energies

MDPI

Article

Scenario Generation Based on Ant Colony Optimization for Modelling Stochastic Variables in Power Systems

Daniel Fernández Valderrama [1,*], Juan Ignacio Guerrero Alonso [2], Carlos León de Mora [2] and Michela Robba [1]

[1] Department of Informatics, Bioengineering, Robotics, and Systems Engineering, University of Genoa, 16145 Genoa, Italy; michela.robba@unige.it
[2] Department of Electronic Technology, Escuela Politécnica Superior, University of Sevilla, 41011 Sevilla, Spain; juaguealo@us.es (J.I.G.A.); cleon@us.es (C.L.d.M.)
* Correspondence: daniel.fernandez@edu.unige.it

Abstract: Uncertainty is an important subject in optimization problems due to the unpredictable nature of real variables in the power system area, which can condition the solution's accuracy. The effective modelling of stochastic variables can contribute to the reduction in losses in the system under evaluation and facilitate the implementation of an effective response in advance. To model uncertainty variables, the most extended technique is the scenario generation (SG) method. This method evaluates possible combinations of complete curves. Classical scenario generation methods are founded on probability distributions or robust stochastic optimization. This paper proposes a novel approach for constructing scenarios using the Ant Colony Optimization (ACO) algorithm, referred to as ACO-SG. This methodology does not require a previous statistical study of uncertainty data to generate new scenarios. A historical dataset and the desired number of scenarios are the inputs inserted into the algorithm. In the case study, the algorithm used historical data from the Savona Campus Smart Polygeneration Microgrid of the University of Genoa. The approach was applied to generate scenarios of photovoltaic generation and building consumption. The low values of the Euclidean distance were used in order to check the validity of the scenarios. Moreover, the error deviation of the scenarios generated with the goal of daily power were 1.77% and 0.144% for the cases of PV generation and building consumption, respectively. The different results for both cases are explained by the characteristics of the specific cases. Despite these different results, both were significantly low, which indicates the capability of the algorithm to generate any kind of feature within curves and its adaptability to any case of SG.

Keywords: ant colony optimization; microgrids; stochastic processes; energy management

Citation: Fernández Valderrama, D.; Guerrero Alonso, J.I.; León de Mora, C.; Robba, M. Scenario Generation Based on Ant Colony Optimization for Modelling Stochastic Variables in Power Systems. *Energies* **2024**, *17*, 5293. https://doi.org/10.3390/en17215293

Academic Editors: Marcin Sosnowski, Jaroslaw Krzywanski, Karolina Grabowska, Dorian Skrobek and Ghulam Moeen Uddin

Received: 16 September 2024
Revised: 16 October 2024
Accepted: 23 October 2024
Published: 24 October 2024

1. Introduction

1.1. Motivation

The European Green Deal fixed numerous objectives to achieve by 2050, including achieving zero net greenhouse gas emissions and decoupling economic growth from resource use [1]. These goals pose significant challenges to the contemporary energy sector. The widespread electrification of aspects such as buildings, transport and industry, driven by renewable energy sources (RESs), carries various complexities in the operation of the power grid. This transition requires additional flexibility and substantial investment to reinforce transmission and distribution networks. Furthermore, the growing integration of variable RESs into the power system, coupled with the shift towards proactive energy consumers who generate and consume energy, presents an increased risk to the grid's reliability and stability of the grid. These variables become a challenge for operational decision-making in the energy system due to their difficulty in terms of prediction and their stochastic nature.

Dealing with uncertainty can significantly affect the objective function value of solutions generated by decision-support techniques in real-world applications. This is a critical aspect in several application areas leading to the development of appropriate solution approaches that incorporate uncertainty. In decision theory, uncertainty is usually represented through scenarios that include various possible outcomes of indeterminate parameters in a problem. These scenarios may or may not have a dedicated probability or probability distribution. Therefore, in addition to expected value methods, non-probabilistic methods can also address uncertainty in decision-making. These methods consider various decision-makers with different risk profiles, ranging from total pessimism to total optimism, including risk-balancing strategies. The scenarios describe possible situations referring to the analyzed variable, which normally is an unpredicted behaviour from the real world. This fact makes scenarios a key tool for stochastic programming. Stochastic programming differs from deterministic optimization in that it accounts for uncertain problem parameters. It has a wide range of applications in fields such as finance, transport, and energy optimization, as many real-world decisions involve uncertainty. In many instances of stochastic programming, discrete distributions known as SG approximate the probability distributions of the uncertain variable [2]. This provides feasible approaches for simplifying large amounts of data to approximate the probability distributions. Specifically, this document is focused on the well-known two-stage stochastic optimization problem [3]. In the first stage, a set of decisions is fixed before any realized uncertainty, which cannot be changed in the second stage. Recourse decisions based on the first stage are made in the second stage.

This paper analyses authentic historical data for the Smart Polygeneration Microgrid (SPM) at Savona Campus, University of Genoa [4,5]. The microgrid was installed at the Savona Campus in 2014 thanks to the Energia2020 project [6] and serves as a living laboratory for the University of Genoa. The microgrid is located in the western part of the Liguria region in Italy and is powered by renewable energy sources such as geothermal energy and solar panels. It also includes storage systems, electric vehicles, gas microturbines, a smart building and an Energy Management System, the layout of which is illustrated in Figure 1. It provides all the necessities such as electrical power, heating, cooling and hot water to the campus buildings. The campus accommodates approximately 2000 individuals, including students, professors, university staff and employees.

1.2. Literature Review

The objective of this section is to offer the reader pertinent background information by presenting and discussing the literature that supports the notion of incorporating evolutionary algorithms for SG. Various methods can be differentiated for scenario generation areas, and [7] categorizes them into three groups: sampling-based, forecasting-based and optimization-based. The predominant approach is sampling-based, which entails fitting a statistical model to the uncertain parameters and generating samples from it. The resulting outcomes are samples that conform to the trained distribution. Some examples are Monte Carlo (MC) [8], Latin Hypercube Samples (LHSs) [9], copula function sampling [10], etc. These approaches are simple and fast, but they present limitations in multivariate scenarios as they cannot capture the temporal and spatial correlations of variables. To achieve a reliable representation of the distribution by the samples and ensure adequate coverage of the space, these approaches require a large number of samples. Consequently, a scenario reduction method is necessary [4]. It is important to note that a high number of scenarios is not practicable for optimization purposes. Forecasting-based methods consist of training models from historical data without taking the distribution into account.

Techniques in this group include Auto Regressive Moving Average (ARMA) [11], Artificial Neural Networks (ANNs) [12] and Generative Adversarial Networks (GANs) [13]. Forecasting-based methods are effective in capturing the characteristics of variables, in particular, correlations and complex nonlinear relationships. Remarkably, this methodology is dependent on data-driven methods, which means that the quality of the scenarios generated is dependent on historical observation samples. Finally, optimization-based methods aim to reduce a large number of possible scenarios. Clustering, Backward Reduction (BR) [14] or Forward Selection (FS) [15] are techniques that belong to the optimization-based group.

They present a significant NP-hard problem that solves intricate complications but proves difficult to implement in large power systems. They can be combined to form new methodologies with some features that enhance traditional techniques. More advanced algorithms derived from the above are available, such as Markov-Chain Monte Carlo (MCMC) [16], LHSs incorporating the correlations between random variables [17] or others, such as those based on decision trees [18]. Yadav et al., in [19], developed a hybrid optimization algorithm, named "Genetic Algorithm-Grey Wolf Optimizer", which combines the GA and Grey Wolf Optimizer algorithms. The objective is to find the global maximum power point of a PV generation under different conditions. However, the proposed algorithm is a complex and resource-intensive method in comparison to both parent algorithms and conventional techniques.

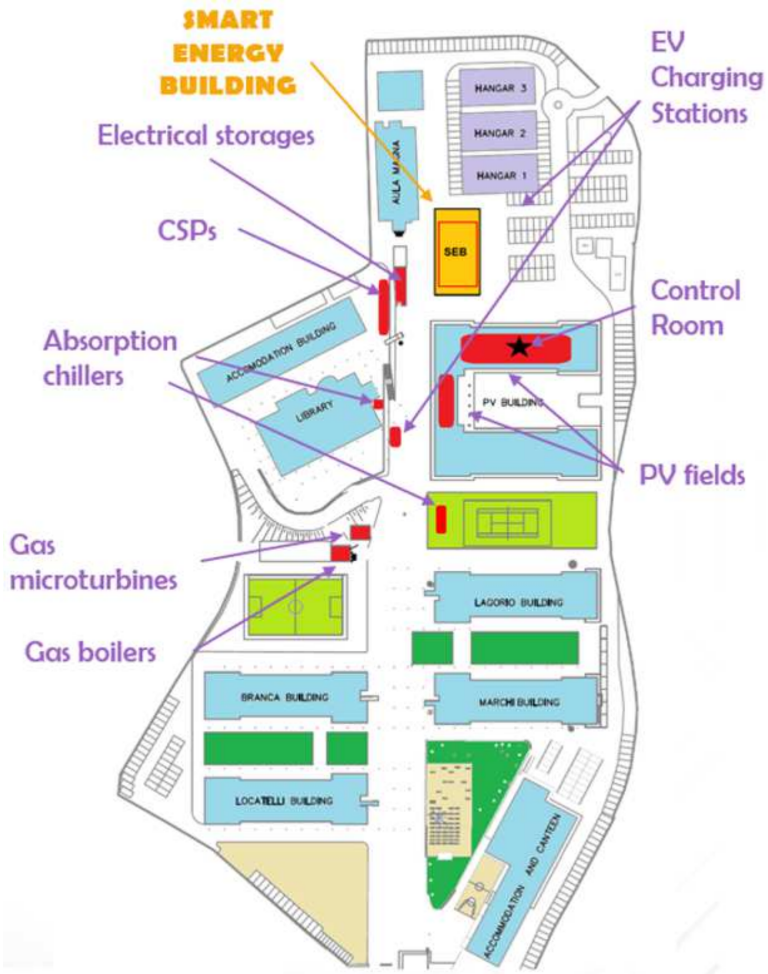


Figure 1. Smart Polygeneration Microgrid layout of Savona.

Generally, SG strategies tend to follow a probability distribution which is associated with an uncertain value. Therefore, the first stage in these approaches is to model the probability distribution, as shown in [20], where the authors implement different SG methods. On the other hand, some works concentrate on a set of scenarios generated by metaheuristic techniques, as in the case of [21] by Oliveira et al., where they propose a Genetic Algorithm (GA) to obtain a diverse number of scenarios. They do not rely on probability distributions, but a metric called "crowding distance" to measure the diversity between scenarios. The GA searches for various scenarios designed to increase the "crowding distance", indicating that they are far apart according to this metric. Nevertheless, the main problem with this method seems to be the execution time, which depends on the number of scenarios chosen.

Several studies have used metaheuristic algorithms for the placement and sizing of different microgrid systems; for example, [22] tests a large number of scenarios to meet the load demand on a tri-objective goal. Indeed, evolutive algorithms proved to be beneficial in solving stochastic multi-objective problems [23]. These examples demonstrate the effectiveness of such algorithms in dealing with large search spaces. However, their purpose is to solve optimization problems by testing various scenarios rather than generating them in the majority of cases. Another interesting algorithm is the musical chair algorithm [24], which consists of removing the worst-performing agents and replacing them with new agents to look for the optimal solution. The author achieved faster convergence with this method. Despite the appeal of the underlying concept, the approach still relies on random computing efforts to solve the issue of convergence. In contrast, ACO employs an intelligent use of the pheromone left by previous agents to guide the subsequent agents and accelerate convergence.

In particular, ACO is a bio-inspired EA that is based on the foraging behaviour of ants. ACO is widely used to solve combinatorial optimization problems such as the Travelling Salesman Problem (TSP) [25], the job shop scheduling problem, the vehicle routing problem, or the knapsack problem. ACO is able to rank possible next steps based on the pheromone deposition strategy inspired by the ability of ants to find the best solution by iteratively constructing and updating a construction graph. The ACO algorithm has proven to be an effective method in numerous applications due to its robustness, versatility and scalability. The adaptability of the problem of SG to identify the optimal path and the adaptability of the ACO to explore the space of the curves are the main reasons for selecting this specific evolutionary algorithm to address this project.

ACO has been used in multiple studies related to power systems, for example, the optimal integration of Distributed Generation (DG) into a distribution system [26] to minimize a multi-objective function based on power loss, voltage deviation and operating costs. In addition, ACO has been implemented in [27] to solve problems of PV module toxicity throughout their lifetime to keep it at an acceptable level, which is an approach that regulates PV module efficiency. These studies adhere to a traditional nonlinear problem where EAs typically perform efficiently in finding a solution. Aghelpour et al. [28] used ACO combined with an adaptive neuro-fuzzy inference system for the daily streamflow prediction of a river. Several studies have used ACO for diverse applications. Nevertheless, there is a lack of an ACO model in the literature that allows scenario generation. Furthermore, few studies have implemented evolutionary algorithms for this purpose. Another advantage of the ACO-SG algorithm is the ability to generate scenarios that are directly representative of the whole space. This eliminates the need to reduce the scenarios after their generation, which is a typical practice in this type of problem, as shown in [4].

This research proposes a novel algorithm for SG based on ACO that generates scenarios from a real dataset, which will be explained in detail in Section 3. The algorithm can generate a set of new scenarios whose patterns are related to the historical dataset. These scenarios must satisfy the criteria of validity and reliability in order to be considered credible and useful, as defined in [29].

1.3. Aims and Contributions

This paper aims to address the potential mismatch between evolutionary algorithms and scenario generation by proposing an efficient tool that can generate a specific number of scenarios accurately representing the range of uncertain variables. Generally, SG methods attempt to approximate the scenarios to a distribution function [30–32], reduce variance to estimate tighter confidence intervals [33] or minimize approximation errors through dynamic sampling [34].

Traditionally, SG methodologies generate scenarios from a probability distribution originally modelled from a statistical study of a historical dataset. Nevertheless, the approach presented in this work attempts to generate scenario set diversity in order to cover the different impacts of the evaluated uncertain variable. In comparison with the traditional methods, one of the main advantages of the approach outlined in this document is the elimination of the necessity to model the probability distribution.

The approach exposed in this document allows the generation of scenarios within the range of the historical dataset. The desired total of scenarios is spread over that space by specifying the daily power of each scenario, each of them with an acceptable margin of error. This methodology avoids the necessity of studying the probability distribution, which is a fundamental aspect of the majority of other methodologies. The algorithm can reproduce scenarios according to spatio-temporal features captured from the historical dataset, ensuring that they are feasible as a function of the total amount of power required. In the case study, scenarios have been implemented for PV generation and building consumption based on historical data. The main contributions of the study are highlighted as follows:

- The ACO algorithm has been adapted for the generation of scenarios, which supposes a novel algorithm in the area. It can generate new scenarios from historical data that accurately represent the range of the case under study.
- The algorithm takes a specific number of scenarios as input to generate them, specified by the user. Those scenarios are spread over the space according to the historical dataset. Subsequently, post-processing for scenario reduction is unnecessary, as is common in this field. Moreover, the necessity for a probability distribution study is negated by this approach.

1.4. Structure

The remainder of the paper is structured as follows: Section 2 describes the theoretical background of the scenario generation by the ACO algorithm, with emphasis on data representation. Section 3 introduces the application problem based on real data obtained from the Savona campus. Finally, Section 4 summarizes the main conclusions, contributions and possible future lines for this new ACO approach.

2. Materials and Methods

This section describes the ACO-SG algorithm developed in this research. The purpose consists of adapting a classical ACO algorithm to find the curves that correspond with scenarios. The algorithm procedure can be divided into a sequence of steps: pre-processing, clustering and applying the modified ACO algorithm. Two separate subsections outline the modifications and updates made to the classical ACO algorithm for scenario generation, differentiating between the classical ACO and the modified ACO.

2.1. Data: Pre-Processing and Clustering

The pre-processing phase consists of identifying the incomplete data within the historical dataset and ensuring data integrity. These identifications have been useful for erasing complete daily samples rather than interpolating them. If data are interpolated, the generated scenarios may generate scenarios that are not directly derived from the original dataset. To analyze and understand the curve characteristics, the data have been categorized into weekdays, weekends and different seasons, following the typical approach in scenario generation. The division must be related to the variable features. For example, for PV production, it does not make sense to divide it into weekdays and weekends because it is not contingent upon the weekday; PV generation is independent. This contrasts with the building's consumption, which can be higher during weekdays because of the use of the installations.

To extract curve features, the use of an automatic divider has been implemented to obtain curve clusters; for this aim, the k-means algorithm has been selected [35]. The elbow method determines how many clusters divide the dataset. K-means is an unsupervised algorithm that is capable of subtracting features without the use of reference to known or labelled results. This methodology has been selected due to its ease of implementation, scalability to sizable datasets and capacity to adapt to novel examples with guaranteed convergence.

2.2. ACO Algorithm

ACO consists of a probabilistic technique modelled on the actions of an ant colony. Artificial ants traverse the search space in search of optimal solutions. Each ant's position is documented as a state, and the simulated ants record their state and the quality of their solutions using an artificial pheromone. In this way, each ant is associated with a solution of a certain quality, which will help to guide future ants to better solutions. ACO formalization is based on a combinatorial optimization problem defined by $P = (\mathbf{S}, \Omega, f)$, in which $\mathbf{S}$ is the space of the search, Ω is the set of constraints among the variables and f corresponds to the objective function.

To implement ACO, the optimization problem must be transformed into a problem of finding the most effective path. Algorithm 1 shows the complete procedure of a general ACO algorithm. It starts with the generation of solutions based on the stochasticity of the problem from a random point. The algorithm rates every solution and updates the pheromone according to their marks. The algorithm ends either when the solution reaches the quality criteria evaluated by the fitness or after a predefined number of iterations.

The possible next movements of the ants are evaluated from their current state, providing different probabilities for each of the options. The ants transit from x to y, where they consider those transitions with $A_k(x)$, which defines the possible next movements according to the probability provided by a mix of the defined heuristic and the pheromone. This probability is defined by Expression (1).

$$p_{xy}^k = \frac{(\tau_{xy}^{\alpha})(\eta_{xy}^{\beta})}{\sum_{z \in allowed_y} (\tau_{xz}^{\alpha})(\eta_{xz}^{\beta})} \tag{1}$$

where η_{xy} is the desirability of state transition xy. The term τ_{xy} represents the quantity of artificial pheromones for the transition from state x to y. Finally, the influence of η_{xy} is regulated by α, while the influence of τ_{xy} is controlled by β.

Trails are updated once all ants have completed their solution, as given by Expression (2). The amount of pheromone deposited by each ant is directly related to the fitness score, which can be used to evaluate the different solutions.

$$\tau_{xy} \leftarrow (1-\rho)\tau_{xy} + \sum_{k}^{m} \Delta\tau_{xy}^k \tag{2}$$

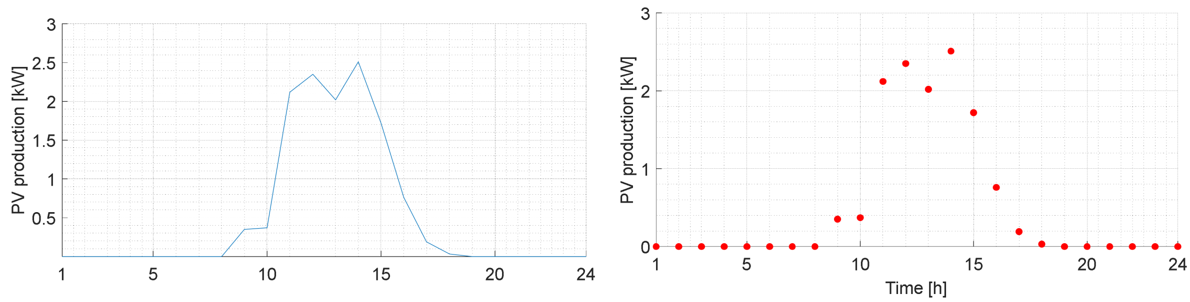
where ρ is the evaporation coefficient of the artificial pheromone, m is the number of ants and $\Delta\tau_{xy}^k$ is the amount of artificial pheromone deposited by ant k.

Algorithm 1: Pseudocode for a Classic ACO

```
While the objective is not satisfied do
    For n from 1 to ant N do
        Ant state ← Initialize randomly
        For p from 1 to point P do
            Compute heuristic
            Ant n chooses point p in the space S
        End for
      Update best solution
      End for
Update pheromone
End While
```

2.3. Problem Adaptation

The graph that ACO uses for constructing solutions can be expressed as $G_C(V, E)$, where V corresponds to the vertices set and E is the set of links between those vertices. Some modifications to the traditional ACO are necessary to produce possible curves. For ACO-SG, V represents the recorded dataset of points, while E is the viable transitions between the evaluated point at a specific time and the available set of points at the subsequent time step. Pre-processing aims to create a space for the ants by adapting the information. The historical dataset contains curves consisting of arrays of 24 values. The ants start from a starting point and traverse the data from the 1st hour to the 24th hour, excluding the starting point at zero. This ensures that the first hour adheres to the heuristic in the algorithm rather than being arbitrarily selected. Figure 2 shows a representation of the available points through which the ants can pass, based on a real curve. The possible points that the ants can select are actual points recorded in the historical dataset. The algorithm considers all the real points in the dataset as possible states for the ants to use. Consequently, this corresponds to defining the search space **S** previously defined.



(**a**) (**b**)

Figure 2. Construction of points based on historical data. Real registered curve (**a**) and its discretized representation (**b**).

The cluster curves obtained from the k-means algorithm in the pre-processing stage can be used as the basis for generating new curves. These new curves can be defined by specific shapes and a designated power for the complete curve. The mathematical operation employed for generating the new centroid is formalized in Expression (3).

$$P^s = P^c - \frac{\sum_{t=0}^{T} P_t^c - P^D}{T} \tag{3}$$

where P^c expresses the obtained cluster in the pre-processing, and T corresponds with the time window considered. The result is the new centroid P^s with the shape of the cluster P^c and the objective daily power.

The heuristic rewards points closer to the new curve P^s obtained from (3) and penalizes the furthest points. Consequently, the attractiveness η carries out this purpose, mathematically expressed in (4).

$$\eta = abs(P_t^a - P_t^s) \tag{4}$$

where P_t^a is the point under evaluation and P_t^s corresponds to the point of the cluster, evaluated according to the time t.

The fitness function minimizes the disparity between the objective power and the candidate power solution, as expressed in Equation (5).

$$f = \min\left(\sqrt{(P^D - P^S)^2}\right) \tag{5}$$

where P^D corresponds to the objective power and P^s corresponds to the candidate power solution of the scenario generated by the algorithm.

The stopping criteria are expressed by Equation (6), which limits the power error desired as a percentage. If this requirement is met, the algorithm stops, and the output is the best scenario according to the specifications. If this condition is not met and the error is higher, another iteration will be computed until the criteria is reached.

$$Error = \frac{P^D - P^S}{P^S} \times 100 \quad (6)$$

To summarize, the algorithm takes as input the historical dataset for modelling the uncertainty and the desired number of scenarios to generate. The clusters of the dataset are obtained with the elbow method, and daily power scenarios are determined in function of the maximum and minimum in the dataset and the number of scenarios to be accordingly distributed. Clusters and daily power will be used in the heuristic of the algorithm to obtain the scenarios. The ants traverse the space from a starting point to the final dimension. Error stopping criteria are applied depending on the desired daily power. From this perspective, Algorithm 2 shows the pseudocode for ACO-SG. Moreover, Figure 3 presents the flowchart of the ACO-SG to provide a comfortable overview of the complete algorithm. In accordance with the aforementioned guidelines, this methodology does not necessitate a previous statistical study to obtain the probability distribution, in contrast to traditional methodologies employed in the SG area.

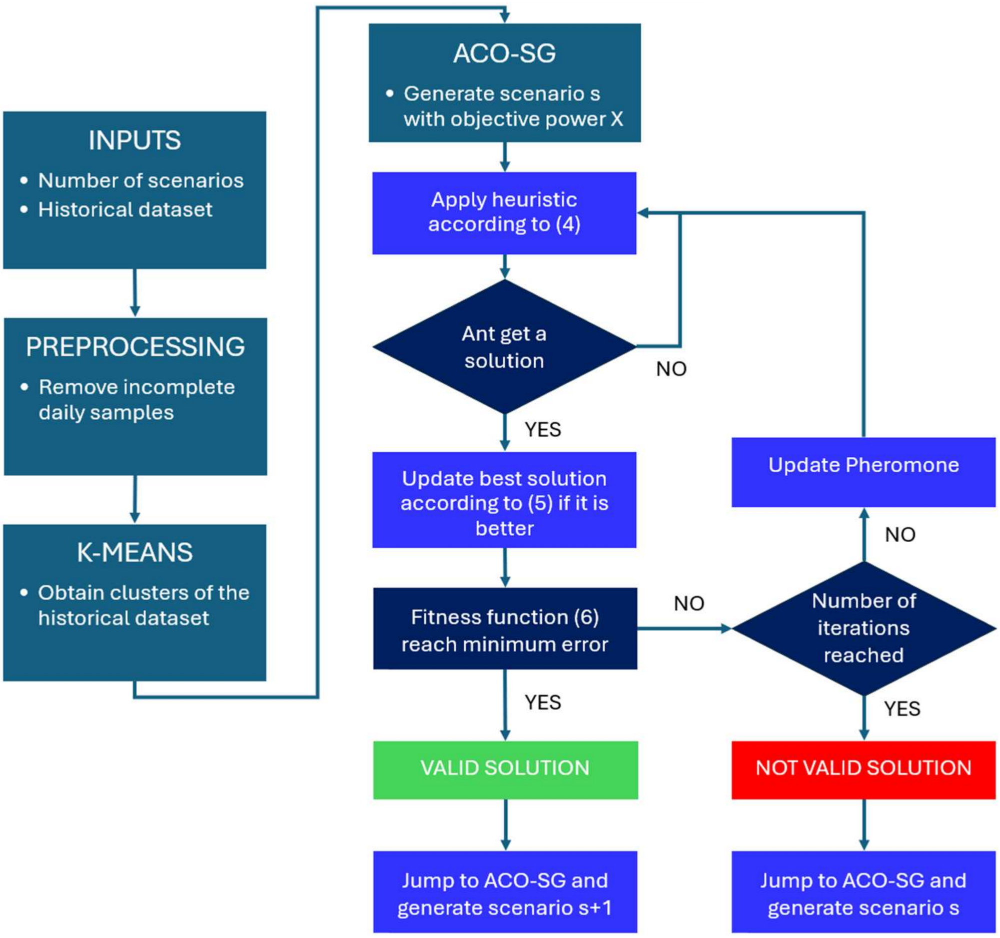


Figure 3. Flowchart of the ACO-SG algorithm.

Algorithm 2: Pseudocode for the Proposed ACO-SG

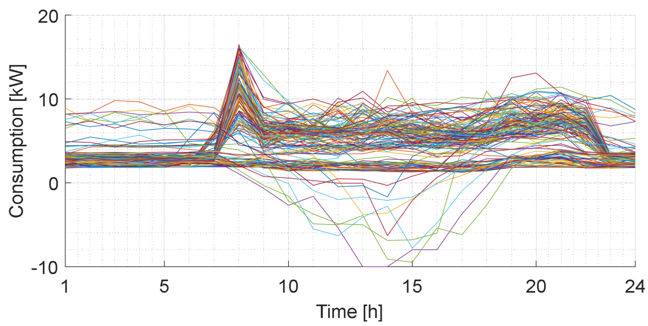
```
Clusters = KMeans_algorithm(historical_dataset)
While the objective is not satisfied do
        For n from 1 to ant N do
                Ant state ← Initialize to point zero
                For t from 1 to time T do
                        For p from 1 to point P(t) do
                                Compute heuristic
                                Ant n chooses point p at time t
                        End for
                End for
        Update best solution
        End for
Update pheromone
End While
```

3. Case Study

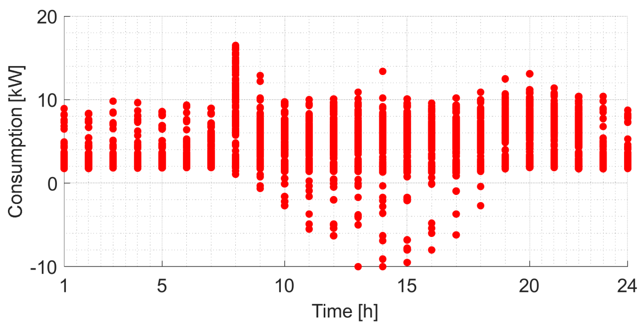
The present section presents the application of the ACO-SG algorithm, where data from the Savona campus have been used to generate both photovoltaic production and building consumption. Historical data are pre-processed, clustered, and finally, ACO-SG is applied.

3.1. Pre-Processing and Clustering of Case Study

For the case study, five years of data were available from 2019 to 2023. In the case of PV data, 1815 daily datapoints were registered, which were reduced to 1795 following pre-processing. On the other hand, the original number for building's consumption data was 1218 and was reduced to 1075 following pre-processing. The discrepancy in the number between the two variables is attributed directly to the available dataset; it affects the density of points. However, the results show a satisfactory performance for both cases. Historical data in SG are commonly divided by season and working/non-working days. For this case study, data from spring and working days were selected. PV production is not affected by whether it is a working or non-working day, so both have been included. Figures 3 and 4 show the curves.



(**a**)



(**b**)

Figure 4. (**a**) Real recorded hourly curves for building consumption; (**b**) its corresponding discretized representation obtained from real hourly curves.

The two variables under consideration in this study are independent of one another. This implies that any scenario involving one variable can be combined with any scenario involving the other variable, resulting in a unique case. All possible combinations between variables will form all possible scenarios to be studied in a hypothetical optimization problem.

After pre-processing the available data, the k-means algorithm was applied. The elbow method was used to determine the number of representative clusters based on the available data. Figures 5 and 6 show the elbow graph and the corresponding clusters according to the results of the elbow method.

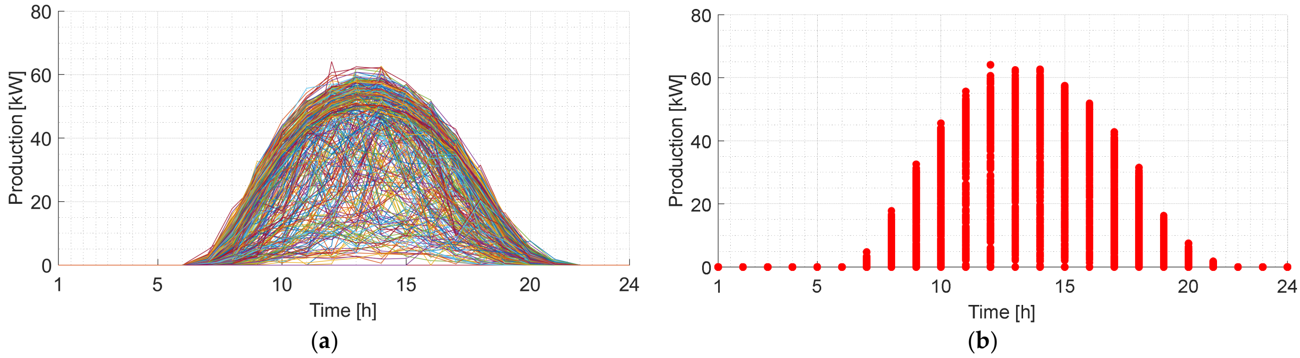


Figure 5. (**a**) Real recorded hourly curves for photovoltaic production; (**b**) its corresponding discretized representation obtained from real hourly curves.

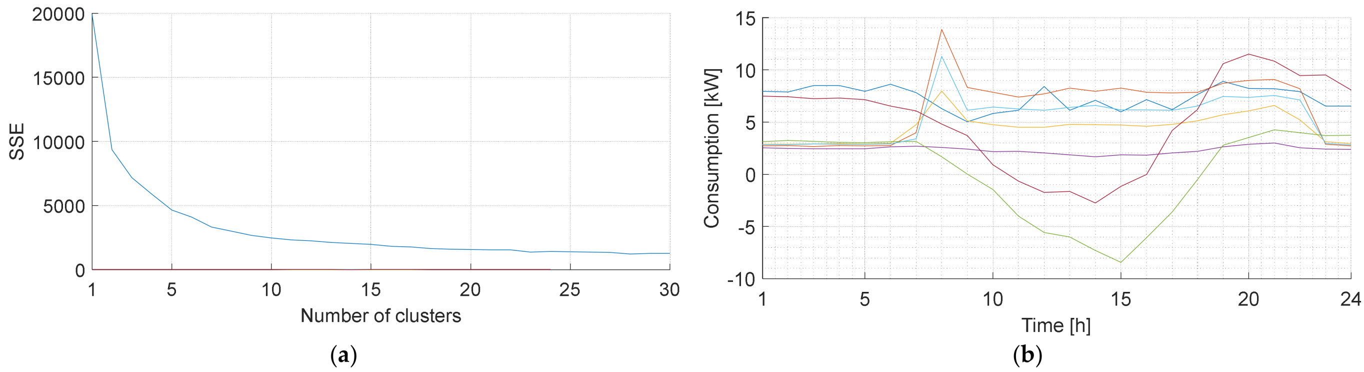


Figure 6. (**a**) Elbow graph for building consumption; (**b**) building consumption clusters.

3.2. Algorithm

The ACO-SG is capable of being configured to produce a predefined number of scenario curves. For instance, it was set to generate 30 scenarios. The scenarios to be generated were distributed within the range of maximum and minimum total consumption of the clusters.

Table 1 summarizes the configuration values that define the ACO-SG. Figures 4 and 5 show the historical datasets and their discretization. Figures 6 and 7 illustrate the clusters and the elbow method of the corresponding cases evaluated in this work. Figure 8 displays the scenarios generated by ACO-SG. A visual comparison shows that the generated scenarios are correlated with the real dataset.

Table 1. Setting values.

Variable	Value
N° of ants	100
N° of iterations	100
α	1
β	1
ρ	0.5
Error	5%

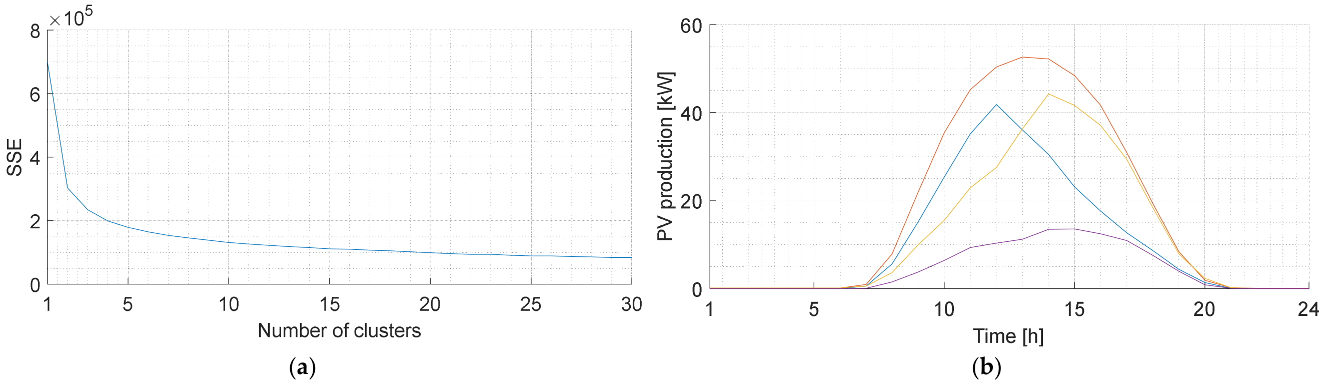


Figure 7. (**a**) Elbow graph for PV production; (**b**) building consumption hourly curve clusters.

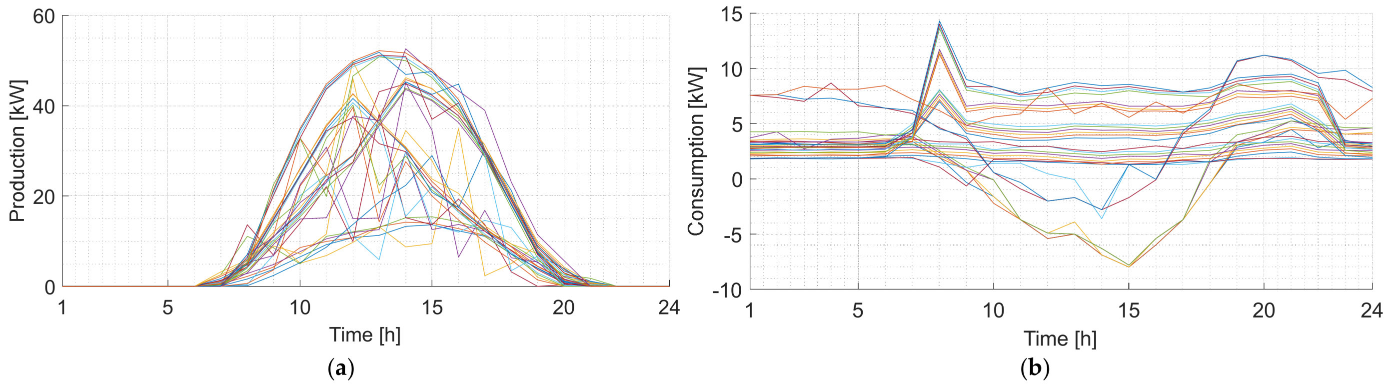


Figure 8. Resulting scenarios: (**a**) PV generation; (**b**) building consumption.

However, an analytic analysis must be carried out to guarantee the new curves keep the correlation with the original dataset.

The basis for comparison is the closest Euclidean distance between the scenarios and their real sample. This results in a list of distances, one per scenario created, which can be used to check the correlation of several scenarios with the real dataset, which allows us to check how truthful several generated scenarios are. On the one hand, for the case of PV generation, the minimum distance is 14.47 and the maximum is 76.24 from the closest registered curve in the dataset. On the other hand, in the case of building consumption, the minimum distance is 2.29 and the maximum is 18.50 from the closest registered curve in the dataset. The aforementioned distances are relatively modest when 24 dimensions are being considered. The discrepancy of the results for both cases can be explained by the particular case of the PV curve shapes. In this case, the curves' extremes remain at zero, while the heuristic relocates the reference curve. This phenomenon results in a decompensation, whereby the generated curve differs from the original shape.

Notwithstanding, the discrepancy is insignificant, and the set of new scenarios exhibits a high degree of alignment with the actual data. Additionally, a 5% margin of error has been established to guarantee that none of the curves exhibit a bias of more than 5% of the total required power. It is important to note that the algorithm has demonstrated its scope in producing precise curves for overall power, with 1.77% and 0.144% for the cases of PV generation and building consumption, respectively. The higher level of total power error in the PV curves can again be explained by the special case itself. This also impacts the intermediate part of the curve, which must now compensate for the shift caused by the extremes.

The low error percentage indicates a precise objective of the generated scenarios; such an error is significantly lower than the established limit of 5%. Additionally, the low Euclidean distance demonstrates the correlation between different curves for both cases, which confirms the validity of the scenarios.

4. Conclusions

The presence of uncertain variables presents a challenge for optimizing power system problems, as the uncertainty directly conditioning the solution's accuracy. In response, a variety of techniques have been developed for modelling variables that are difficult to predict, such as renewable energies or consumption. Consequently, the effective modelling of stochastic variables can contribute to the maximization of energy exploitation in the system under evaluation and the provision of an effective response in advance.

The paper proposes an ACO-based SG method to find the set of scenarios for a two-stage stochastic optimization problem. The focus of the work is on the development of ACO applied to SG, which has great potential for generating new sets of scenarios based on the input dataset. Moreover, this approach does not require a scenario reduction process as it generates the predefined number of scenarios directly. In the case study, this approach has been applied to both building consumption and PV generation, which have completely different hourly curves to prove the model for different cases.

To verify whether the generated scenarios replicate the seasonal patterns presented in the real dataset, both are compared using the Euclidean distance. This indicates the correlation between several curves and demonstrates the adaptability of this approach to any other variable with any characteristic of the hourly curve shape. Real data from the Smart Savona Campus of the University of Genoa have been used for the case study.

Several avenues for future research can be explored based on this novel approach for generating scenarios. For example, one possible area of research is to expand the scope of the search to the continuous space or explore other encoding techniques. Another potential area of research could consist of including physical factors, such as weather conditions, in scenario generation.

Author Contributions: Conceptualization, D.F.V. and J.I.G.A.; methodology, D.F.V.; software, D.F.V.; validation, D.F.V.; formal analysis, D.F.V. and J.I.G.A.; investigation, D.F.V. and J.I.G.A.; resources, M.R.; data curation, D.F.V.; writing—original draft preparation, D.F.V.; writing—review and editing, D.F.V., J.I.G.A., C.L.d.M. and M.R.; visualization, D.F.V.; supervision, J.I.G.A., C.L.d.M. and M.R.; project administration, C.L.d.M. and M.R.; funding acquisition, C.L.d.M. and M.R. All authors have read and agreed to the published version of the manuscript.

Funding: This research was partially funded by the "Network 4 Energy Sustainable Transition—NEST" project (MUR project code PE000021, Concession Decree No. 1561 of 11 October 2022), in the framework of the NextGenerationEu NRRP plan (CUP D33C22001330002). And the paper is part of the TED2021-129702B-I00 project, financed by MCIN/AEI/10.13039/501100011033 and by the European Union "NextGeneration EU"/PRTR.

Data Availability Statement: The original contributions presented in the study are included in the article, further inquiries can be directed to the corresponding author.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. European-Comission. A European Green Deal. January 2020. Available online: https://commission.europa.eu/strategy-and-policy/priorities-2019-2024/european-green-deal_en (accessed on 24 September 2024).
2. Hewitt, M.; Ortmann, J.; Rei, W. Decision-based scenario clustering for decision-making under uncertainty. *Ann. Oper. Res.* **2021**, *315*, 747–771. [CrossRef]
3. Bounitsis, G.L.; Papageorgiou, L.G.; Charitopoulos, V.M. Data-driven scenario generation for two-stage stochastic programming. *Chem. Eng. Res. Des.* **2022**, *187*, 206–224. [CrossRef]
4. Di Somma, M.; Buonanno, A.; Caliano, M.; Graditi, G.; Piazza, G.; Bracco, S.; Delfino, F. Stochastic Operation Optimization of the Smart Savona Campus as an Integrated Local Energy Community Considering Energy Costs and Carbon Emissions. *Energies* **2022**, *15*, 8418. [CrossRef]
5. Bracco, S.; Delfino, F.; Laiolo, P.; Morini, A. Planning & Open-Air Demonstrating Smart City Sustainable Districts. *Sustainability* **2018**, *10*, 4636. [CrossRef]
6. University of Genoa Energia 2020 Project. Available online: http://www.energia2020.unige.it/en/home/ (accessed on 15 September 2024).

7. Li, J.; Zhou, J.; Chen, B. Review of wind power scenario generation methods for optimal operation of renewable energy systems. *Appl. Energy* **2020**, *280*, 115992. [CrossRef]
8. Bilici, S.; Kamislioglu, M.; Guclu, E.E.A. A Monte Carlo simulation study on the evaluation of radiation protection properties of spectacle lens materials. *Eur. Phys. J. Plus* **2023**, *138*, 80. [CrossRef]
9. Bourcet, J.; Kubilay, A.; Derome, D.; Carmeliet, J. Representative meteorological data for long-term wind-driven rain obtained from Latin Hypercube Sampling—Application to impact analysis of climate change. *Build. Environ.* **2023**, *228*, 109875. [CrossRef]
10. Jun, H.; Wei, J.; Wenjie, P.; Haoyuan, C.; Jia, Z.; Chao, C.; Zhenjian, X.; Jian, D.; Na, W. A Data-driven Distribution System Scenario Generation Method with Probabilistic Assessment of PV Station Generation. In Proceedings of the 2020 12th IEEE PES Asia-Pacific Power and Energy Engineering Conference (APPEEC), Nanjing, China, 20–23 September 2020. [CrossRef]
11. Singh, B.; Pozo, D. A Guide to Solar Power Forecasting using ARMA Models. In Proceedings of the 2019 IEEE PES Innovative Smart Grid Technologies Europe (ISGT-Europe), Bucharest, Romania, 29 September–2 October 2019. [CrossRef]
12. Chen, J.; Zhao, J. Synthetic Wind Speed Scenarios Generation using Artificial Neural Networks for Probabilistic Analysis of Hybrid Energy Systems. In Proceedings of the 2021 IEEE 30th International Symposium on Industrial Electronics (ISIE), Kyoto, Japan, 20–23 June 2021. [CrossRef]
13. Dong, W.; Chen, X.; Yang, Q. Data-driven scenario generation of renewable energy production based on controllable generative adversarial networks with interpretability. *Appl. Energy* **2022**, *308*, 118387. [CrossRef]
14. Ji, X.; Li, C.; Xie, B.; Wang, Y.; Wang, Q. A Wind Power Scenario Simulation Method Considering Trend and Randomness. In *The Proceedings of the 16th Annual Conference of China Electrotechnical Society*; Lecture Notes in Electrical Engineering; Springer Nature: Singapore, 2022; pp. 1043–1050. [CrossRef]
15. Bode, A.; Lamasigi, Z.Y.; Drajana, I.C.R. The K-Nearest Neighbor Algorithm using Forward Selection and Backward Elimination in Predicting the Student's Satisfaction Level of University Ichsan Gorontalo toward Online Lectures during the COVID-19 Pandemic. *Ilk. J. Ilm.* **2023**, *15*, 118–123. [CrossRef]
16. Gruffaz, S.; Kim, K.; Durmus, A.O.; Gardner, J.R. Stochastic Approximation with Biased MCMC for Expectation Maximization. *arXiv* **2024**. [CrossRef]
17. Cai, D.; Shi, D.; Chen, J. Probabilistic load flow computation with polynomial normal transformation and Latin hypercube sampling. *IET Gener. Transm. Distrib.* **2013**, *7*, 474–482. [CrossRef]
18. Goyal, M.K.; Ojha, C.S.P. Evaluation of Rule and Decision Tree Induction Algorithms for Generating Climate Change Scenarios for Temperature and Pan Evaporation on a Lake Basin. *J. Hydrol. Eng.* **2014**, *19*, 828–835. [CrossRef]
19. Yadav, K.; Kumar, B.; Guerrero, J.M.; Lashab, A. A hybrid genetic algorithm and grey wolf optimizer technique for faster global peak detection in PV system under partial shading. *Comput. Inform. Syst.* **2022**, *35*, 100770. [CrossRef]
20. Ruan, Y.; Qian, F.; Sun, K.; Meng, H. Optimization on building combined cooling, heating, and power system considering load uncertainty based on scenario generation method and two-stage stochastic programming. *Sustain. Cities Soc.* **2023**, *89*, 104331. [CrossRef]
21. Oliveira, B.B.; Carravilla, M.A.; Oliveira, J.F. A diversity-based genetic algorithm for scenario generation. *Eur. J. Oper. Res.* **2022**, *299*, 1128–1141. [CrossRef]
22. Khezzane, K.; Doumbia, M.L.; Khoucha, F. Multi-objective Sizing of Hybrid Generation Energy System for Remote Area Using Genetic Algorithm. In Proceedings of the 2021 Sixteenth International Conference on Ecological Vehicles and Renewable Energies (EVER), Monte-Carlo, Monaco, 5–7 May 2021. [CrossRef]
23. Liu, B.; Zhang, Q.; Fernandez, F.V.; Gielen, G.G.E. An Efficient Evolutionary Algorithm for Chance-Constrained Bi-Objective Stochastic Optimization. *IEEE Trans. Evol. Comput.* **2013**, *17*, 786–796. [CrossRef]
24. Eltamaly, A.M. Musical chairs algorithm for parameters estimation of PV cells. *Sol. Energy* **2022**, *241*, 601–620. [CrossRef]
25. Wang, Y.; Han, Z. Ant colony optimization for traveling salesman problem based on parameters optimization. *Appl. Soft Comput.* **2021**, *107*, 107439. [CrossRef]
26. Godha, N.R.; Bapat, V.N.; Korachagaon, I. Ant colony optimization technique for integrating renewable DG in distribution system with techno-economic objectives. *Evol. Syst.* **2022**, *13*, 485–498. [CrossRef]
27. Dong, B.; Luzin, A.; Gura, D. The hybrid method based on ant colony optimization algorithm in multiple factor analysis of the environmental impact of solar cell technologies. *Math. Biosci. Eng.* **2020**, *17*, 6342–6354. [CrossRef]
28. Aghelpour, P.; Graf, R.; Tomaszewski, E. Coupling ANFIS with ant colony optimization (ACO) algorithm for 1-, 2-, and 3-days ahead forecasting of daily streamflow, a case study in Poland. *Environ. Sci. Pollut. Res.* **2023**, *30*, 56440–56463. [CrossRef] [PubMed]
29. Mirza, N.; Cinel, J.; Noyes, H.; McKenzie, W.; Burgess, K.; Blackstock, S.; Sanderson, D. Simulated patient scenario development: A methodological review of validity and reliability reporting. *Nurse Educ. Today* **2020**, *85*, 104222. [CrossRef]
30. Kaut, M. Scenario generation by selection from historical data. *Comput. Manag. Sci.* **2021**, *18*, 411–429. [CrossRef]
31. Buonanno, A.; Caliano, M.; Somma, M.D.; Graditi, G.; Valenti, M. A Comprehensive Tool for Scenario Generation of Solar Irradiance Profiles. *Energies* **2022**, *15*, 8830. [CrossRef]
32. Memari, M.; Karimi, A.; Hashemi-Dezaki, H. Reliability evaluation of smart grid using various classic and metaheuristic clustering algorithms considering system uncertainties. *Int. Trans. Electr. Energy Syst.* **2021**, *31*, e12902. [CrossRef]
33. Cui, M.; Ke, D.; Sun, Y.; Gan, D.; Zhang, J.; Hodge, B.-M. Wind Power Ramp Event Forecasting Using a Stochastic Scenario Generation Method. *IEEE Trans. Sustain. Energy* **2015**, *6*, 422–433. [CrossRef]

34. Islam, T.; Pruyt, E. Scenario generation using adaptive sampling: The case of resource scarcity. *Environ. Model. Softw.* **2016**, *79*, 285–299. [CrossRef]
35. Bahri, R.S.; Sudirman, I.D.; Utama, I.D.; Susanto, R.H. Data Mining Techniques To Uncovering Customer Segments: K-Means Clustering Using The Elbow Method Approach In Medium-Scale Grocery. In Proceedings of the 2023 International Conference on Computer Science, Information Technology and Engineering (ICCoSITE), Jakarta, Indonesia, 16 February 2023. [CrossRef]