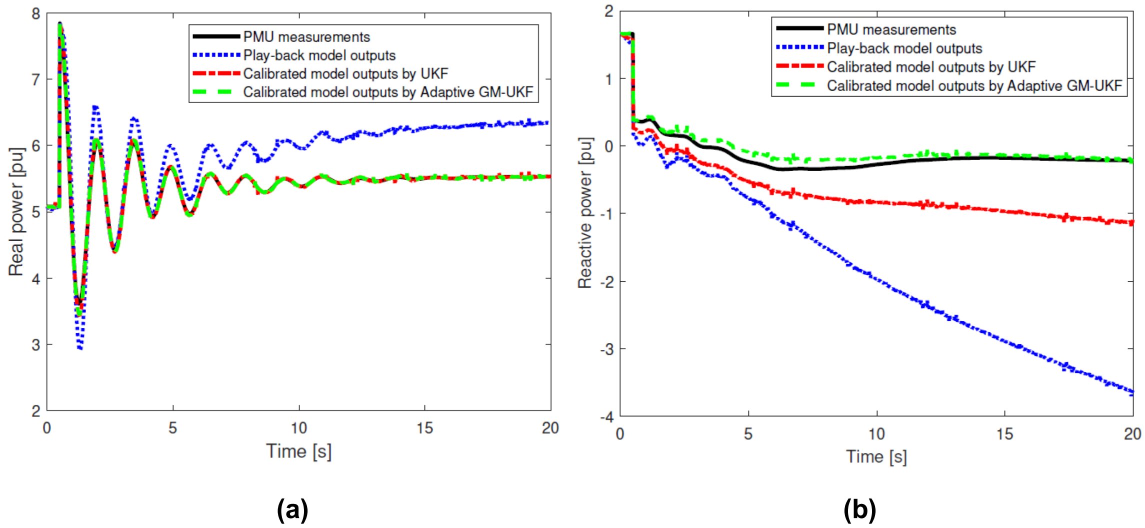


Figure 3-1. Dynamic responses under non-Gaussian noise. (a) Real power response using calibrated parameters with non-Gaussian noise; (b) Reactive power response using calibrated parameters with non-Gaussian noise.

In the presence of non-Gaussian noise, as shown in Fig. 3-1, although UKF does not diverge, its results are strongly biased. The results are expected because UKF is derived based on Gaussian assumption and the non-Gaussian noise significantly deteriorates the estimation efficiency. By contrast, thanks to the adaptiveness and robustness of GM-UKF, it can filter out the non-Gaussian noise and its performance is slightly affected.

3.1.4 SVD Based Methods

Further analysis on the trajectory sensitivity matrix can reveal the information of the magnitude and dependency of parameter sensitivity. A parameter with high sensitivity magnitude and low dependency with other parameters would be a great candidate to alter the simulation response to match the real measurement. In contrast, a parameter with low sensitivity magnitude or high dependency with other parameters would be a poor candidate for the model calibration purpose.

Singular value decomposition (SVD) is a useful tool for extracting such information based on the trajectory sensitivity matrix. Let **A** be the *Nt*-by-*Np* trajectory sensitivity matrix, where *Nt* is the number of time samples, *Np* is the number of parameters.

$$A = S_y = \left[\frac{\partial y}{\partial \theta}\right]_{NtXNp} = \left[\frac{\partial y(t_1)}{\partial \theta_1}\ \frac{\partial y(t_1)}{\partial \theta_2} \cdots \frac{\partial y(t_1)}{\partial \theta_{Np}}\ \frac{\partial y(t_2)}{\partial \theta_1}\ \frac{\partial y(t_2)}{\partial \theta_2} \cdots \frac{\partial y(t_2)}{\partial \theta_{Np}} \vdots\vdots\vdots\vdots\ \frac{\partial y(t_{Nt})}{\partial \theta_1}\ \frac{\partial y(t_{Nt})}{\partial \theta_2} \cdots \frac{\partial y(t_{Nt})}{\partial \theta_{Np}}\right] \tag{4}$$

The SVD of **A** is,

$$A = USV^T \tag{5}$$

where the *Nt*-by-*Nt* matrix **U**'s columns are the left singular vectors u_i's (i = 1…*Nt*) of **A**; the *Nt*-by-*Np* matrix **S**'s upper left diagonal elements are the singular values σ_i's (i = 1…*Np*) of **A**; the *Np*-by-*Np* matrix **V**'s columns are the right singular vectors v_i's (i = 1…*Np*) of **A**.

After applying SVD to the trajectory sensitivity matrix **A**, the magnitude of the parameter sensitivities can be calculated as

$$Ranking\ Factor = M_{sen} = \sum_{i=0}^{Np} \sigma_i^2 V_i^2 \tag{6}$$

The dependency of parameter sensitivities can also be calculated from the result of the SVD. As indicated in [39], the dependency of parameter sensitivities is contained in the right singular vectors corresponding to zero singular values (i.e. null space of sensitivity matrix). For each column of the null space, if there is more than one element larger than a predefined threshold, the parameter sensitivities corresponding to those elements have dependencies. This can be explained by the definition of null space. Every right singular vector in the null space represents a zero-mode, the value of elements in that right singular vector indicate the contribution of parameter sensitivities to that zero mode. If there are multiple large elements, they contribute to the zero mode by offsetting each other as a linear combination.

A dependency index can be defined by counting the large elements in the right singular vectors in the null space. A spiral graph [47], an example of which is provided in Fig. 3-2, can be used to visualize both the magnitude and dependency of parameter sensitivities. The parameters at the vertices of the spiral are arranged in such a way that the sensitivity magnitude decreases in counter-clockwise direction. Also, the connectivity between any two parameters suggests their sensitivities have dependencies.

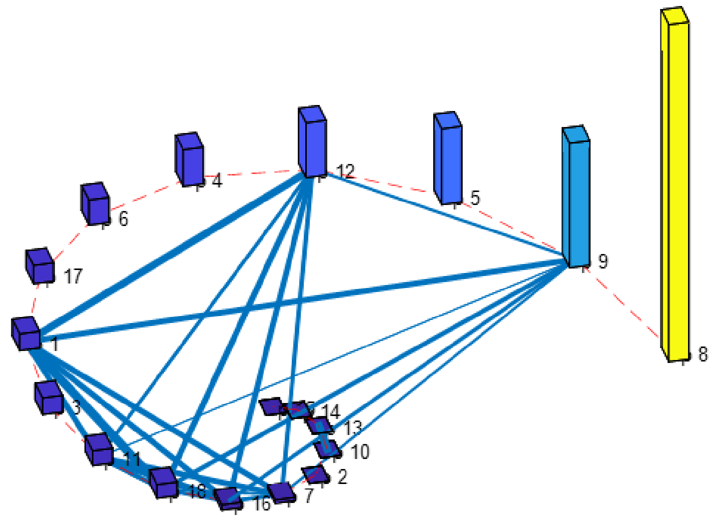


Figure 3-2. Spiral graph showing magnitude and dependency of parameter sensitivity [47].

In reference [39], a similar idea was used to select a subset of the model parameters for tuning purpose. By partitioning the eigenvalues of the Hessian Matrix into "active" columns and "redundant" columns, the parameters corresponding to the "active" columns, also called "well-conditioned" parameters, are subject to model tuning. In an exemplified case study with 9 parameters for a synchronous generator, the 9 machine model parameters were partitioned into 7 well-conditioned parameters and 2 ill-conditioned ones. The estimation results by fixing the 2 ill-conditioned parameters to their default values lead to time saving by a factor of 10 with a more accurate parameter results compared to the full-order cases (all 9 parameters are estimated). On the other hand, a huge mean deviation from the "true" values are observed showing estimating all model parameters can yield extremely unreliable results.

3.1.5 Similarity Based Methods

The sensitivity trajectory-based approach exploits the effect of the model parameter to the model response at the operating point of the given event. Considering the motivation of the model parameter screening is to reduce the residual between the simulation response and the measurement output, it is desirable to leverage the geometric relationship between model parameter sensitivity to the actual measurement. Specifically, it is reasonable to expect model parameters whose sensitivity with less "orthogonal angles" to the residual direction are more likely to reduce the residual [15].

In contrast with the traditional trajectory sensitivity, a cosine similarity between the response residual r(t) and the parameter sensitivity vector (Jacobian column) is used to represent the significance of each parameter.

$$\alpha = \{\frac{r(\theta_k)\cdot S_{yk}}{\|r(\theta_k)\|\ \|S_{yk}\|}\} \tag{7}$$

The parameters will be ranked based on their magnitude of the angle α. The α for an ideal model parameter θ_k is zero, but that is generally unachievable. The parameter to be selected for further tuning can be specified as the top N ranked parameter or can be those whose angle is smaller than a predefined value such as 5°.

3.1.6 Empirical Gramian Based Method

The trajectory sensitivity method is only locally defined for one operating point and the nonlinear behavior of the generator model is inevitably lost. In [50] the empirical Gramian approach is adapted to analyze the sensitivity of the outputs to parameters. Specifically, an empirical parameter sensitivity Gramian (EPSG) is defined which perturbs the parameters to reveal the parameter-output behavior. Note that EPSG is defined for a reasonable region of the parameters directly for the original nonlinear dynamical model and can thus better reflect the identifiability of the parameters. The following sets are defined for EPSG:

$$\begin{gathered} T = \{T_1, \cdots, T_r;\ T_r \in R^{v \times v}, T_j^T T_j = I_v, j = 1, \cdots, r\} \\ M = \{c_1, \cdots, c_s;\ c_m \in R, c_m > 0, m = 1, \cdots, s\} \\ E = \{e_1, \cdots, e_v;\ standard\ unit\ vectors\ in\ R^v\}, \end{gathered} \tag{8}$$

where T defines the initial parameter perturbation direction, r is the number of matrices for perturbation directions, M defines the perturbation sizes, s is the number of perturbation sizes for each direction, I_v is an identity matrix with dimension v, and E defines the parameter to be perturbed.

For the ith parameter, α_i, in the nonlinear power system model, with fixed initial states x_0, EPSG can be defined as:

$$W^{con}(\alpha_i) = \sum_{j=1}^{r} \sum_{m=1}^{s} \frac{1}{rsc_m^2} \int_0^\infty \Phi^{jm}(t) dt, \tag{9}$$

where $\Phi^{jm}(t) \in R^{o \times o}$ is given as $\Phi^{jm}(t) = (y^{jm}(t) - y_0^{jm})(y^{jm}(t) - y_0^{jm})^T$, y_0^{jm} refers to the outputs corresponding to the unperturbed initial parameter α_0, and $y^{jm}(t)$ is the output of the nonlinear system under parameter $\alpha^{jm} = \alpha_0(1 + c_m T_j e_i)$.

The discrete form of EPSG can be defined as:

$$W(\alpha_i) = \sum_{j=1}^{r} \sum_{m=1}^{s} \frac{1}{rsc_m^2} \sum_{k=1}^{K} \Phi_k^{jm} \Delta t_k, \tag{10}$$

where $\Phi_k^{jm} \in R^{o \times o}$ is given by $\Phi_k^{jm} = (y_k^{jm} - y_0^{jm})(y_k^{jm} - y_0^{jm})^T$, y_k^{jm} is the output of the nonlinear system at time step k corresponding to the parameter $\alpha^{jm} = \alpha_0(1 + c_m T_j e_i)$, and $\Delta t_k$is the time step size.

The identifiability of the ith parameter can be indicated by the trace of $W(\alpha_i)$. Through numerical experiments, it is found that trace as a metric that measures the overall identifiability of a parameter is a good choice among several available metrics [50].

Note that some parameters do not change much during the generator lifetime based on engineering experience. If that is the case those parameters will not be calibrated even if they are identified as critical parameters or prior distributions for those parameters can be properly assigned based on this information (i.e., for uniform prior distribution the width of the interval can be reduced for those parameters and for normal prior distribution the standard deviation can be reduced).

3.2 Advanced Model Parameter Tuning

A power system simulation engine can be represented by an Ordinary Differential Equation (ODE) model:

$$\begin{gathered} \dot{x} = f(x, u, p) \\ y = h(x, p) \end{gathered} \tag{11}$$

where x, u, p, and y are state, input, parameters, and outputs, respectively. In playback mode, recorded excitation or input data u such as voltage and frequency are played into an ODE or a Differential-Algebraic system of Equations (DAE) model of the power system component and the model response or output y such as real and reactive power is to be computed. The problem is to seek a procedure to determine a set of parameters p such that the playback simulation of the model when excited with measured input data u_m, produces a simulated response that is close in some sense to the recorded response y_m.

Note that one typical approach is to treat the model parameters p as state variables to cast the problem as a standard nonlinear state estimation problem, where w is process noise that accounts for input noise and model mismatch, and v is the measurement noise:

$$\begin{gathered} (p, x\ ) = (0\ f(x, u, p)\ ) + w \\ y = h(x, p) + v \end{gathered} \tag{12}$$

To solve this nonlinear estimation problem, particle filtering, extended, ensemble and/or an unscented Kalman filter may be used.

According to another approach, the measured input/output data (u, y^m) may be used by a power system component model and an optimization-based approach to create the estimation parameter ($p*$). In this case, the following optimization problem may be solved:

$$\|y^m - \hat{Y}(p)\|^2 \tag{13}$$

The system may then compute output as compared to parameter Jacobian information and iteratively solve the above optimization problem by moving parameters in directions indicated by the Jacobian information.

Machine learning methods have also been applied for parameter calibration. Its main idea is to extract the hidden relationship between dynamic response and parameters. The following sections will provide some examples for the three approaches.

3.2.1 Estimation-Based Approach

The model parameter calibration or tuning can be cast as a state estimation problem. The model parameters are treated as state variables, with a process noise accounting for input noise and model mismatch. Reference [49] has presented the pioneering work to apply Kalman Filter technique (including Extended Kalman Filter) to estimate the generator model parameters using the field test data.

The Kalman filter method is a recursive data processing algorithm, which estimates the states of a dynamic system with minimum mean-square error recursively over time using incoming observations (measurements) and a dynamic system model. The Kalman filter method typically has two recursive steps, a prediction step and a correction step, as shown in Fig 3-3. In the prediction step, the measurements at step k are predicted based on their values at the previous step, k−1, using the system dynamic equations. In the correction step, the Kalman gain is computed. With the information from the latest obtained measurements and the Kalman gain, the estimated states can be updated. Once the updated states are obtained, the algorithm will move to the next time step. Note that the critical element for the Kalman Filter method is how to compute the Kalman gain, and thus the Kalman filter method has many variation methods based on the different approach to compute the Kalman gain, such as extended Kalman Filter method (EKF), Ensemble Kalman Filter (EnKF) method, etc.

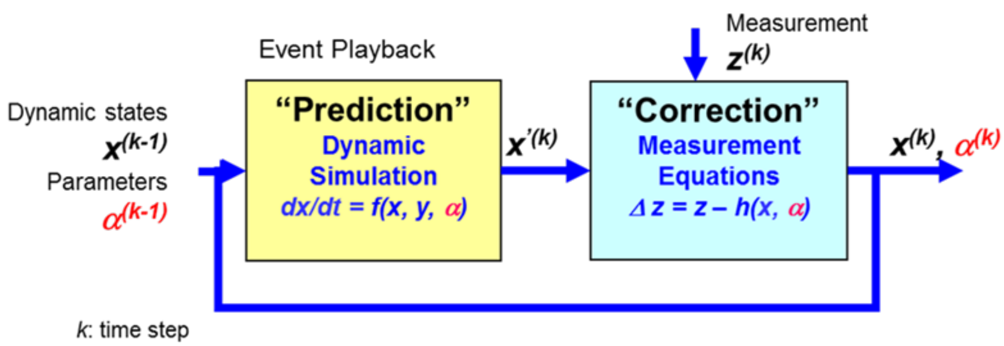


Figure 3-3. The basic Kalman Filter algorithm

Reference [44] proposed an EKF-based and sensitivity trajectory-based method to calibrate the parameters of the generation plant. It utilizes a perturbation-based method to compute the linearized system transition matrix for the dynamic system of the generation plant, and the Kalman gain is computed based on the linearized system transition matrix.

The method in [44] is applied to calibrate parameters of a WECC generation plant. Table 3-2 lists the original values and the estimated values for the calibrated parameters. Fig. 3-4 shows a comparison between the real and reactive power measurements generated from the generation plant model with calibrated parameters and the real PMU measurements. It is seen that the model outputs based on the calibrated parameters match the PMU measurements very well.

Table 3-2 The original values and the estimated values for the calibrated parameters of the generation plant.

		Before calibration	After calibration
Machine	Inertia constant (H)	3.50	4.92
Exciter	Gain (Ka)	360	309
PSS	Gain (Kqs)	5	0.25

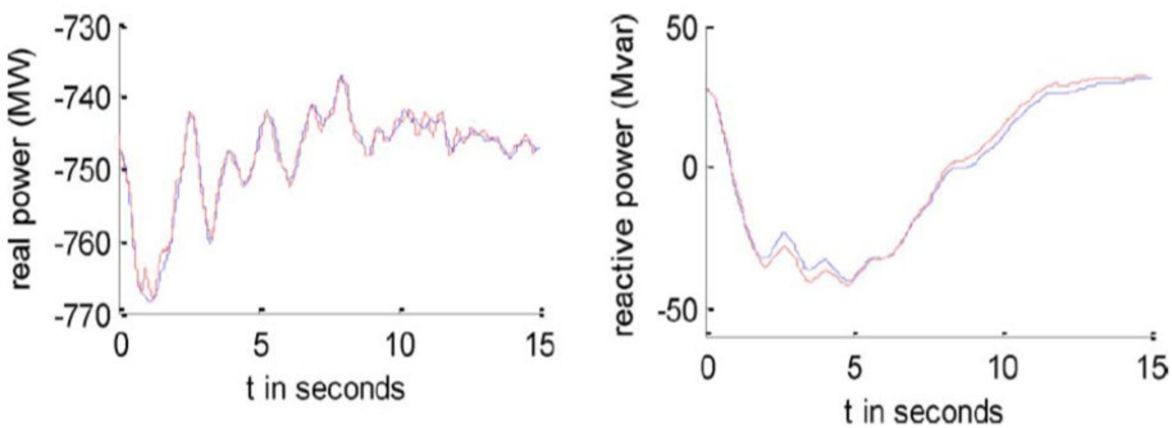


Figure 3-4. Comparison between the real and reactive power measurements generated from the generation plant model with calibrated parameters and the real PMU measurements.

Reference [45] proposed an EnKF-based and sensitivity trajectory-based method to calibrate the parameters of the generation plant. The EnKF is a sequential Monte Carlo implementation of the Kalman Filter method, which introduces an ensemble of samples to represent and propagate the probability density function (PDF) of the state variables. The PDF can be approximated with high accuracy by using a large number of samples. The Kalman gain can be also computed efficiently based on the ensembles of the state variables and the system dynamic equations. Compared with Kalman Filter algorithms, the EnKF algorithm has the advantages that it eliminates the need to derive Jacobian matrices for nonlinear models, and thus simplifies the complexity of algorithm implementation, and accelerates computation speed.

The method in [45] is applied to calibrate parameters of a WECC generation plant. Table 3-3 lists the original values and the estimated values for the calibrated parameters. Fig. 3-5 shows a comparison for the real and reactive power measurements generated from the generation plant model with uncalibrated parameters, calibrated parameters, and the real PMU measurements. It is seen that the model outputs based on the calibrated parameters match the PMU measurements very well.

Table 3-3 The original values and the estimated values for the calibrated parameters of the generation plant.

Para-meter	Before Calibration	After Calibration	Para-meter	Before Calibration	After Calibration
L_d	0.55	0.53	R_{perm}	0.057	0.065
L_q	0.47	0.32	R_{temp}	0.24	0.54
T_{pdo}	4.3	6.17	T_r	1.2	2.0
H	4	5.22	T_5	15	10
T_c	2.1	1.85	T_6	0.035	0.037
T_b	2.5	3.5	A_1	0.035	0.036
K_a	60	73.5	K_s	32	21

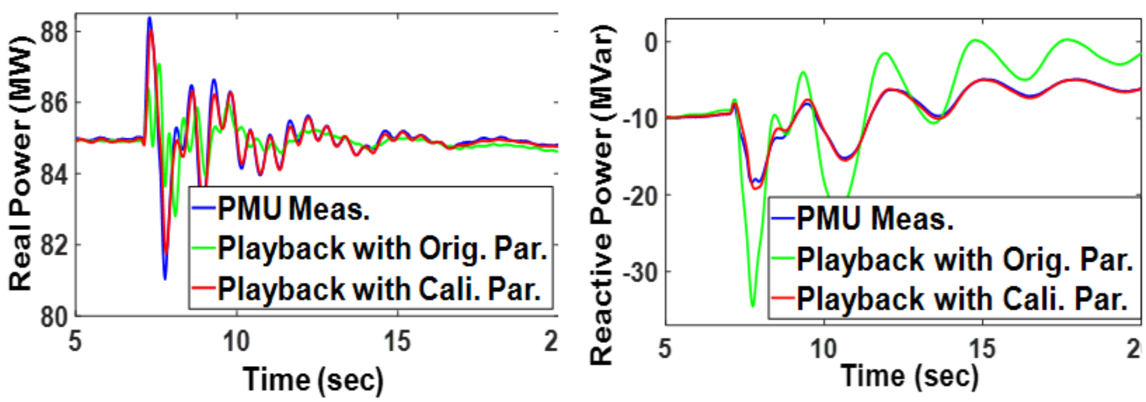


Figure 3-5. Comparison between the real and reactive power measurements generated from the generation plant model with original and calibrated parameters, as well as the real PMU measurements.

3.2.2 Optimization Based Approach

Reference [50] has presented the pioneering work to apply the weighted least square approach to estimate the generator model parameters for an aircraft generator, based on short circuit test. Then Juan [42] and Stephen [43] combined the trajectory sensitivity into the least square method for model calibration of a simulated synchronous generator together with exciter.

3.2.2.1 Efficient Trust Region Approach

Fig. 3-6 describes a Gauss-Newton Trust Region algorithm. S820 has definitions and initializations, S830 determines a step based on the trust region approach, and S850 provides the updated Jacobian information. One pass thru S820 and S850 is termed as an iterate. Each model evaluation involves calling a standalone external solver that performs the play-back simulation.

These calls are expensive, typically taking from 0.5 to 1 second to evaluate. If the system has 30 parameters to tune, about half a minute may be required just for a forward difference Jacobian call.

Thus, the number of function evaluations may be reduced, especially reducing the number of expensive Jacobian calls. To this end, reusing the Jacobian between successive iterates is a natural option, *i.e.* reuse the last computed Jacobian in determining the step so long as such steps produce sufficient decrease in objective function.

A further refinement can be done by performing certain rank-one updates to improve the accuracy of the Jacobian along the iterates (as opposed to just reusing the last computed Jacobian). The implemented algorithm might, according to some embodiments, perform a Broyden rank-one update on the last computed Jacobian and use these for calculating the search direction until the resulting steps provide sufficient reduction in objective. These iterations may be termed as "inexact" since they are not based on the exact Jacobian. If the reduction in objective from inexact steps is insufficient, a computation of Jacobian at the current point may be trigged and an "exact" iteration follows.

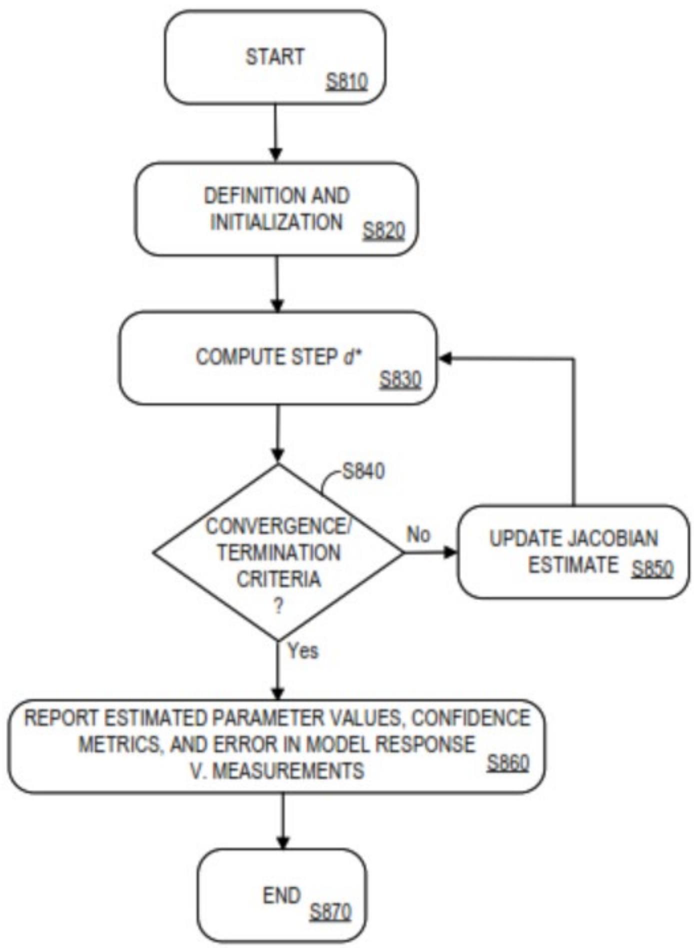


Figure 3-6. Trust region approach for model calibration.

3.2.2.2 Black-Box Optimization Based Approach

In [51] a synchrophasor measurement-based generator parameter calibration method is proposed by a black-box optimization approach with a stochastic radial basis function (RBF) surrogate model. Based on comparison between the outputs of the generator model with estimated parameters and the PMU measurements, an L_1 norm based objective function is defined for the black-box optimization problem, which is approximated by an RBF surrogate model. The prior information of the parameters is treated as constraints in the black-box optimization problem. The formulated black-box optimization problem is then solved by a Stochastic Response Surface Method (MSRSM). This method does not require an explicit objective function, can solve non-convex problems by building a global model of the objective function, and guarantees convergence to the global optimum from a theoretical standpoint if the number of iterations is large enough.

The method in [51] is applied to calibrate eight critical parameters of a generator. The prior of these parameters is assumed to follow uniform distributions. The mean values of the uniform distribution are set as 10% greater than their true values. The lower/upper bounds of the uniform prior distributions are chosen as 30% smaller/greater than the mean values. Table 3-4 lists the estimated values and the percentage errors. It is seen that the algorithm has good accuracy under high dimension. Fig. 3-7 shows a comparison between the real and reactive power measurements generated from the true parameters and the model outputs from estimated parameters. The model outputs based on the estimated parameters match the measurements very well, indicating that the estimated parameters are very accurate.

Table 3-4 Calibration of eight critical parameters [51]

Parameter	True value	Estimated value	% Error
K_{A}	125	125.47	0.4
T_{b}	3.86	3.84	0.5
a_{23}	1.102	1.110	0.7
T_{pdo}	5.40	5.45	0.9
X_{d}	0.57	0.575	0.6
T_{c}	0.90	0.90	0
H	5.40	5.38	0.4
T_{R}	2.40	2.43	1.5

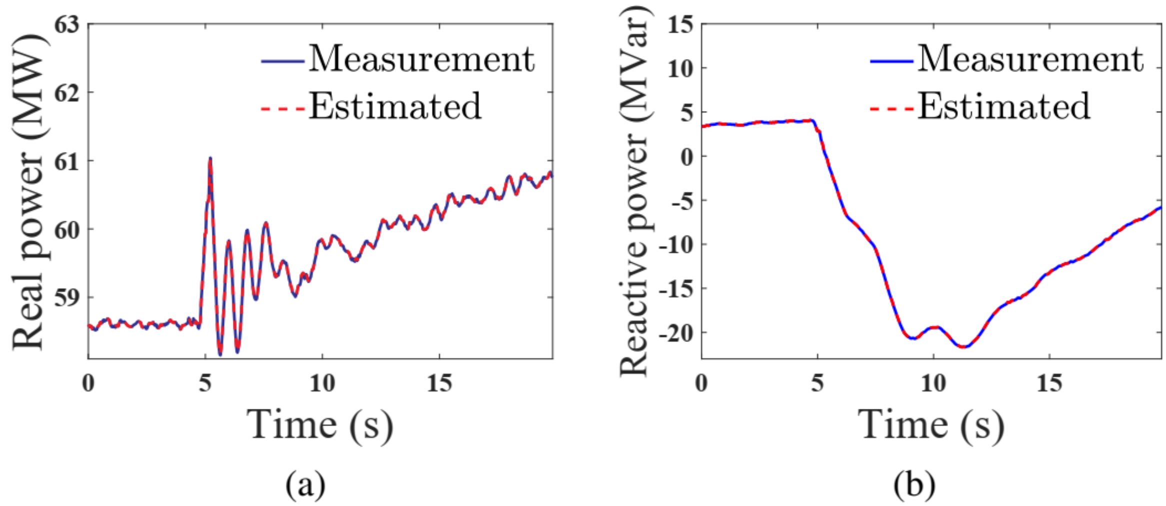


Figure 3-7. Measurements and the outputs for the model with the estimated parameters: (a) Real power; (b) Reactive power.

3.2.2.3 Approximate Bayesian Computation-Based Approach

In [50], a generator parameter calibration approach is proposed by adaptive Approximate Bayesian Computation with sequential Monte Carlo sampler (A-ABC-SMC). It is a likelihood-free method that does not directly explore the likelihood surface of the parameters but instead estimates the posterior distributions of the parameters by a simulation-based procedure, significantly improves the computational efficiency of ABC SMC through adaptive threshold sequences and perturbation kernel function, and carefully chosen distance function.

In the A-ABC-SMC algorithm, an L_1 distance function is used which is shown to have better performance than the other distance functions such as chi-squared distance function or the root mean square (RMS) distance function. Since dynamic simulation is the most computationally expensive part of the A-ABC-SMC algorithm, a properly chosen threshold sequence should reduce the total number of dynamic simulations. This requires a careful balance between the number of simulations in each iteration and a good approximation of the posterior distribution. Thus, an adaptive threshold sequence that considers the impact of all previous iterations is proposed. This addresses the inefficiency issue for the existing methods especially for later iterations and greatly improves the efficiency.

Also, it is noticed that utilizing the existing kernels ABC SMC may waste a lot of time in sampling the areas of low likelihood, and the acceptance rate could be very low and the algorithm may be stuck in local modes. For existing kernels, the algorithm will search around the particles located in low densities. Therefore, an adaptive perturbation kernel is proposed for which the focus is more on the space with higher densities, thus improving time efficiency.

The A-ABC-SMC approach provides the posterior distribution of the parameters based on the N particles in the last iteration, not just one point estimation. The maximum-a-posteriori (MAP), the mean of the N particles in the last iteration for each parameter is chosen as the estimated parameter. A-ABC-SMC is applied to calibrate fourteen critical parameters of a generator. The priors of these parameters are assumed to follow Gaussian distributions whose means are 10% greater than their true values and whose standard deviations are 20% of the mean values. Table 3-5 lists the estimated values, which indicate that the algorithm has good accuracy.

In real application, true model parameters are never known. It is possible that for different events the algorithm may provide different parameters. In addition to engineering judgment and experience that can help choose reasonable parameters, techniques can also be developed based on the available multiple events to help select the best parameters. For example, one can consider one event to estimate the parameters and use the other events to cross-validate the estimated parameters. Assume PMU measurements for three events are available. The parameters are estimated with each one of the three events and the L_1norm error for all three events is then calculated. The parameter set with the smallest average L_1 norm error is selected. Fig. 3-8 shows the model validation results under three events with the original parameters and the estimated best parameters. It is seen that the mismatch between model outputs and PMU measurements is significant under original parameters while the model outputs from the estimated parameters can match the PMU measurements very well for all three events.

Table 3-5 Calibration of fourteen critical parameters [50]

Parameter	True value	Estimated value	Error (%)
K_{S}	20	20.176	0.9
T_6	10	10.081	0.8
T_5	10	10.124	1.24
K_{A}	125	123.12	1.5
H	5.40	5.332	1.3
a_{23}	1.102	1.101	0.1
T'_{do}	5.40	5.352	0.9
T_{b}	3.86	3.970	2.9
T_3	0.15	0.150	0
T_1	0.15	0.151	0.6
T_{c}	0.90	0.909	1
X_{d}	0.57	0.566	0.7
X'_{d}	0.25	0.24	4
R_{T}	0.42	0.428	1.9

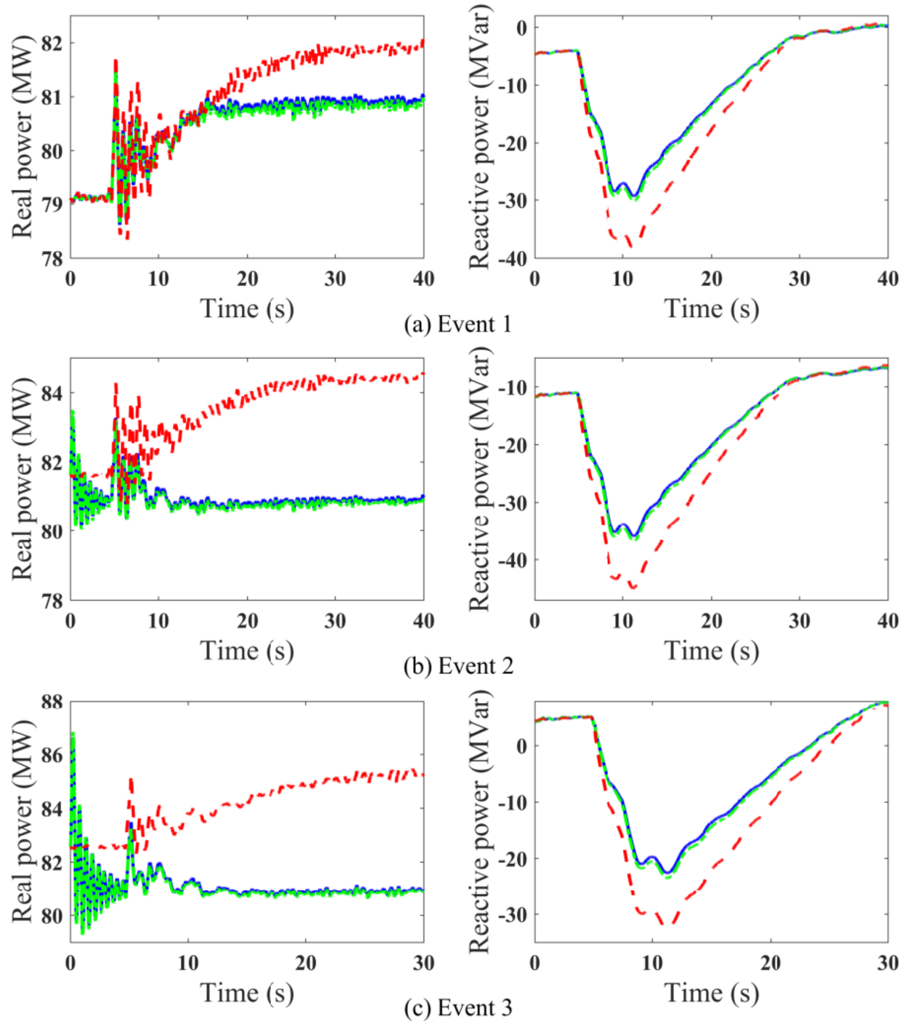


Figure 3-8. Model outputs before and after parameter calibration: (a) Event 1; (b) Event 2; (c) Event 3. Black curve is for PMU measurements, red curve is for model outputs before calibration, and green curve is for model outputs after calibration.

In terms of time efficiency, the calibration of two and fourteen critical parameters using A-ABC-SMC takes 1 and 10 minutes, respectively. By contrast, it takes at least 20 minutes for the two-parameter case and more than 2 hours for the fourteen-parameter case if using the existing threshold sequence schemes and perturbation kernel functions.

3.2.3 Machine Learning Based Approach

Machine learning methods have also been applied for parameter calibration. The main idea is to extract the hidden relationship between dynamic response and parameters. [53] is the first work to generate extensive simulation data to train a multi-output convolutional neural network model and predict a small number of generator parameters. [54] uses Q-learning-based method but only works for cases where a few parameters need to be calibrated. [55] further employs deep Q-learning-based method and performs well in low-dimensional cases under different events. However, the q-value affects the policy significantly, and a small change in the Q-value affects the policy a lot. [56] proposes soft actor-critic-based method and has a good performance for calibration, but it requires a lot of hyper-parameters tuning to converge. To address these issues, [57] proposes conditional variational autoencoder to find the model of conditional probabilistic distribution between synchrophasor measurements the critical parameters.

3.2.3.1 Q-Learning Based Approach

In [54] a Q-learning based method is proposed for parameter calibration. The implementation of this method is much more straightforward and reliable without the many complications in deep Q-learning, and it works well for the parameter calibration problem with a small number of parameters.

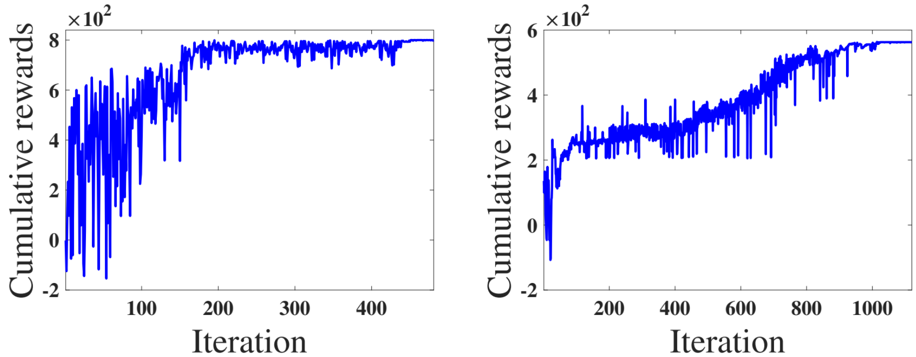
Q-learning is a model-free reinforcement learning (RL) algorithm with the goal of learning a policy to tell an agent what action to take under what circumstances. In Q-learning, an agent takes sequential actions at a series of states based on a state-action value matrix, Q-table, until reaching an ultimate goal. Let A and S be the action space and state space respectively. At each episode t, the agent observes a state $s_t \in S$ and chooses an action $a_t \in S$ based on policy π, which is a function that maps states into actions. As a consequence of taking action a_t, the agent receives a reward Rt defined as $R_t = R(s_t, a_t, s_{t+1})$ and observes the next state s_{t+1} of the environment. The RL framework considers the Markov decision process assumption, i.e. s_{t+1} is only conditioned by s_t and a_t and is sampled according to the transition probability $p(s_{t+1}|s_t, a_t)$. The above process is continued until the agent reaches the last episode, called the terminal state.

Let α_c be the vector of critical parameters and ϵ_s a discrepancy function of the simulated measurements and real measurements. The pseudocode of the Q-learning based parameter calibration algorithm is shown below [54].

Algorithm : Q-Learning based method for estimating the parameter $\boldsymbol{\alpha}_c$

```
Set hyper parameters λ, γ, ε, N
Initialize experience pool 𝓞 as an empty set
Discretize the parameter space
𝓠 ← 0
for 1 ≤ t ≤ N do
    Start with an un-searched state with parameter α_c^*
    Generate data z from α_c^*: z ~ Model(α_c^*)
    Calculate discrepancy ϵ_s(z, z^*)
    Assign a reward to the state
    With probability ε, select a random action a_t;
  otherwise select a_t = arg max_{a∈A} 𝓠_t(s, a)
    Update 𝓠-value matrix
    Update 𝓞-value matrix
end for
```

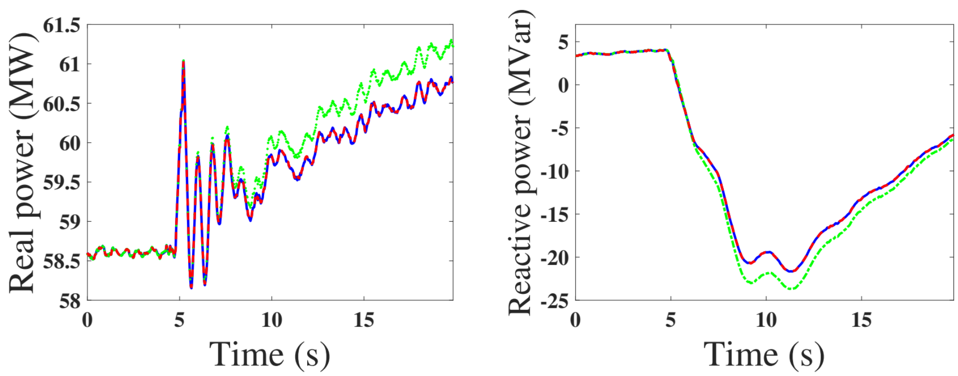
The Q-learning based method is applied to a power system to calibrate two and four critical parameters of a generator. Fig. 3-9a shows the cumulative rewards for calibrating two parameters. It is seen that the training converges after 421 iterations, which takes 5 hours. Fig. 3-9b shows the cumulative rewards for calibrating four parameters. The training converges after 1000 iterations which takes 8 hours.



(a) (b)

Figure 3-9. Cumulative rewards for two and four-parameter case: (a) Two parameter case; (b) Four parameter case [54].

Fig. 3-10 shows the results for real and reactive power under the estimated parameters and the parameters before the calibration. The parameters before calibration are 10% greater than the true values. Before calibration there is obvious discrepancy between the model outputs and the PMU measurements while with the estimated parameters the model outputs match the PMU measurements very well.



(a) (b)

Figure 3-10. Model performance before and after parameter calibration: (a) Real power; (b) Reactive power. Black curve is PMU measurements, green curve is for model before calibration, and red curve is for model after calibration [54].

3.2.3.2 Conditional Variational Autoencoder based Approach

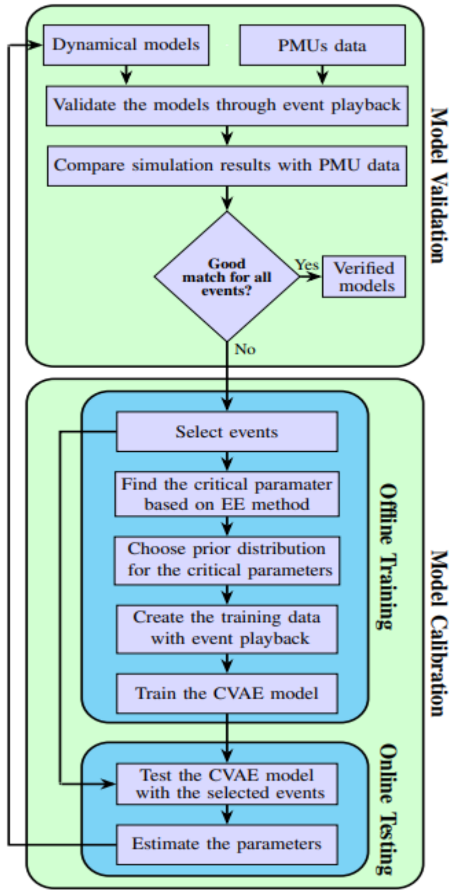


Figure 3-11. The overall framework for CVAE based model validation and parameter calibration [57].

Based on synchrophasor measurements, [57] has proposed a method using conditional variational autoencoder (CVAE) to calibrate the parameters of power plant model, including synchronous machine, power system stabilizer, exciter and governor models. The overall framework is shown in Fig. 3-11. By employing elementary effects approach, the critical parameters with the highest identifiability can be accurately identified for the nonlinear power plant model, while the traditional sensitivity-based method cannot handle the nonlinear behavior well. Subsequently, with the distribution projection of CVAE, the conditional probabilistic distribution between synchrophasor measurements and the critical parameters can be fitted. Consequently, the critical parameters can be calibrated accurately.

Due to the application of CVAE, a likelihood function or state-space model of the generator is not required. Besides, thanks to the distribution projection function of CVAE, the high-dimension parameters can be easily handled and no prior probabilistic distribution of the parameters is required.

The proposed method is tested using PSS/E software. A PMU is installed at the 230-kV side of the substation with a sampling rate of 30 samples/s. Table 3-6 lists the top eighteen critical parameters identified by elementary effects approach and sensitivity-based approach for the same event. It can be seen that elementary effects approach can distinguish critical parameters more easily due to its capability of handling nonlinearity.

Table 3-6 Top Eighteen Critical Parameters Identified by Elementary Effects Approach and Sensitivity-based method

No.	Elementary Effects		Sensitivity	
	Parameter	$\mathrm{EE}(\alpha_i)$	Parameter	$S(\alpha_i)$
01	K_{S}	1	T_5	1
02	T_5	0.99	K_{S}	0.99
03	T_6	0.91	T_6	0.71
04	K_{A}	0.59	a_{23}	0.48
05	T'_{do}	0.48	K_{A}	0.45
06	H	0.41	T_{b}	0.43
07	T_{b}	0.37	T'_{do}	0.41
08	a_{23}	0.33	X'_{d}	0.38
09	T_c	0.32	H	0.37
10	X_{d}	0.31	T_{c}	0.33
11	T_1	0.18	A_1	0.33
12	T_3	0.17	X'_{q}	0.33
13	A_1	0.08	T_3	0.32
14	X'_{d}	0.07	T_1	0.31
15	X'_{q}	0.07	T_2	0.30
16	R_{T}	0.04	X''_{d}	0.29
17	R_{P}	0.01	X_{d}	0.29
18	T_{R}	0.01	T''_{d0}	0.28

In Table 3-7, π denotes the distribution; H^t and K_A^t are respectively the true value of inertia and AVR steady state Gain; U and N are respectively the uniform distribution and normal distribution; Γ and Λ are respectively the Gamma distribution and Weibull distribution. It is seen that the proposed method can accurately estimate the parameters under different prior distributions, and the largest errors for inertia H and AVR steady state Gain K_A are, respectively, 0.3 %, and 0.4 %.

Table 3-7 Parameter Calibration under Different Distributions

$H(H^{\mathrm{t}} = 5.40)$		$K_{\mathrm{A}}(K_{\mathrm{A}}^{\mathrm{t}} = 125)$	
$\boldsymbol{\pi}(\mathbf{H})$	**Estimated** (% Error)	$\boldsymbol{\pi}(\mathbf{K}_{\mathrm{A}})$	**Estimated** (% Error)
$\mathcal{U}(0.60,\ 12)$	5.401 (0)	$\mathcal{U}(13.8,\ 275)$	125.4 (0.3)
$\mathcal{N}(6,\ 1.8^2)$	5.41 (0.2)	$\mathcal{N}(137.5,\ 41.3^2)$	124.7 (0.2)
$\Gamma(5,\ 2)$	5.38 (0.3)	$\Gamma(5,\ 20)$	125.5 (0.4)
$\Lambda(4,\ 4)$	5.42 (0.3)	$\Lambda(3,\ 100)$	124.6 (0.3)

Table 3-8 lists the true and estimated values of 18 parameters. The prior distributions of these parameters for training the model are assumed to follow a uniform distribution. The estimation errors verify the well-trained model has acceptable accuracy for high dimensional parameters. The more detailed analysis can be seen in [57].

Table 3-8 Calibration under 18 Parameters

Param.	**True value**	**Estimated** (% Error)	**Param.**	**True value**	**Estimated** (% Error)
K_{s}	20	19.75 (1.1)	X_{d}	0.57	0.57 (0)
T_5	10	9.97 (0.3)	T_1	0.15	0.14 (4.1)
T_6	10	9.63 (0.3)	T_3	0.15	0.14 (4.3)
K_{A}	125	122.8 (1.7)	A_1	0.035	0.035 (0)
T'_{do}	5.4	5.34 (1)	X'_{d}	0.25	0.25 (0)
H	5.4	5.4 (0)	X'_{q}	0.32	0.32 (0)
T_{b}	3.86	3.81 (1)	R_{T}	0.42	0.42 (0)
a_{23}	1.102	1.107 (0.4)	R_{P}	0.01	0.01 (0)
T_{c}	0.9	0.87 (2.2)	T_{R}	1	0.99 (1)

3.3 Performance Validation Process and Metrics

Model Validation and Calibration (MV&C) is key to ensure the generator model compliant to the NERC standards with adequate accuracy. GE Grid Solution's current Model Validation module has only the playback simulation function and response comparison for user to determine whether the model is "acceptable" or not. No domain knowledge or intelligence is imbedded so far and no means to verify if the model or parameter is valid or compliant to NERC standard or case study metrics. The current Model Calibration Module using numerical curve fitting without adequate engineering guidance tends to provide overfitted parameter result, which should be avoided at all costs. Issue of nonunique set of parameters (leading to same curve fitting performance) may cause invalid model parameter values, which may even lead to an unstable system at some operating conditions. Without properly integrating power system and control system domain knowledge, the result from MV&C is difficult to gain industrial acceptance and trust.

GE proposed a new MV&C framework: Model Validation, Model Calibration and Post evaluation. The New Model Calibration includes dynamic features (phase shift, amplitude, and damping ratio) in the objective functions and parameter value reasonable constraints; the Post Evaluation includes another layer of check on model, parameter, response, and control stability check. The corresponding adaptive parameter or bounds adjustment is designed at the end of Post Evaluation so that post evaluation results can readjust the iteration with model calibration [58].

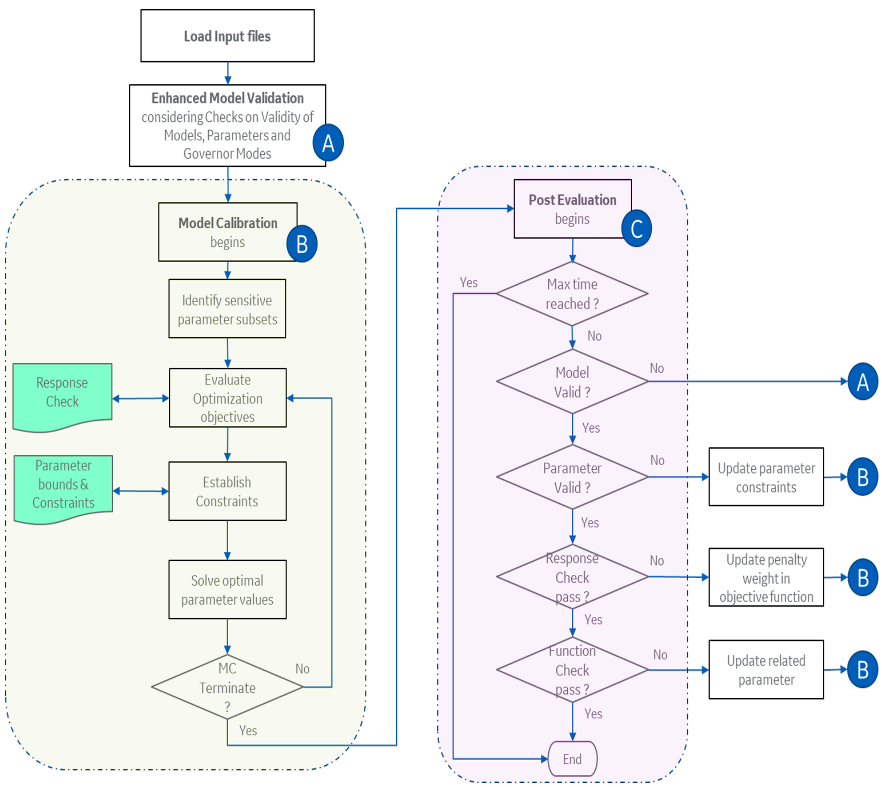


Figure 3-12. model validation and calibration including post processing [58].

Fig. 3-12 shows an exemplified flowchart of recommended solution to enhance MV&C with all four-reasonableness check. There are four major blocks in the new scheme: data ingestion, enhanced model validation (Block A), enhanced model calibration (Block B) and post evaluation (Block C).

The data flow starts from input data file ingestion including event data file, generator dynamic model file, generator's network file (netmom) and subsystem definition. The enhanced model validation is conducted after the data ingestions. Compared to current practice, the enhanced model validation incorporates model validity check, NERC case study metric related parameter check and governor mode evaluation, which has been fully describe in [58].

The core of the current model calibration is a Non-linear Least Square (NLS) Optimization without specifying parameter bounds. In the enhanced model calibration, the Parameter Check is added as constraints; while the Response Match is added as part of objective functions (see [58] for details).

After completing the model calibration, the Post Evaluation will automatically evaluate all four-reasonableness check, as shown in the right part of Fig. 3-12 (Block C), this is the last defense layer to safeguard the model meet the predefined requirements. If the model does not pass the Validity check, the code will go back to Model Validation (Block A) to re-verify. If the model does not pass the Parameter Check, then the corresponding constraints for the out-of-spec parameters will be updated and go back to Model Calibration (Block B). If the Response Check fails, then the corresponding penalty weight for that specific response feature will be increased. If the Function Check fails, depending on which model or function fails, the corresponding parameter affecting that model or function will be updated before restarting the model calibration. One example of the user interface for a post evaluation checklist is given in Fig. 3-13.

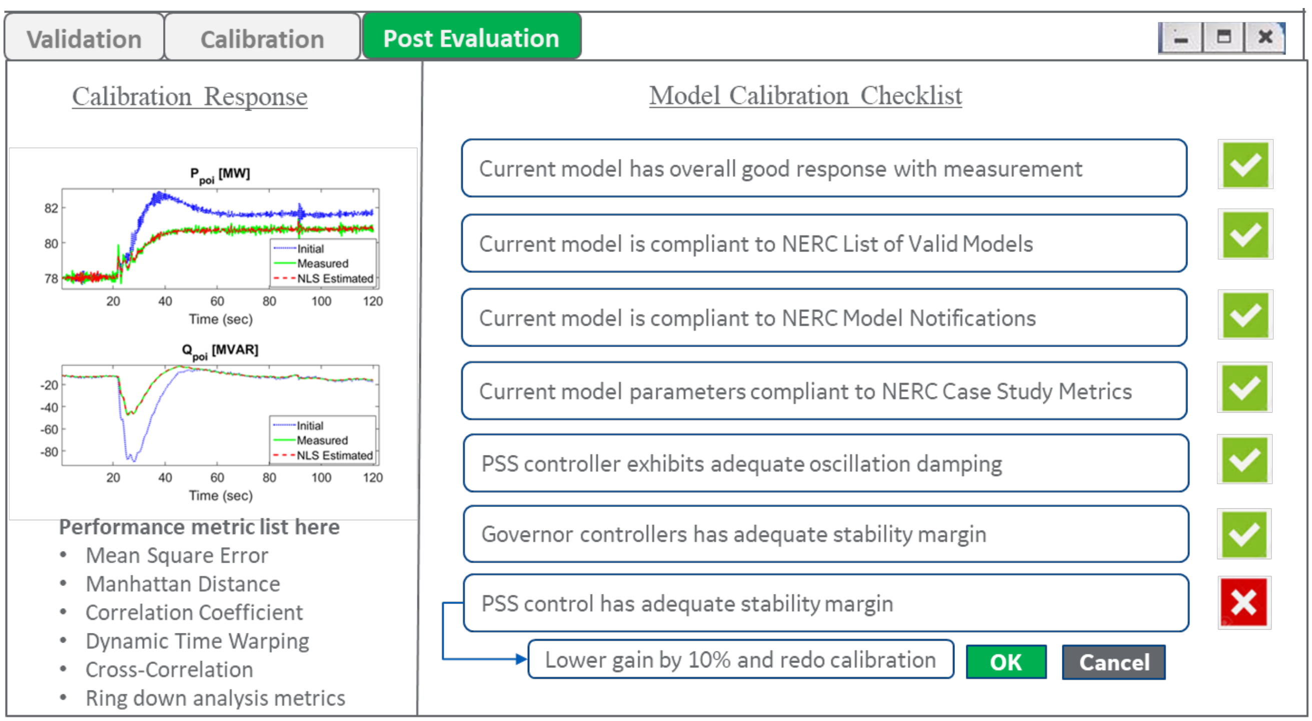


Figure 3-13. model validation and calibration including post processing [58].

The enhanced MV&C can make the generator model and parameter compliant to relevant NERC standards and notifications, ensure the calibrated control system stability at reasonable operating range, and allow better match on control dynamics.

4 Multiple Event Based Model Validation & Calibration

One of the key objectives for disturbance-based model validation is to keep the power plant models and interconnected system models updated so that they can replicate the system events in the evolving power grid. The need for using multiple disturbance events for model validation has been given in Reference [3]: "*The goal of that validation should not be to mimic just one response but rather to provide the best match of response to a number of system conditions*." Using multiple disturbance events for model validation can improve the model consistency [38] and avoid overfitting the model parameters to a certain disturbance [45].

How to best leverage multiple disturbance events for disturbance-based model validation? Thus far, the four primary questions in the community could be summarized as: which subset events to select, what metric to use for evaluation of aggregated performance over multiple events, what parameters to calibrate, and how to calibrate.

4.1 Motivation for Using Multiple Events

The performance of the calibration tool is expected to improve by running it simultaneously over multiple events [48]. The underlying idea is that quality of the estimate improves as the "amount of information" increases. A simple way to see this is to consider three jointly Gaussian random variables (x, y_1, y_2) with covariance matrix $\left[\sigma_{xx}^2\ \sigma_{xy_1}\ \sigma_{xy_2}\ \sigma_{xy_1}\ \sigma_{y_1y_1}^2\ 0\ \sigma_{xy_2}\ 0\ \sigma_{y_2y_2}^2\ \right]$. Say x is the unknown parameter we wish to estimate based on independent observations $y_1$and y_2. By applying the formula for conditional distribution, one can see that the variance of the estimate based on both independent observations $(P[x|y_1, y_2])$ is $\sigma_{xx}^2 - \frac{\sigma_{xy_1}^2}{\sigma_{y_1y_1}^2} - \frac{\sigma_{xy_2}^2}{\sigma_{y_2y_2}^2}$, which is at most equal to the smaller of the estimates based on individual observations $\sigma_{xx}^2 - \frac{\sigma_{xy_1}^2}{\sigma_{y_1y_1}^2}$ and $\sigma_{xx}^2 - \frac{\sigma_{x2}^2}{\sigma_{y_2y_2}^2}$ (for $P[x|y_1]\ and\ P[x|y_2]$ respectively). A smaller variance is being considered as a metric for estimator performance here. Indeed, the statement that 'combining all available information always leads to a better estimate' holds true more generally and can be argued in several ways (eg. using information theoretic quantities).

For the purpose of testing the algorithms, a full simulation platform and a playback simulation platform are needed. The first one should be able to simulate the power system responses subject to different disturbances to build a diverse library of measurement data. The second one should be able to simulate the responses of certain device model driven by the recorded measurement data. To test algorithms in a controllable environment, open source software is preferred because power system simulation software may have built-in parameter limits and corrections which should be transparent to the calibration algorithm. For the full simulation, the full Power System Toolbox (PST) setup is used. For the playback simulation, the PST is modified to allow the system outside certain device being replaced by recorded measurement data as shown in Fig. 4-1.

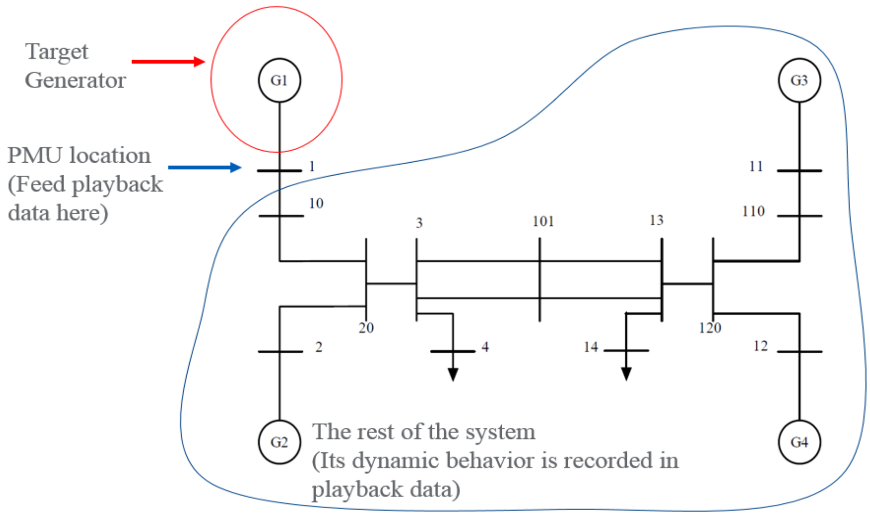


Figure 4-1. Playback simulation using Kundur 2-area system.

To use this setup, there are two steps, data preparation step (using full simulation and true parameter value) and algorithm testing step (using playback simulation and default/corrupt parameter value). For example, if calibrating the parameters in generation plant 1 is of interest, several full simulations will first run using the true parameter value. For different simulation runs, different types of faults will be applied in the system and the PMU measurement data will be collected at bus 1 to build a library of event data. Once enough event data is recorded, the rest of the system outside bus 1 will be ignored and use the recorded event data to drive the playback simulation of generation plant 1. Both identifiability analysis algorithm and model calibration algorithm can call the playback simulation.

To demonstrate the need for sequential estimation, different trials were randomly defined where each trial has different sequence of event (12 events in total, as shown in Table 4.1). The single event mode means the parameter calibrated by the first event will be used to predict the other events afterwards. The multi-event mode means the model parameter will be calibrated in a sequential way, and the prediction error across all events is given. The mean absolute errors between measured and simulated real and reactive (P and Q respectively) power responses is reported.

Table 4.1. Sequence of Events setup

	Event 1	**Event 2**	**Event 3**	**Event 4**	**Event 5**	**Event 6**	**Event 7**	**Event 8**	**Event 9**	**Event 10**	**Event 11**	**Event 12**
Trial -1	4	36	3	2	5	26	30	31	24	16	17	1
Trial -2	36	5	4	1	3	24	2	31	30	26	17	16
Trial -3	17	4	16	31	2	3	24	36	1	5	26	30
Trial -4	24	31	5	36	3	4	30	2	26	17	16	1

Fig. 4-2 shows a typical result of the trials. The multi-event-based calibration results in a significant lower prediction error, compared to those used only one single event. Note the single event mode shows the current practice in the industry. When needed, the generator owner hires testing team to validate and calibrate some parameter, and then keep it in the dyd or dyr file for use for several years.

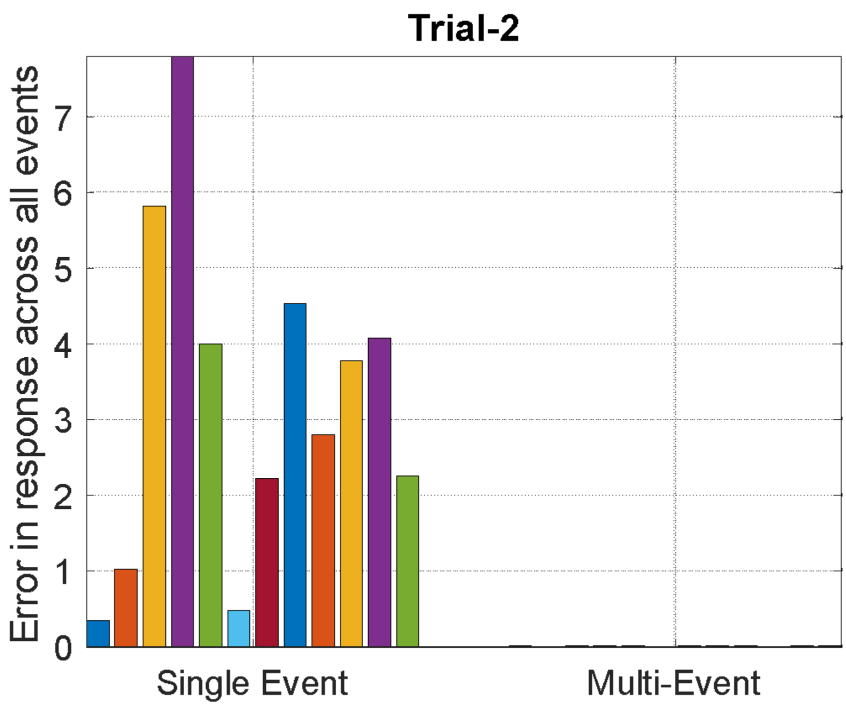


Figure 4-2. Comparison of response error: single event based vs. multi-event based.

4.2 Event Selection

Multiple events over different time can be perceived by PMUs at each generator/load terminal (POI). Using every perceived event to initiate the model validation and calibration process can be very time consuming and may not be cost effective. There is a need to automatically determine which event to be used for the model validation and calibration process.

Some grid events may happen frequently around a PMU. These events may carry similar dynamic modality information and using them to conduct MVC may not be able improve the model performance. Other events may happen infrequently or with little observability at the PMU, but they may carry valuable dynamic modality information for the MVC purpose. There is a further need to ensure the event selected can increase the diversity of the dynamic modality.

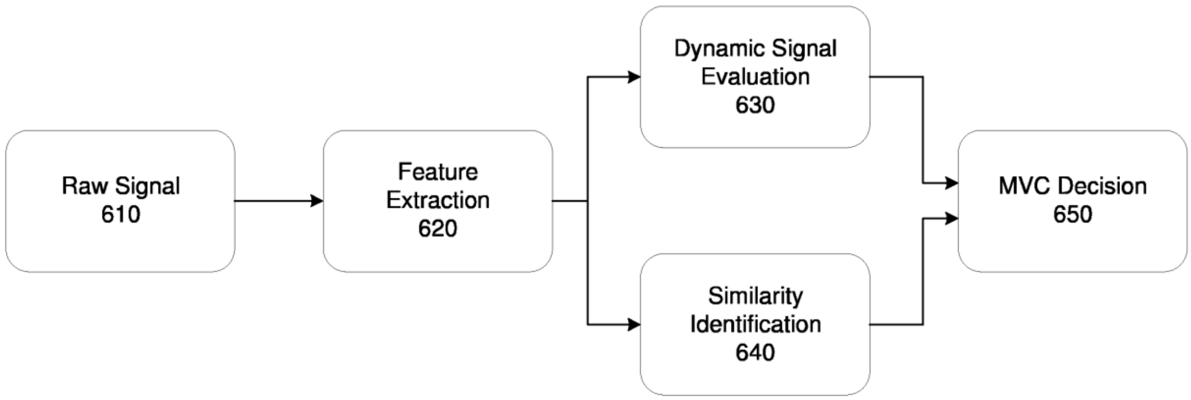


Figure 4-3. Automatic event selection algorithm [59].

An automatic system for event selection for mode validation and calibration [59] would include:

1. Feature extraction module, wherein features related to the system dynamics are extracted from the time series during event, including peak value, rising time, settling time, damping ratio, 2^{nd} largest deviation over the 1^{st} largest deviation of frequency, voltage, power, and reactive power, ROCOF, energy function, cumulative deviation in energy, etc.
2. Dynamic modes evaluation module, wherein the overall magnitude and diversity of the dynamic modes excited are identified and evaluated, including residual analysis based on auto-associated models, such as Auto-Encoder (AE); the input and output of the AE are used to generate the overall magnitude and the diversity (variance of the residuals).
3. Similarity based evaluation module, wherein the identified features are compared with the instance in the existing feature database, the similarity index (like cosine or distance based) is generated to determine how diverse the newly identified event is from the event identified before.
4. Decision making module, wherein the result of the dynamic modes evaluation results and the similarity-based evaluation results are synthesized or fused to determine whether the new event will be used for MVC or not. The decision fusion process can use max, min, or weighted sum on both outputs.
5. The user interface, as shown in Fig 4-4, indicates some details in the event detection.

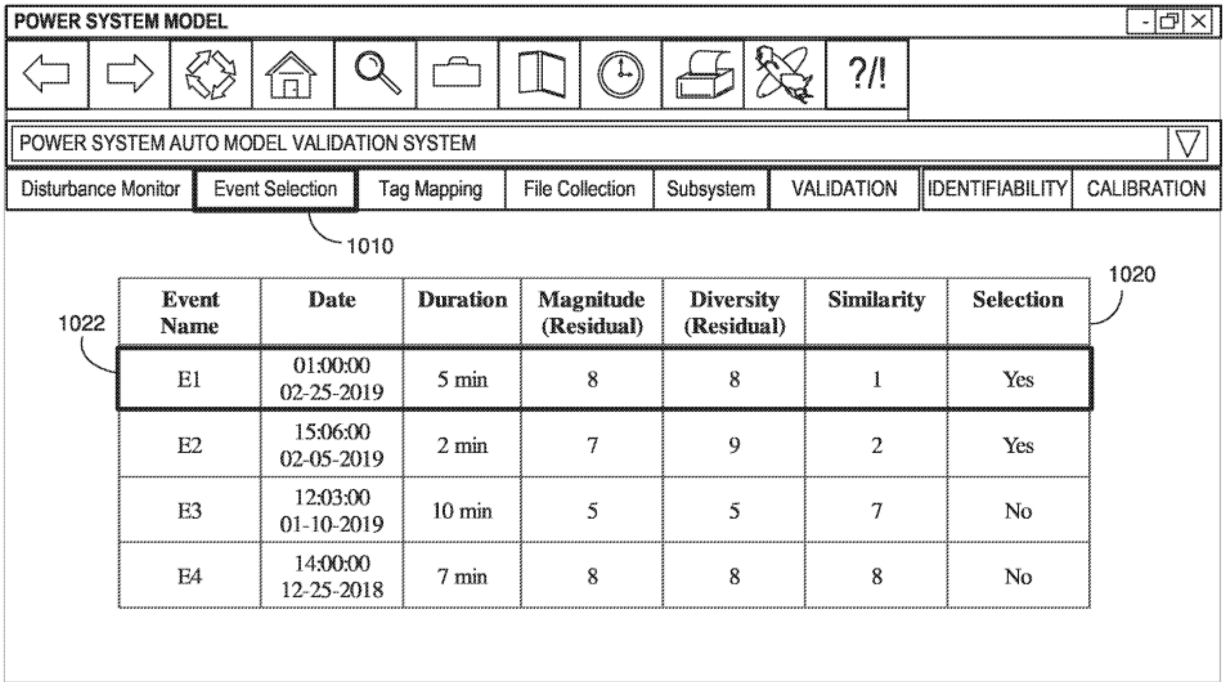


Figure 4-4. Automatic event selection user interface [59].

4.3 Multiple Event Model Calibration

4.3.1 Simultaneous Calibration

For a single event, the model calibration problem can be formulated as a minimization problem with objective function:

$$f_i(x,w) = \sum_{t=1}^{T} w_p(t) * \left(\frac{y_p^m(t) - y_p(x,t)}{y_p^{base}}\right)^2 + w_q(t) * \left(\frac{y_q^m(t) - y_q(x,t)}{y_q^{base}}\right)^2 \qquad (1)$$

where t represents each point of time in the event, T is the event time length, w_p(t) is a weight vector assigned along the time axis to the active power p, and w_q(t) is a weight vector assigned along the time axis to the reactive power q. This weight variable allows the user to emphasize a certain segment in the event with abundant transient information. $y_p^m(t)$ represents the measured active power at time stamp t, $y_p(x,t)$ represents the simulation result at time stamp t with parameter x, y_p^{base} represents the base value of the active power p, which could be 100 MVA for example. x_l, x_u represents the low bound and high bound for parameter x.

For multiple events, the model calibration problem is the minimization of the below equation. The equation keeps x_0 as an input variable to highlight the model calibration is a non-convex problem wherein the parameter searching depends on the initial parameter value x_0.

$$\sum_{i=1}^{N} f_i(x, w, x_0) \tag{2}$$

One straightforward way is simply adding multiple response error terms along multiple data segments, each segment as one event, to form the objective function in (1). This may require multiple calls to each simulation case since each event playback simulation has its own initial condition. The aggregation of response errors from multiple events may come directly from response curve stitched together if those events are sequential events with the same initial condition [60].

The generation plant to be calibrated is the generation plant 1 in the standard Kundur 2-area 4 machine test system. The first step is generating the measurement data. To do that, line faults are applied with different locations and different fault durations then record the terminal bus voltage, phase angle, real and reactive power at generation plant 1. After the measurement data is recorded, replace the system outside generation plant 1 with a voltage source represented by the measurement data. Then the generation plant response driven by the measurement data can be acquired from playback simulation. If the real and reactive power output from the playback simulation matches the measurement data, the model of generation plant 1 is consider accurate. The generation plant has one 6th order generator, one 1st order exciter and one 3rd order governor. The total number of parameters is 24. It is worth mentioning that the methodology and platform are not limited to Kundur 2-area system and generation plant with specific types of generator and control devices, it can be applied to more complex systems and models. By applying 3-phase line to ground fault at different location and with different duration (0.1 s, 0.15 s, 0.2 s), 39 events are created. Using the K-medoids clustering approach, 10 characteristic events are chosen. A wide variety of events are covered in the selected 10 events (local oscillation, inter-area oscillation, generation-load imbalance, etc.). In real life, the information of the location and duration of those events are not needed. They are presented in Table 4.2 to show that the capability of the framework to reduce the event number merely based on the signature of P, Q measurement.

The 5 most identifiable parameters are chosen using the identifiability analysis approach. After calibrating the 5 most identifiable parameters simultaneously for all 10 characteristic events, the P, Q responses from the model match better to the measurement as shown in Fig. 4-5. Since not all corrupt parameters are identifiable, it is difficult to calibrate all corrupt parameters to their true value. However, with the framework, the calibration results simultaneously satisfy multiple events. When a new event with similar characteristic as the chosen events occurs, the calibrated parameters will still provide accurate response. The limitation of the approach is that its performance relies on the diversity of the available events. When the number of available events is limited, it is still possible that the framework could fail when a characteristically unique event happens. However, it is expected that as the library of events becomes more diverse, the framework will become more robust.

Table 4.2. Characteristics of Events.

Event No.	Fault Location	Fault Duration (s)
2	Line 2-20	0.1
3	Line 3-4	0.1
4	Line 3-20	0.1
16	Line 3-4	0.15
17	Line 3-20	0.15
24	Line 13-101	0.15
26	Line 13-120	0.15
30	Line 3-20	0.2
31	Line 3-101	0.2
36	Line 13-101	0.2

Table 4.3. Estimation of the most identifiable parameters.

	X'd (p.u.)	X"d (p.u.)	T'do (s)	Xq (p.u.)	X'q(p.u.)
Def	0.36	0.24	9.60	1.80	0.66
True	0.30	0.25	8.00	1.70	0.55
Tuned	0.30	0.24	8.05	1.58	0.55

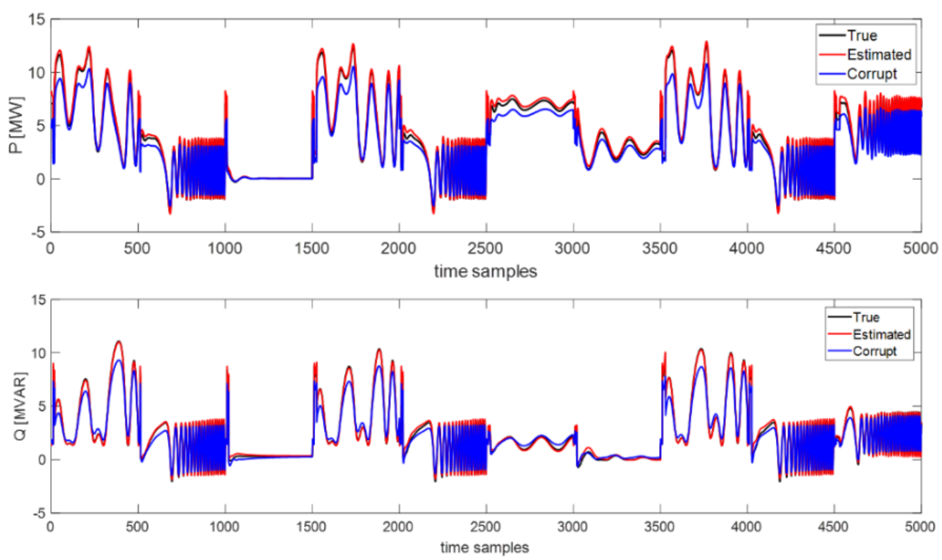


Figure 4-5. Simultaneous calibration results of 10 characteristic events.

4.3.2 Sequential Calibration

Since grid disturbances occur intermittently, the user of the calibration tool would need to re-calibrate model parameters in a sequential manner as new disturbances come in. In this scenario, the user has a model that was calibrated to some observed grid disturbances to start with and observes a larger than acceptable mismatch with a newly encountered disturbance. The task now is to tweak the model parameters so that the model explains the new disturbance without detrimentally affecting the match with earlier disturbances. Of course, an obvious solution would be to run calibration simultaneously on all events of interest strung together but this comes at the cost of computational expense and engineering involved in enabling running a batch of events simultaneously. It would be far more preferable if some essential information from the earlier calibrations runs can be used guide the subsequent calibration run that helps explain the new disturbance without losing earlier calibration matches.

Fig. 4-6 shows an example of how the user could calibrate the model as multiple events come in a sequential manner at different time of history. Each new calibrated model parameter set will be based on the previous calibrated model parameter set and the newly arrived event. To avoid the parameter traps in the local optimality for the newly arrived event, each newly calibrated model parameter set would be evaluated against all the available events to check the overall performance. The best parameter set will be selected based on the highest model validation performance. Note the order of the event, the forgetting factor of the past events, and the performance metrics will affect the performance of this algorithm.

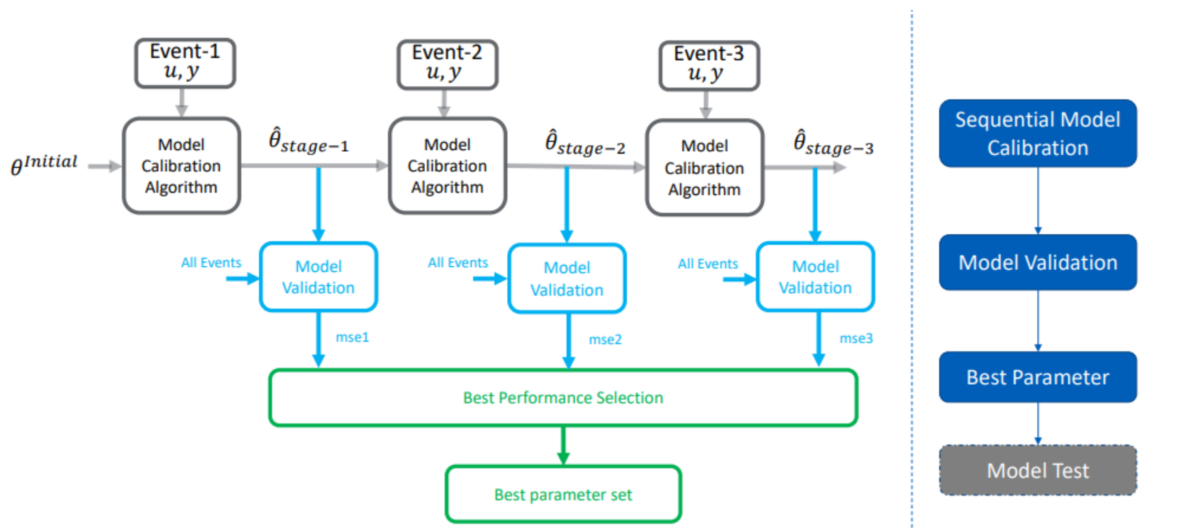


Figure 4-6. Illustration of sequential model calibration algorithm.

The framework of Bayesian estimation has been used to develop a sequential estimation capability into the existing calibration framework. The true posterior distribution of parameters (assuming Gaussian priors) after the calibration process can be quite complicated due to the nonlinearity of the models. The typical approach in sequential estimation is to consider a Gaussian approximation of this posterior as is done in Kalman filtering approaches to sequential nonlinear estimation. In the nonlinear least squares approach, this boils down to a quadratic penalty term for deviations from the previous estimates, and the weights for this quadratic penalty come from a Bayesian argument.

Given a set of events in the pool for training and a nonlinear parameter estimation algorithm for curve fitting, model calibration procedure can be performed in the following way.

The training starts by randomly arranging all the training events to create multiple sequences of events. An event appears at least once in each sequence. The multi-stage training is performed with each of the sequences. During multi-stage training, an event data is given as input to curve fitting algorithm in each stage depending on its order in the sequence. The parameters are calibrated in each stage to minimize the mismatch between the event true response to that of the simulation engine. After getting acceptable accuracy on the curve fitting results at the end of each stage, these calibrated parameters are passed on to the next stage of training and becomes the starting point for the algorithm while training with the next event in that sequence.

Multiple staging ensures that multiple disturbance events participate in the process of model calibration. Randomness in the arrangement of events in each sequence ensures that all the events were given equal priority for training. The parameter sets calibrated during multiple stages of every training sequence can then be compared against each other to find out the best set of calibrated parameters which minimizes the training error also called as mean response error (MRE) uniformly for all events in the training set. The decision on selection of final set of calibrated parameters x can be taken based on minimum over total mean response error (TMRE) which is the MRE averaged over all K training events for each calibrated x. When the true parameters are known, the mean parameter error (MPE) can be calculated as below equation, where x_{true} represents the true parameter set.

$$MRE(x,k) = \frac{1}{T}\sum_{t=1}^{T}|y^m(t,E_k) - y(x,t,E_k)| \tag{3}$$

$$TMRE(x) = \frac{1}{K}\sum_{k=1}^{K} MRE(x,k) \tag{4}$$

$$MPE(x) = \frac{1}{T}\sum_{t=1}^{T}|x - x_{true}| \tag{5}$$

The framework for multi-event model calibration is tested on playback simulation engine platform created using Power System Toolbox (PST). A single generation plant of six order generator model with a first order exciter and a third order governor in Kundur's 4-machine 2-area system is considered for model calibration in this work. Total 24 parameters of the generation model are being calibrated using the approach. Disturbance events such as 3-phase line fault at different locations and with different duration (0.1 s, 0.15 s, 0.2 s), 39 events disturbances were created in the system with a known set of generator parameters. The measurements such as voltage, angle, active power P, and reactive power Q at the point of common coupling of generation plant 1 were recorded. Among these, 12 events which are representative of generator load imbalance, local and inter-area oscillations in the system were considered in the training pool events. Total 8 generator parameters out of 24 were then corrupted.

Table 4.4. Initial, True and Tuned Parameter for Kundur System.

Parameter Names	**Initial**	**TRUE**	**Tuned**
leakage reactance (p.u.)	0.21	0.2	0.1998
d-axis synchronous reactance (p.u.)	1.5	1.8	1.8011
d-axis transient reactance (p.u.)	0.36	0.3	0.2999
d-axis subtransient reactance (p.u.)	0.3	0.25	0.2501
d-axis open-circuit time constant (sec)	9.6	8	8.0084
q-axis synchronous reactance (p.u.)	2.04	1.7	1.7001
q-axis transient reactance (p.u.)	0.66	0.55	0.5501
q-axis open-circuit time constant(sec)	0.5	0.4	0.4002

The calibration is performed in multiple stages wherein each stage uses a randomly selected event from the pool of 12 training events. The multiple stage based training is repeated for different sequence of training events. Fig. 4-7 shows the results obtained with 5 sequences with 2 sub-figures as the TMRE obtained with the calibrated parameter set in each stage. It's clear from Fig. 4-7 that for all the sequences, both MRE and MPE reduces reasonably in each stage. With more training events used for calibration, the parameters fall reasonably close to true set of parameters. Note that the process of sequential model calibration exploits the parameter set obtained from previous training stage, which can be both advantageous/disadvantageous for some sequences. At any stage, if the tuned parameters are close to true solution, further training leads to better solution. However, if an event utilized in a training stage led to a solution which turned out to be a local minimizer of the response errors for all events, the future training stages may not improve the situation. A simple way to tackle this problem is to calibrate with different order of events in different sequence. This phenomena can be observed in Fig. 4-7(b) which shows the evolution of error of the calibrated parameters during training stages carried out with 5 sequences. All the sequences except sequence-1 and 2 lead to lesser parameter error. For those two sequences, the MPE does not improve from the first stage of training. The best set of parameters was obtained from sequence-3 in the 12th stage of training. Table 4.4 lists the true, corrupted and best set of calibrated vales of the parameters obtained throughout all the stages based on TMRE index. All the parameters converge to very close to true parameters. Thus, a sequential way of multiple event model calibration with multiple sequences has the potential to drive towards true parameter set [48].

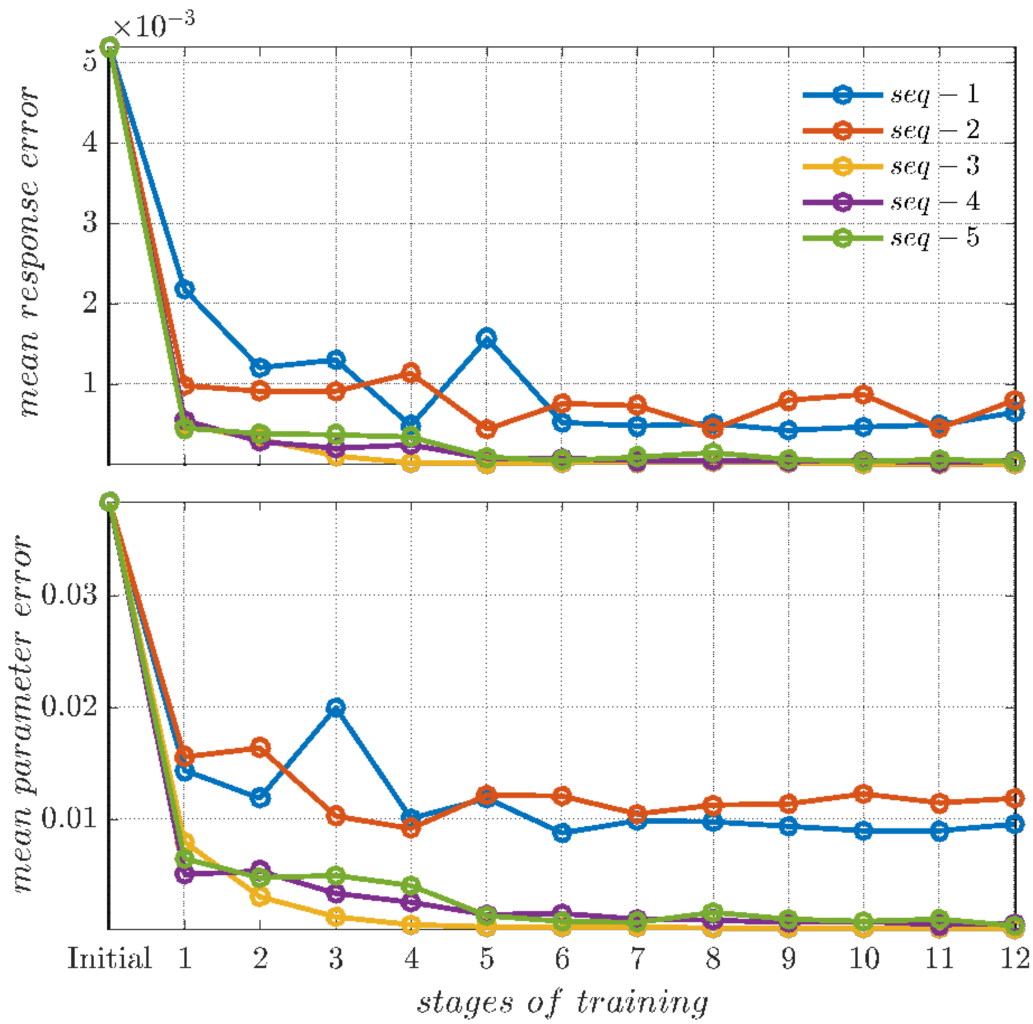


Figure 4-7. Fitting errors for parameter and response across 12 events with 5 different sequences.

4.3.3 Distributed Calibration

As more events and/or larger systems are analyzed, there is a need to improve the scalability of the model calibration. In addition, there is a need for calibration results to be robust against defective datasets and/or events.

A master-client computation system has been designed to handle multiple event data, as shown in Fig. 4-8. The modified MVC engine 502 deploys model validation and calibration tasks in a distributed and collaborative manner using the master-local nodes configurations described herein. The master node 508 allocates and transmits event data 308, system configurations (dynamic model set-up), consensus parameters, best parameter, and penalty gain to the local nodes 510. The local nodes 510 transmit updated search parameters and integration errors to the master node 508 based on the received datasets. The collaboration between the master node and client nodes will improve the convergence, speed, and accuracy of the model calibration. The master node 508 and each local node 510 may be configured as separate processor cores, virtual machines, multi-thread computation with concurrency at a single computer, and/or heterogeneous computation across a network of computers.

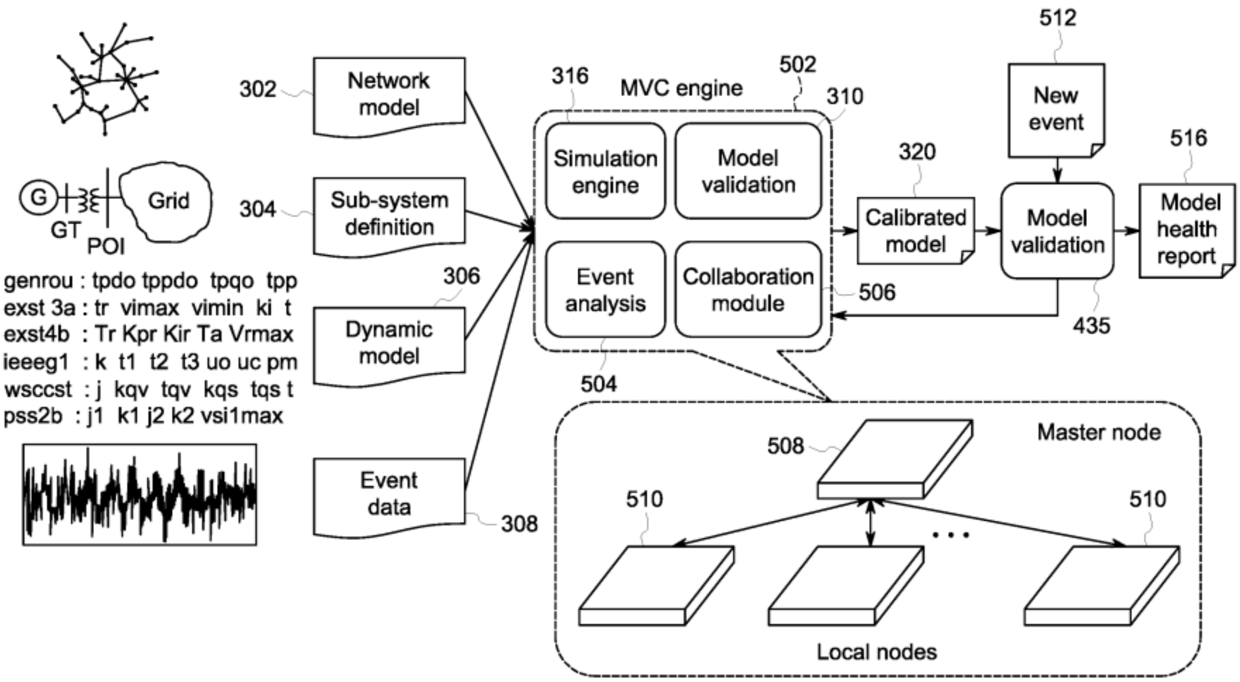


Figure 4-8. System architecture for distributed power system model calibration [61].

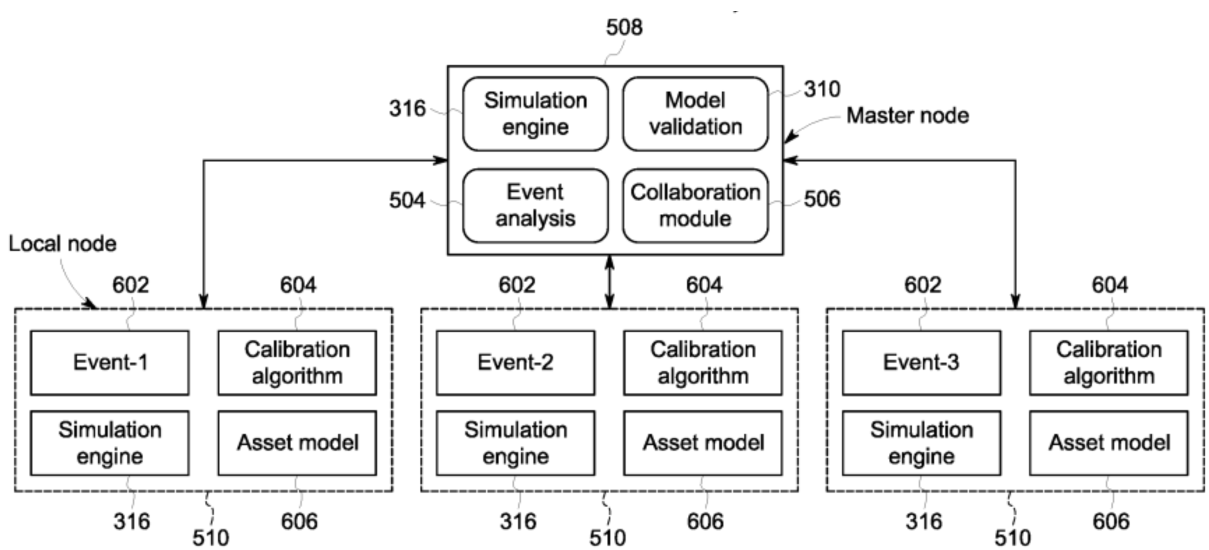


Figure 4-9. Master/Client Nodes for distributed power system model calibration [61].

As shown in Fig. 4-9, each local node 510 includes an event 602, a calibration algorithm 604, the simulation engine 316, and an asset model 606, each local node searches for optimal parameters separately and independently. The search parameters are exchanged and aggregated by the master node 508. The exchange and aggregation are performed at a certain time interval. The master node 508 then distributes the aggregated information to each local node 510 for further search. This process is repeated until a convergence is reached. Variant alternating direction method of multipliers (ADMM) algorithms may be employed for processing multiple events in a distributed system, as shown in Fig. 4-10. The interested reader is referred to [61] for more details.

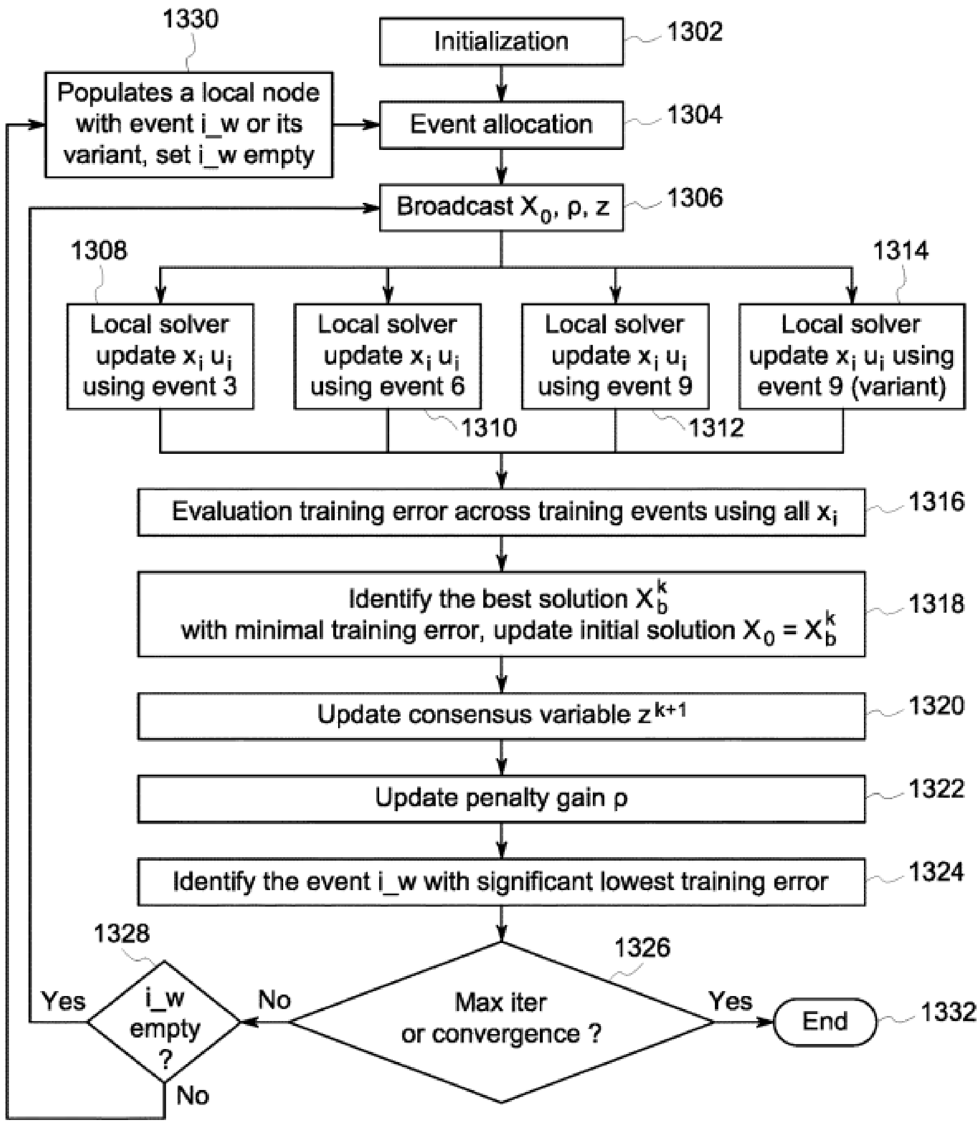


Figure 4-10. Master/Client Nodes for distributed power system model calibration [61].

5 Conclusions

This white paper summarizes the latest progress on the power system model validation and calibration using PMU data, including:

1. **Comprehensive and automated model validation procedure** (model type, parameter compliance to NERC valid lists) beyond a simple playback response comparison.
2. **Systematic performance metrics for model validation** (time domain, frequency domain, separation of oscillatory and governor response).
3. **Advanced model calibration algorithms**.
    - Parameter selection approaches including SVD, Similarity based, Empirical Gramian based methods.
    - Parameter tuning approach including ensemble Kalman filter, trust region optimization, black box optimization, Bayesian optimization, machine learning based tuning algorithms.
4. **Trustworthy MVC framework** by including a post verification step for system stability, model & parameter validity.
5. **Multiple Event based MVC**.
    - Automatic event selection algorithm (feature extraction, transient richness, and diversity evaluation modules).
    - Multi-event calibration algorithms, including simultaneous calibration, sequential event calibration and distributed calibration.
6. **Needs for public dataset** for benchmarking MVC toolsets.

It is important to point out the limitations of the PMU-based parameter model validation and calibration: 1) its sampling frequency is limited; and 2) observations from only POC data are also limited [38]. Therefore, some model errors cannot be corrected through parameter calibration using measurements from PMUs. Also, parameters cannot be estimated reliably if they have small trajectory sensitivity, which is also referred to as "ill-conditioned."

Some future directions include but not limited to:

1. The lack of public available dataset inevitably hinders the development progress of the MVC toolset. So far, the only benchmark dataset for evaluating the MVC algorithm is from the 2017 NASPI workshop [62]. There is still a need to provide public dataset for evaluation of the model validation and calibration algorithm.
2. There is a need to make available open-source MVC toolset to allow developers from different domains (machine learning, complex systems, control engineering, power system, and optimizations) to advance this area.
3. The developed algorithms, software toolset should be extended to the renewable system models with higher resolution event data such as point on waves or digital fault recorder.

6 References

[1] A. Silverstein, E. Andersen, F. Tuffner et al., “Model validation using phasor measurement unit data,” NASPI Technical Report, March 2015.

[2] Wang, H., Wang, P., Menon, A., Parashar, M., Srinivasan, K., Chen, S. and Markham, R., 2020, August. Feature Adaptive Generator Model Calibration. In 2020 IEEE Power & Energy Society General Meeting (PESGM) (pp. 1-5). IEEE.

[3] Power System Model Validation, A White Paper by the NERC Model Validation Task Force of the Transmission Issues Subcommittee, Dec. 2010.

[4] Reliability Guideline: Power Plant Model Verification using PMUs, NERC, Sep. 2016.

[5] NERC, MOD-026-1, Verification of Models and Data for Generator Excitation Control Systems or Plant Volt/Var Control Functions. https://www.nerc.com/pa/Stand/Reliability%20Standards/MOD-026-1.pdf

[6] NERC, MOD-027-1, Verification of Models and Data for Turbine Governor and Load Control or Active Power Frequency Control Functions. https://www.nerc.com/pa/Stand/Project%20200709%20%20Generator%20Verification%20%20PRC0241/MOD-027-1.pdf

[7] NERC, MOD-032-1, Data for Power System Modeling and Analysis. https://www.nerc.com/pa/Stand/Reliability%20Standards/MOD-032-1.pdf

[8] NERC MOD-033-1, Steady State and Dynamic Model Validation. http://www.nerc.com/pa/Stand/Reliability%20Standards/MOD-033-1.pdf

[9] NERC, N., Reliability Functional Model, Version 6, June 2016.

[10] NERC, N., Reliability Functional Model Technical Document, Version 6, July 2016.

[11] Wang, D., Wilson, D.H. and Clark, S., 2014, April. Defining constraint thresholds by angles in a stability constrained corridor with high wind. In 2014 IEEE PES T&D Conference and Exposition (pp. 1-5). IEEE.

[12] GE Energy Consulting Grid Code Testing. https://www.geenergyconsulting.com/practice-area/global-power-projects/grid-code-testing (accessed Jan 3 2021)

[13] Li, Y., Diao, R., Huang, R., Etingov, P., Li, X., Huang, Z., Wang, S., Sanchez-Gasca, J., Thomas, B., Parashar, M. and Pai, G., 2017, July. An innovative software tool suite for power plant model validation and parameter calibration using PMU measurements. In 2017 IEEE Power & Energy Society General Meeting (pp. 1-5). IEEE.

[14] Menon, A., Baone, C.A., Dokucu, M., Thomas, B., Sanchez-Gasca, J., Acharya, N., Parashar, M., Srinivasan, R. and Jampala, A., 2018, August. Towards a commercial-grade tool for disturbance-based model validation and calibration. In 2018 IEEE Power & Energy Society General Meeting (PESGM) (pp. 1-5). IEEE.

[15] Wang, H. Operationalizing Synchrophasors for Enhanced Grid Reliability and Asset Utilization. United States: N. p.1-101, Jan 2020. Web. doi:10.2172/1601103.

[16] Power Plant Parameter Derivation (PPPD) software user’s manual: version 12.1, EPRI, Palo Alto, CA: 2020. Product ID # 3002019311.

[17] P. Pourbeik, R. Rhiner, S-M. Hsu, B. Agarwal, R. Bisbee, “Semiautomated model validation of power plant equipment using online measurements,” IEEE transactions on Energy Conversion, June 2013, pp 308-316.

[18] P.Pourbeik, C. Pink, E. Bisbee, “Power plant model validation for achieving reliability standard requirements based on recorded online disturbance data,” Proc. of IEEE power systems conference and exposition, March 2011,

[19] P. Pourbeik, “Automated parameter derivation for power plant models form system disturbance data,” Proc. of IEEE PES GM, Calgary, Canada, July 2009.

[20] D. Kosterev et al., (2010). Power Plant Model Validation Tool. https://svn.pnl.gov/PPMV

[21] Huang, R., Diao, R., Li, Y., Sanchez-Gasca, J., Huang, Z., Thomas, B., Etingov, P., Kincic, S., Wang, S., Fan, R. and Matthews, G., 2017. Calibrating parameters of power system stability models using advanced ensemble Kalman filter. IEEE Transactions on Power Systems, 33(3), pp.2895-2905.

[22] Etingov P.V., J.D. Follum, U. Agrawal, H. Wang, F.K. Tuffner, L. Newburn, and R. Huang, et al. 2020. Open Source Suite for Advanced Synchrophasor Analysis. PNNL-30492. Richland, WA: Pacific Northwest National Laboratory. Available Online: https://www.pnnl.gov/main/publications/external/technical_reports/PNNL-30492.pdf

[23] U. Agrawal, P. Etingov and R. Huang, "Initial Results of Quantification of Model Validation Results Using Modal Analysis," 2020 IEEE Power & Energy Society General Meeting (PESGM), Montreal, QC, 2020, pp. 1-5.

[24] Agrawal, Urmila, Pavel Etingov, and Renke Huang. "Advanced Performance Metrics and their Application to the Sensitivity Analysis for Model Validation and Calibration." *IEEE Transactions on Power Systems* (2021), pp. 4503-4512.

[25] MathWorks. https://www.mathworks.com/matlabcentral/fileexchange/74912-power-plant-model-validation-simscape-design-solution

[26] N. Nayak, H. Chen, W. Schmus, R. Quint, "Generator Parameter Validation and Calibration Process based on PMU Data," IEEE PES Transmission & Distribution Conference & Exposition, 2016.

[27] GMV Website: http://electricpowergroup.com/gmv.html

[28] GMV Brochure: http://electricpowergroup.com/downloads/GMV.pdf

[29] GMV Webinar: https://electricpowergroup.app.box.com/v/EPGWebinar4GPV

[30] W. Ju, N. Nayak, C. Vikram, H. Silva-Saravia, K. Sun and G. Zu, "Indices for Automated Identification of Questionable Generator Models Using Synchrophasors," 2020 IEEE Power & Energy Society General Meeting (PESGM), Montreal, QC, Canada, 2020, pp. 1-5, doi: 10.1109/PESGM41954.2020.928216

[31] Wang, H. Production Grade MV&C Application V2.0. NASPI Work Group Meeting, October 28-30 - Richmond, Virginia. https://www.naspi.org/sites/default/files/2019 10/02_FOA1492_ProductionGrade_Wang_20191030.pdf

[32] https://www.ge.com/digital/applications/transmission/phasoranalytics

[33] H. Wang (2019), "Production Grade MV&C Application V2.0 - Recent enhancements to overcome practical challenges from customer demos," [Online]. Available:

https://www.naspi.org/sites/default/files/2019-10/02_FOA1492_ProductionGrade_Wang_20191030.pdf

[34] Wang, H., Parashar, M., Srinivasan, R., Shao, M., Rao, S.D. and Sanchez-Gasca, J., General Electric Co, 2020. Systems and methods for enhanced power system model validation. U.S. Patent Application 16/717,474.

[35] Rezaei, E. and Venkatasubramanian, V., 2018, May. Quantitative Indicators for Quality of Fit Assessment in Power System Model Validation Problems. In 2018 IEEE International Symposium on Circuits and Systems (ISCAS) (pp. 1-5). IEEE.

[36] NERC, “Technical Report: NERC Modeling Improvements Initiative Update”, pp.1-28, May 2018.

[37] NERC, “Case Quality Metrics, Annual Interconnection-wide Model Assessment”, pp 11-12, September 2017.

[38] Nayak, N., Chen, H., Schmus, W. and Quint, R., 2016, May. Generator parameter validation and calibration process based on PMU data. In 2016 IEEE/PES Transmission and Distribution Conference and Exposition (T&D) (pp. 1-5). IEEE.

[39] M. Burth, G.C. Verghese, M. Velez-Reyes, "Subset selection for improved parameter estimation in on-line identification of a synchronous generator", Power Systems IEEE Transactions on, vol. 14, no. 1, pp. 218-225, 1999.

[40] C. Baone, J. Lim, S. Bose. “Systems and methods for analyzing model parameters of electrical power systems using trajectory sensitivities.” U.S. Patent Publication No. US20150149128A1, May. 28, 2015.

[41] I. Hiskens, “Nonlinear dynamic model evaluation from disturbance measurements,” IEEE Transactions on Power Systems, vol. 16, no. 4, pp. 702-710, November 2001.

[42] Sanchez-Gasca, J.J., Bridenbaugh, C.J., Bowler, C.E.J. and Edmonds, J.S., 1988. Trajectory sensitivity based identification of synchronous generator and excitation system parameters. IEEE Transactions on Power Systems, 3(4), pp.1814-1822.

[43] S.M. Benchluch, J.H. Chow, "A trajectory sensitivity method for the identification of nonlinear excitation system models", Energy Conversion IEEE Transactions on, vol. 8, no. 2, pp. 159-164, 1993.

[44] Huang Z, Du P, Kosterev D, et al. Generator dynamic model validation and parameter calibration using phasor measurements at the point of connection. IEEE Transactions on Power Systems, 2013, 28(2): 1939-1949.

[45] Huang, R., Diao, R., Li, Y., Sanchez-Gasca, J., Huang, Z., Thomas, B., Etingov, P., Kincic, S., Wang, S., Fan, R. and Matthews, G., 2017. Calibrating parameters of power system stability models using advanced ensemble Kalman filter. IEEE Transactions on Power Systems, 33(3), pp.2895-2905.

[46] Zhao, Junbo, et al. “Robust Adaptive Nonlinear Kalman Filter for Synchronous Machine Parameter Calibration.” in *Proceedings of the 54th Hawaii International Conference on System Sciences*, 2021.

[47] Baone, C.A., Duan, N., Menon, A. and Dokucu, M.T., General Electric Co, 2020. Power system model parameter conditioning tool. U.S. Patent 10,809,683.

[48] Mahapatra, K. and Wang, H., 2020, February. Generator Dynamic Model Calibration using Multiple Disturbance Events. In 2020 IEEE Power & Energy Society Innovative Smart Grid Technologies Conference (ISGT) (pp. 1-5). IEEE.

[49] M. Namba, T. Nishiwaki, S. Yokokawa, K. Ohtsuka and Y. Ueki, "Identification of Parameters for Power System Stability Analysis Using Kalman Filter", IEEE Trans. Power Apparatus and Systems, vol. PAS-100, no. 7, pp. 3304-3310, July 1981.

[50] S. R. Khazeiynasab and J. Qi, “Generator parameter calibration by adaptive approximate Bayesian computation with sequential Monte Carlo sampler,” *IEEE Trans. Smart Grid*, vol. 12, no. 5, pp. 4327–4338, Sept. 2021.

[51] S. R. Khazeiynasab and J. Qi, “PMU measurement based generator parameter calibration by black-box optimization with a stochastic radial basis function surrogate model,” 2020 North American Power Symposium (NAPS), Apr. 2021.

[52] C. C. Lee and O. T. Tan, "A Weighted-Least-Squares Parameter Estimator For Synchronous Machines", IEEE Trans. Power Apparatus and Systems, vol. PAS-96, no. 1, pp. 97-101, January/February 1977.

[53] R. Huang, R. Fan, T. Yin, S. Wang, and Z. Tan, “Parameters calibration for power grid stability models using deep learning methods,” arXiv preprint arXiv:1905.03172, May 2019.

[54] S. R. Khazeiynasab, J. Qi, and I. Batarseh, “Generator parameter estimation by q-learning based on pmu measurements,” in *Innovative Smart Grid Technologies (ISGT)*, Feb. 2021, pp. 01–05.

[55] S. Wang, R. Diao, T. Lan, Z. Wang, D. Shi, G. N. America, H. Li, and X. Lu, “A drl-aided multi-layer stability model calibration platform considering multiple events,” in *PESGM*, pp. 1–5, 2020.

[56] S. Wang, R. Diao, C. Xu, D. Shi, and Z. Wang, “On multi-event cocalibration of dynamic model parameters using soft actor-critic*,” IEEE Trans. Power Syst*., 2020.

[57] S. R. Khazeiynasab, J. Zhao, I. Batarseh and B. Tan, “Power Plant Model Parameter Calibration Using Conditional Variational Autoencoder,” *IEEE Transactions on Power Systems*, 2021.

[58] Wang, Honggang, Manu Parashar, Radhakrishnan Srinivasan, Miaolei Shao, Shruti Dwarkanath Rao, and Juan Sanchez-Gasca. "Systems and methods for enhanced power system model validation." U.S. Patent Application 16/425,797, filed December 3, 2020.

[59] Weizhong Yan, Wang, Honggang. " Event selection for power grid disturbance." U.S. Patent Application PCT/US2019/018913, filed February 21, 2019.

[60] Duan, N., Baone, C., Menon, A. and Dokucu, M., 2018, February. Synchrophasor based dynamic model validation leveraging multiple events. In 2018 IEEE Power & Energy Society Innovative Smart Grid Technologies Conference (ISGT) (pp. 1-5). IEEE.

[61] Wang, H., Wang, Y.S., Zhou, J. and Wang, P., General Electric Co, 2022. Systems and methods for distributed power system model calibration. U.S. Patent 11,347,907.

[62] NASPI & NERC Synchronized Measurement Subcommittee Model Verification Tools Technical Workshop, [online] Available: https://www.naspi.org/naspi/sites/default/files/2017-03/naspiworkshop_model_verification_20161018.pdf .