Empowering Data Model Convergence in Power System Planning Study & Control Design by Machine Learning with Utility Data

May 30, 2019

Xiaoyuan Fan
Senior Energy Research Engineer, Ph.D.
2019 WECC JSIS Meeting
Salt Lake City, UT
Ongoing DOE project: Adaptive RAS/SPS System Settings for Improving Grid Reliability and Asset Utilization through Predictive Simulation and Controls

▶ Main Objective
- To develop innovative mathematical and advanced computing methods for adaptively setting RAS/SPS parameters based on realistic and near real-time operation conditions, powered by HPC.

▶ Resources
- Abundant expertise in power grid modeling and simulation;
- Dedicated industry support from Western Utility collaborators with details of active RAS models as well as multi-year data;
- HPC/Cloud platforms and commercial packages for high-fidelity simulations.

▶ Deliverables
- Prototype design and development in commercial platform;
- Technical report and research publications.

Figure 1. An illustration of “ABCDE” design concept for transformative remedial action scheme tool (TRAST) [1].
OE’s Advanced Modeling Grid Research Program objectives [2] are to:

- Support the transformation of data to enable preventative actions, rather than reactive responses to changes in grid conditions;
- Direct the research and development of advanced computational and control technologies to improve the reliability, resiliency, security, and flexibility of the nation’s electricity system;
- Help system operators and utilities prevent blackouts and improve reliability by expanding wide-area real-time visibility into the conditions of the grid;
- Support improvement of the performance of modeling tools and computations that are the basis of the grid operations and planning; and
- Support the tracking and expansion of the use of quantitative risk and uncertainty methods by federal- and state-level energy system decision makers regarding energy infrastructure investments.

DOE RAS project directly supports AGM program vision and goals, by:

- Developing innovative mathematical methods for determining RAS parameters in near real-time;
- Enhancing RAS modeling accuracy and, therefore, system resilience following severe disturbances;
- Building software prototypes for automating study procedures;
- Leveraging high-performance computing techniques to achieve speed gain.
Overview of Transformative RAS Tool (TRAST)

Realistic Scenario Generation
- Base Cases
- Path stress patterns
- EMS Cases

Massive Simulations (HPC)
- Unified Fault model for multi-section line
- Dynamic models and parameters
- Contingency definition
- Massive Simulations
- New algorithms for calculating RAS settings in near real time

Validation of RAS settings
- Adaptive RAS/SPS settings for operation
- RAS Event replay in TSAT

Data correlation analysis
- Automated Case generation
- Unified ctg definition
- Power flow case analysis & validation

Smart Sampling
- Parallel Computing & Cloud Application for power system
- Machine Learning Framework
- Validation of RAS settings
- Comparison & Validation

: Implemented Functionality
★: Functionality in progress

Figure 2. Overview of Data-driven analytical functionalities in TRAST [1].
Machine Learning for RAS Design and Calculation

RAS Control Feature Analysis and Selection by Machine Learning techniques

2017 WECC planning cases
Heavy summer/spring

Base Cases
Path stress patterns
EMS Cases

Realistic Scenario Generation

Contingency definition
Dynamic models and parameters
RAS models

Massive Simulations (HPC)

System status
Arming Levels for (S,N) pair

New algorithms for calculating RAS settings in near real time

Validation of RAS settings

Adaptive RAS/SPS settings for operation

4 events of RAS action

ePMU data

M: Synthesized data/cases
O: Data/Cases from IPC, PacifiCorp & Peak
A: RAS model design (PacifiCorp)
C: RAS event record (PacifiCorp)

Machine Learning Framework for RAS Coefficient Prediction

Figure 3. Overview of Utility data interface with Machine Learning block in TRAST [1].
TRAST Use Case: Jim Bridger RAS

Jim Bridger RAS Logic

- **AL**: Arming Level Calculation (33 ALs)
- **GT**: Generation Tripping Amount Calculation (33 Gs)
- **CCL**: Capacitor Control Logic
- **GUS**: Unit Selection & Tripping

System Input

- **System Output**

Arming Level Calculation

- **Stage 1**: Read in all RAS Coefficients
- **Stage 2**: Locate Active S State
- **Stage 3**: Is this S State defined?
  - Locate RAS Coefficients by Season, in total 8 * 33 = 264 coefficients
- **Stage 4**: Calculate 33 Arming Levels
  - Report 33 AL results
  - End

Every 200 ms

Figure 4. Jim Bridger RAS and its operation logics [3][4].
TRAST Use Case: Jim Bridger RAS studies

- Transformative RAS Tool (TRAST) enables a statistical and efficient way to identify a list of appropriate scenarios to represent the system conditions of utility study interests.

- For Jim Bridger RAS studies:
  - 2017 full-year Jim Bridger related SCADA data has been provided by PacifiCorp and Idaho Power.
  - 2017 full-year WECC Path related SCADA data has been provided by Peak Reliability to support Jim Bridger RAS studies.

- Automated utility planning case generation provides a powerful, yet flexible, way for generating a reasonable case pool for RAS studies based on 2017 WECC planning model.

- For the targeted RAS in our project, there are roughly about $365 \times 648 \times 33 \approx 7.8$ million dynamic simulations to be evaluated.

- TRAST provides an systematic and automated/semi-automated solution for RAS validation and assessment.
  - Parallel computing in Cloud environment.
  - Machine Learning tool assisted control feature analysis and selection, as well as arming level prediction.
Utility Data Analysis in TRAST
(1) Jim Bridger RAS Offline Arming-Level Calculation

2017 full-year SCADA data

Figure 5. Offline Arming Level Calculation for Jim Bridger RAS [4].
Utility Data Analysis in TRAST

(2) Review Control Input in existing RAS Design

► Time period of study:
12/01/2016 0:00 – 04/30/2018 23:30.

► Measurements are recorded and pre-processed at a 30-min resolution, with a total of 24768 data points for each variable.

► The following 7 variables are included:
  - **Gen** is the power plant real power generation;
  - **Path1** is the first WECC path real power flow;
  - **Path2** is the second WECC path real power flow;
  - **Path3** is the first internal path real power flow;
  - **Path4** is the second internal path real power flow;
  - **Gvar** is the power plant reactive power generation;
  - **AvaiComp** is one equipment status indicator.

Figure 6. Normalized utility data visualization (top) and initial analysis by season (bottom) [5].
**Utility Data Analysis in TRAST**

(2) Review Control Input in existing RAS Design (Cont’)

**Conclusion:** Correlation exists between current RAS input data.

**Recommendation:** Dimension reduction can be performed for the original RAS input data.

Figure 7. Pairwise correlation coefficients of all seven variables.

Figure 8. Scatter plot between Gen and Path 1 (top) and box plots between path 1 and AvaiComp (bottom) [5].
The classification and regression trees (CART) analysis was also performed to rank the other six physical features and two temporal features [6].

- Further identify the variables of most importance for the reactive compensation level P7 (AvaiComp)
- The physical features P3 (Path1) and P1 (Gen) are the most important predictors,
- The temporal factors (e.g., month and season) are also significant

Combination of features have also been evaluated, for one, two and six features correspondingly.

<table>
<thead>
<tr>
<th></th>
<th>MLR_1</th>
<th>MLR_2</th>
<th>MLR_6</th>
</tr>
</thead>
<tbody>
<tr>
<td>2017</td>
<td>74.7%</td>
<td>77.4%</td>
<td>78.5%</td>
</tr>
<tr>
<td>Spring</td>
<td>70.5%</td>
<td>71.2%</td>
<td>73.1%</td>
</tr>
<tr>
<td>Summer</td>
<td>84.4%</td>
<td>86.6%</td>
<td>86.6%</td>
</tr>
<tr>
<td>Autumn</td>
<td>74.4%</td>
<td>78.2%</td>
<td>81.4%</td>
</tr>
<tr>
<td>Winter</td>
<td>73.5%</td>
<td>77.3%</td>
<td>76.9%</td>
</tr>
</tbody>
</table>
Utility Data Analysis in TRAST
(4) Intercorrelation between Jim Bridger System and WECC Systems

**Conclusion:** No significant correlation exists between Jim Bridger system and other WECC systems, considering path flow.

**Recommendation:** Case generation can be performed considering Jim Bridger system operating conditions without loss of generality.
Power System Model in TRAST: (1) Automated Power System Planning Case Generation

Objectives:
- Simplify the planning case generation process for RAS studies;
- Provide a generic and automated way to create a reasonable pool of candidate cases, starting with a few utility planning cases.

Approach:
- Leverage PowerWorld’s Optimal Power Flow (OPF) engine to redispatch generation and load while enforcing select system constraints;
- Make the smallest change to generation and/or load possible to meet system constraints.

Solution:
- An iterative process to generate the traversal of utility planning cases;
- Verified power flow solution in PowerWorld and PSS/E.

Results: More than 11,500 “new” planning cases have been generated based on WECC 2017 Heavy Summer planning case.

<table>
<thead>
<tr>
<th>Procedure</th>
<th>Total # of Planning Cases</th>
</tr>
</thead>
<tbody>
<tr>
<td>Starting point</td>
<td>1 for 2017 HS</td>
</tr>
<tr>
<td>Step 1 &amp; 2</td>
<td>6</td>
</tr>
<tr>
<td>Step 3 &amp; 4</td>
<td>80+</td>
</tr>
<tr>
<td>Step 5 &amp; 6</td>
<td>1,200+</td>
</tr>
<tr>
<td>Step 7 &amp; 8</td>
<td>11,500+</td>
</tr>
<tr>
<td>Smart Sampling</td>
<td>295 out of 365 (ideal)</td>
</tr>
</tbody>
</table>
Power System Model in TRAST: (1) Automated Power System Planning Case Generation (Cont’)

- Observation:
  - Correlation between Path1 and Gen is preserved
  - The 11,500+ auto-generated power flow cases provides solid foundation for Jim Bridger RAS studies
Data Model Convergence in TRAST

(1) Smart Sampling for Faster RAS Studies Considering Automated Planning Case Generation

- Originates from 2017 full year’s SCADA data, using 30-min resolution
  - In total, **17520** data points for each variable
- The objectives [5]:
  - For each variable, represent the probability distribution according to the original data using much fewer samples
  - Consider the data-dependency among the variables
- Solution: Customized Latin Hypercube Sampling (LHS)
  - Resolved unknown PDF issues
  - Account for original correlation with Cholesky decomposition
- Results:
  - **365** sampled points to guide automated utility planning case generation in TRAST
  - Significantly reduce the dynamic simulation efforts

![Figure 15. Transformation of LHS sampled points from CDF to sample percentiles.](image1)

![Figure 16. Accuracy of samples (difference between original and sampled histogram curves) Vs the number of samples.](image2)

![Figure 17. Results of Gen samples considering data dependency.](image3)

![Figure 18. Samples of Gen and Path1 displayed in 2D space considering correlation.](image4)
Data Model Convergence in TRAST
(1) Smart Sampling for Faster RAS Studies Considering Automated Planning Case Generation (Cont’)

Observation:
- Correlation between Path1 and Gen is preserved
- Relationship between Path1 and Path 2 is preserved
- Significantly narrow down the study list of power flow cases (11,500+ → 295), speed gain is about 38 times

Figure 19. 2D & 3D visualization for the smart sampled 295 power flow cases in the initial pool of automated case generation process.
Power System Model in TRAST:
(2) Customized Dynamic Simulation for Arming Level Derivation

Figure 20. Visualization for 295 streams of power flow cases that created for power system transient stability analysis [8].

<table>
<thead>
<tr>
<th></th>
<th>N1</th>
<th>N2</th>
<th>...</th>
<th>N33</th>
</tr>
</thead>
<tbody>
<tr>
<td>S1</td>
<td>aa</td>
<td>bb</td>
<td>...</td>
<td>yy</td>
</tr>
<tr>
<td>S2</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
</tr>
<tr>
<td>S648</td>
<td>xx</td>
<td>...</td>
<td>...</td>
<td>...</td>
</tr>
</tbody>
</table>

Derived Arming Level
Figure 21. Visualization for the dynamic screening results in PSS/E for 295 streams of power flow cases that created for power system transient stability analysis.
Power System Model in TRAST: (4) Dynamic Initialization Diagnosis

- Major initialization issue shows up when dispatching some generators to their maximum active power generation ($P_{\text{max}}$):
  - Some generators should not be dispatched to $P_{\text{max}}$ due to a slight model implementation difference between PSS/E and PSLF (e.g., HYGOV, HYG3 in PSLF [9] and WSHYGP in PSS/E, HYGOV4 in PSLF and IEEEG3 in PSS/E [10])
  - Some generation should not be dispatched to $P_{\text{max}}$ as it will cause inner loop converge issue due to model deficiency, typically wind generation models (e.g., WT2G1, REECBU1)
  - Some generators should not be dispatched to $P_{\text{max}}$ as the governor has a smaller turbine MW capability compared with the generator MVA in PSS/E (e.g., 7XXX "HXXXX", $P_{\text{max}} = P_{\text{gen}}$=MVA = 202 in PSS/E, $MW\, CAP = 184$ for Governor in PSLF)
  - Some generation should not be dispatched to $P_{\text{max}}$ as it will cause limitation issues for the excitation system
  - Possible workaround: in the automated case generation procedure, generators should only be dispatched up to 0.95$P_{\text{max}}$
Data Model Convergence in TRAST
(2) Machine Learning Framework for RAS Coefficient Prediction

- Identify coefficients for the RAS Arming Level calculation based on results of large-scale dynamic simulation:
  - Currently, the monitored parameters and coefficients are determined mainly based on engineers’ experience;
  - Conservative, time-consuming, and hard to maintain;
  - Moving forward, a systematic, data-driven approach is preferable.

- Investigate methods for improving the Arming Level calculation and prediction
  - Linear or non-linear regression: will more complicated linear or non-linear calculation methods help? (In Progress)
Data Model Convergence in TRAST

(3) Next Step

- Complete the Jim Bridger RAS use case study
- Future scenario studies
  - High penetration of renewable energy (3000 MW Wind)
    1) Is Jim Bridger RAS still needed?
    2) If needed, how to improve/upgrade existing RAS platform for new control requirement?
    3) How to identify the input for the new RAS control logic?
  - Major system topology change (TransWest Express)
    1) How should the Jim Bridger RAS be revised?
    2) How to bridge the gap between new transmission line planning and critical RAS operation without interruption?
- Grid control scheme such as RAS must be reviewed in face of new operational frontier
- TRAST could be the answer!
  - New data, Different model
  - Same data model convergence study in TRAST
  - PNNL Prototype available for testing, ready for more RAS!

Figure 25. An illustration of electricity grid. Wikipedia, Own work Originally derived from de:Datei:Stromversorgung.png, CC BY 3.0.
Acknowledgement:
Department of Energy

The Project team at PNNL:
Xiaoyuan Fan (xiaoyuan.fan@pnnl.gov)
Renke Huang, Qiuhua Huang,
Emily Barrett, Xinya Li, Jason Hou,
James O’Brien, Huiying Ren

Peak Reliability Collaborators:
Hongming Zhang, Slaven Kincic

Western Utility Advisors:
Song Wang (PacifiCorp),
Orlando Ciniglio (Idaho Power Company)

Former Team Member:
Ruisheng Diao
Reference & Web-links


