Sensitivity analysis of the Power Demand Uncertainties on the electrical power system optimization models

Sara Fakih[a], Mohamed Tahar Mabrouk[b], Mireille Batton-Hubert[c] and Bruno Lacarrière[d]

[a] IMT Atlantique, Nantes, France, s.fakih995@gmail.com, CA
[b] IMT Atlantique, Nantes, France, Mohamed-tahar.mabrouk@imt-atlantique.fr,
[c] Ecole des Mines de Saint-Etienne, Saint-Etienne, France, batton@emse.fr
[d] IMT Atlantique, Nantes, France, Bruno.lacarriere@imt-atlantique.fr

Abstract:
Accurate estimations of future energy consumption are crucial for decision-makers to better plan the future design and operation of production and distribution systems. The existence of uncertainties in the inputs of the planning process can affect the quality of the model's outcomes and potentially can lead to sub-optimal solutions. This work proposes an investigation of input uncertainties of an Electrical Power System Model (EPSM) based on dynamic linear optimal power flow. The considered model inputs are the electricity demand curves of the buses of the grid. A set of forecasted demand time series is used to generate a probability model of their variability at each time step of the period studied. This probabilistic model is applied to generate the uncertainty of the demand curves and the associated input macro indicators (IMI) which are used to define the experimental design of the sensitivity analysis. The results show that input uncertainties on the demand have significant effects on the results in terms of Levelized Cost of Energy (LCOE) and system design like the installed capacity of Wind Turbines (WT), extracted energy from Classic Generators (CGs), and Battery Energy Storage (BES) location and sizing. In addition, it is shown that the input demand uncertainties can affect the results on the distribution performance parameters like the level of saturation of the different grid branches.
Keywords:
Battery energy storage; Demand Uncertainties; Electrical Grid; Renewable energy sources; Sensitivity analysis.

1. Introduction

Load forecasting models are used to predict future demand behaviors to be used as input for optimal planning models. They rely on long-term expectations for the parameters influencing the load profile (e.g., temperatures, occupancy, behavior, etc.). Therefore, they are likely to give estimations of low accuracy [1]. This could affect the accuracy is the model's output. Thus, there is a need for robust and rigorous techniques that can provide quantifiable information about the impact of uncertainty on model outputs. Sensitivity Analysis (SA) technics incorporate a set of methods that aim to identify the most important input parameters driving the model output variability in addition to the non-influential parameters whose uncertainty can be safely ignored.

In the literature, sensitivity analysis in energy models considers different uncertain input parameters e.g., the intermittent renewable generation, final energy demand, primary energy prices, economic growth, etc. [2]. These uncertain parameters are generally associated with randomness and temporal variability [3]. However, uncertainty in demand estimation introduces additional complexity alongside energy planning scenarios, including Renewable Energy Sources (RES) and Battery Energy Storage (BES) sizing and placement. Under such conditions, it is crucial to understand how demand curves uncertainty affects the optimal design and to identify the most influential uncertain parameters [4].

For the SA of demand curves, two different approaches exist. The first one is the total demand uncertainty applied to the demand patterns directly measured or extracted from the load forecasting models. To assign uncertainty on energy demand profiles, the most commonly used approach is applying Probability Density Function PDF to each time step of the demand time series, like normal distribution [5], uniform distribution [6], etc. The second one is the model-based uncertainty in which the uncertainty is assigned to the input parameters of the load forecasting tool and is propagated to obtain a series of demand patterns. The model-based uncertainty approach is applied in the literature for different types of demand patterns. In Mavromatidis et al. [7], using a building performance simulation tool, the uncertainties are attributed to the building material properties, occupancy patterns, hot water services, ventilation, and climate parameters, etc. Probability distributions are attributed to these parameters (normal, triangular, etc.), and the distribution parameters are then estimated. Thousands of profiles are then extracted and represented in the form of probability distribution to sample them to generate random energy demand profiles for the following study steps. The same approach is applied for the load forecasting model called MOSAIC in [8] developed by the principal French DSO (Enedis [9]), where the input variables are classified into four types: quantitative continuous (local height, coefficient of

https://doi.org/10.52202/069564-0308

performance for heater, temperatures, etc.), qualitative binary (e.g., presence of air conditioner or not), quantitative discrete (e.g., number of occupants) and qualitative nominal (e.g., thermal performance level). The variability range of the parameters are defined, and probability density distributions are attributed to continuous variables. Multiple demand patterns are then extracted from MOSAIC.

Various methods of sensitivity analysis on energy models are performed in the literature while considering the energy demand uncertainty; a local sensitivity analysis is used to study the demand effect on life cycle cost and loss of power supply in Sadeghi et al.[10]. A two-stage elementary effect - variance based technique is developed by Mavromatidis et al. [4] to study the effect of energy demand on the total cost of the system. Many other sensitivity analysis models are developed in [11], [12], [13] and [14], etc. Basically, SA models are classified into three major groups : *1- screening models* that coarse sorting of the most influential inputs from a large number (e.g. Morris method [15]), *2-measure of importance* or quantitative sensitivity indices (e.g. linear regression [16]) and *3- deep exploration of model behavior* that measure the effects of inputs over their total range of variation (e.g., Metamodels-based SA [17]). The selection of the most appropriate method is based on the number of input parameters and the model's complexity.

This paper uses the total demand uncertainty characterization for the extracted curves using the probabilistic techniques presented in [18]. Then Input Macro Indicators IMI are defined to describe these demand curves. The electric demand curves are used as input for the energy planning model that aims to optimally size and place the RES and storage in the existing electrical distribution network. Different indicators are then defined to describe the resulted scenarios aiming to perform the sensitivity analysis. In the studied case, features are assigned to the input curves as macro-parameters indicators and their correlation to the output indicators is studied. Based on the size (the number of model simulations is about 10 times the number of input parameters) and type of the data set, the linear regression (Pearson) is selected for the sensitivity analysis.

2. Methodology

The methodology summarized in Figure 1 aims to investigate the uncertainties in the Electrical Power System Model (EPSM). The EPSM used is the Dynamic Linearized Optimal Power Flow (DLOPF) model detailed in [18]. The model inputs considered are the demand curves. The uncertainties and variability of the demand are first characterized at each time step according to probability models which are used to generate additional demand curves. This set of generated demand profiles are used as inputs of the EPSM on a predefined distribution network to create different planning scenarios. Besides the IMI describing the patterns, output indicators are then determined to help measuring the sensitivity of model's outputs to inputs uncertainties implied by the variability of the demand. Finally, the sensitivity analysis is implemented in order to identify the most effective correlations between the input and output parameters of the model.

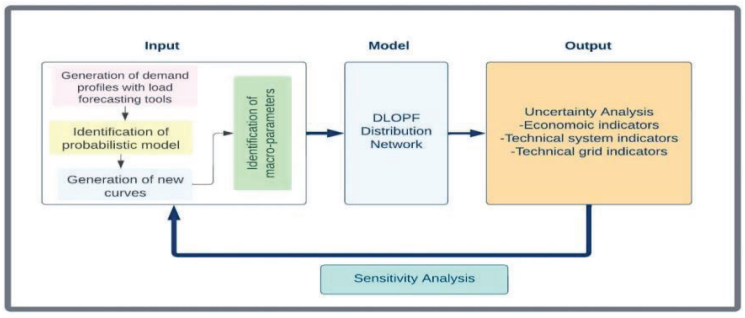


Figure 1. General workflow for uncertainty/sensitivity investigation in EPSM

2.1. Input Data

The electricity demand curves are the input data on which this uncertainty study is based. These curves are generated using a bottom-up load forecasting model called MOSAIC [19]. This tool is based on the characteristics of French loads built from a crossing of different databases: INSEE's residences database (French National Institute on Statistics and Economical Studies; French Distribution System Operators databases).

The total load curves by HTA/BT transformer are generated. For the studied area, MOSAIC simulations are repeated 25 times and therefore 25 curves per HTA/BT substation covering a whole year are obtained. The difference between curves is due to the attributed assumptions from the input data (building parameters,

https://doi.org/10.52202/069564-0308

weather data, etc.). These assumptions contain values and associated probability laws [20]. MOSAIC uses these values and makes random draws, according to the assumptions, to calculate a possible scenario.

The time-dependent uncertainties of the load curves are then characterized using the probability laws tests as explained in [18]. The Gamma law is proved to be the most appropriate law to describe these uncertainties. Then, based on the identified parameters describing the Gamma law at each time step, N=300 load curves are generated for each HTA/BT substation.

The demand curves are generated for one year. Hence, to reduce the simulation time, a clustering method is applied to choose 9 typical days (3 periods of 3 days) representing three different demand levels (high, medium, and low). The variation of the total demand over the 300 cases at each time-step (t) is shown in Figure 2.

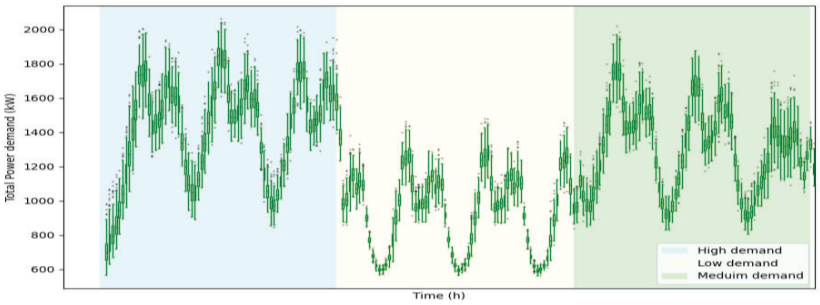


Figure 2. Ranges of variation of the generated energy demand curves

Several input macro indicators can be defined to describe load variation. The chosen parameters are detailed in the followings:

2.1.1.Peak demand values

The peak demand represents the higher demand value recorded in a time frame. The peak demands in each bus (i) are collected for the N_c cases (e) using the following equation (1) where the number of buses $\mathcal{N} = 14$, the time steps $\mathcal{T} = 216$, and the number of cases $N_c = 300$.

$$P_{i,e}^{max} = \max_{t} (P_{i,t,e}) \qquad i \in [1:\mathcal{N}] \quad t\epsilon [1:\mathcal{T}] \quad e\epsilon [1:N_c] \tag{1}$$

2.1.2.Standardized Variances

The variances $(Var_{i,e})$ in load demand curves designates the spread between demand values in each curve. More specifically, variance measures how far each value number in the curve is from the mean (average). The variances are calculated using Eq. (2)

$$Var_{i,e} = \frac{1}{\mathcal{T}} \sum_{t=1}^{\mathcal{T}} (P_{i,t,e} - \frac{\sum_{t=1}^{\mathcal{T}} P_{i,t,e}}{\mathcal{T}}) \tag{2}$$

2.1.3.Duration of maximum loads

The maximum demands are defined here as demands higher than 95% of the total peak demand. The number of hours during which these demands are encountered are defined as maximum durations Du_e.

2.1.4.Total energy demand variation

For that, three different parameters are defined summarizing the total demand in each clustered period (low, medium, and high).

These parameters are calculated using the following formulas Eq. (3)

$$E_e^{lev} = \sum_{t \in \Gamma_{lev}} \sum_{i=1}^{\mathcal{N}} P_{i,t,e} \tag{3}$$

Where $\boldsymbol{lev} \in \{\boldsymbol{Low}, \boldsymbol{Medium}, \boldsymbol{High}\}$ is the demand level and $\boldsymbol{\Gamma_{lev}}$ is the time period corresponding to the demand level $\boldsymbol{lev}$.

2.1.5. Simultaneity factor (Load coincidence)

The simultaneity factor represents how much consumers tend to consume simultaneously. The maximum demand for each bus may not occur at the same time. The ratio of the aggregated maximum demand of the whole network during a particular time to the sum of the maximum demand of individual consumers is called simultaneity factor eq. (4).

$$SF_e = \frac{P_e^{MAX}}{\sum_{i=1}^{Nbus} P_{i,e}^{max}} \tag{4}$$

Where P^{MAX} is the total load maximum value and P_i^{max} is the maximum value of bus i. SF_e ranges between 0 and 1 if all peak demands occur at the same time.

2.2. DLOPF model used.

Performing uncertainty and sensitivity analysis requires many simulations (300 in our case). An electrical planning model is used. This DLOPF model takes as initial conditions the Photovoltaic (PV), Wind Turbines (WT), and BES available surfaces and possible locations. The main outputs are the optimal scenario for sizing and placement of PV, WT, and BES in addition to the network simulation with time series variables.

$$min\, F_{obj} = min \frac{\left(\Gamma_{op} + \Gamma_{o\&m}\right) + \Gamma'_{inv}}{\sum_{i=1}^{\mathcal{N}} \sum_{t=1}^{\mathcal{T}} \frac{P_{i,t}^{l}}{(1+ra)^t}} \tag{5}$$

Such that

$$\Gamma'_{inv} = \mathcal{I}^{wt} + \mathcal{I}^{pv} + \mathcal{I}^{st} \tag{6}$$

$$\mathcal{I}^{st} = \sum_{s=1}^{\mathcal{N}^{st}} I^{st}.\boldsymbol{E_{max}^{st}} \tag{7}$$

$$\mathcal{I}^{pv} = \sum_{s=1}^{\mathcal{N}^{pv}} I^{pv}.\boldsymbol{A_s^{pv}} \tag{8}$$

$$\mathcal{I}^{wt} = \sum_{w=1}^{\mathcal{N}^{wt}} I^{wt}.\boldsymbol{R_w^{wt}} \tag{9}$$

$$\Gamma_{op} = \mathcal{C}^{cg} + \mathcal{C}^{vg} + \mathcal{C}^{wt} + \mathcal{C}^{pv} \tag{10}$$

$$\mathcal{C}^{cg} = \sum_{t=1}^{\mathcal{T}} \sum_{c=1}^{\mathcal{N}^{cg}} \frac{C^{cg,p}.\boldsymbol{P_{c,t}^{cg}}.\Delta t + C^{cg,q}.\boldsymbol{Q_{c,t}^{cg}}.\Delta t}{(1+ra)^t} \tag{11}$$

$$\mathcal{C}^{vg} = \sum_{t=1}^{\mathcal{T}} \sum_{i=1}^{\mathcal{N}^{vg}} \frac{(C^{vg,p}.\boldsymbol{P_{i,t}^{vg}}.\Delta t + C^{vg,q}.\boldsymbol{Q_{i,t}^{vg}}.\Delta t)}{(1+ra)^t} \tag{12}$$

$$\mathcal{C}^{wt} = \sum_{t=1}^{\mathcal{T}} \sum_{w=1}^{\mathcal{N}^{wt}} \frac{C^{wt}.\boldsymbol{P_{w,t}^{wt}}.\Delta t}{(1+ra)^t} \tag{13}$$

https://doi.org/10.52202/069564-0308

$$\mathcal{C}^{pv} = \sum_{t=1}^{\mathcal{T}} \sum_{s=1}^{\mathcal{N}^{pv}} \frac{C^{pv}.\boldsymbol{P}_{\boldsymbol{s,t}}^{\boldsymbol{pv}}.\Delta t}{(1+ra)^t} \quad (14)$$

$$\Gamma_{o\&m} = \frac{\sum_{i=1}^{\mathcal{T}}(O^{pv}.\boldsymbol{A}^{\boldsymbol{pv}} + O^{wt}.\boldsymbol{R}^{\boldsymbol{wt}} + O^{cg}.P^{cgT})}{(1+ra)^t} \quad (15)$$

The optimization model used aims to minimize the Levelized Cost of Energy (LCOE) of the system. The objective function is presented by (eq.(5)). It includes the total investment cost (Γ'_{inv}), production cost (Γ_{op}) and operation & maintenance cost ($\Gamma_{\mathrm{o\&m}}$). This sum is divided by the sum of demand ($\mathrm{P}_{\mathbf{i,t}}^{\mathrm{l}}$) (in all buses ($\mathcal{N}$)/ i: bus index $\epsilon\, \mathcal{N}$) for the defined period ($\mathcal{T}$) in years / t: time index $\epsilon\, \mathcal{T}$) over a function of the discount rate of the project (ra).

The total investment cost (Γ'_{inv}) presented by (eq. (6)) considers:

- The batteries investment cost ($\mathcal{I}^{st}$) in (eq. (7) where $\boldsymbol{E}_{\boldsymbol{max}}^{\boldsymbol{st}}$is the maximum energy and I^{ST} is the battery investment cost per unit of capacity.
- The PVs investment cost ($\mathcal{I}^{pv}$) in (eq. (8) where $\mathbf{A}^{\mathbf{pv}}$ is the installed area of PVs is an optimization and I^{pv} is the PV investment cost per unit of surface .
- The WTs investment cost ($\mathcal{I}^{wt}$) in (eq. (9) where $\mathbf{R}^{\mathbf{wt}}$ is the continuous design variable that defines the ratio of installed power capacity to a maximum installable capacity ($\mathbf{R}^{\mathbf{wt^{max}}}$) of local wind energy and I^{wt} is WT the investment cost of $\mathbf{R}^{\mathbf{wt^{max}}}$

The total actualized operational costs (Γ_{op}) is represented by Eq. (10). The $\mathcal{C}^{cg}$ is the actualized operational cost of the Classic Generators (CG)s and $\mathcal{C}^{vg}$ is the actualized operational cost of Virtual Generators[1] (VG)s. Solar and wind power operational costs (C^{wt}, C^{pv}) are supposed to have very low values since they are clean energy sources. The total operational cost for CGs ($\mathcal{C}^{cg}$) is given by eq.(11) where both the sum of active ($\boldsymbol{P}^{\boldsymbol{cg}}$) and reactive ($\boldsymbol{Q}^{\boldsymbol{cg}}$) produced power are multiplied by their corresponding costs ($C^{cg,p}$ for active and $C^{cg,q}$ for reactive). Like Eq.(11), Eq.(12) gives the virtual generation cost (C^{vg}). In each bus, virtual generators have generation costs significantly higher than other generators (both $C^{vg,p}$ and $C^{vg,q}$). The total operational costs of WTs ($\mathcal{C}^{wt}$) and PVs ($\mathcal{C}^{pv}$) are represented respectively by Eq.(13) and Eq.(14). Γ_{op} is variable according to the energy consumed.

Actualized operation and maintenance cost $\Gamma_{o\&m}$ is given by Eq.(15). It is dependent on the sizing of WTs, PVs and the existing CGs. and their O&M costs (O^{wt}, O^{pv}) in addition to the existing CGs maintenance costs (O^{cg}) multiplied by the total installed power of CGs (P^{cgT}). This cost is a fixed cost for the year regardless of the amount of production. The other equations that form the EPSM model are the linearized power flow equations of the DLOPF. These equations are the same presented in [21] and [18].

3. Case study

A theoretical case study is selected. The chosen topology is a Medium-voltage rural distribution benchmark network shown in

Figure 3. The benchmarking network comprises two separate subnetworks supplied by classic generators located in buses 1 and 12. These subnetworks are connected by buses 8 and 14. In this case, WTs, PVs and BESs could be added to specified buses concerning available surfaces for PVs and the existing potential of WTs. BES systems could also be added within the limit of maximum allowed capacities. This information is detailed in Table 1 with the corresponding costs.

3.1. Output

For the uncertainty analysis step, 300 simulations of the EPSM are performed. Four different types of output indicators are studied:

- Technical indicators of energy production and storage technologies: energy extracted from CG – Installed PVs, WTs capacities – Total produced energy – Batteries capacities.
- Technical grid indicators: branches' saturation levels (Δ_{ij}) – Duration in which this percentage is less than 10% (D_{ij})

[1] [1]Virtual Generators are defined as producers supposed to be added to all the buses to guarantee a feasible solution

- Losses

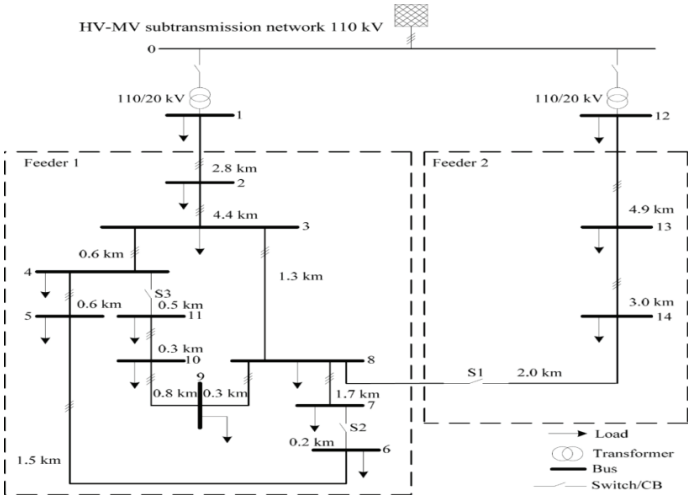


Figure 3: Medium voltage rural distribution benchmark network [22]

Table 1. Summary of buses specifications

Production type	Buses	Maximum capacity	Costs [23]		
			Production ($/MWh fuel)	Investment	O & M ($/MW/yr)
PV	3,4,7,11,13	$A^{pv^{max}} = 800\ \mathrm{m}^2$ $\mathrm{PV_t}$ (peak load) =205 W/m²	$\mathrm{C^{pv}} = 0$	J^{pv} =178 ($/m²)	$\mathrm{O^{pv}} = 8000$
WT	2,3,4,5,10,14	$R^{wt^{max}} = 10$ $\mathrm{WT_t}$ (peak load)=1000 KW	$\mathrm{C^{wt}} = 0$	$\mathrm{J^{wt}} = 997\ 000$ $/MW	$\mathrm{O^{wt}} = 33000$
CG	1, 12	20 MW	$\mathrm{C^{cg,p}}, \mathrm{C^{cg,q}} =$ 36	0 (already installed)	$\mathrm{O^{cg}} = 10500$
VG	all buses	10000 KW	$\mathrm{C^{vg,p}}, \mathrm{C^{vg,q}} = 10^9$	0	0
BES	2,3,4,5,7,9, 10,11,13,14	100 kWh	0	I^{ST}=350 $/kWh	0

3.1.1.Impact on the economic indicator: LCOE

The impact of load data uncertainty on the economic indicator is important since the objective function here is initially based on minimizing the LCOE. The violin plot of Figure 4 represents the shape of the LCOE results from the 300 cases generated above.

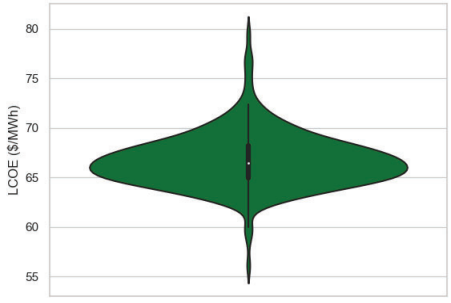


Figure 4. Violin plot for the variation of the levelized cost of energy in the different cases

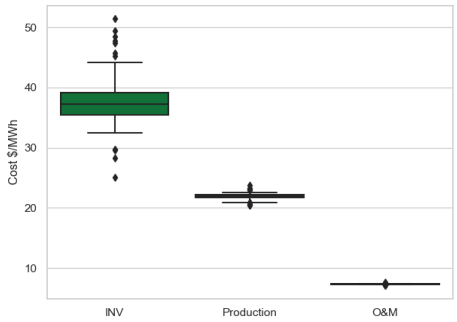


Figure 5. Cost details over the lifetime

https://doi.org/10.52202/069564-0308

The results are centered around the median value (white point) and are almost symmetrically distributed in within -15% and +20% of the variation around the median value. The LCOE analysis revealed that the randomness of the input variables significantly impacts the LCOE and results in a variation of 28 $/MWh which represents millions of dollars over the project lifespan.

The variation in the LCOE is a combination of variations between the investment, production, and O&M costs. These variations are presented in Figure 5. The LCOE variation is mainly due to the investment cost (between 25$/MWh and 51$/MWh) since this includes the investment in WTs, PVs, and batteries. This LCOE is affected to a minor extent by the production cost which is based on the CG production. The O&M cost variation is insignificant compared to others and has a negligible effect on the LCOE.

3.1.2.Technical system indicators

The impact of uncertainties affects the technical indicators of energy production and storage technologies differently. These indicators include the extracted energy from the existing CGs, the installation surface of PVs, the fraction of installed WTs, and the total energy produced in addition to the total installed capacity of the batteries. The results are summarized in the parallel coordinate plots of Figure 6.

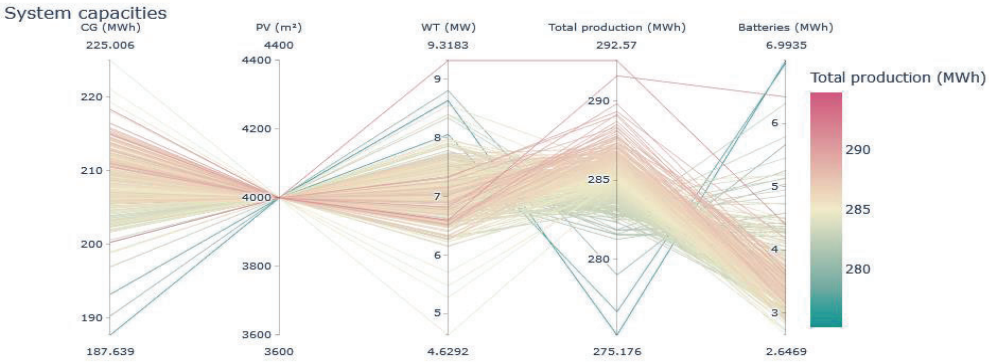


Figure 6. Parallel coordinate plots for the different cases results

The quality of the input data has no effect on the investment cost of the PVs. The installed area is constant at the maximum level for all cases (4000 m²), which means that the model always prioritizes the investment in the cheapest production technology compared to the other sources (CGs and WTs) so that the variations in investment and production amounts will be concentrated in the CGs and WTs. Moreover, substantial variations (16.6%) are shown in the energy extracted from the classic generators (up to 40 MWh) as well as in the installation of WTs (50.2%). The variation in the total production (5.94%) is less important than in the WT and CG since the sum contains both, and this balances the total variation because tracking the extreme points, we notice a maximum energy extracted from CG (225 MWh) that corresponds to a minimum installed WT power (4.6 MW) and a minimum CG (187.6 MWh) corresponds to a significant WT (8.6 MW). The variation in the WTs and PVs that produce intermittent energy induces the variation in the dimensioning of the batteries that aim to compensate for this intermittence. As long as the installed power of PV does not vary, the installed power of WT causes this important variation in the battery installation (62.13%).

3.1.3.Technical grid indicators

To focus on the network itself, technical indicators are used related to the grids, especially the amount of power flowing in the branches. The first indicator is the percentage of saturation of the branches (Δ_{ij}) and the second is the time (D_{ij}) during which the saturation is low and does not exceed 10%.

- The percentage of saturation (Δ_{ij}) in apparent power is calculated using eq.(16)

$$\Delta_{ij}\ (\%) = \frac{S_{ij}}{S_{ij}^{max}} \times 100 \qquad (16)$$

The power flow behavior is depicted in **Erreur ! Source du renvoi introuvable.Erreur ! Source du renvoi in trouvable.** for four different branches of one of the 300 simulated cases during the three typical days. In the high demand period, some branches are almost fully saturated, like branch 3-4; this branch also reaches a complete saturation in specific hours in medium (72h to 75h) and low (190h-195h) demand periods. Since this branch transmits power to several other branches, it affects the power delivery to the succeeding buses during these saturation periods, and therefore, it may influence the installation of generation sources to compensate for the lack of power delivery.

Other branches do not reach a saturation point that exceeds 60% of their capacity (e.g., branches 11-4). The maximum Δ_{ij} reached are identified for each branch in the 300 cases. The boxplots of Figure 7 show the repartition of the 300 values attributed to each branch. The uncertainties in input demand do not affect the maximum saturations of the first four branches. The first 3 branches are saturated most of the time because they are responsible for the delivery of energy to other branches. Full saturation is noticed in branches 6-7 also, and it comes back to the non-possibility of installation of RES and BES on bus 6. However, the maximum saturation rates achieved in the other branches vary by ranges of about 5% (12-13, 13-14, etc.).

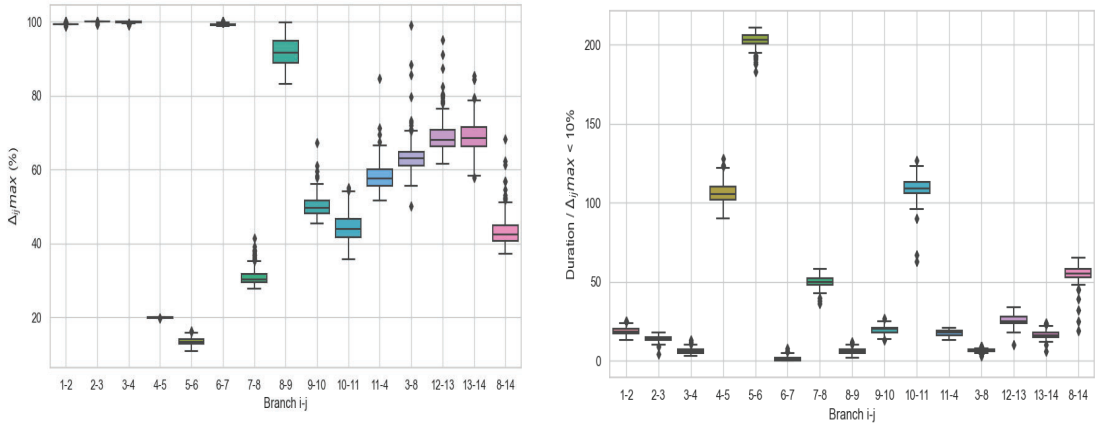


Figure 7. Boxplots of the apparent power saturations (a) and the duration of low saturation (less than 10 %) (b) in each branch

- The duration with very low branch saturation is expressed in eq.(17)

$$D_{ij} = duration\ of\ \ \Delta_{ij} < 10\ \% \qquad (17)$$

The time during which the branches are within 10% of their saturations over the defined duration is considered as a technical indicator for the grid since it concerns the grid's branches dimensioning and their maximum apparent power. The results of the 300 cases are presented in the boxplots of Erreur ! Source du renvoi i ntrouvable. for each branch.

D_{ij} is varying more in the branches where its values are more significant than in those where D_{ij} is about few hours that corresponds to the branches with higher maximum Δ_{ij}. The more the branches are saturated, the less they reach saturation less than 10%. For better understanding, we select the most critical branches (with Δ_{ij}> 80% and D_{ij} < 25 hours). The selected branches are, therefore (1-2, 2-3, 3-4, 6-7 and 8-9). The first branch (1-2) is critical since it is one of the main branches importing power from the classic generators to an important part of the network. The other two branches (2-3 and 3-4) are part of the branch (1-12) that will feed an important part of the network; in addition, they contain RES and BES that satisfy their demands first and transmit as much as possible for the rest of the network. Branches 6-7 and 8-9 are critical because they are connected to buses 6 and 8, where no installation is possible (neither RES nor BES) as shown in **Table *1***. Therefore, these branches deliver power with their maximum capacity to be able to saturate the demand for these buses.

- The grid losses

The grid losses are estimated by calculating the difference between production and demand. The results are presented in the following **Erreur ! Source du renvoi introuvable.**. In most cases, losses have less than 1% o f differences since they vary between 7.1 % and 7.8%. In some cases, losses are at the level of 9.8 %, while minimum observed losses are around 4%.

4. Sensitivity Analysis

The objective of sensitivity analysis is to determine which model parameters are important and their relative impact on the results. This gives insights about the efforts or resources needed to reduce the total uncertainty of the system's forecasts. Among the different SA technics presented in the literature, quantitative tests are made for sensitivity analysis using correlation analysis.

Tests based on correlation analysis.

https://doi.org/10.52202/069564-0308

The Pearson correlation coefficient is widely used in the literature. It measures the association between each input considered separately and the output. A number between -1 and 1 measures the strength and direction of the relationship between two variables. The correlation is given by the eq.(18).

$$r = \frac{\sigma_{xy}}{\sigma_x . \sigma_y} = \frac{\sum_i (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_i (x_i - \bar{x})^2} \sqrt{\sum_i (y_i - \bar{y})^2}} \qquad i = 1, \dots, N \tag{18}$$

Where N is the number of samples.

The values of r are interpreted as follows [24]:

- Between 0 and 1: a positive correlation exists between variables (strong r > 0.5, moderate 0.3< r < 0.5 and weak 0 < r < 0.3) – when one variable changes, the other variable changes in the same direction.
- 0: No correlation (there is no relationship between the variables)
- Between 0 and -1: a negative correlation exists between variables (strong r < -0.5, moderate -0.3> r > -0.5 and weak 0 > r > -0.3) – when one variable changes, the other variable changes in the opposite direction.

To perform the Pearson correlation test, both variables should be quantitative. They should be normal or a little non-normally distributed. In the studied case, this distribution is visualized by the diagonal of Figure 10. In Pearson correlation tests, choosing a sample size of up to 258 variables means a correlation test with a power 90%, an error 5% and an alternative correlation 0.2 as detailed in the sample size guideline for correlation analysis [25].

The input parameters are therefore $\boldsymbol{P_{i,e}^{max}}$, $\boldsymbol{Var_{i,e}}$, $\boldsymbol{Du_e}$, $\boldsymbol{E_e^{Low}}$, $\boldsymbol{E_e^{Meduim}}$, $\boldsymbol{E_e^{High}}$ and $\boldsymbol{SF_e}$. The output indicators are $LCOE_e$, WT_e, PV_e, BES_e, $losses_e$, $\Delta_{ij,e}$ and $D_e = \sum D_{ij,e}$ with i: bus index, e: case index and ij: branch index. The results of the test of linear relationships between the parameters ($\mathrm{Var_{i,e}}$, $\mathrm{P_{i,e}^{max}}$) and indicators are presented respectively in the following **Erreur ! Source du renvoi introuvable.** and **Erreur ! Source du renvoi introuvable.**. The Pearson correlation index r is calculated for the remaining parameters and represented in Figure 10. The indicator of PVs is excluded since it is demonstrated that the uncertainties do not affect this investment.

From **Erreur ! Source du renvoi introuvable.**, it can be noticed that 4 buses (1,3,10 and 14) have marked an effect of variances on the output indicators. The increased standardized variances of buses 1, 10 and 14 cause an increase in the installed WT power and a decrease in the extracted CG energy. The batteries are affected by the variances in buses 3 and 14. The LCOE increases with the increase of standardized variances of buses 3, 10 and 14. Even if the correlation exists, it is considered weak (r < 0.3). The most influencing bus based on variances is the last one (bus 14) since the increase of load variance in bus 14 also causes an increase in the duration of minor saturations, and a decrease in losses.

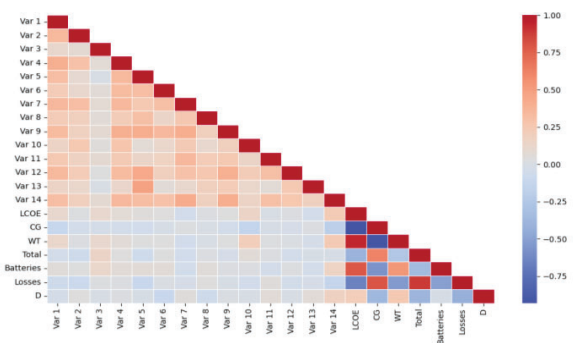


Figure 8. r values resulted from Pearson correlations tests of standardized variances ($Var_{i,k}$) on each bus with the output indicators

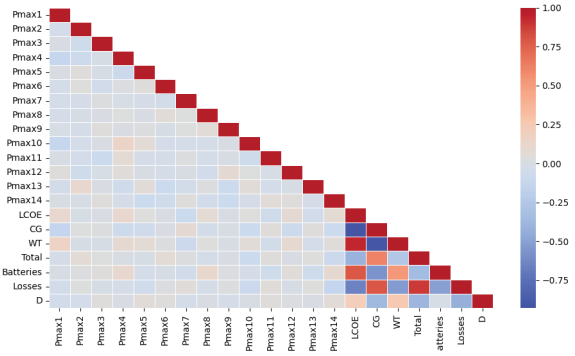


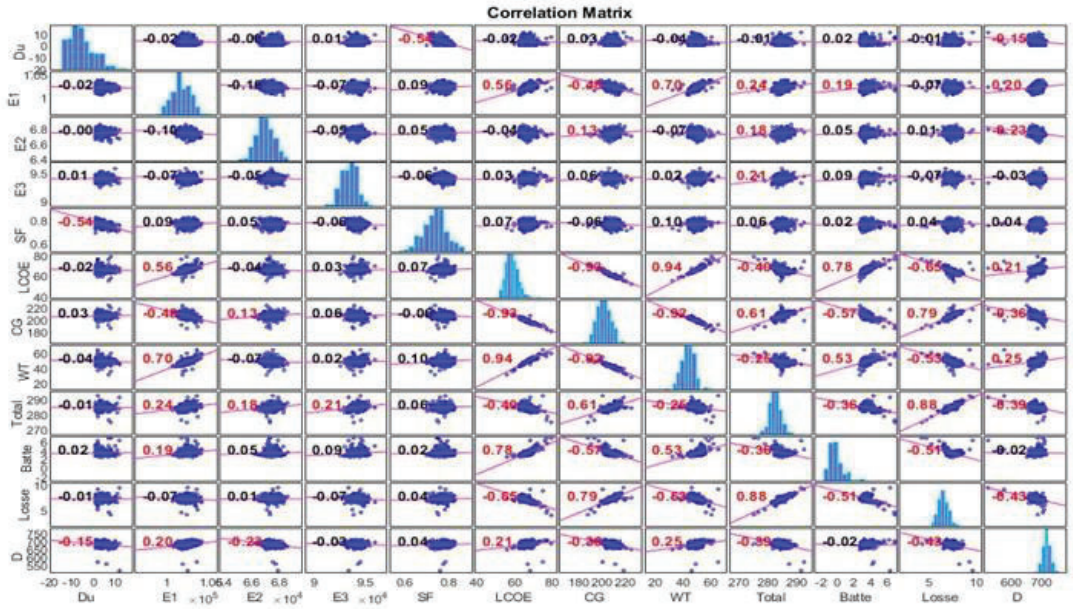
Figure 9. r values resulted from Pearson correlations tests of P^{max} on each bus with the output indicators

In **Erreur ! Source du renvoi introuvable.**, the results show that 4 buses (1,4,8 and 14) could affect the indicators but also with weak correlations (r < 0.15). An increase in the peak power of bus 1 leads to an increase in WT investment and a decrease in CG extraction because this bus conveys power to other buses; when the demand increases, the saturation effect of the branch decreases the quantity conveyed, and consequently the

system invests more in WTs. Increasing the peak power of bus 8 leads to an increase in battery investment since this bus does not contain a RES or BES and is fed from the grid. This increase in peak demand will be compensated by the storage of the power injected into the network, and then more investment in batteries.

The five input macro-parameters correlations with output indicators are displayed in Figure 10. The increase of the first parameter (the maximum duration parameter (Du)) leads to a decrease in the duration with very low branch saturation D_k (r = -0.15). By the fact that as long as the maximum demand values are reached, the low saturations of the branches appear less.

Figure 10. Tests on the different input parameters correlations with the output



The total energy demands at each demand level calculated by eq.(3) are expressed in Figure 10 respectively by $\mathrm{E1}$ for $\mathrm{E_e^{High}}$, $\mathrm{E2}$ for $\mathrm{E_e^{low}}$ and $\mathrm{E3}$ for $\mathrm{E_e^{Medium}}$. The values of LCOE increase with the high demand total energies $\mathrm{E_e^{High}}$ increasing. Thus, E1 has a strong effect (r = 0.56) on the LCOE, while others have moderate effects (r = -0.04) for E2 and (r = 0.03) for E3. When the total low demand energy $\mathrm{E2}$ increases, the extracted energy from CGs increases also (r = 0.13) which means that this low-level energy is mainly extracted from CGs. This returns to the grid behaviour in the low-demand period when the branch limits are not reached, and the energy from the classic generators can be continuously transmitted. In addition, E2 coincides with a period where renewable energy is less present and therefore increased use of batteries which is less competitive with classic generators. The increasing in $\mathrm{E2}$ also results in a decrease in the $\mathrm{D_k}$ ($\mathrm{r} = -0.23$) because this demand period is the one with a high occurrence of the hours of less than 10% of Δ_{ij}. So, a decrease in the demand increases these durations.

High total energy demand $(E1)$ is the most impactful parameter. An increase in $E1$ leads to a strong increase in investment in WTs ($r = 0.7$) and, therefore, an increase in battery dimensioning ($r = 0.19$). This leads to a decrease in CGs energy extraction ($r = -0.48$) and a significant decrease of the LCOE ($r = -0.5$). An increase in duration D_k is observed ($r = 0.2$). This is because of decentralization of producers (WTs and BES) and the limitation of the transition of the power from CGs, which reduces the saturation of the nearest branches to CGs after recourse to renewables and batteries.

The simultaneity factor (SF) has very weak influences on the indicators ($\mathrm{r} < 0.1$).

The distribution of the parameters is shown in the diagonal of Figure 10. The Gamma test is applied to the output indicators distribution, and the results give p-values lower than 0.05, so the hypothesis of this law is rejected this means that the probabilities of accepting the hypothesis (the data are gamma-distribut) are less than 5% [26].The difference in distributions between the input parameters and the output indicators means that the model does not propagate uncertainties in a linear way between inputs and outputs. The distribution of most parameters is close to a normal distribution as shown in the distributions of E1, E2, E3, SF, CG, Losses, etc. Therefore, the use of Pearson correlations is validated.

5. Conclusions

https://doi.org/10.52202/069564-0308

This paper uses an EPSM model to perform uncertainty and sensitivity analyses. Input macro-parameters were defined to describe an quantify the model's inputs uncertainties. A sensitivity analysis is performed using the correlations between the input parameters and the output indicators to identify the most influential parameters. The results show that the uncertainties have considerable effects on the results. In economic terms, uncertainties lead to a 35% variation range of the LCOE. In terms of system design, there is a need to increase the installed capacity of wind turbines, extract more energy from power plants, and/or install more BES depending on the case. Uncertainties also affect grid saturation in the branches. The analysis enabled the identification of the most important parameter which is the total energy demand during the high demand period (winter), as this parameter strongly affects the defined indicators. This makes this parameter a key parameter in the sizing of RES, BES and thus in the network planning.

Nomenclature

Variables	
P	Active power [MW]
Q	Reactive power [MVAr]
F	Objective function
$\mathcal{C}$ / $\mathcal{I}$	Total operational/investment cost by production
Γ	Total cost by category
CC	Capital Cost
A	Installed surface of PV [m²]
R	Ratio of installed WT capacity [MW]
E	Active energy
r	Correlation parameter
Var	Variances
Du	Duration of maximum loads
SF	Simultaneity factor
Δ	Percentage of brunch saturation
D	Duration of less than 10% of brunch saturation
Parameters	
$\mathcal{C}$	Elementary operational cost
$\mathcal{J}$	Elementary investment cost
O	Elementary O&M cost
LT	Lifetime
ls	Lifetime of the overall system
ra	Discount rate of the project
Indices and sets	
$\mathcal{N}$	Set of buses, $i, j \in \mathcal{N}$
$\mathcal{T}$	Set of all times, $t \in \mathcal{T}$
$\mathcal{B}$	Set of batteries
Nb	Number of batteries
c / u / s / w	Indices of buses with CG/VG/PV/WT
k	Iteration number in DLOPF
inv	Investment
$o\&m$	Operation and maintenance
op	Operational
e	Case index
Upper-scripts	
cg / vg	Classic generator/virtual generator
pv / wt / st	Photovoltaic/wind turbines/Storage
b	BES index
l	Load index at a bus
p / q	Active/reactive power
max	Maximum value for the upper limit
low	Low energy levels
$medium$	Medium energy levels

$high$	High energy levels

Acknowledgments

This work was carried out under the auspices of the ValaDoE chair at IMT Atlantique, in partnership with Télécom Paris and Mines Saint Etienne, and was supported by Enedis, Région Pays de la Loire, Nantes Métropole and Akajoule.

References

[1] S. Moret, V. Codina Gironès, M. Bierlaire, et F. Maréchal, « Characterization of input uncertainties in strategic energy planning models », *Appl. Energy*, vol. 202, p. 597-617, sept. 2017, doi: 10.1016/j.apenergy.2017.05.106.
[2] S. F. Santos *et al.*, « Novel Multi-Stage Stochastic DG Investment Planning with Recourse », *IEEE Trans. Sustain. Energy*, vol. 8, nº 1, p. 164-178, janv. 2017, doi: 10.1109/TSTE.2016.2590460.
[3] A. Ehsan et Q. Yang, « State-of-the-art techniques for modelling of uncertainties in active distribution network planning: A review », *Appl. Energy*, vol. 239, p. 1509-1523, avr. 2019, doi: 10.1016/j.apenergy.2019.01.211.
[4] G. Mavromatidis, K. Orehounig, et J. Carmeliet, « Uncertainty and global sensitivity analysis for the optimal design of distributed energy systems », *Appl. Energy*, vol. 214, p. 219-238, mars 2018, doi: 10.1016/j.apenergy.2018.01.062.
[5] R. Mena, M. Hennebel, Y.-F. Li, C. Ruiz, et E. Zio, « A risk-based simulation and multi-objective optimization framework for the integration of distributed renewable generation and storage », *Renew. Sustain. Energy Rev.*, vol. 37, p. 778-793, sept. 2014, doi: 10.1016/j.rser.2014.05.046.
[6] A. Schmoldt, H. F. Benthe, et G. Haberland, « Digitoxin metabolism by rat liver microsomes », *Biochem. Pharmacol.*, vol. 24, nº 17, p. 1639-1641, sept. 1975.
[7] G. Mavromatidis, « Model-based Design of Distributed Urban Energy Systems under Uncertainty », p. 386.
[8] M. R. Font, « ANALYSE DE SENSIbILITE SUR UN MODELE STOCHASTIQUE Pour la PREVISION DEs CONSOMMATIONS ELECTRIQUES », p. 72.
[9] « Enedis | Gestionnaire du réseau de distribution d'électricité ». https://www.enedis.fr/ (consulté le 15 mars 2023).
[10] D. Sadeghi, A. Hesami Naghshbandy, et S. Bahramara, « Optimal sizing of hybrid renewable energy systems in presence of electric vehicles using multi-objective particle swarm optimization », *Energy*, vol. 209, p. 118471, oct. 2020, doi: 10.1016/j.energy.2020.118471.
[11] « Global sensitivity and uncertainty analysis of the levelised cost of storage (LCOS) for solar-PV-powered cooling | Elsevier Enhanced Reader ». https://reader.elsevier.com/reader/sd/pii/S0306261921000854?token=19FF74DD5BF3BF4079BBB8375E018B4BD68EBD20735709B67DA2969AE5C4D5F306631561DBB25A6BD56B206F4D4DAFC6&originRegion=eu-west-1&originCreation=20220822090205 (consulté le 22 août 2022).
[12] M. Ma, H. Huang, X. Song, F. Peña-Mora, Z. Zhang, et J. Chen, « Optimal sizing and operations of shared energy storage systems in distribution networks: A bi-level programming approach », *Appl. Energy*, vol. 307, p. 118170, févr. 2022, doi: 10.1016/j.apenergy.2021.118170.
[13] C. Lythcke-Jørgensen, A. V. Ensinas, M. Münster, et F. Haglind, « A methodology for designing flexible multi-generation systems », *Energy*, vol. 110, p. 34-54, sept. 2016, doi: 10.1016/j.energy.2016.01.084.
[14] A. Mian, « Optimal design methods applied to solar-assisted hydrothermal gasification plants », EPFL, 2016. doi: 10.5075/epfl-thesis-6945.
[15] A. Saltelli, Éd., *Global sensitivity analysis: the primer*. Chichester, England ; Hoboken, NJ: John Wiley, 2008.
[16] G. Saporta, *Probabilités, analyse des données et statistique*. Editions TECHNIP, 2006.
[17] L. L. Gratiet, S. Marelli, et B. Sudret, « Metamodel-Based Sensitivity Analysis: Polynomial Chaos Expansions and Gaussian Processes », in *Handbook of Uncertainty Quantification*, R. Ghanem, D. Higdon, et H. Owhadi, Éd., Cham: Springer International Publishing, 2015, p. 1-37. doi: 10.1007/978-3-319-11259-6_38-1.
[18] S. Fakih, M. T. Mabrouk, M. Batton-Hubert, et B. Lacarriere, « Impact of Uncertainties in Power Demand Estimation on the Optimal Design of Renewable Energy Sources and Storage Systems », in *2022 IEEE 10th International Conference on Smart Energy Grid Engineering (SEGE)*, Oshawa, ON, Canada: IEEE, août 2022, p. 68-73. doi: 10.1109/SEGE55279.2022.9889769.
[19] N. Kong *et al.*, « Long-term forecast of local electrical demand and evaluation of future impacts on the electricity distribution network », *CIRED - Open Access Proc. J.*, vol. 2017, nº 1, p. 2401-2405, oct. 2017, doi: 10.1049/oap-cired.2017.0743.
[20] M. R. Font, « ANALYSE DE SENSIbILITE SUR UN MODELE STOCHASTIQUE Pour la PREVISION DEs CONSOMMATIONS ELECTRIQUES », p. 72.
[21] S. Fakih, M. T. Mabrouk, M. Batton-Hubert, et B. Lacarriere, « Optimal Allocation and Sizing of Renewable Energy Sources and Storage Systems to Support Over-Solicited Electricity Grid », in *2022 6th International Conference on Green Energy and Applications (ICGEA)*, Singapore, Singapore: IEEE, mars 2022, p. 14-21. doi: 10.1109/ICGEA54406.2022.9791941.
[22] « pandapower ». [En ligne]. Disponible sur: https://pandapower.readthedocs.io/en/v2.1.0/networks/cigre.html
[23] Eurostat, « Natural gas price statistics ». [En ligne]. Disponible sur: https://ec.europa.eu/eurostat/statistics-explained/index.php?title=Natural_gas_price_statistics#Natural_gas_prices_for_non-household_consumers
[24] S. Turney, « Pearson Correlation Coefficient (r) | Guide & Examples », *Scribbr*, 13 mai 2022. https://www.scribbr.com/statistics/pearson-correlation-coefficient/ (consulté le 10 octobre 2022).
[25] M. A. Bujang et N. Baharum, « Sample Size Guideline for Correlation Analysis », *World J. Soc. Sci. Res.*, vol. 3, nº 1, p. 37, mars 2016, doi: 10.22158/wjssr.v3n1p37.
[26] S. Mcleod, PhD, « P-Value And Statistical Significance: What It Is & Why It Matters ». 1 mai 2023. [En ligne]. Disponible sur: https://www.simplypsychology.org/p-value.html

 https://doi.org/10.52202/069564-0308