Empirical Analysis of the Optimal Capacity Investment Solutions in Distribution Grids

Luis Lopez, Alvaro Gonzalez-Castellanos, David Pozo
Center for Energy Science and Technology
Skolkovo Institute of Science and Technology (*Skoltech*)
Moscow, Russia
{luis.lopez, alvaro.gonzalez, d.pozo}@skoltech.ru

***Abstract*—This paper presents an analysis of the stability and quality of the distributed generation planning problem's investment solution. A two-stage stochastic programming model is used to find the optimal distributed generators' installed capacities. We emphasize the design of scenarios to represent the stochasticity of power production from renewable sources. For scenario generation, a method is proposed based on the clustering of real measurements of meteorological variables. The quality and stability of the optimal investment solutions are thoroughly analysed as a function of the number of selected scenarios. The results show that a reduced selection of scenarios can give an inadequate solution to distributed generators' investment strategy.**

***Index Terms*—Clustering, Distributed generators, Investment solution, Stochasticity.**

NOMENCLATURE

Indices and Sets

$(n, m) \in L$	Index/set of lines	
$\omega \in \Omega$	Index/set of scenarios	
tech	Set of installable technologies {PV, WT, CG}	
$n \in N$	Index/set of power nodes	
$t \in T$	Index/set of time blocks	

Production Models Parameters

α	Temperature coefficient power	[%/°C]
G_T	Solar radiation incident on the PV array	[kW/m^2]
T_a, T_c	Ambient and PV cell temperature	[°C]
$T_c^{\text{NOCT/STC}}$	PV cell temperature under NOCT/STC	[°C]
v_i, v_r, v_o	Cut-in, rated, and cut-off wind turbine speeds	[m/s]
$Y_{\text{PV}}, Y_{\text{WT}}$	PV array and wind turbine rated capacity.	[kW]
$G_T^{\text{NOCT/STC}}$	Solar radiation at NOCT/STC	[kW/m^2]

Parameters

β_n^{tech}	Binary parameter for buses available to install DG	
$\lambda^{\text{+/-,tech}}$	Technologie's lead/lagging power factor	[p.u.]
$\overline{P_n}$	Maximum installable active power at n	[p.u.]
$\text{pf}^{\text{+/-,tech}}$	Technologies' lead/lagging power factor angle	[rad]
$R_{n,m}$	Resistance line from bus n to bus m	[p.u.]
S_b	Apparent power base	[kVA]
$X_{n,m}$	Reactance line from bus n to bus m	[p.u.]
$Z_{n,m}$	Impedance of branch n, m	[p.u.]

This work was funded by the Skolkovo Institute of Science and Technology as a part of the Skoltech NGP Program (Skoltech-MIT joint project).

Variables

Variables are defined within the text in section II-A.

I. INTRODUCTION

In modern distribution networks, users can inject active power into the grid via small capacity generation [1]. The generators connected near demand buses are called distributed generation units (DG units) [2]. Large amounts of power are being injected through DG units due to policies that promote renewable energies in different countries [3]–[5]. With DG's widespread deployment, the distribution system operator needs to plan and coordinate the new DG units' installation capacity. DG planning can reduce operating costs and provide support during the operation regime [6].

The investment solution in distribution networks refers to determining the installed capacities and locations of DG units. When DG units are power-based on renewable technologies, they behave as non-controllable stochastic negative load. Thus, we need to capture the uncertainty associated with meteorological measurements [7]. Modeling the stochasticity of renewable generation sources has been widely confronted by several authors [2], [6], [8], [9]. Jooshaki *et al.* [10] propose a tool for integrating DG units using a mixed-integer linear stochastic model and perform a case study on a 24-node distribution network. In [11], the authors proposed a methodology using mixed-integer stochastic programming to find the best reinforcement plan for mitigating greenhouse gas emissions. In [12], a stochastic model is proposed to address the problem of distribution system expansion with uncertainties of DG units and issues related to CO2 emissions [13].

Stochastic programming is a mathematical framework that allows modeling the uncertainty of power production from non-conventional renewable sources [14], [15]. It has been proposed in [16] to use Sample Average Approximation (SAA) to generate scenarios in the planning problem with stochastic parameters. Nevertheless, scenario generation techniques are limited because they are a discrete approximation of (most times unknown) probabilistic distributions. Therefore, the stochastic model relies on scenario representation, and if the scenario representation is deficient, information about the actual probability distribution may be lost. This work proposes a two-stage stochastic programming model that provides an

Model 1 Sitting and sizing of distributed generation with non-conventional renewable energies

Objective:

$$\min\ (\pi^{\text{inv}} + \pi^{\text{OM}}) \tag{1}$$

Constraints:

$$\pi^{\text{inv}} = \sum_{n,\text{tech}} \left(\pi^{\text{inv,tech}} x_n^{\text{tech}}\right) \tag{2}$$

$$\pi^{\text{OM}} = \sum_{\tau} \rho_\tau \pi_\tau^{\text{OM}} \tag{3}$$

$$\pi_\tau^{\text{OM}} = \pi_\tau^{\text{loss}} + \pi_\tau^{\text{SS}} + \pi_\tau^{\text{DG}} \tag{4}$$

$$\pi_\tau^{\text{loss}} = \pi^{\text{loss}} \sum_{n,m\in L} S_b R_{n,m} i_{n,m,\tau}^2 \tag{5}$$

$$\pi_\tau^{\text{SS}} = \pi_\tau^{\text{SS}} S_b p_\tau^{\text{SS}} \tag{6}$$

$$\pi_\tau^{\text{DG}} = S_b \sum_{n,\text{tech}} \pi^{\text{OM,tech}} p_{n,\tau}^{\text{tech}} \tag{7}$$

$$\gamma_\tau^{\text{D}} P_m^{\text{D}} = \sum_{n,m\in L} (p_\tau^{n,m} - p_\tau^{m,n}) - R_{n,m} i_{n,m,\tau}^2 + \sum_{\text{tech}} p_{m,\tau}^{\text{tech}} + p_{m,\tau}^{\text{SS}} \tag{8}$$

$$\gamma_\tau^{\text{D}} Q_m^{\text{D}} = \sum_{n,m\in L} (q_\tau^{n,m} - q_\tau^{m,n}) - X_{n,m} i_{n,m,\tau}^2 + \sum_{\text{tech}} q_{m,\tau}^{\text{tech}} + q_{m,\tau}^{\text{SS}} \tag{9}$$

$$2\left(R_{n,m} p_\tau^{n,m} + X_{n,m} q_\tau^{n,m}\right) = v_{n,\tau}^2 + |Z_{n,m}|^2\, i_{n,m,\tau}^2 + v_{m,\tau}^2 \tag{10}$$

$$w_{n,m,\tau} \geq \underline{V} i_{n,m,\tau}^2 + v_{n,\tau}^2 \underline{I}_{n,m}^2 - \underline{I}_{n,m}^2 \underline{V} \tag{11}$$

$$w_{n,m,\tau} \geq \overline{V} i_{n,m,\tau}^2 + v_{n,\tau}^2 \overline{I}_{n,m}^2 - \overline{I}_{n,m}^2 \overline{V} \tag{12}$$

$$w_{n,m,\tau} \leq \overline{V} i_{n,m,\tau}^2 + v_{n,\tau}^2 \underline{I}_{n,m}^2 - \overline{V} \underline{I}_{n,m}^2 \tag{13}$$

$$w_{n,m,\tau} \leq v_{n,\tau}^2 \overline{I}_{n,m}^2 + \underline{V} i_{n,m,\tau}^2 - \underline{I}_{n,m}^2 \overline{I}_{n,m}^2 \tag{14}$$

$$\underline{V}^2 \leq v_{n,\tau}^2 \leq \overline{V}^2 \tag{15}$$

$$i_{n,m,\tau}^2 \leq \overline{I}_{n,m}^2 \tag{16}$$

$$\overline{P}^{\text{PV}} x_n^{\text{PV}} + \overline{P}^{\text{WT}} x_n^{\text{WT}} + \overline{P}^{\text{CG}} x_n^{\text{CG}} \leq \overline{P}_n \tag{17}$$

$$0 \leq p_{n,\tau}^{\text{tech}} \leq \gamma_\tau^{\text{tech}} \overline{P}^{\text{tech}} x_n^{\text{tech}} \beta_n^{\text{tech}} \tag{18}$$

$$\lambda^{\text{tech},+} p_{n,\tau}^{\text{tech}} \leq q_{n,\tau}^{\text{tech}} \leq \lambda^{\text{tech},-} p_{n,\tau}^{\text{tech}} \tag{19}$$

$$\lambda^{\text{tech,+/-}} = \mp \tan(\cos^{-1}(\text{pf}^{\text{+/-}})) \tag{20}$$

$$\pi^{\text{inv}} \leq \Pi^{\text{bgt}} \tag{21}$$

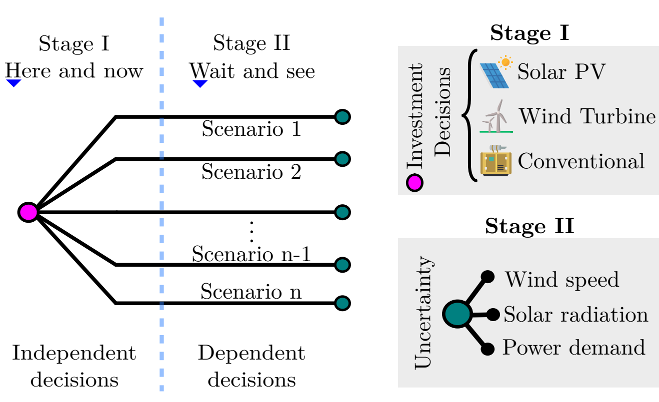

Figure 1. Two-stage stochastic approach.

investment solution considering short-term uncertainty in a long-term planning problem, analyzing the quality of the investment solution as a function of the number of scenarios used. We employ the popular k-means clustering technique for scenario generation to reduce the problem's dimensionality and capture the underlying correlation between the random variables.

This work's main contribution is the analysis of the quality and stability of the investment solution in the DG planning problem using empirical measurements. We assess how the investment solution deviates from its ground-truth value when we use an inadequate description of the problem's stochasticity (few numbers of scenarios). The remainder of this work is organized as follows: Section II describes the two-stochastic programming model and the estimation of the upper and lower boundaries. Section III introduces the case study and the scenario generation technique. Section IV shows the findings and simulations performed on a test distribution system with real measurements. Section V provides the discussions and conclusions of the observed empirical stability.

II. Methodology

Stochastic programming provides solutions using scenarios to represent possible uncertainty realizations. This section describes our methodology for solving the problem of investment in DG units using stochastic programming. In Section II-A, we briefly describe the employed model, and in Section II-B, we describe the metrics to evaluate the quality of the solution obtained.

A. Two-stage problem formulation

This article addresses DG planning's problem through a two-stage stochastic programming approach (Fig. 1). The first stage consists of the sizing and location of the investment on DG units. Three technologies of DG units are considered: solar photovoltaic (PV), wind turbines (WT), and conventional fuel-based generators (CG). The first-stage variables are integers since the power plant units are manufactured in discrete modules of installed power. The first stage considers the investment costs, while the second stage consists of the computation of the operation and maintenance cost for every scenario. The second stage calculates the expected operation cost of the power produced by the newly installed DG units. The uncertainty of power production and energy balance is associated with the meteorological variables of solar radiation, wind speed, temperature, and energy demand. Evaluating the expected value of power production given an investment decision requires numerous second-stage optimization problems representing the scenarios.

The objective function (1) minimizes the system's energy cost over the analyzed time horizon. The total cost is the sum of the investment cost and the cost of operation and maintenance. The investment cost (2) is equal to the sum of the installation costs per technology in each node. The expected operation and maintenance cost π^{OM} (3) is calculated as the sum of the products of the operation and maintenance costs per scenario and the scenario's probability of occurrence, ρ_τ. Each scenario's operation and maintenance costs depend on the costs of active power losses, the energy imported from the power grid, and the new DG units' power production costs. Demand profiles and weather conditions only depend on the scenario and not on the system nodes since distribution networks cover the same area.

The constraints of the stochastic DG planning approach are divided into physical law constraints (8)-(10) and engineering constraints (15)-(20). The physical law constraints are the active and reactive power balance constraints (generated power must supply the demanded power) and the power flow constraints through the transmission lines. The power flow constraints are modeled through the DistFlow equations [17], [18]. The McCormick envelope (11)-(14) is used as a relaxation technique to solve the product of two bounded variables for calculating the apparent power of the DistFlow equations. Engineering constraints are set by the distribution system operator and include: node voltage limits, line loadability limits, installed DG capacity limits, reactive power DG limits set by power factors, and investment limits determined by the available budget (21).

B. Quality and stability of the optimal solution

Our investment problem formulation described in Model 1 can be compactly summarizes as a classical two-stage stochastic optimization problem (22).

$$\begin{aligned} z^* = \min \quad & c^T x + \mathbb{E}_P \left[d^T \mathbf{y} \right] && \text{(22a)} \\ \text{s.t.:} \quad & x \in X && \text{(22b)} \\ & \mathbf{y} \in \mathbf{Y}(x) && \text{(22c)} \end{aligned}$$

The vector x is representing investment decisions at the first stage while the random vector $\mathbf{y}$ represents the operational decisions at the second stage. The objective (22a) aims to minimize the capital and expected operational costs. At the same time, the budget-limit constrains and power grid operational feasibility constraints are represented by (22b) and (22c), respectively. The symbol $\mathbb{E}_P$ is the expected operator over the random parameter distribution P. If P represents a continuous distribution vector, this problem is infinite-dimensional, and different approaches have been proposed for solving it.

In the rest of this subsection, we describe the metrics to evaluate the investment solution's quality and stability. We will use the sample-average approximation (SAA) method, [16] for approximating problem (22).

A particular feature of this work is that data is collected from an actual grid. Thus, instead of inferring continuous parametric distributions, we use directly observed data in the investment problem addressed here. Still, data can potentially have massive amounts of data points, so we need to find means to reduce the problem's computational complexity. We denote by N the total number of collected data points, i.e., the original set of scenarios.

In this work, a set of SAA scenarios are generated using the k-means clustering technique explained in Section III-A. The solution of Model 1 using SAA has the following properties.

1) Lower bound estimation: Using the SAA algorithm, we estimate the lower bound value for the DG planning problem's investment solution. To evaluate the lower bound, we solve m replicas of the two-stage problem (22) with n scenarios (where $n < N$). We initially generate m sample sets independently with n scenarios and then solve the approximated sample-based optimization problem (23). The optimal objective of this problem is a lower bound of the original problem (22), i.e., $\mathbf{LB}_m(n) \leq z^*$ for any replica m. Because the n-drawn scenarios are random, the $\mathbf{LB}$ is also a random parameter.

$$\begin{aligned} \mathbf{LB}_m(n) = \min \quad & c^T x + \frac{1}{n} \sum_{k=1}^{n} d^T y_k && \text{(23a)} \\ \text{s.t.:} \quad & x \in X && \text{(23b)} \\ & y_k \in Y_k(x) && \text{(23c)} \end{aligned}$$

2) Upper bound estimation: Given a trial (not necessarily optimal) solution for the first stage decision variables denoted by $\hat{x}$, we can compute an upper bound of the original problem (22) by (24), i.e., $z^* \leq \text{UB}_m(\hat{x})$.

$$\text{UB}_m(\hat{x}) = c^T \hat{x} + \mathbb{E}_P \Big[\min_{\mathbf{y} \in \mathbf{Y}(\hat{x})} \quad d^T \mathbf{y} \Big] \tag{24}$$

Optimization problem (24) is scenario-decomposable due to the fixed value of the first-stage decision variables. When the probability distribution function P is discrete, the expected value can be computed exactly for each possible random state that can be observed (scenarios). However, if the number of discrete values of the probability distribution P is large or if P is continuous, we can approach the upper bound by the law of large numbers (25).

$$\begin{aligned} \mathbf{UB}_m(\hat{x}, \widehat{N}) = c^T \hat{x} + \frac{1}{\widehat{N}} \sum_{k=1}^{\widehat{N}} + \min \quad & d^T y_k && \text{(25a)} \\ \text{s.t.:} \quad & y_k \in Y_k(\hat{x}) && \text{(25b)} \end{aligned}$$

The first observation is that $\mathbf{UB}_m(\hat{x}, \widehat{N})$ is random when $\hat{N}$ random scenarios are drawn. The second observation is that for discrete distributions, as in this paper, the random $\mathbf{UB}_m(\hat{x}, \widehat{N})$ should be approaching to the deterministic $\text{UB}_m(\hat{x})$ when $\widehat{N} \to N$.

Finally, we can estimate the optimal solution gap between the lower and upper bounds (26), which gives the statistical information about the problem's stability.

$$\mathbf{gap}_m(\hat{x}, n, \widehat{N}) = \mathbf{UB}_m(\hat{x}, \widehat{N}) - \mathbf{LB}_m(n) \tag{26}$$

III. Scenario generation and test case

There are several methods for generating scenarios from a known probability distribution or a large historical data set. In Section III-A we present the scenario generation technique based on clustering. Besides, we present a description of the case study for the computational tests in Section III-B.

A. Scenario generation

In the stochastic programming model, we analyze four parameters of uncertainty: solar radiation (W/m2), wind speed (m/s), temperature (°C), and active power consumption (W). We use a set of historical data measured with a weather station a power meter. The database has hourly measurements of the uncertainty parameters over one year of recording. The technique used for the generation of scenarios is the k-means clustering technique [19]. The k-means technique is a method to create representative clusters of a data group whose partitions are given in k clusters. All k clusters have a centroid representing the mean value of the uncertainty parameters contained in that set, minimizing variances within each cluster.

The generation of scenarios is done using the historical record of uncertainty parameters (Fig. 2). Initially, we generate k clusters containing representative data of solar radiation, wind speed, temperature, and power demand. Then we calculate the probability of occurrence of that scenario depending on the cluster's size (amount of data it represents) over the total of registered empirical scenarios. Then, the weather variables are the input to the power production models (27)-(29) of the DG units. The power production model of the PV units depends on solar radiation and ambient temperature, as worked in [20], [21]. The power production model of WT depends only on wind speed. All variables are scaled for the distribution system.

$$P^{\text{PV}} = Y^{\text{PV}} \left(\frac{G_T}{G_T^{\text{STC}}} \right) \left[1 - \alpha \left(T_c - T_c^{\text{STC}} \right) \right] \tag{27}$$

$$T_c = T_a + \frac{G_T}{G_T^{\text{NOCT}}} \left(T_c^{\text{NOCT}} - T_a^{\text{NOCT}} \right) \tag{28}$$

$$P^{\text{WT}} = \begin{cases} Y^{\text{WT}} \frac{v - v_i}{v_r - v_i}, & v_i \leq v < v_r \\ Y^{\text{WT}}, & v_r \leq v < v_o \\ 0, & \text{otherwise} \end{cases} \tag{29}$$

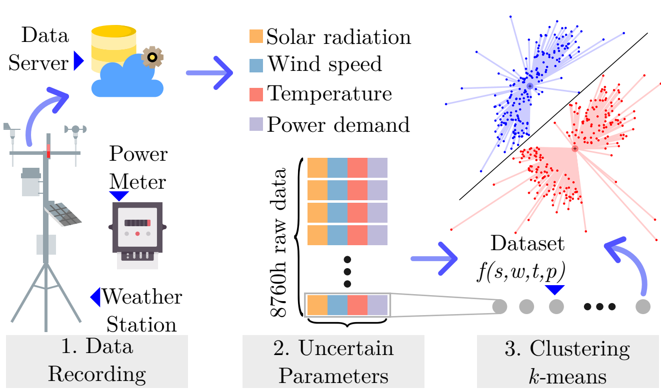

Figure 2. Scenario generation methodology.

B. Case study

The stability analysis of the investment solution is applied to the 34-node distribution system with the topology presented in [22]. The total installed system demand is 5.4 MW with an average power factor of 0.85 in the lag. Historical data was recorded from January 1 to December 31, 2018, with a weather station with an elevation of 36m at 11.02°N - 74.85°W. The two-stage stochastic programming problem was formulated in Julia v1.6 using JuMP v0.21.3 and Gurobi v9.1.1. The test machine features Microsoft Windows Server 2016 Standard, Intel Xeon Gold 6148 CPU @ 2.40GHz, 2394 Mhz, 20 Cores, 40 Logical Processors, Total Physical Memory 256 GB.

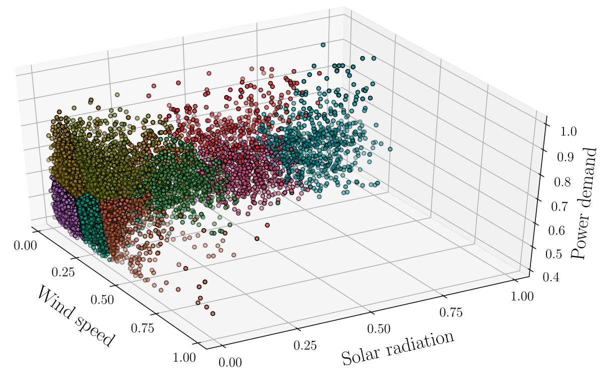

Figure 3. Empirical distribution clustering ($n = 10$).

IV. Results and simulations

We apply stability tests for the investment solution with different numbers of generated scenarios for the DG planning problem. For scenario size n, we solve the optimization problem a total of 10 times (replications). The reference value for the solution that we call ground truth is calculated with the maximum number of scenarios that we computationally manage to solve (5000).

Fig. 4 shows the optimal solution's value solution from the investment problem (Model 1), the estimated lower and upper bounds. We can see that the bounds vary with the number of generated scenarios. The optimal solution's value improves, and the optimality gap size narrows when we increase the size of the generated scenarios' set. This mainly results from the fact that the lower bounds variance decreases as we approach the full empirical distribution. This occurs because

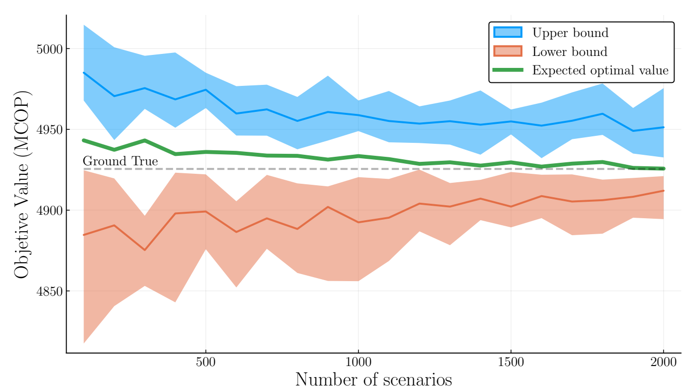

Figure 4. Optimal objective function, estimated lower bound, estimated upper bound.

the generated scenarios are clustered, and their values may be outside the initial set. The previous problem can be solved with much higher replication values, but it would considerably increase the simulation time, Fig. 5.

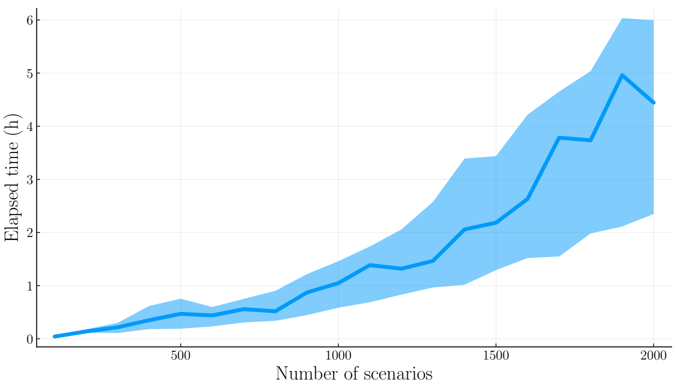

Figure 5. Solution time for the optimization problem.

The results show that using a few scenarios to solve a stochastic programming problem can lead to substantial errors and sub-optimal solutions. Additionally, the actual probability distribution and stochasticity may not be properly represented in the generated scenarios.

Fig. 6 shows the in-sample stability calculated as the optimal solution's relative value in the n scenario vs. the optimal ground actual value. Also, Fig. 6 shows that in-sample stability is improved when we increase the number of scenarios used significantly. On the other hand, Fig. 7 shows the out-of-sample stability for different numbers of scenarios. The in-sample stability is calculated using the equations and based on our previous notation; the optimal derived values are calculated using different scenarios (M sets of scenarios with N scenarios each). On the other hand, to calculate out-of-sample stability, we will insert the fixed first-stage solution of each sample m with n size into an optimization problem using the N-scenarios, representing the true distribution. From the figures, we can conclude that high variability in in-sample stability is correlated with high out-of-sample variability.

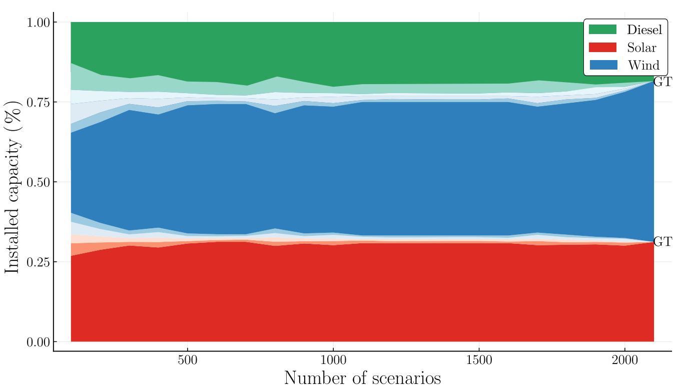

Figure 8. Normalized DG units' installed capacities mix.

Finally, we analyse the optimal technology mix resulting from the capacity investment problem under different numbers of scenarios. Fig. 8 shows the mix of installed capacities when there is no budget constraint. Total new capacity is normalized. We can see that the installed capacities highly fluctuate when we have a small number of scenarios. At the same time, that variability becomes smaller when we have a more significant number of scenarios.

V. Conclusions

This paper applies the sample average approximation (SAA) technique and stability tests to evaluate the optimal distributed network investment solution's quality obtained from a two-stage stochastic optimization model. We show that an investment solution based on a few scenarios can lead to misestimation and deviations from the true solution. On the other hand, the representation of stochasticity and scenarios' use affects the solution's quality when we have uncertainty parameters. From the stability perspective, we can conclude that the solution satisfies the criteria of in-sample and out-of-sample stability. We can conclude that the model has an excellent out-of-sample stability performance when the number of scenarios generated surpasses 1000 data points for this particular distribution grid. The deviations from the optimum relative value are less than 10% for all scenarios. For the in-sample stability test, we can conclude that for several scenarios greater than 500, the optimal value deviations are less than 10%.

The optimal share of installed technologies depends on the number of scenarios used; few scenarios lead to high variations in the optimal energy mix. A poor representation of the scenarios can lead to the oversizing of fuel-based generation, resulting in higher operational costs for the distribution network operator. Finally, we recommend using scenario generation techniques to adequately capture and represent the uncertainty parameters' real distributions. Besides, using as large a number of scenarios as is computationally feasible is highly recommended to find stable and quality solutions to stochastic DG planning.

References

[1] S. Kakran and S. Chanana, "Smart operations of smart grids integrated with distributed generation: A review," *Renewable and Sustainable Energy Reviews*, vol. 81, pp. 524–535, 1 2018.

[2] B. Singh and J. Sharma, "A review on distributed generation planning," *Renewable and Sustainable Energy Reviews*, vol. 76, pp. 529–544, 9 2017.

[3] P. Pereira da Silva, G. Dantas, G. I. Pereira, L. Câmara, and N. J. De Castro, "Photovoltaic distributed generation – An international review on diffusion, support policies, and electricity sector regulatory adaptation," *Renewable and Sustainable Energy Reviews*, vol. 103, pp. 30–39, 4 2019.

[4] M. Q. Duong, N. T. Nam Tran, G. N. Sava, and M. Scripcariu, "The impacts of distributed generation penetration into the power system," in *2017 International Conference on Electromechanical and Power Systems (SIELMEN)*. IEEE, 10 2017, pp. 295–301.

[5] J. Dong, T. Feng, H. Sun, H. Cai, R. Li, and Y. Yang, "Clean distributed generation in China: Policy options and international experience," *Renewable and Sustainable Energy Reviews*, vol. 57, pp. 753–764, 5 2016.

[6] S. Jain, S. Kalambe, G. Agnihotri, and A. Mishra, "Distributed generation deployment: State-of-the-art of distribution system planning in sustainable era," *Renewable and Sustainable Energy Reviews*, vol. 77, pp. 363–385, 9 2017.

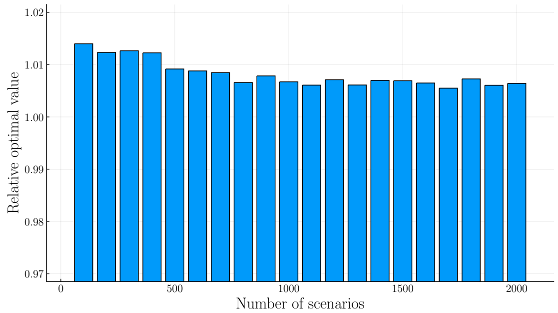

Figure 6. In-sample stability.

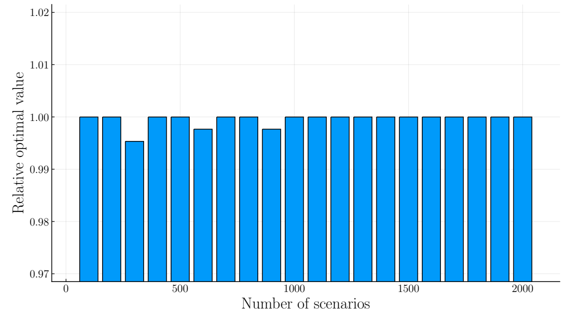

Figure 7. Out-of-sample stability.

[7] G. Notton, M. Nivet, C. Voyant, C. Paoli, C. Darras, F. Motte, and A. Fouilloy, "Intermittent and stochastic character of renewable energy sources: Consequences, cost of intermittence and benefit of forecasting," *Renewable and Sustainable Energy Reviews*, vol. 87, pp. 96–105, 5 2018.

[8] A. Zakaria, F. B. Ismail, M. H. Lipu, and M. Hannan, "Uncertainty models for stochastic optimization in renewable energy applications," *Renewable Energy*, vol. 145, pp. 1543–1571, 1 2020.

[9] A. Gonzalez-Castellanos, D. Pozo, and A. Bischi, "Stochastic Unit Commitment of a Distribution Network with Non-ideal Energy Storage," in *2019 International Conference on Smart Energy Systems and Technologies (SEST)*. IEEE, 9 2019, pp. 1–6.

[10] M. Jooshaki, H. Farzin, A. Abbaspour, M. Fotuhi-Firuzabad, and M. Lehtonen, "A Model for Stochastic Planning of Distribution Network and Autonomous DG Units," *IEEE Transactions on Industrial Informatics*, vol. 16, no. 6, pp. 3685–3696, 6 2020.

[11] J. M. Home-Ortiz, O. D. Melgar-Dominguez, M. Pourakbari-Kasmaei, and J. R. S. Mantovani, "A stochastic mixed-integer convex programming model for long-term distribution system expansion planning considering greenhouse gas emission mitigation," *International Journal of Electrical Power and Energy Systems*, vol. 108, pp. 86–95, 6 2019.

[12] T. D. De Lima, A. Tabares, N. B. Arias, and J. F. Franco, "A Stochastic Programming Model for the Planning of Distribution Systems Considering Renewable Distributed Generation and CO2 Emissions," in *2019 IEEE PES Conference on Innovative Smart Grid Technologies, ISGT Latin America 2019*. Institute of Electrical and Electronics Engineers Inc., 9 2019.

[13] J. López, D. Pozo, J. Contreras, and J. R. Mantovani, "A convex chance-constrained model for reactive power planning," *International Journal of Electrical Power and Energy Systems*, vol. 71, pp. 403–411, 10 2015.

[14] A. Narayan and K. Ponnambalam, "Risk-averse stochastic programming approach for microgrid planning under uncertainty," *Renewable Energy*, vol. 101, pp. 399–408, 2 2017.

[15] L. Lopez, I. Oliveros, A. Gonzalez-Castellanos, and D. Pozo, "Distributed generation allocation under uncertainty," in *2020 IEEE Power & Energy Society General Meeting (PESGM)*. IEEE, 2020, pp. 1–5.

[16] P. Seljom and A. Tomasgard, "Sample average approximation and stability tests applied to energy system design," *Energy Systems*, 10 2019.

[17] M. Baran and F. Wu, "Optimal sizing of capacitors placed on a radial distribution system," *IEEE Transactions on Power Delivery*, vol. 4, no. 1, pp. 735–743, 1989.

[18] M. E. Baran and F. F. Wu, "Optimal capacitor placement on radial distribution systems," *IEEE Transactions on Power Delivery*, vol. 4, no. 1, pp. 725–734, 1989.

[19] A. K. Jain, "Data clustering: 50 years beyond K-means," *Pattern Recognition Letters*, vol. 31, no. 8, pp. 651–666, 6 2010.

[20] Y. Atwa, E. El-Saadany, M. Salama, and R. Seethapathy, "Optimal Renewable Resources Mix for Distribution System Energy Loss Minimization," *IEEE Transactions on Power Systems*, vol. 25, no. 1, pp. 360–370, 2 2010.

[21] Y. Riffonneau, S. Bacha, F. Barruel, and S. Ploix, "Optimal Power Flow Management for Grid Connected PV Systems With Batteries," *IEEE Transactions on Sustainable Energy*, vol. 2, no. 3, pp. 309–320, 7 2011.

[22] M. Chis, M. Salama, and S. Jayaram, "Capacitor placement in distribution systems using heuristic search strategies," *IEE Proceedings - Generation, Transmission and Distribution*, vol. 144, no. 3, p. 225, 1997.