Preprints of the 21st IFAC World Congress (Virtual)
Berlin, Germany, July 12-17, 2020

Probabilistic Look-ahead Contingency Analysis Integration with Commercial Tool and Practical Data ⋆

Yousu Chen * **Huiying Ren** ** **Jun Wu** *** **Zhangshuan Hou** ****

* *Pacific Northwest National Laboratory, Richland, WA 99352 USA (e-mail: yousu.chen@ pnnl.gov).*
** *Pacific Northwest National Laboratory, Richland, WA 99352 USA (e-mail: huiying.ren@pnnl.gov)*
*** *GE Grid Solutions, Redmond, WA, (e-mail: jun.wu@ge.com)*
**** *Pacific Northwest National Laboratory, Richland, WA 99352 USA (e-mail: zhangshuan.hou@pnnl.gov)*

Abstract: This paper presents an initial effort of integrating a smart sampling-based probabilistic look-ahead contingency analysis algorithm with a commercial energy management system (EMS) tool as a proof-of-concept for a seamless research tool integration using real world large-scale grid data. With the increasing impact of random forces such as variable generation and load, their stochastic behaviors cannot be ignored. However, the current practices are still dominated by deterministic tools. They are becoming increasingly inadequate for the future grid. The developed look-ahead contingency analysis algorithm incorporates forecast errors of variable energy and load to address the challenges brought by the increasing uncertainty of power system. The algorithm can reveal the potential violations caused by the variance of variable energy and load that are not normally detected by traditional deterministic approaches. To test its performance under practical environments (practical data with commercial tool), significant efforts have been made to prepare test cases, modify the commercial tool to interface with the probabilistic algorithm, and adapt an extreme value distribution algorithm to analyze the commercial tool's violation-only outputs. The test results clearly demonstrate the effectiveness of the developed algorithm as new transformer violations that were not previously detected have been identified. This performance provides better situational awareness to engineers for their decision-making process under uncertainty. Moreover, with the discussion of computational performance and future work, this paper has shown a clear path for integrating the probabilistic algorithm with commercial tools to make us better equipped for the changing power system.

Keywords: Modeling and simulation of power systems, constraint and security monitoring and control, look-ahead contingency analysis, energy management system.

1. INTRODUCTION

Contingency analysis (CA) is a critical energy management system (EMS) function that is widely used to evaluate the electric grid's status if various combinations of component failures occur in the grid. The outputs of CA provide information to assess grid health to ensure that there are no operating limit violations in the system. In North America, power companies are required to perform CA for grid operations and planning studies to meet the North American Electric Reliability Corporation (NERC) operation standards. Today's CA functions in the control centers are dominated by deterministic methods. It uses state estimation outputs to perform power flow and contingency analysis study. When there are violations, operators are required to take actions to maintain a system's reliability according to NERC standards. With the fast growth of power grids, power companies are under pressure to take more contingencies into consideration as the demand of power supply pushes boundary constraints of the grid. Therefore, parallel processing and/or multi-threading techniques are utilized to perform CA to reduce the simulation time. For instance, multi-threading technology has been used in General Electric (GE) Grid Solutions' EMS tools.

In addition to examining the current operational cycle, some companies also conduct look-ahead contingency analysis functions to foresee the system's health condition in next operational intervals; e.g., in next 1 hour with a 15-minute interval. Commercial vendors, such as GE Grid Solutions' Multiple-timepoint study (SMTNET) tool, have the look-ahead contingency analysis capability developed. However, exact forecast values, as well as generation schedules, load forecast, and outage plans, are used as the inputs to perform study. Therefore, it is still a deterministic

⋆ This work is supported by the U.S. Department of Energy Office of Electricity (DOE-OE) Advanced Grid Modeling (AGM) program and the Bonneville Power Administration Technology Innovation (BPA TI) program. Pacific Northwest National Laboratory (PNNL) is operated by Battelle for the DOE under Contract DOE-AC05-76RL01830.

approach. The simulation results might not be accurate because they do not consider the stochastic nature of the forecast in the simulation. Some critical violations caused by variance of variance energy could be missed if traditional deterministic approach is used.

With increasing impacts of variable generation and load uncertainties, the traditional contingency analysis needs to be extended to cover the difference between forecast value and actual value. Thus, these impacts should be studied to maintain system reliability, which requires the function of probabilistic look-ahead contingency analysis to provide a more complete picture to operators. The methodology of the probabilistic approach to contingency analysis has been studied by other researchers, such as Brønmo et al. (2015) and Yue and Wang (2016). However, to the authors' best knowledge, there are not other papers that explore the path forward for integrating a probabilistic algorithm with a commercial EMS tool like this one.

This paper is based on the authors' previous effort in Chen et al. (2016). It presents a probabilistic look-ahead contingency analysis algorithm using smart sampling and probabilistic analysis approaches, and shows the initial effort on integrating the developed algorithm with commercial SMTNET. A practical prototype tool has been developed to validate the performance of this algorithm on the Bonneville Power Administration (BPA) system. With the finding of new violations that are missed by deterministic methods, the advantages of the algorithm are also demonstrated.

The remainder of this paper is organized as follows: Section 2 introduces the architecture of the probabilistic look-ahead contingency analysis algorithm. Section 3 presents how the study cases were prepared with real-world utility data and fed into the SMTNET tool. Section 4 presents the case study results to show the advantage of the proposed algorithm. Section 5 concludes the paper with a discussion of future work.

2. PROBABILISTIC LOOK-AHEAD CONTINGENCY ANALYSIS FRAMEWORK

The framework of the probabilistic look-ahead contingency analysis is shown in Fig. 1. This framework is developed based on an early version (Chen et al. (2016)) that was implemented with an in-house massive contingency analysis (MCA) tool (Huang et al. (2009)), (Chen et al. (2012)). In this paper, the algorithm has been enhanced to integrate with the GE Grid Solutions' SMTNET tool.

The core function in this framework is a smart sampling approach, which factors forecast errors into look-ahead contingency analysis applications. The inputs for smart sampling techniques are historical forecast information and historical actual information. The outputs of smart sampling techniques are a reduced and representative set of realizations (scenarios) that consider the forecast errors efficiently, which leads to a reduced computational burden.

In addition to help address forecast uncertainties, the smart sampling algorithm can also be used on contingency lists, (i.e., a reduced set of contingencies, instead of full list, is assigned to each realization). These reduced and representative sets of realizations are then written in GE SMTNET format. SMTNET then reads this information in, and can perform multi-thread contingency analysis computations. Because the SMTNET only outputs contingency violations, instead of the full solutions, the extreme value distribution (EVD) algorithm is adapted and applied to this framework for branch overflow violations.

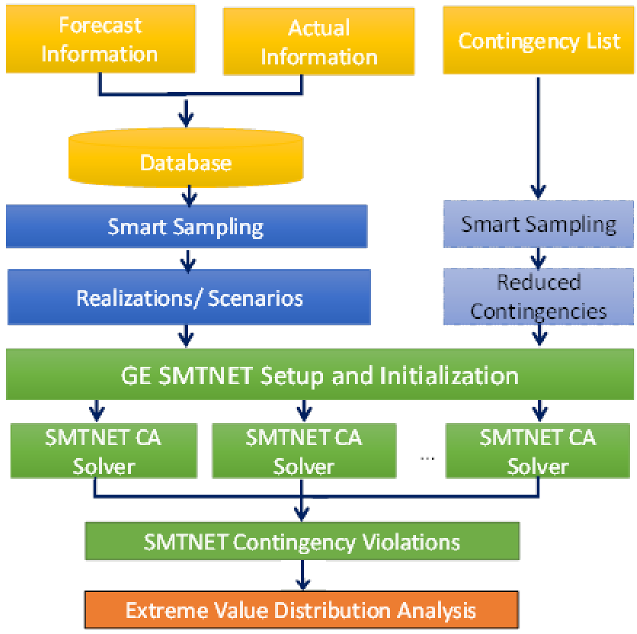

Fig. 1. Look ahead contingency analysis framework

2.1 *Smart Sampling Algorithm*

The purpose of the smart sampling algorithm is to address slow convergence of the traditional Monte Carlo methods with good coverage and less redundancy. This helps to generate a reduced and representative dataset that covers the forecast errors for look-ahead type of analysis.

Traditional Monte-Carlo-based sampling, such as general random sampling (GRS) (Davis (1987)), generates pseudo-random sequences as needed for the trials. It is often computationally expensive with slow convergence rate. Therefore, efficient sampling is needed to minimize the number of required Monte Carlo runs but can fully explore the system characteristics. Quasi-Monte Carlo (QMC) (Niederreiter (1978)) sampling utilizes sequences of quasi-random numbers that can be deterministically-chosen based on equally distributed sequences to help minimize sampling errors. One main difference between the QMC method and GRS method is that pre-designed deterministic points are used in QMC, instead of random points. The advantage of QMC is that the elements are well dispersed and spatially covered which would enable a more rapid convergence for either Monte-Carlo integration or for ensemble system representation. The QMC method is the smart sampling algorithm that has been integrated in the look-ahead contingency analysis framework.

The comparison of samples generated by GRS (left) and QMC (right) is shown in Fig. 2. The red points are sampling points obtained for two input variables, and the blue ones are the projections of these sampling points onto the two dimensions. An even distribution of blue dots along

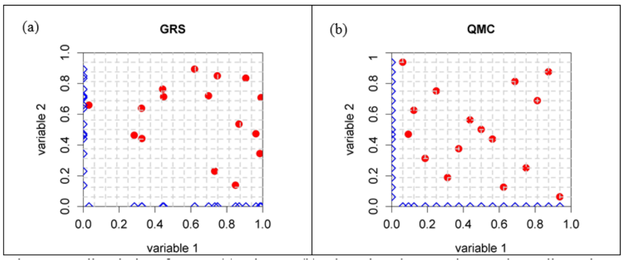

Fig. 2. Comparison of GRS (a) and QMC (b)

the variable axis (space-filling) is expected for a good sampling approach. Moreover, without prior information on the relative weights of the different parameters (or dimensions), it is vital that no two design points coincide when projected onto a lower number of dimensions (i.e., non-collapsing). The "space-filling" and "non-collapsing" properties are important criteria to evaluate the performance of sampling algorithm. Fig. 2 clearly shows the QMC outperforms the GRS method.

2.2 Parallelization Setup

The smart sampling outputs are a set of realizations (scenarios) and a list of contingencies. The implementation of the original algorithm uses a dynamic load-balancing-scheme-based to execute each contingency at different realizations in parallel (Chen et al. (2012)). However, the SMTNET's built-in multi-thread option can only be applied to contingency level at each time step and the computation of each time point is still sequential. Therefore, a series of SMTNET runs need to be setup to mimic the original implementation because the number of realizations is greater than the optimal number of time points in one SMTNET run. To achieve scalability comparable to the original implementation in Huang et al. (2009), further modifications on SMTNET are required to allow parallelization at the time point level.

2.3 Extreme value distribution

A probability density function (PDF)-based algorithm was used to post-process the look-ahead CA outputs in Chen et al. (2016) because it uses an in-house massive contingency analysis tool that can provide full solutions to all contingencies at all scenarios. However, GE SMTNET only outputs violation information, the PDF-based approach would not work.

The violations can be considered as extreme values that occur when a risk takes values from the tail of its possibility distribution. Extreme value theory studies extreme deviations from the median of probability distributions in statistics. It has been widely applied in a variety of areas of risk analysis. In this paper, the extreme value distribution (EVD) algorithm has been innovatively adapted to case studies.

Let's introduce the basis of EVD algorithm. Suppose X_1, X_2, X_3 are independent random variables with the common cumulative distribution function, $\mathcal{F}$, and a new dataset, $M_n = \max(X_1, X_2, X_3 ... X_n)$, can be created by including only the first n maximum values. This dataset can only be represented by EVD models. In statistics, the Fisher-Tippett-Gnedenko theorem is a general result in extreme value theory for the behavior of the maxima of independent, identically distributed sequences. The continuous probability distribution based on this theory is a generalized extreme value (GEV) distribution, shown in (1). This continuous version is developed by combining Gumbel Gumbel (2012), Fréchet and Weibull distribution families (Coles et al. (2001)):

$$\mathcal{F}(x; \mu, \xi, \varphi) = exp\left\{-(1+\xi\frac{x-\mu}{\varphi})^{-1/\xi}\right\} \tag{1}$$

where μ is a location parameter, φ is a scale parameter and ξ is the shape parameter. The single GEV distribution can be represented using these three parameters. When shape parameter ξ=0, the GEV corresponds to the Gumbel distribution, also called type I extreme value distribution. The ranges of interest for the three extreme value distributions are different; Gumbel distribution is unlimited, while Fréchet distribution has a lower limit.

In this study, the parameter ξ was estimated in GEV for differnt scenarios, most of them are close to zero. So the Gumbel distribution is used to analyze the distribution of the overflow violation. Each power flow containing contingencies is considered as an extreme event. Because the probability distribution of violation is estimated with uncertainties, a 95% confident interval is used in the GEV distribution fitting to represent the model uncertainties.

3. CASE PREPARATION FOR SMTNET INTEGRATION WITH REAL-WORLD DATA

This section discusses what has been done to test the developed look-ahead CA on GE SMTNET with real-world data provided by BPA.

3.1 Real-world Data Pre-processing

The real-world data used in this study include:

(1) The BPA operational node-break models in a 15-minute interval for the year of 2014. It is used to extract historical actual load and wind generation information.
(2) BPA historical forecast data (wind/load) in 2014 (BPA (2014)). The wind data is hourly and the load data is at 5-minute interval.
(3) A BPA planning model in PSS/E PTI format in 2014.
(4) Location of BPA wind farms.
(5) BPA-solved STNET (steady state network study) and RTCA (real-time contingency analysis) save cases on the selected days in 2014, including all key modules, outage files, and resource plans.

There are two issues associated with the actual and forecast wind/load data: (1) inconsistent sampling rate in actual data and forecast data; and (2) incomplete data due to unsolved state estimation cases. To tackle these issues, down-sampling techniques with gap filling were performed. Extensive effort has been expanded to prepare a mapping table between the operational models and the planning model to identify the actual values at each wind unit. The initial study with the planning model in Chen et al. (2016) has shown very promising results and provided meaningful

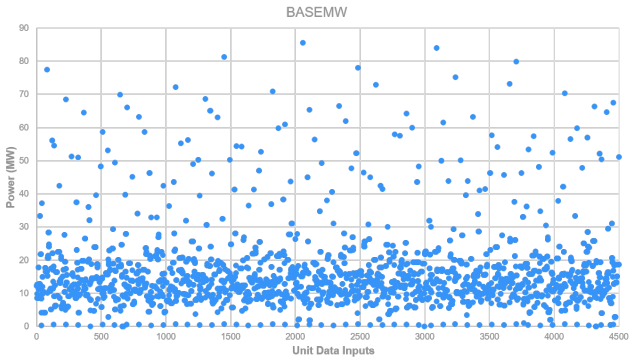

Fig. 3. The range of all wind unit generations in 144 smart sampling realizations

value to the users. Therefore, further requests were sent to BPA to have BPA-solved operational case in GE Grid Solutions' EMP 2.6 format to test the algorithm with real data on GE EMS tools. Because SMTNET is only available at EMP 3.1 and above versions, the database definition and size of all records were also requested to help transition BPA's database to the SMTNET function-capable EMP.

3.2 Realization creation

After matching and pre-processing all of the BPA wind and load data, 32 variables, including 31 wind farms and the total BPA load, are included in the smart sampling realization generation. These realizations are converted to the format that is suitable for the GE SMTNET model. A sensitivity study based on the number of realizations and the efficiency of the QMC was conducted to find a reasonable number of realizations generated by the QMC that can represent forecast errors. Based on the sensitivity study results, 144 realizations were generated and divided into three SMTNET runs (48 realizations each) to facilitate task-parallel computing and optimally use the computing resources. To have a better idea of the wind generation range, all 4,608 (144*32) wind generations (MW) are shown in Fig. 3.

3.3 Commercial Tool Modifications

The GE SMTNET is a platform to perform a look-ahead security analysis for future time points given a set of inputs. It can run multiple time points with multiple network models and summarize violations for each time step. In this study, the 144 realizations created by smart sampling algorithms can be considered as a new network model with different sizes of renewable generation and load.

To utilize SMTNET, the BPA operational cases were converted from EMP 2.6 to EMP 3.1. A resizing process was conducted to let GE's SMTNET/STNET application run with the provided BPA real-time cases.

To integrate with the PNNL probabilistic look-ahead contingency analysis function, the SMTNET application has been modified to read a data set of BPA wind, load, contingency combination and allow user to perform the SMTNET run with each data set. The SMTNET outputs format were also modified to output branch flow violations for each contingency at each scenario for the EVD analysis.

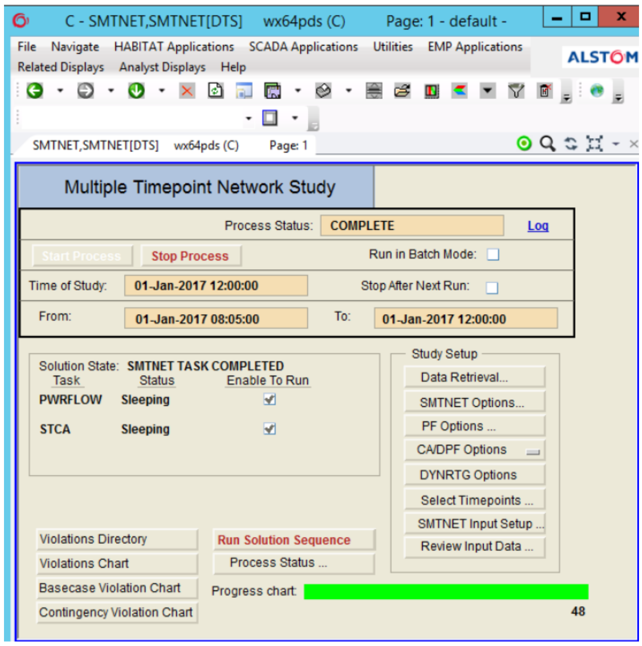

Fig. 4. The screenshot of SMTNET results with 48 time points

4. PERFORMANCE EVALUATION

The performance evaluation was conducted on the GE Grid Solutions Habitat virtual machine (VM). The VM was installed on Windows Server 2012 R2 standard with a processor of Intel Xeon X7550 @ 2.0 GHz(four cores) and 24 GB memory. Each realization was set up to run with a range of contingencies from 20 to 1,241 active contingency. A comparison between the original STNET CA results and the proposed look-ahead CA algorithm results is the focus of this evaluation.

4.1 STNET base case contingency results

STNET is the GE EMS tool function to study network analysis in steady state. To have a fair comparison, the BPA base case with the original contingency setup was performed in GE Grid Solutions' STNET application. This study has focused on the contingency analysis results, instead of base case results because the same base case is used to both STNET and SMTNET, except the changes of wind units and BPA load introduced by the smart sampling algorithm. In our case study, there are 26 violations total, including 18 branch violations. These 18 violations happened with 8 contingencies. For data protection reason, the detailed violation information is not provided.

4.2 Commercial Tool Results

SMTNET is a GE EMS tool function to perform a multiple time point study given a set of inputs. This function was leveraged for the probabilistic look-ahead CA algorithm by assigning different wind units and load realizations to different time points, and performing probabilistic analysis (EVD) on the violation only outputs of SMTNET. A screenshot of SMTNET graphical user interface is shown in Fig. 4.

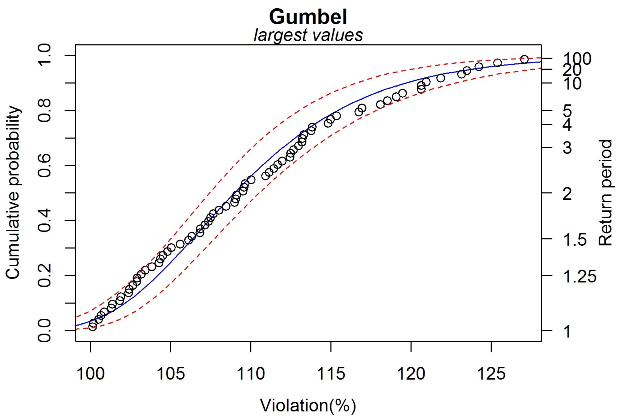

Fig. 5. EVD Distribution for the Line ID 1061

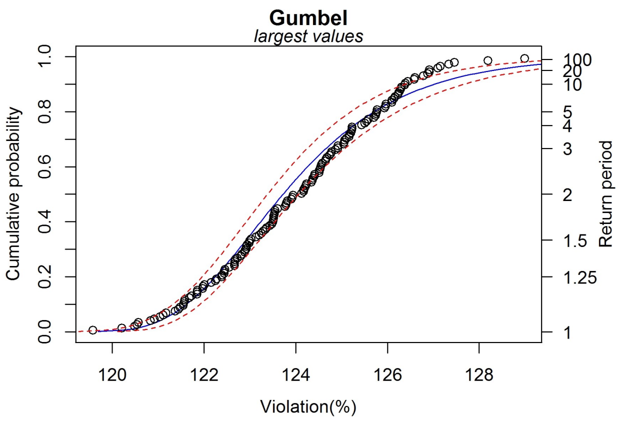

Fig. 6. EVD Distribution for the Line ID 2475

In our study, one SMTNET simulation runs 48 realizations with 1,241 contingencies for each time point. There are a total of 59,568 simulations for one SMTNET run. As a result, the probabilistic look-ahead contingency analysis function produces more contingency violations, such as branch flow violations and low voltage violations. For example, for the scenario at the time point of 08:00:00, the number of total contingency violations is 29, including 20 branch flow violations, while the base case contains 26 violations, including 18 branch flow violations. This difference is caused by the variance of wind unit capacity and load as a result of forecast errors.

Every SMTNET run creates an output file that contains branch violations for each realization and contingency. An R script has been developed to read this file and perform EVD analysis to present the probabilistic contingency violations to users. Two examples of EVD curves for the line ID 1061 and 2475 are shown in Fig. 5 and Fig. 6, respectively, where the x-axis is the branch overflow violation in percentage and y-axis is the cumulative probability. The blue solid line is the Gumbel fitting curve and the red dash lines are the uncertainty bounds associated with the Gumbel model. These figures provide users additional information on the distribution of the branch violations.

The R script also extracted the frequency and worst violations for the branches that contain violations. Table 1 summarizes the frequency of branch overflow violations, as well as the worst branch violations happened in all contingency simulations.

Table 1. The frequency and worst violations for the branches that contain violations

Line ID	Frequency	Worst Violation (%)
1061	72	127.1
455	144	120.0
1815	288	135.8
1817	288	132.7
3101	144	105.9
3017	287	119.2
3099	267	117.0
1445	144	129.0
1485	144	115.4
1488	144	120.5
2474	144	124.3
2475	144	129.0
3602	144	154.4
3019	91	105.6
581	80	105.0

Some new transformer violations that belong to the line ID 1061 in Table 1) were identified by the probabilistic look-ahead CA function. These violations. The violations were caused for two contingencies, with a worst violation percentage of 113% and 112.5%, respectively. The number of the transformer violation occurrences is 72 total. Therefore, this type of violations cannot be overlooked considering the frequency and the severity level. This finding demonstrates the advantages of probabilistic CA algorithm as it can capture the potential violations caused by variations of renewable energy forecast that are ignored by the deterministic method. The application of this developed algorithm can help utilities to better prepare for more complex and dynamic grids.

4.3 Computational performance

Computational performance is another critical factor to meet the challenges of more complex and dynamic grid. The original development with in-house massive contingency analysis function was able to achieve near-linear speedup with more than 10,000 cores on a super cluster (Chen et al. (2012)). However, such a capability and computation platform was not readily available to the GE Grid Solutions EMS tool as it requires a significant effort the revise the GE's legacy code. Therefore, only the SMTNET built-in multi-thread feature is available for scalability testing. As mentioned earlier, this multi-threading capability only affects contingency-level analysis. This means that the simulation with 48 time points (realizations) is still sequential.

To determine the optimal number of threads for the case study, a scalability test was conducted with different numbers of threads for 1,241 BPA contingencies. The test results are shown in Table 2.

Table 2 shows no additional gain by using more than six threads. Therefore, six threads were selected for SMTNET runs. The improvement in time over a single thread is about 40%. Compared against the near-linear scalability with the in-house massive contingency analysis tool, the scalability with multiple-thread techniques in the Windows environment has room for improvement.

Table 2. The scalability testing with different numbers of threads for 1,241 contingencies

# of threads	Computational time (second)
1	550
2	435
4	388
6	321
8	323

Table 3 summarizes the contingency analysis performance study based on the number of contingency cases at one time point using six threads.

Table 3. Computational time vs. number of contingencies using six threads

# of contingency	Computational time (second)
20	29
65	40
188	96
322	143
391	174
508	220
849	228
1027	329
1171	370
1232	385
1443	461
1837	587

Based on Table 3, the estimated computational time for 48 time points would be around 5 hours, which would be too long for look-ahead CA function. This emphasizes the need of applying more efficient parallelization mechanism to both levels of contingency tasks and time points like what has been done in Chen et al. (2016).

5. CONCLUSION AND FUTURE WORK

This paper has shown a clear path for integrating the probabilistic look-ahead contingency analysis algorithm with GE Grid Solutions' SMTNET tool. The work reported here represents a preliminary study with the real model data. The advantages of the developed algorithm have been demonstrated through case studies: it can help reveal violations not previously detectable by traditional deterministic approaches, providing better situational awareness for grid operation for enhanced reliability and security.

The original plan was to deploy the entire package, including GE Habitat EMS tool, on a cluster machine to accelerate the time intensive look-ahead contingency analysis computation. This was expected to lead to parallelizing the contingencies and time point through message passing interface (MPI)-based mechanisms, which are much more efficient than the Windows built-in multi-thread based mechanism. Therefore, the potential next step is to work with GE to identify the path forward to implement MPI-based approach for parallelizing both contingencies and time points to meet the real-time operation requirements.

Visualizing the probabilistic contingency output is also a key item for future work. Recently, with GE Grid Solutions' help, PNNL linked GE Habitat data base to PI data base. As a result, when SMTNET creates new outputs, the outputs will be stored in PI data base automatically. Therefore, another future work is to develop a web-based visualization tool that can manage multiple dimension data from PI and other data resources and visualize look-ahead contingency analysis results in a probabilistic manner intuitively.

REFERENCES

BPA (2014). BPA forecast in 2014. URL `http://transmission.bpa.gov/business/operations/Wind`.

Brønmo, G., Vergnol, A., Gamov, N., Bukhsh, W.A., Bell, K., Andreev, A., Zahov, V., and Zahov, B. (2015). Generally accepted reliability principle with uncertainty modelling and through probabilistic risk assessment: Functional analysis of system development process.

Chen, Y., Etingov, P., Ren, H., Hou, Z., Rice, M., and Makariov, Y. (2016). A look-ahead probabilistic contingency analysis framework incorporating smart sampling techniques. In *2016 IEEE Power and Energy Society General Meeting (PESGM)*, 1–5. doi:10.1109/PESGM.2016.7741449.

Chen, Y., Huang, Z., and Rice, M. (2012). Evaluation of counter-based dynamic load balancing schemes for massive contingency analysis on over 10,000 cores. In *2012 SC Companion: High Performance Computing, Networking Storage and Analysis*, 341–346. doi:10.1109/SC.Companion.2012.53.

Coles, S., Bawa, J., Trenner, L., and Dorazio, P. (2001). *An introduction to statistical modeling of extreme values*, volume 208. Springer.

Davis, M.W. (1987). Production of conditional simulations via the lu triangular decomposition of the covariance matrix. *Mathematical Geology*, 19(2), 91–98. doi:10.1007/BF00898189. URL `https://doi.org/10.1007/BF00898189`.

Gumbel, E. (2012). *Statistics of Extremes.* Courier Corporation.

Huang, Z., Chen, Y., and Nieplocha, J. (2009). Massive contingency analysis with high performance computing. In *2009 IEEE Power Energy Society General Meeting*, 1–8. doi:10.1109/PES.2009.5275421.

Niederreiter, H. (1978). Quasi-monte carlo methods and pseudo-random numbers. *Bull. Amer. Math. Soc.*, 84(6), 957–1041. URL `https://projecteuclid.org:443/euclid.bams/1183541461`.

Yue, M. and Wang, X. (2016). An enhanced probabilistic contingency analysis for development of risk-informed transmission planning. In *2016 IEEE Power and Energy Society General Meeting (PESGM)*, 1–5.