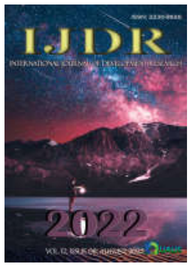
Available online at http://www.journalijdr.com

IJDR

ISSN: 2230-9926

International Journal of Development Research
Vol. 12, Issue, 08, pp. 58418-58425, August, 2022
https://doi.org/10.37118/ijdr.25010.08.2022



RESEARCH ARTICLE **OPEN ACCESS**

TOOL FOR AUTOMATED PLANNING OF MEDIUM VOLTAGE NETWORKS CONSIDERING THE EXPANSION OF PHOTOVOLTAIC DISTRIBUTED GENERATION

Daniel Perez Duarte[2], Ana Gabriela Bezerra Benitez[2], Sérgio Mishima dos Santos Barbosa[2], Reginaldo Prestes[1] and Cleverson Luiz da Silva Pinto[1]

[1]COPEL-DIS – Companhia Paranaense de Energia
[2] e-Amazonia

ARTICLE INFO

Article History:
Received 30th June, 2022
Received in revised form
26th July, 2022
Accepted 25th July, 2022
Published online 30th August, 2022

Key Words:

Network Planning, Distributed Generation, Machine Learning, Hosting Capacity.

Corresponding author: Daniel Perez Duarte

ABSTRACT

This paper deals with the general status of the rural Indian women and the Awareness of women Policies in rural areas and impact of women Policies on rural women. Women Plays vital role in our Indian Society. Rural women play a key role in supporting their households and communities in achieving food and nutrition security, generating income, and improving rural livelihoods and overall well-being. They contribute to agriculture and rural enterprises and fuel local and global economies. In concern Government has taken so many Programmes and Policies for women empowerment**.** The National Policy on Empowerment of Women adopted in 2001 states that All forms of violence against women, physical and mental, whether at domestic or societal levels, including those arising from customs, traditions or accepted practices shall be dealt with effectively with a view to eliminate its incidence. However, some positive intentions of the ground-level officials and the awareness of the beneficiaries can really bring in a revolution in the lives of Indian rural women. The Study is Practical and Theoretical oriented. The Study is based on Primary and secondary data sources. The necessary information about the Policies for Women Empowerment and its various components are collected from Books, Journals, Internet Source or related topic. The Researcher is going to study about Awareness of Women Policies in rural areas. The Research Work includes I. Introduction II. Methodology III. Women Specific Policies in India IV. Conclusion.etc

Citation: ***Daniel Perez Duarte, Ana Gabriela Bezerra Benitez, Sérgio Mishima dos Santos Barbosa, Reginaldo Prestes and Cleverson Luiz da Silva Pinto.*** "Tool for automated planning of medium voltage networks considering the expansion of photovoltaic distributed generation.*", International Journal of Development Research*, 12, (08), 58418-58425.

INTRODUCTION

This paper presents a tool to support expansion planning studies, aiming at automating the proposition and allocation of enhancement equipment from a list of predefined options. In addition to ensuring compliance with classical technical power supply criteria, the incorporation of new diagnostic parameters such as the networks' hosting capacity is proposed, defining a new paradigm for power grid planning. To reduce the computational burden associated with simulating all interventions in the portfolio, a prior clustering of networks by similarity, through machine learning techniques, was proposed. Thus, for each cluster, a representative network is elected, to which all reinforcement works are applied exhaustively, allowing the definition of a best-options ranking, through the combination of multicriteria methods, and the extrapolation of the results to the rest of each clusters' networks. In the clustering process, several attributes that characterize feeders are used, such as number of consumers, network length extension, maximum load, maximum voltage drop, load factor, SAIDI/SAIFI, among others. In an innovative way, one of the parameters used is the PVDG adopter penetration projection, based on a diffusion method inspired by the Bass model. In the developed solution, all input data (including user inputs) and results are stored in a database, which is accessible for parameters changes and tables exports, and allows reporting in Power BI software. The overall tool's flowchart is presented in Figure 1, which is detailed in the following sections, guided by screenshots and results of a case study applied to COPEL-DIS. This paper is organized as follows: Section 2 introduces the theoretical guidelines used in this work and describes the most relevant models used for the proposed technical solution. Section 3 presents the results of the innovative analytical tool to execute the presented methodology of networks evaluation. Finally, Section 4 reports this work's conclusions.

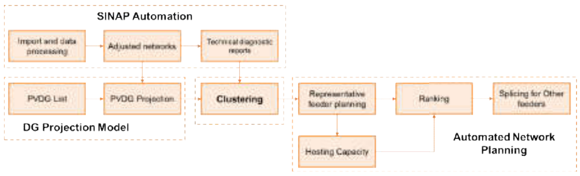


Figure 1. General tool flowchart

MATERIALS AND METHODS

Methodological guidelines: The planning of the expansion of electricity networks until the 1990s, both in Brazil and in other countries, was based on meeting established voltage criteria and ensuring demand capacity. The primary objective of network expansion, which was meeting demand and its growth in an economically viable manner, was added to technical criteria (on voltage transgression, thermal overload and power losses) that would ensure power supply quality and tariff moderation. In the 2000s, the use of automation elements began to be analyzed as a planning resource.

As pointed out by Duarte (2008), whenever there is a technical criterion transgression, automation elements should be proposed and analyzed as alternatives of structural reinforcement works, seeking the most economical solution that meets the planning criteria. In this sense, the automation proposed works, while meeting the expansion planning criteria, also fulfills voltage transgression and thermal overload technical criteria. In analysis of the international scenario, the new challenges faced by the distribution planning area are mainly linked to the increasing distributed generation (DG) penetration in the networks. Nyseg (2018) presents in his work a series of aspects that should be considered in the planning process, in addition to the traditional methodology already established, to address this expansion:

- Adoption of methods to forecast the growth of distributed generation, such as those that are traditionally performed for loads projection.
- Conducting hosting capacity analyses to assess the network's ability to support distributed generation expansion, including the evaluation of works to expand this capacity, according to the need.
- Evaluation of Non-Wires Solutions (NWS), including possible planning solutions that allow power utilities to postpone or avoid investment in conventional infrastructure.

In the methodology developed in this work, it is proposed the use of a PVDG diffusion projection model and hosting capacity analyses as inputs for the characterization of networks and evaluation of new reinforcement works. Another element also explored by the literature in the context of alternatives to planning is the use of clustering techniques, allowing the evaluation of the network from representative feeders, whose results are later extrapolated to the entire system. In this process, works such as Duarte (2014) developed methodologies for determining the representative individual of each group using physical, operational, and intelligent electrical network characteristics, for further analysis of scenarios and cost-benefit assessment. In the following subsections, the keys concepts which were incorporated into the work are presented: diffusion models applied to PVDG, clustering, and hosting capacity.

DG diffusion models: As in the case of the introduction of a any new technology, which requires changing consumer behavior patterns for its introduction, it is possible to study the PVDG from the perspective of innovation theories. On Brazil's national scene, Konzen (2014) was the first to project the incorporation of PVDG into the residential consumer class according to socioeconomic characteristics of the population, based on Rogers' Innovation Diffusion Theory (Rogers, 2003), mathematically represented by the Bass model (Bass, 1969).

According to Rogers' Innovation Diffusion Theory (Rogers, 2003), any population can be divided into different segments that stipulates the individual's propensity to adopt a new technology. On a cumulative basis over time, the adoption curve can be described as a sigmoid distribution, known as the S-curve, presented in Figure 2.

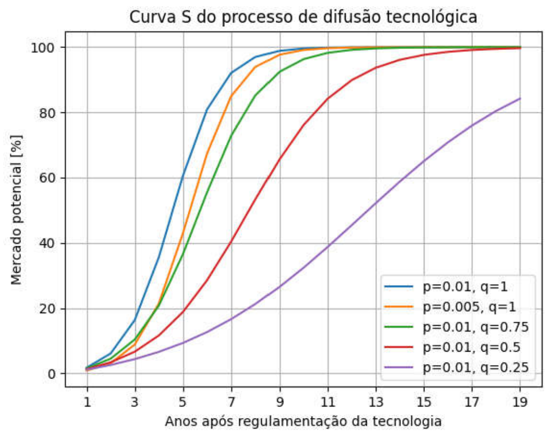


Figure 2: Technological diffusion Curve S for different p and q values

The mathematical representation of the S curve is given by the accumulated distribution function of the Bass model ($F(t)$), being guided by parameters that seek to capture the trends of innovation (p) and imitation (q) of the analyzed population:

$$F(t) = \frac{1 - e^{-(p+q)t}}{1 + \frac{q}{p}e^{-(p+q)t}}$$

Typically, the Bass curve represents the percentage of the potential market that has adopted a particular technology after$t$ years of its regulation. It is important to highlight that the Bass model estimates the technology's diffusion rate regardless of the market potential (m). The accumulated number of adopters ($N(t)$)can be determined by:

$$N(t) = m \times F(t)$$

One of the possible improvements to Bass' model is the consideration of a potential market that varies over the years ($m(t)$), as a function of product characteristics, such as the payback period (Beck, 2009; Denholm et al., 2009; Konzen, 2014; Sigrin et al., 2016). From Figure 2, one can also identify the impacts of p and q parameters on the diffusion speed – i.e., the slope of the S-curve. In the literature, there are different ways of estimating these parameters, whether they are based on historical data, analogies with similar products, or via opinion surveys of consumer intent (Lilien et al., 2007). In this work, due to the availability of adopter's historical data since 2013, the parameters p and q were stipulated based on the following objective function (F_{obj}):

$$F_{obj} = \sum_{i=1}^{n} (N(t) - A_{acum}(t))^2$$
$$= \sum_{i=1}^{n} \left(m(t) \times \frac{1 - e^{-(p+q)t}}{1 + \frac{q}{p}e^{-(p+q)t}} - A_{acum}(t) \right)^2$$

Where(A_{acum}) represents the actual historical accumulated PVDG adopters. In addition, for the execution of Bass's parameters optimizing process based on historical data, as well as to project the number of new adopters in future scenario, it is necessary to establish criteria to determine the evolution of the potential market over time. In adaptation of the 4MD methodology (EPE, 2021), the historical potential market (2013-2020) is extrapolated for the projection years (2021-2030) as it follows:

$$m(t) = UC(t) \times CSE \times FAL \times FMM(t)$$
$$m(t) = UC(8) \times (1 + ECC)^t \times CSE \times FAL \times FMM(t)$$

- UC: is the number of consumers collected from Consumption and Distribution Revenue Reports (ANEEL, 2021a) in December of each year. For the latter equation, $UC(8)$ represents the number of consumers in 2020, the last year of historical data considered. In addition, consumers who are under special tariff regime, such as low-income residential consumers, were excluded to homogenize the individuals of each group.
- CSE: is a socioeconomic criterion for market limitation, if consumers with incomes below 3 minimum wages are not able to bear the costs of a photovoltaic system (EPE, 2020).
- FAL: it is a local aptitude factor depending on the condition of occupation and the type of property (homeownership) and restrictions on roof use. While the condition of property occupancy was aggregated at regional state level based on 2010 Census data, roof restrictions are characterized by a single value (85%) (Konzen, 2014).
- ECC: is the number of consumers' growth expectation. It is assumed as an estimate for this parameter the variation in the number of consumers in the last 5 years (2015-2020).
- FMM: is the maximum fraction of the market, that is, the percentage of the market able technically and economically and willing to make the investment according to its economic attractiveness.

It is noteworthy that the values of CSEandFALraised for the residential consumer class were extrapolated to the other low voltage groups (rural, commercial, and industrial), in the absence of specific reference for its determination.
For medium voltage consumers, no limitations related to socioeconomic criterion and local aptitude factor were considered – therefore,$CSE = FAL = 1$ Based on empirical formulations by Kastovich (1982) and Navigant Consulting (2007), Beck's formulation (2009) was adopted to determine the FMMdependent investments' payback period (TPB):

$$FMM(t) = e^{-0,3 \times TPB(t)}$$

It is reinforced that the result of the payback calculation varies year by year, due to technology price drop, and the current tariff rule. A distinction is also drawn between the methodology used for the residential market (simple payback) and for the other groups (discounted payback). Regarding the price of photovoltaic systems, the values collected by the Greener Group (Greener, 2019, 2020) were incorporated. The PVDG installed capacity is based on COPEL-DIS's list of PVDG adopters and the current MMGD Electric Power Compensation System (ANEEL, 2012, 2015) was applied with tariffs from the PCAT tables (ANEEL, 2021b) –2020 tariff values were kept constant for the projection years. In addition, photovoltaic panel productivity (measured in kWh/kWp/year) was extracted from Simone (2019) as a statewide value, considering a performance rate of 80%. Payback parameters such as average photovoltaic panel degradation and maintenance cost, exchange inverter's costs, discount rate, low voltage consumers' availability cost and photovoltaic panels price reduction were directly extracted from the 4MD methodology (EPE, 2021). In a simplified way, the Bass model developed by EPE is based on the analysis of technology historical adoption rate for the construction of a base curve of diffusion for future years, considering technical and economic limitations of the population, as well as the attractiveness of the investment, marked by the value of the tariff, the credit compensation system and the photovoltaic panels price evolution. A result of a temporal projection at the level of the concession areas of the Brazilian power utilitiesis obtained.

Although the broad temporal projection already presents a relevant panorama for power utilities, studies of spatial diffusion in higher granularity are necessary to analyze the impact of PVDG on the operation and planning of expansion in local systems. While the application of time diffusion metrics of new technologies presents itself as a traditional methodology (Bass, 1969), the domain of the study of spatial diffusion models is still a little explored field (Dharshing, 2017). In the international literature, there are studies that are characterized using extremely detailed databases, often georeferenced, associated with a history of adoption of at least 10 years. Despite different methodologies – such as the application of econometric methods (Balta-Ozkan et al., 2015), the use of cellular automata (Zhao et al., 2017), or the estimation of diffusion equation via spatial auto regression (Dharshing, 2017) – the models converge to highlight the impact of knowledge transfer between users and non-users in the diffusion process, concept defined as spillover effect (Balta-Ozkan et al., 2015; Dharshing, 2017; Ferreira, 2015; Zhao et al., 2017). Another aspect observed is the objective of presenting results by neighborhoods, districts, states, and geographic areas, and not at the level of networks or substations, as intended and developed in the present work. To ensure adherence to the project methodology, a model capable of distributing the time-projected PVDG in granularity of COPEL-DIS feeders was developed, exploring the spillover effect and spatial distribution characteristics already captured by the previously established temporal diffusion. The new adopters are allocated in proportion to the penetration rate (PR) of each feeder$i$ (in terms of adopters (A) per consumer (UC)):

$$PR_i = \frac{A_i}{UC_i}$$

For the diffusion process itself, a weight proportional to the previously defined penetration rate is assigned for each feeder (W_i). The sum of the weights of the feeders results in the unit value, according to the following equation:

$$W_i = \frac{PR_i}{\sum PR}$$
$$\sum W_i = \sum \frac{PR_i}{\sum PR_i} = 1$$

The number of adopters in the year (t) for each feeder (A_i), is obtained by:

$$A_i(t) = round(W_i \times A_{copel}(t); 0)$$

Where $A_{copel}(t)$ represents the total number of new adopters expected for the entire COPEL-DIS concession area in the year$t$ (Bass model). It is noted that due to the application of the round function, the results obtained for each feeder are always integers, ensuring consistency between the stipulated parameter and the physical variable. This measure, however, does not ensure that the sum of the adopters of each feeder results in the total number of adopts stipulated by the Bass model. To cope with this problem, the missing adopters are distributed in the feeders with the lowest number of adopters in an iterative process, and in case of a tie, the highest penetration rate is privileged.

This method ensures, even in a marginal and complementary way, that adopters are inserted in zero penetration feeders, introducing the innovation variable characteristic of diffusion processes. In the proposed methodology, it is observed that $A_i(t)$ is as higher as the feeder's penetration rate in the last year of historical data. In other words, feeders that have higher PVDG penetration tend to concentrate the new adopters, which is consistent with the imitation parameters obtained for the Bass model and the concept of previously defined spillover effect. It is also noteworthy that, by the adopted model, feeders with low number of consumer and significant penetration rates tend to have higher growths, limited to the factor $CSE \times FAL$ previously defined.

Clustering: Clustering analysis is part of a set of multivariate techniques, whose objective is to group individuals considering their attributes characteristics. Clustering is an unsupervised classification method, in which a cluster is understood as a group of individuals with similar attributes and dissimilar to individuals from another group, as illustrated in Figure 3. There are several strategies in the literature to cluster individuals, and one of more explored in the k-means.

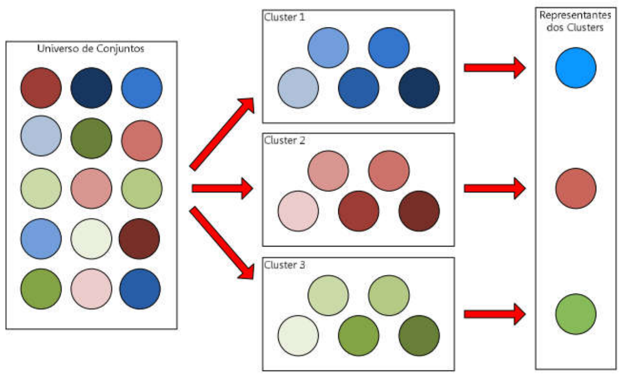


Figure 3: Illustration of the clustering concept

The purpose of this technique is to partition the groups so that their elements are as close as possible to the cluster average value(minimizing intra-cluster distance) and away from the average of the other clusters (maximizing inter-cluster distance).

As the dimensionality of the problem increases, that is, the number of variables evaluated becomes larger, clustering algorithms that use distance formulas for elements classification (such as k-means) tend to struggle with obtaining significant clusters, since the points tend to have very similar distances from each other. In this case, preprocessing techniques such as disregarding highly correlated variables (identified via Pearson index), which add redundancies to the dataset (Signoretti, 2019), can be applied, as well as dimensionality reduction algorithms such as the Principal Component Analysis (PCA) or the T-Distributed Stochastic Neighbor Embedding (t-SNE) (Mysiak, 2020). Additionally, to guarantee clusters' homogeneity, itis possible to use statistical processes to determine the optimal number of clusters for a given data set. One of the most used methods is the Sum of Squared Errors (SSE), which seeks to measure the mean sum of intra-cluster quadratic errors for different amounts of clusters. The SSE value decreases with increasing the number of clusters and is minimal for clusters composed of a single element. The existence of an optimal point (known as elbow) from which the SSE decreases with less speed with the increase in the number of clusters is identified from a graphical analysis.

Hosting Capacity: The concept of hosting capacity (HC) can be defined as the penetration rate of distributed energy resources that can be introduced to a network without violating the normal operating conditions of the system or causing damage to powerquality (Bollen and Häger, 2005; Ismael et al., 2019). It is noteworthy that the calculation of HC is not a fixed procedure, with a single result. In effect, HC is calculated for different performance indexes such as voltage and frequency variations, network thermal overload limits, technical power losses, power quality and protection systems malfunctioning. In a survey conducted with more than 100 power utilities from 23 countries, Accenture Consulting identified system operators' concern about distributed generation development and its impacts on distribution networks. In fact, about half of the respondents indicated the perception that the hosting capacity of their networks should be reached as early as the next decade (Accenture, 2017). Although many of the utilities interviewed have already incorporated measures to simplify distributed resource connection processes, few have a forecasting approach to hosting capacity needs in planning.

According to a survey conducted by Palmintier et al. (2016), many distribution system operators often use the 15% rule defined by EPRI, which considers that penetration levels greater than 15% of peak load should be avoided and investigated by further studies. The correct network HC evaluation may allow the adoption of improvement measures to increase the absorption capacity of DG. Regarding the calculation of the HC, uncertainties related to location, installed capacity and variable generation of future distributed resources are required. In this process, analytical, stochastic, or simplified methods can be applied, depending on the complexity of the system studied and the level of detail required (Ismael et al., 2019). In general, the calculation models are distinguished by the precision and time of computational processing, ranging from the bar-to-bar study of the system (analytical) to the estimation of multiple generation scenarios (stochastic) and to the establishment of correlations of electrical and consumer characteristics – from power flow and short circuit analyses – between the target network and previously evaluated feeders (simplified). The latter model, due to its characteristic, allows estimating the hosting capacity of a system with very reduced processing time, being explored by commercial planning tools such as EPRI DRIVE (Rylander et al., 2015).

Tool presentation: The tool's operation flow can be divided into 4 large blocks that must be executed in sequence, that is: SINAP Automation; DG Projection Model; Characterization and Evaluation of Medium Voltage Networks; and Automated Network Planning. Figure 4presentsa screenshot of the tool's home page, which shows the buttons dedicated to the configuration and execution of each of the mentioned blocks. In the upper-right corner, one can access the Power BI reports that synthesize the tool results (pie-chart button).

Figure 4. Tool's home screen

SINAP Automation: In SINAP Automation, the import, treatment, and generation of technical diagnostics of the utility's medium voltage networks (COPEL-DIS, as in the case study) is performed, using tool calls to the SINAP grid platform, through parameters configuration by the user (
Figure 5). With the use of SINAP grid software, after networks importing process from the company's database, the feeders' output demand values are adjusted, the annual growth rates are incorporated in the horizon of the next decade and the reliability indicators are calibrated, based on the occurrences history. After the import and processing steps, the technical reports at feeder level (SDMT) and substations (SED) are exported to the database, allowing their access through the following operation blocks.

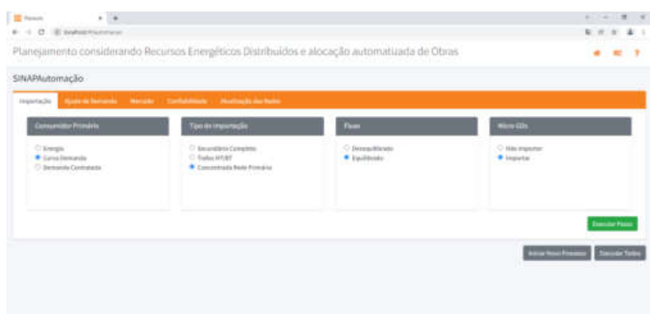


Figure 5. Network import parameter configuration screen

In further details, the feeders' output demand values are adjusted according to treated curves to characterize the situation of maximum demand (90% percentile curve hourly levels, considering 2021 data). A bottom-up adjustment is performed, i.e., the energy of the loads is adjusted proportionally until the result of the power flow is as close to the measurement value. Furthermore, the reliability indicators are adjusted, from the company's occurrences file import combined to the values of SAIDI and SAIFI calculated in 2021.

DG Projection Model: For the DG Projection Model, the stages of temporal diffusion and spatial allocation of projected DGs are distinguished. For the first, an adaptation of the EPE's 4MD methodology, based on the Bass model (EPE, 2021), to obtain amounts of new adopters and installed capacity of PVDG year by year for the next decade, with results segmented into consumer classes, in five distinct scenarios of evolution of DGregulation is performed. In terms of spatial diffusion, a model based on the spillover effect was proposed, and the new adopters were allocated proportionally to the PVDG penetration rate in the initial year of the study. The user can set parameters such as the rate of taxes on the energy tariff, the photovoltaic systems prices by power range, as well as consumer market limitation factors, maintenance, and equipment exchange costs (
Figure 6). At the user's command, the DG Projection Model block calls a Python script that runs the model by updating database tables.

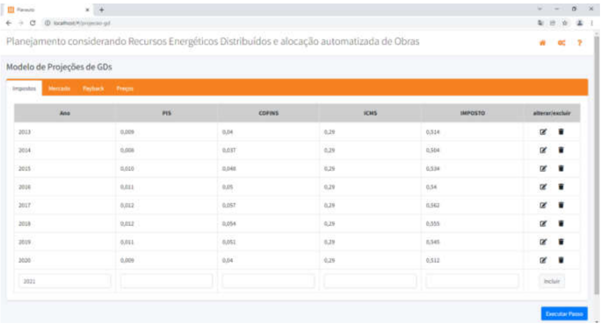


Figure 6. Price tax setup screen for PVDG projection

Characterization and Evaluation of Medium Voltage Networks: In the stage of Characterization and Evaluation of Medium Voltage Networks, a feeders clustering is performed by the k-means method implemented in Python, using user-defined classification attributes. The user should also select whether the networks will be clustered by voltage level and/or regional area and what the purge criteria are applied to the evaluated feeders (
Figure 7). Among the available classification attributes are parameters of technical diagnoses and PVDG projection. As a result of the clustering step, the cluster that each feeder belongs to and the representative network of each group (the one closest to the average cluster value) is identified.

Figure 7. Cluster parameters configuration screen

Automated Network Planning: After the clustering process, a file that aggregates the record of representative feeders and a set of standard features used to simulate reinforcement works is sent to SINAPgrid software for the allocation of the network possible interventions. With the definition of specific rules for the creation of each of the interventions in the portfolio (allocation of voltage regulators, rewiring, feeder reconfiguration, circuit reclosers, and self-healing strategies) a software module that automatically creates those enhancement works was developed. With that allocation, it is possible to determine for each of the resulting networks the hosting capacity in a calculation integrated with SINAPgrid. The amount of HC and the result of new technical diagnostics of the feeders are sent to the database allowing the execution of a prioritization algorithm, implemented in Python, for the creation of a works ranking in terms of the technical benefit brought by each improvement alternative.

The prioritization algorithm is based on the combination of multicriteria methods, using 4 weights elicitation techniques - equal weights, rank order centroid (ROC), entropy method (Mussoi, 2013) and ideal point (Ma et al., 1999) - and 2 aggregation functions - SMART (Mussoi, 2013) and TOPSIS (Mussoi, 2013). The results of the 8 rankings created are aggregated by an un weighted additive function generating a final ranking for each representative network, which is extrapolated to all feeders of the groupings that require enhancement works. In the tool, the user can select specific networks and configure whether they should receive all the possible alternatives simulated or only the best work indicated by the ranking, as well as adjust the resources adopted in each of the allocations (Figure 8). Once the user selection is finished, a new aggregation file is created, which is sent to SINAPgrid to perform the automatic enhancement works allocation on the networks, concludinga tool execution round and enabling the visualization of reports in Power BI.

Figure 8: Configuration screen of the resources adopted in the voltage regulator allocation work

RESULTS

Networks evaluation: As a case study, the results obtained by applying the methodology developed based on COPEL-DIS are presented in the context of the PD-2866-0527/2020 project. In the Power BI report, analyses of technical diagnostics at substations and feeders, results of the PVDG projection model and clustering study are compiled (Figure 9).

Figure 9. Tool results report start screen

At the level of substations (SED), dashboards were developed to the evaluation of installed power compared to the maximum power demand verified, with identification of the number of critical transformers; diagnosis of power losses; maximum loading and criticality in terms of number of feeders with load or voltage drop transgressions (Figure 10). At the feeders (SDMT) level, the analysis includes the evaluation of the maximum voltage drop as a function of the length of the network and the number of large distributed generators, as well as the evaluation of the distribution of reliability indicators within the COPEL-DIS concession area (Figure 11).

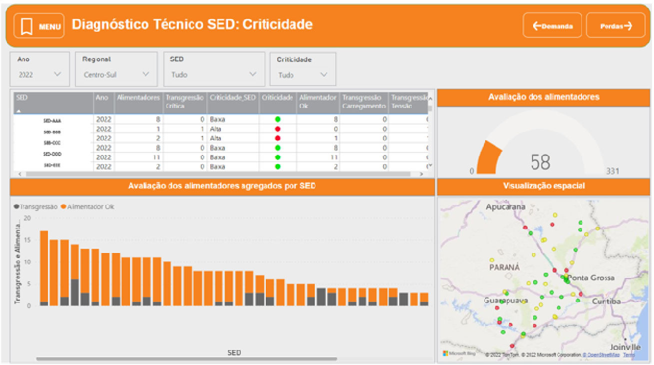


Figure 10: Technical criticality diagnostic screen at sub-stations level

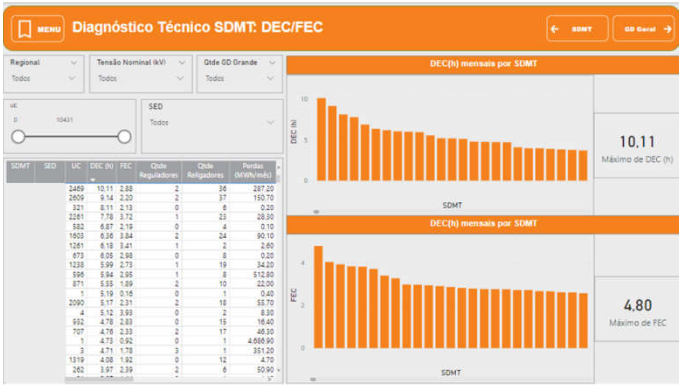


Figure 11: Technical diagnostic screen of feeder reliability indicators

PVDG projection: In the PVDG diffusion projection model, the results can be analyzed by consumer classes for each of the evaluated regulatory scenarios. The main scenarios highlight the *Superior* scenario, referring to the compensation rules of REN 482/2012(ANEEL, 2012, 2015), and the *Fio B Gradual* scenario, based on PL 5829/2019 (Camara, 2019). The comparison of scenarios can be seen in Figure 12

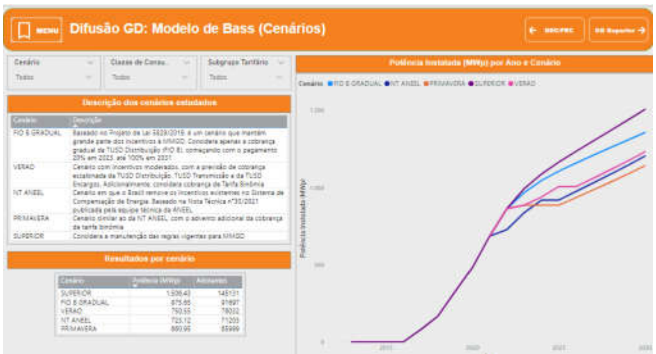


Figure 12: PVDG projection model screen in COPEL-DIS

Prioritization Plan: Based on the results of technical diagnostics at feeder and substations level and the PVDG projection model, it is possible to perform the network characterization block, according to the user's settings. The consistency evaluation of the clustering output in relation to the selected parameters is typically performed by reducing the dimensionality of the attributes to the 2D plane, by applying algorithms such as the T-Distributed Stochastic Neighbor Embedding (t-SNE). Thus, it is possible to visualize whether the groupings formed are coherent to the distribution of feeders, as it is the case in Figure 13. Executing the algorithm of works prioritization applied to representative networks, the extrapolation of the results to the other feeders is performed, which can be checked via Power BI report (Figure 14). If a work does not bring technical benefit in terms of improving reliability indicators, loading levels, voltage drop, technical power losses and hosting capacity, it is excluded from the list of suggestions, implying in fewer viable alternatives. Finally, after the correct execution of all tool blocks, the automated allocation of works in medium voltage networks is suggested when the SINAPgrid planning module is opened, as indicated in the screenshot shown in Figure 15.

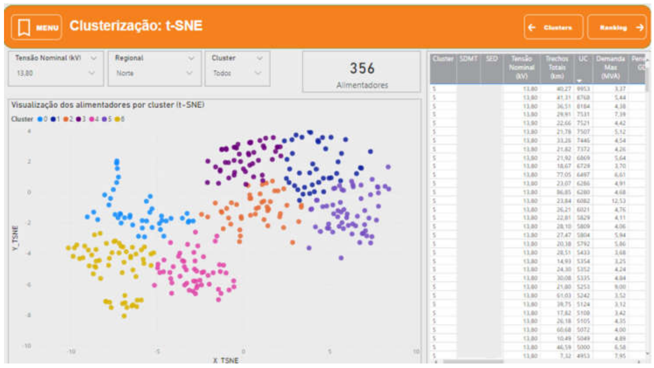


Figure 13. Network clustering result screen

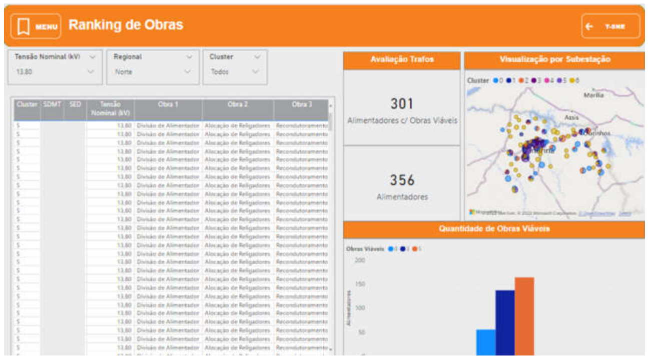


Figure 14. Display screen of the work suggestions for each feeder

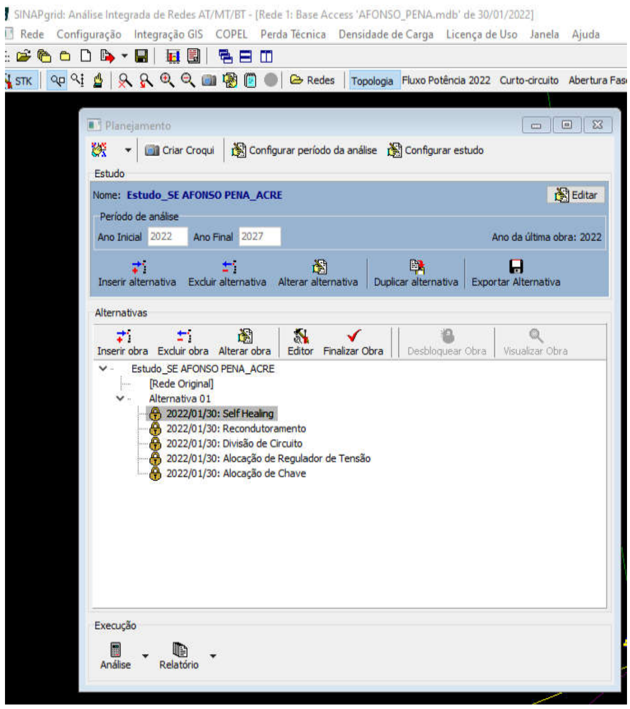


Figure 15: SINAP grid planning module with the suggestion of predefined planning enhancement works

DISCUSSION

In response to the need of power utilities to optimize the planning process, this work aimed to present a network planning tool, capable of automating the proposition and allocation of equipment from a list of predefined options, with parameters configured by the planner. Thus, a more efficient use of information of the resources is allowed – which can be directed to the analysis of specific cases that escape to the algorithms – with the potential to achieve significant technical-economic return. Methodologically, the project innovates by incorporating a PVDG growth projection model as a classification attribute in a in medium voltage networks clustering stage, as well as by conducting hosting capacity studies in selected representative networks, applying this parameter as an input to the multicriteria prioritization, thus creating a new paradigm for the distribution expansion planning. It is also noteworthy that although the tool adds a set of advanced machine learning techniques, innovation diffusion models, task automation and prioritization methods, the user interface remains intuitive and user friendly, facilitating the adoption of the solution by planners. Finally, the presentation of the results obtained in a case study applied to the COPEL-DIS base allows the validation of each of the stages developed, ratifying the potential of the proposed solution, that may be replicated to other distributors.

Acknowledgment

The authors acknowledge the financial support of Copel – Companhia Paranaense de Energia through the ANEEL R&D Program for the development of the research project entitled: "PD-02866-0527/2020 Methodologies and tooling for intelligent automated planning of distribution networks considering distributed energy resources".

REFERENCES

Accenture 2017. *Power Surge Ahead: How Distribution Utilities Can Get Smart with Distributed Generation*. Accenture consulting.

ANEEL 2012. National Electric Energy Agency. *Normative Resolution No. 482/2012*. ANEEL, Brasilia, Brazil, p. 9.

ANEEL 2015. National Electric Energy Agency. *Normative Resolution No. 687/2015*. ANEEL, Brasilia, Brazil, p. 25.

ANEEL 2021a. National Electric Energy Agency. Consumption and Distribution Revenue Reports. ANEEL, Brasilia, Brazil.

ANEEL 2021b. National Electric Energy Agency. *Calendar and Results of Distribution Tariff Processes*. ANEEL, Brasilia, Brazil

Balta-OzkanN, Yildirim J, Connor P 2015. *Regional distribution of photovoltaic deployment in the UK and its determinants: A spatial econometric approach*. Energy Economics, v. 51, p. 417-429, 2015. 10.1016/j.eneco.2015.08.003

Bass FM 1969. *A new product growth for model consumer durables*. Management Science, v. 15, n. 5, Theory Series, p. 215-227, 1969.

Beck R 2009. *Distributed Renewable Energy Operating Impacts and Valuation Study*. Prepared for Arizona Public Service by R.W. Beck, Inc. 2009.

Bollen M, HÄGER M 2005. *Power quality: interactions between distributed energy resources, the grid, and other customers*. Leonardo Energy, 2005.

Camara S 2019. *Projeto de Lei 5829/2019*. Brasília, Brazil

Denholm P, Drury E, Margolis R 2009. *The Solar Deployment System SolarDS Model: Documentation and Sample Results*. TechnicalReport NREL/TP-6A2-45832, 2009. Availableat: <www.nrel.gov/docs/fy10osti/45832.pdf >.

Dharshing S 2017. *Household dynamics of technology adoption: A spatial econometric analysis of residential solar photovoltaic PV systems in Germany*. Energy research & social science, v. 23, p. 113-124, 2017. 10.1016/j.erss.2016.10.012.

Duarte DP 2008. *Automação como recurso de planejamento de redes de distribuição de energia elétrica*. MSc Thesis, Department of Electrical Energy Engineering and Automation, Polytechnic School of the University of São Paulo, São Paulo,Brazil,doi:10.11606/D.3.2008.tde-11082008-134751. Availableat:<www.teses.usp.br>.

Duarte DP 2014. Modelo de avaliação de políticas de implementação de funcionalidades de redes elétricas inteligentes em sistemas com alto grau de heterogeneidade. PhD thesis, Department of Electrical Energy Engineering and Automation, Polytechnic School of the University of São Paulo, São Paulo, Brazil,doi:10.11606/T.3.2013.tde-18092014-163603. Available in: <www.teses.usp.br>.

EPE 2020. Energy Research Company. *Technical Note EPE DEA-SEE 010/2020. Micro Market Model and Distributed Minigeneration 4MD*: Methodology - PDE Version 2030. Brasilia, p. 28, 2020.

EPE 2021. Energy Research Company. *Technical Note EPE DEA-SEE 009/2021. Micro Market Model and Distributed Minigeneration 4MD: Methodology- PDE Version 2031*. Brasilia, p. 27, 2021.

Ferreira ZR 2015. *Determinantes da irrigação no Brasil: uma análise do Spillover espacial*. MScThesis, Universidade Federal de Viçosa, Viçosa, Brazil.

Greener 2019. *Strategic Study - Distributed Generation Photovoltaic Market: 1st Semester - 2019 - Brazil*. Greener, Sao Paulo, Brazil, 2019.

Greener 2020. *Strategic Study - Distributed Generation: 1st Semester - 2020 - Brazil*. Greener, São Paulo, Brazil, 2020.

Ismael SM, Abdel Aleem SHE, Abdelaziz AY, Zobaa AF 2019. State-of-the-art of hosting capacity in modern power systems with distributed generation. Renewable Energy, vol. 130, p. 1002–1020, Jan. 2019, doi: 10.1016/j.renene.2018.07.008.

Kastovich JC, Lawrence RR, Hoffman RR, Pavlak C 1982. *Advanced electric heat pump market and business analysis*. ORNL/Sub/79-2471/1. Prepared under subcontract for Oak Ridge National Laboratory by Westinghouse Electric Corp. Oak Ridge, TN: Oak Ridge National Laboratory.

Konzen G 2014.*Difusão de sistemas fotovoltaicos residenciais conectados à rede no Brasil: uma simulação via modelo de* Bass. MSc Thesis, Institute of Energy and Environment, University of São Paulo, São Paulo, 2019.

Lilien GL, Rangaswamy A, Bruyn AD 2007. *The Bass model: marketing engineering technical note*. Trafford Publishing, 2007.

Ma J, Fan ZP, Huang, LH 1999. *A subjective and objective integrated approach to determine attribute weights*. European journal of operational research, v. 112, n. 2, p. 397-404, 1999.

Mussoi FLR 2013. *Modelo de decisão integrado para a priorização multiestágio de projetos de distribuição considerando a qualidade da energia elétrica*. PhD Thesis, Federal University of Santa Catarina, Florianópolis, Brazil, 2013.

Mysiak K 2020. *Explaining K-Means Clustering*. Towards Data Science, July 13. 2020. Available at: <https://towardsdatascience.com/explaining-k-means-clustering-5298dc47bad6>. Access: 21July 2021.

Navigant Consulting 2007. *Arizona Solar Electric Roadmap Study – Full Report*. Prepared for: Arizona Department of Commerce. Navigant Consulting, Burlington, MA, USA, 2007.

Nyseg R 2018. *Distributed System Implementation Plan*. Distributed System Design, p. 185.

Palmintier B, Broderick R, Mather B, Coddington M, Baker K, Ding F, Reno M, Lave M, Bharatkumar A 2016. *On the path to sunshot: emerging issues and challenges in integrating solar with the distribution system*. National Renewable Energy Lab. NREL, Golden, CO United States, 2016.

Rogers E 2003. *The Diffusion of Innovations*. The Free Press, New York, USA, 5th edition, 2003.

Rylander M, Smith J, Sunderman W 2015. *Streamlined method for determining distribution system hosting capacity*. In: 2015 IEEE Rural Electric Power Conference. IEEE, 2015a. p. 3-9.

Signoretti G 2019. *Digital Tribes: customer clustering with K-Means*. Towards Data Science, 03 Dec. 2019. Available at: <https://towardsdatascience.com/digital-tribes-customer-clustering-with-k-means-6db8580a7a60>. Access: 21July 2022.

Sigrin B, Gleason M, Preus R, Baring-Gould I, Margolis R 2016. *Distributed generation market demand model dGen:*

Documentation. National Renewable Energy Lab. NREL, Golden, CO United States, 2016.

Simone LFC 2019. Inserção da micro e minigeração distribuída solar fotovoltaica: impactos na receita das distribuidoras e nas tarifas dos consumidores. MSc Thesis, Department of Electrical Energy Engineering and Automation, Polytechnic School of the University of São Paulo, São Paulo, Brazil, 2019.

Zhao T, Zhou Z.; Zhang Y, Ling P, Tian Y 2017. Spatio-Temporal Analysis and Forecasting of Distributed PV Systems Diffusion: A Case Study of Shanghai Using a Data-Driven Approach. IEEE Access, v. 5, p. 5135-5148, 2017. 10.1109/ ACCESS. 2017.2694009.

****