Energy

AIMS Energy, 6(1): 70–96.
DOI: 10.3934/energy.2018.1.70
Received: 20 November 2017
Accepted: 02 January 2018
Published: 08 January 2018

http://www.aimspress.com/journal/energy

Research article

Optimal sizing and operation of energy storage systems considering long term assessment

Gerardo Guerra and Juan A. Martinez-Velasco*

Departament d'Enginyeria Electrica, Universitat Politecnica de Catalunya, Barcelona, Spain

*** Correspondence:** Email: martinez@ee.upc.edu; Tel: +34934016725.

Abstract: This paper proposes a procedure for estimating the optimal sizing of Photovoltaic Generators and Energy Storage units when they are operated from the utility's perspective. The goal is to explore the potential improvement on the overall operating conditions of the distribution system to which the Generators and Storage units will be connected. Optimization is conducted by means of a General Parallel Genetic Algorithm that seeks to maximize the technical benefits for the distribution system. The paper proposes an operation strategy for Energy Storage units based on the daily variation of load and generation; the operation strategy is optimized for an evaluation period of one year using hourly power curves. The construction of the yearly Storage operation curve results in a high-dimension optimization problem; as a result, different day-classification methods are applied in order to reduce the dimension of the optimization. Results show that the proposed approach is capable of producing significant improvements in system operating conditions and that the best performance is obtained when the day-classification is based on the similarity among daily power curves.

Keywords: clustering technique; energy storage system; genetic algorithm; optimization; parallel computing; photovoltaic generation

1. Introduction

Distributed Generation (DG) and Energy Storage (ES) are two technologies that fall under the concept of Distributed Energy Resources (DER) [1]. Both technologies can support voltage, reduce losses, provide backup power and ancillary services, improve local power quality and reliability, and defer distribution system upgrade [2–5]. Furthermore, the combined operation of DG and ES can enhance their potential benefits to the grid; for example, ES units can be charged when there is a generation excess, and provide support to the grid when generation decreases [6].

Modelling of renewable generation raises several challenges to distribution load flow calculations since capabilities for representing intermittent generators, voltage-control equipment, or multi-phase unbalanced systems are required. In addition, the study of systems with intermittent non-dispatchable resources will usually require a probabilistic approach and calculations performed over an arbitrary time period that may range from minutes to years. The evaluation of ES presents additional challenges: ES operation can be conducted from the perspective of a utility (i.e., aiming to improve network conditions) or operated by a private operator (i.e., with the objective of generating profits) [7]. In both cases an optimal operation strategy must be defined in order to maximize the benefits provided by ES; see [8–18].

This paper uses a General Parallel Genetic Algorithm (GPGA) to estimate the optimal ratings of multiple fixed-location ES units and photovoltaic (PV) generators in order to minimize distribution system energy losses, annual energy and peak power supplied by the substation transformer. The goal is to explore the potential improvement on the overall operating conditions of the distribution system when ES units and PV generators are operated from the perspective of a utility. The developed methodology should be adequate for performing a technical evaluation of DG and ES as an alternative within the utility's planning scheme.

The optimal sizing of ES units has been conducted from different perspectives; e.g., improving network conditions [19,20], minimizing power and operation costs [21–23], meeting a predefined operation schedule [24–26], and multi-objective optimization [27,28]. Furthermore, various optimization techniques, such as particle swarm optimization [20,23], linear, dynamic, and cone programming [22,24,27], stochastic optimization [21,25], and genetic algorithms [26,28] have been employed.

The optimization of a combination of generation and energy storage has also been the subject of several recent works; see [29–33]. In [29,30], and [31] the optimization is conducted to minimize operation costs in a microgrid, while [32] presents a techno-economical procedure for sizing DG and ES in smart households. Reference [33] explores the optimization of pumped hydro storage that is operated in conjunction with wind generation to minimize operation costs or maximize generation and storage penetration. As with optimal sizing of ES, different optimization algorithms have been used; e.g., differential evolutionary algorithms [30,31], mix-integer linear programming [32], and genetic algorithms [33]. Furthermore, references [22,26,27] make use of clustering techniques in order to reduce the number of scenarios considered for optimization.

A strategy for operating ES units based on the variation of load and photovoltaic generation over the evaluation period (i.e., one year) is proposed here. This strategy is simultaneously optimized along with the ES units and PV generators ratings and it assumes that the operation curve is the same for all ES units while the days of the evaluated period are grouped according to different classification methods in order to reduce the dimension of the optimization problem. ES operation curves are validated through the simulation of the system under study for the selected evaluation period; only those curves that produce grid compliant solutions are considered as valid operation strategies.

The simulation tools used for this purpose are MATLAB and OpenDSS, a distribution system simulator whose capabilities allow users to represent the most important distribution components and perform multiphase calculations. This environment takes advantage of OpenDSS capabilities, which can be expanded with an external link to other tools, in this case MATLAB [34,35]. The use of OpenDSS as a simulation tool can provide an accurate estimation of the network conditions and avoid the use of simplification techniques (such as DC power flow or linearization of the power flow equations).

System behavior over the evaluation period is defined by node load and PV generation yearly curves which have been synthetically generated with different applications previously developed by the authors; see [36]. New applications based on clustering techniques have been implemented in this work for classifying and grouping the days of the evaluation period. This work was carried out using a modified version of the built-in OpenDSS storage model [37].

The paper has been organized as follows. The main characteristics of the ES model are summarized in Section 2. The procedure followed to obtain the operation curves of ES units using different day-classification methods is presented in Section 3. Section 4 details the GPGA procedure implemented for this work. The application to an introductory case study is presented in Section 5, while the optimization of multiple ES units and PV generators is analyzed in Section 6. A discussion regarding the procedure's main points and future work is presented in Section 7. Finally, the main conclusions drawn from this study are summarized in Section 8.

2. Modeling of energy storage units

OpenDSS provides a technology-independent ES model that allows users to perform planning studies under different operation modes [37]: Peak Shaving, Follow, Support, Load Level, and Time. Additionally, different charge and discharge strategies can be combined in order to produce the desired effect. The use of this ES model allows the user to set limits for charge/discharge power and minimum/maximum stored energy; moreover, it automatically calculates ES losses and updates the stored energy value.

The OpenDSS ES model can be used in the different solution modes available. Furthermore, the StorageController object allows the user to apply the same control strategy to a fleet of ES units.

To adequate the OpenDSS storage model to the optimization process proposed in this work, new features have been added to its native capabilities:

- ES efficiency is not assumed constant but defined by a function that depends on ES loading with respect to its rated power (in kVA).
- A minimum charge/discharge power as a percentage of the ES rated power (in kVA) has been defined. Regardless of the dispatched power, the ES unit will not absorb/inject power unless this value is greater than the minimum charge/discharge power.
- The model defines a maximum charge/discharge power based on the stored energy and simulation time-step. The objective is aimed at ensuring that the minimum/maximum stored energy values are not violated. For example, assume the present stored energy is 100 kWh and the minimum acceptable value is 80 kWh; given a 1-hour time-step and a discharge efficiency of 100%, the ES unit should only be able to inject a power of 20 kW for a period of 1-hour. Regardless of the predefined dispatched power, the model will enforce the 20 kW limit.

3. Operation of energy storage units

3.1. Determination of energy storage operation curves

The operation of multiple ES units in a power distribution system can be optimized by establishing coordination among all units; that is, it is advisable for all ES units to follow the same control strategy and dispatch in order to ensure total cooperation since individual operation strategies based on local measurements may not accurately reflect the global system requirements [34,37].

An ES dispatch strategy based on the principles of the Peak Shaving and Load Level operation modes has been developed for this work; it is aimed at minimizing daily power peaks and flattening the active power profile served by the substation transformer. This work assumes that the same operation curve is used for all ES units and it is derived from the power curve that would be measured at the substation terminals with DG but without ES.

Two limits are defined for charge and discharge operations: ES discharge can take place if the served power is above the discharge limit, whereas ES charge will occur if the served power is below the charge limit. In both cases the behavior will mirror the daily power curve pattern (see Figure 1). Additionally, charge and discharge curves will be modified by applying pertinent correction factors.

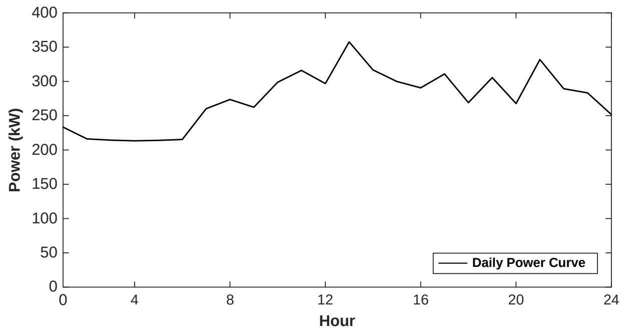
The operation curve of ES units for a specific day is built as follows:

(1) Obtain the daily power curve served by the substation transformer, including PV generation (see Figure 1a).
(2) Calculate the mean value of this daily power curve.
(3) Subtract the mean value from the daily power curve (see Figure 1b).
(4) Set power limits for ES charge/discharge (see Figure 1b).
(5) Find the power values below and above the charge/discharge limits.
(6) Subtract average charge/discharge power limit from the power values found in the previous step (see Figure 1c).
(7) Apply correction factors to the new power values found in step 6.
(8) The corrected power curve is normalized using the total rated ES power connected to the system. Negative values will correspond to ES charge, while ES discharge will be performed following the positive values in the normalized curve (see Figure 1d).

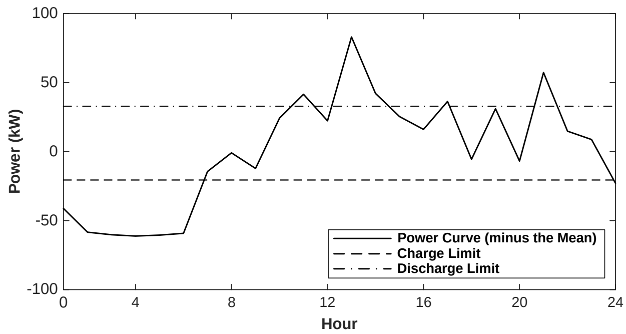
With this approach four parameters are required to determine the operation curve to be followed by all ES units: charge power limit, discharge power limit, charge curve correction factor, and discharge curve correction factor. Charge and discharge power limits are defined as a multiplier of the daily power's standard deviation; for the example presented in Figure 1 charge and discharge limits were set as –0.5 and 0.8 times the standard deviation, respectively.

3.2. Classification of days using hierarchical clustering techniques

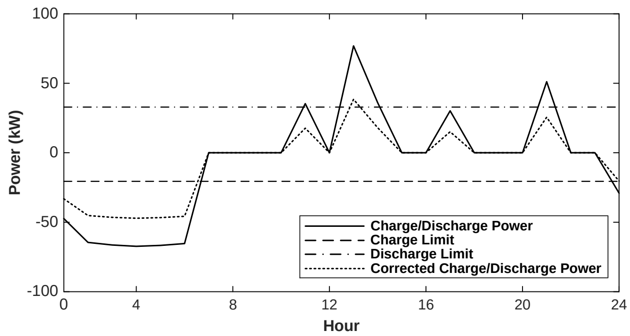
The operation of ES units is specified in this study for an evaluation period of one year. Since using only one set of parameters for constructing the yearly operation curve is not enough to guarantee an optimal operation and defining a set of parameters for every day of the year leads to a very high number of parameters, different methods for classifying days have been implemented.



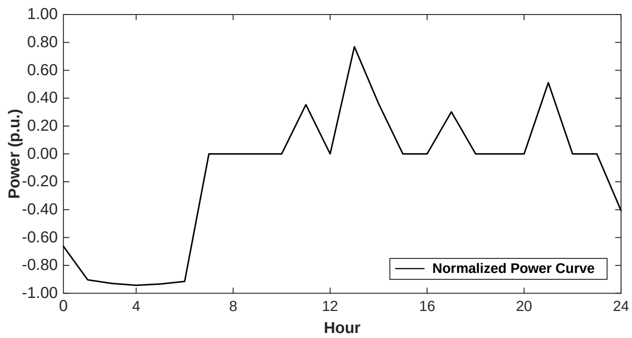
(a) Substation daily power curve—With PV generation.



(b) Substation power curve after subtracting mean.



(c) ES charge/discharge power curve.



(d) Normalized corrected ES charge/discharge power curve.

Figure 1. Proposed ES dispatch mode.

A hierarchical clustering algorithm groups a collection of individuals by forming a cluster tree [38]. The cluster tree does not separate the individuals into clusters but rather shows the connections among them at several hierarchical levels.

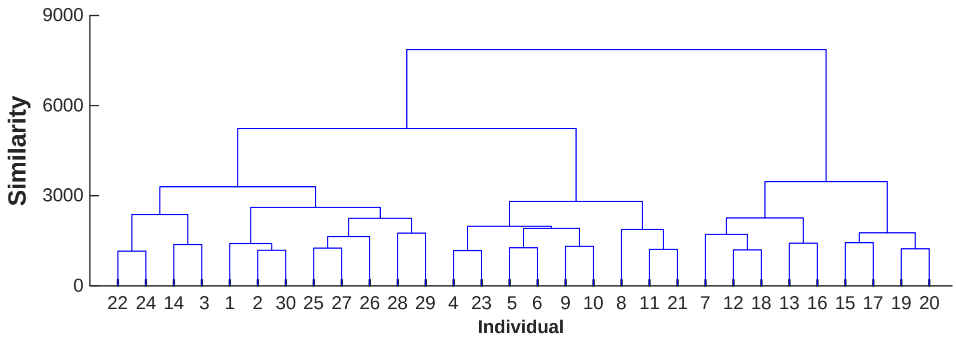
Hierarchical clustering is conducted using the following steps:

(1) Define a similarity measurement (e.g., Euclidean distance).

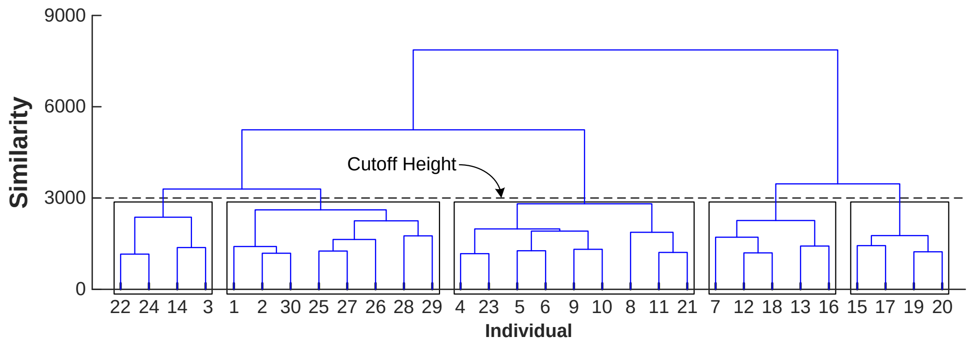
(2) Calculate similarity for every pair of individuals.

(3) Build the cluster tree.

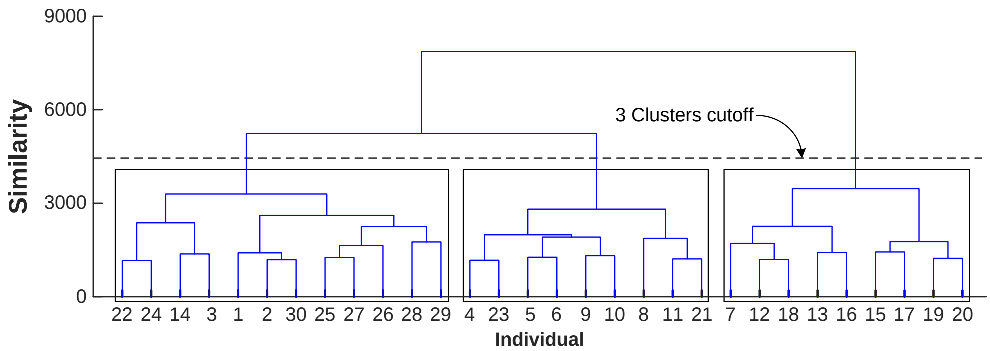
(4) Form clusters by cutting the cluster tree. The number of clusters can be defined as a fixed parameter or as a height h at which the cluster tree is cut.



(a) Cluster tree.



(b) Cluster generation using tree height as criterion.



(c) Cluster generation using fixed number of clusters as criterion.

Figure 2. Different approaches for generating clusters.

The similarity for every pair of individuals is expressed as a row vector of length $N \times (N-1)/2$, where N is the number of individuals (points) in the collection. In this manner, the duplication of similarities is avoided. The cluster tree is built using the similarity vector as input. The algorithm links similar individuals into binary clusters (i.e., with two elements); moreover, these new binary clusters are linked to other binary clusters and individuals. The result is a multilevel tree, where all individuals are connected among them and also to the new binary clusters, see Figure 2a. Once the cluster tree has been built, it must be cut in order to group individuals into clusters. Depending on how the tree is separated, different clusters will be generated. In Figure 2b the cluster tree is cut at a height of 3000 producing 5 clusters, while in Figure 2c the total number of clusters has been previously chosen and is equal to 3.

Classification methods can be used in different applications. Reference [39] presented a two-stage pattern recognition of load curves, in which the first stage requires grouping daily load curves into clusters, using different clustering algorithms, and determining the optimal number of typical days (i.e., optimal number of clusters); clusters are constructed based on the Euclidean distance among curves. In [40] different similarity indices (including Dynamic Time Warping, DTW) were used as measurements for time series clustering in order to discover building energy patterns. Classification methods based on field indices (e.g., ratio of average to maximum daily load) are also used for categorizing the behavior of distribution customers [41].

Three different methods aimed at classifying the days of the evaluation period have been implemented in MATLAB using the Statistics and Machine Learning Toolbox.

(1) Daily served energy and PV generation: In this scenario the information related to the energy daily served by the substation transformer (without PV generation) and energy daily produced by the PV generator is gathered for every day of the year. Next, all days are assigned to a quartile according to the daily served and daily produced energy values. Each quantity (served and produced energy) produces four sub-groups; the combination of these sub-groups creates 16 groups. Finally, each day of the year is placed in one of the 16 groups based on the sub-groups to which they were originally assigned.

(2) Time-series clustering: This method is based on the similarity among daily power curves (including PV generation). Similarity among daily curves is calculated according to the Euclidean distance between two different curves. However, since a simple Euclidean distance may yield an incorrect evaluation of similarity; the minimum distance between two daily curves is calculated using the DTW algorithm [42], which grants a better assessment of similarity among the evaluated curves. After obtaining the distances among all daily curves, the day classification is conducted using MATLAB's *linkage* and *cluster* functions; the distances calculated with the DTW algorithm are used as inputs for the hierarchical clustering algorithm. The number of groups or clusters can be defined as a fixed parameter or as a height h at which the cluster tree is cut. In this work the number of clusters is defined as a fixed parameter.

    Taking into account that the total number of operation parameters depends on the number of clusters, but a too low number of clusters may lead to a poor ES performance, the optimal number of clusters is determined by means of the so-called CH index as follows [43]:

$$CH(K) = \frac{B(K)}{W(K)} \cdot \frac{N-K}{K-1} \tag{1}$$

$$W(K) = \sum_{k=1}^{K} \sum_{i=1}^{N_k} \|X_i - \overline{X_k}\|^2 \quad (2)$$

$$B(K) = \sum_{k=1}^{K} N_k \|\overline{X_k} - \bar{X}\|^2 \quad (3)$$

where K is the total number of clusters, N_k is the number of points that belong to cluster k, $\overline{X_k}$ is the average of points in cluster k, $\bar{X}$ is the overall average, and $\| \quad \|^2$ represents the square of the Euclidean distance between two points. In order to make use of this index, hierarchical clustering must be evaluated for different numbers of clusters; afterwards the CH index is calculated for every evaluation. The optimal number of clusters will be that with the largest CH index.

(3) Daily-values clustering: For this scenario four measurements are calculated: the energy daily served by the substation transformer and the standard deviation of the daily power curve, with and without PV generation. Each day of the year is then defined by these four measurements and hierarchical clustering is performed based on the Euclidean distance among days. As in the Time-series clustering, the optimal number of clusters can be determined by means of the CH index.

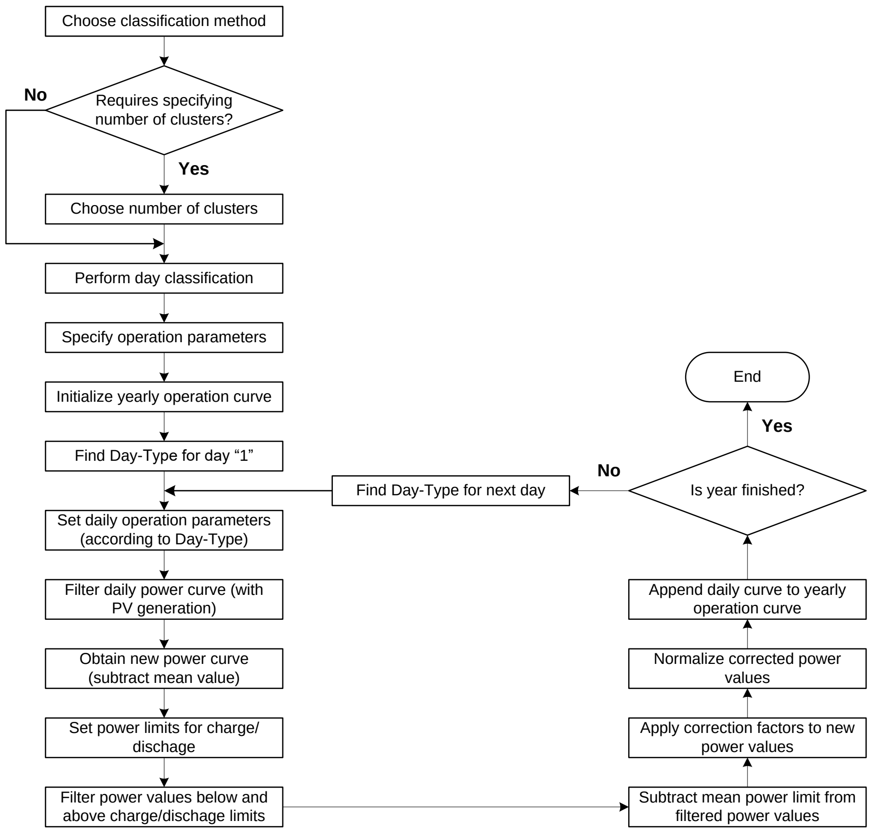


Figure 3. Flow chart-construction of ES operation curves.

The three methods rely on different system quantities: the first method uses a fixed number of groups, whereas the other two can be executed considering different numbers of classification groups. The Time-series approach was implemented in a similar manner to the approaches presented in [39] and [40], while the other two approaches could be seen as classification methods based on field indices. With these classifications, all days belonging to the same group will share the same operation parameters so the total number of required parameters for defining ES operation will consequently be reduced. To obtain ES operation curves during the evaluation period (i.e., one year), the day-classification is first conducted (according to the selected method); next, a set of operation parameters is specified for every classification group: each day is first defined by its day-type (i.e., the group/cluster number to which it belongs) and then its daily operation curve is constructed according to the procedure presented above. Figure 3 presents a flow chart with the complete process implemented to obtain the operation curve of each ES unit.

4. General parallel genetic algorithm

A genetic algorithm (GA) is an optimization technique that evaluates more than one area of the search space and can discover more than one solution to a problem [44]. The application of a GA starts with the evaluation of a randomly-generated initial population. After the first evaluation, the fittest individuals are selected to become parents for a new generation, where different methods can be applied for the parent-selection process (e.g., fitness proportionate selection, tournament selection, truncation selection). The crossover genetic operator is used to mate the selected parents in order to create the members of a new generation; finally, the mutation operator introduces a random variation to the members of the new generation.

A complete GA procedure must include the following characteristics [45]: (i) a criterion to create the initial population; (ii) an evaluation function to obtain the fitness of all possible solutions; (iii) a genetic operator to obtain the next generation; (iv) a stopping criterion.

GAs are prone to parallelization, depending on the approach used to evaluate the fitness function and how the genetic operators are applied. Several types of General Parallel Genetic Algorithms (GPGA) have been proposed: global single-population master-slave GAs, single-population fine-grained, multiple-population coarse-grained GAs [46]. In global single-population GAs only one population is created and the evaluation of each member of a given generation is independent from the rest of the members; therefore, the evaluation of a generation can be seen as an embarrassingly parallel problem. This approach yields a reduction of the total execution times by distributing the evaluation of individuals in a generation among available cores in a multi-core installation.

GAs have been applied to a wide range of optimization problems related power systems (e.g., [26,28,33,46–48]).

A global single-population GPGA has been implemented in this work to determine the optimal values of the following inputs:

- Nominal PV generator power rating (in kVA).
- Nominal ES unit power rating (in kVA).
- Nominal ES unit energy rating (in kWh).
- Security factor for interconnection transformer (in pu).
- ES unit operation strategy parameters.

The number of inputs in the GPGA will be equal to:

$$N_{IN} = 4 \times N_U + 4 \times K \tag{4}$$

where N_{IN} is the number of inputs to be optimized by the GPGA, K is the number of day groups (or clusters), and N_U is the number of PV generators and ES units under consideration.

The parameters used in the optimization procedure are shown in Table 1.

The GPGA seeks to minimize the following operational variables:

- distribution system energy losses, without including losses in substation transformer;
- annual energy supplied from the MV side of the substation transformer;
- peak power supplied by the substation transformer and measured at MV-side terminals;
- standard deviation of the power values measured at the MV-side of the substation.

Table 1. GPGA parameters.

Maximum number of generations	200
Population size	$10 \times N_{IN}$
Epsilon (ε)-Tolerance	0.0001
Delta (δ)-Number of generations for best fitness checking	20
Crossover probability (p_c)	0.80
Mutation probability (p_m)	0.02
Elitism	Yes
Maximum rating PV generation[1]	1.00
Minimum rating PV generation[1]	0.01
Maximum power rating ES unit[1]	1.00
Minimum power rating ES unit[1]	0.01
Maximum energy to power ratio[2]	10.0
Minimum energy to power ratio[2]	1.00
Maximum security factor interconnection transformer[3]	2.00
Minimum security factor interconnection transformer[3]	1.01
Maximum correction factor	2.00
Minimum correction factor	0.01
Maximum power limit factor[4,5]	2.00
Minimum power limit factor[4,5]	–1.00

Note: [1] It is a factor of the maximum generation power that can be connected to the node where the device is located; [2] ES unit rated energy is equal to the energy to power ratio multiplied by the ES unit rated power; [3] It is a factor of the maximum value when comparing the PV generator and ES unit rated power; [4] It is a factor of the daily power curve's standard deviation, including PV generation; [5] A negative sign is assumed for the set power limit during ES charge.

Two stopping criteria have been established:

(1) The *generation count* (i.e., the number of executed generations) is equal to the established maximum number of generations.

(2) The best fitness changes less than ε over δ generations.

The GPGA has been implemented in MATLAB and the evaluation of each member in a population is conducted by simulating the test system during one year using a 1-hour time-step. Individual executions are distributed among 60 cores using the library developed by M. Buehren [49]. For more details about the multi-core installation, see [50]. The GPGA complete general procedure is depicted in Figure 4.

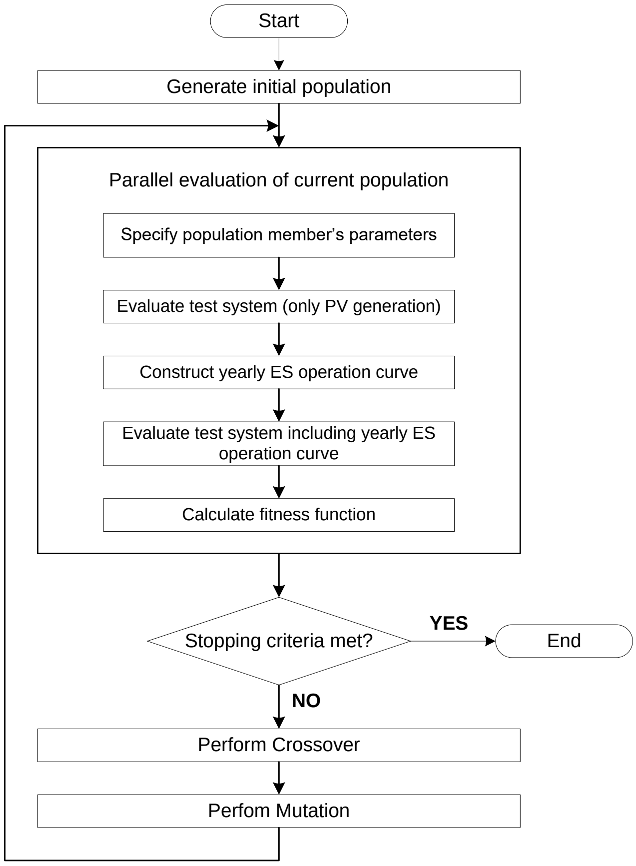


Figure 4. Flow chart of the implemented GPGA procedure.

The fitness function calculated for every evaluation will be equal to the Euclidean distance obtained from the reduction of the operational variables:

$$Fitness = \sqrt{(1 + \Delta P_L)^2 + (1 + \Delta P_{MAX})^2 + (1 + \Delta P_{STD})^2 + (1 + \Delta E)^2 + (\Delta R_{PF})^2} \tag{5}$$

$$\Delta R_{PF} = \begin{cases} 1, \; P_{N\_MIN} > 0 \\ 1 + \frac{P_{N\_MIN}}{P_{MAX}}, \; P_{N\_MIN} \leq 0 \end{cases} \tag{6}$$

where ΔP_L is the reduction of energy losses, ΔP_{MAX} is the reduction in the peak power supplied by the substation transformer, ΔP_{STD} is the reduction of the standard deviation of the yearly power measured at the MV terminals of the substation, ΔE is the reduction of the annual active energy supplied by the substation transformer, P_{N_MIN} is the minimum yearly power delivered by the substation transformer, and P_{MAX} is the maximum yearly power in the base case (i.e., without PV generation or energy storage). ΔR_{PF} has been introduced in the fitness function as penalty for those cases in which reverse power flow occurs. All reductions are calculated in pu with respect to the case without generation and energy storage.

The procedure adds "1" to the reduction values in order to avoid any inconsistencies due to the appearance of negative reductions. For every evaluation, the procedure checks that all voltages are within accepted limits (between 0.95 and 1.05 pu), no network element experiences overload, and the iterative algorithm converges for all solution steps. If one or more of these conditions are not fulfilled, the procedure will assign a fitness value close to zero, which will difficult its participation in the crossover stage of the GPGA.

The following aspects must be taken into account when executing the optimization procedure:

- The resulting ratings for PV generators and ES units will be rounded in steps of 5 kVA, while the rating of the interconnection transformers will be rounded in steps of 10 kVA.
- PV generators and ES units only inject active power, and ES units only absorb active power.
- PV generators are allowed to produce as much energy as possible. Moreover, the PV generator model implemented in OpenDSS assumes that the generator is capable of reaching the Maximum Power Point (MPP) for every simulation step.
- ES units absorb/inject power only when the dispatched power is above 10% of the rated power; in the same manner, the PV generators only inject power if the generated power is above 10% of the rated power.
- The minimum stored energy for ES units is equal to 20% of the ES rated energy.
- Although the procedure relies on a technology-independent ES model, the present application and characteristics of the ES unit (i.e., rated power, rated energy, and energy to power ratio) indicate that Battery Systems are the technology that would best fit a real-life implementation.
- The ES operation curve (see Section 3) is constructed from the active power profile served by the substation transformer when PV generation is present. This curve depends on the actual PV generation and cannot be estimated before the optimization procedure is executed. Therefore, in order to construct the operation curve, for every member of a generation the test system must be first evaluated taking only into account the presence of PV generation; the resulting active power profile will be used by the procedure to construct the yearly ES operation curve.
- The day-classification methods also rely on the active power curve served by the substation transformer, including PV generation. Although a similar approach to that presented in the previous bullet could be used, this could lead to discrepancies in the procedure. The system quantities used as criterion for day-classification depend on the actual value of PV generation and different ratings for PV generators could generate different classifications. If each generation member was to produce different day-classifications, the optimization of the operation parameters would be incorrect because the set of parameters assigned to a specific cluster would not be applied to the same groups of days. Furthermore, the CH index could not be used to predefine the optimal number of clusters to be used in the optimization procedure, since each day-classification method could have a different optimal number of clusters. In order to circumvent this inconvenient, a preliminary study will be conducted; in this study, only the sizing of PV generators will be optimized, while the fitness function is defined by the following equation:

$$Fitness_{PV} = \sqrt{(1 + \Delta P_L)^2 + (1 + \Delta P_{MAX})^2 + (1 + \Delta E)^2 + (\Delta R_{PF})^2} \tag{7}$$

The GPGA will use the parameters presented in Table 1; however, the security factor of the interconnection transformer will be constant and equal to 1.01. The day-classification will be conducted using the optimized PV generator ratings found by this study.

- The maximum rated power for a connection node and the initial values for PV generator and ES unit power ratings are generated according to rules presented in [51].

- The term “with/without PV generation” makes reference to the PV generation to be defined by the procedure. Should the system under study have previously installed DG, it would be considered as part of the original design and not included in the optimization. Moreover, the active power profile “with and without PV generation” is obtained through a yearly simulation using the synthetically generated system curves.
- The locations of the PV generators and ES units have been arbitrarily chosen. Although the locations could have been derived from a procedure aimed at optimizing the impact on the grid, there are several non-technical factors that can have a greater impact on selecting the DER locations than the results obtained from such procedure. Land price, environmental permits, availability of space, and the so-called “not-in-my-backyard” stance are some of the aspects that can limit the options for the location of PV generators and ES units.

5. Case study 1: optimization of a single energy storage unit

5.1. Test system configuration

The test system is a three-phase 60-Hz overhead system serving two spot loads from a High-Voltage/Medium-Voltage (HV/MV) substation transformer; see Figure 5. The PV generator and ES unit are connected to the MV network through a distribution Medium-Voltage/Low-Voltage (MV/LV) transformer.

The main characteristics and operating conditions are as follows:

- HV rated voltage: 230 kV
- HV actual voltage: 1.05 pu
- MV rated voltage: 4.16 kV
- LV rated voltage: 0.4 kV
- Substation transformer rated power: 1000 kVA
- Total rated active load: 550 kW.

System loads have been modelled as voltage-independent and characterized by curve shapes derived with the procedure presented in [36]. The efficiency curve used for the ES unit is shown in Figure 6. The same curve has been assumed for both charge and discharge cycles.

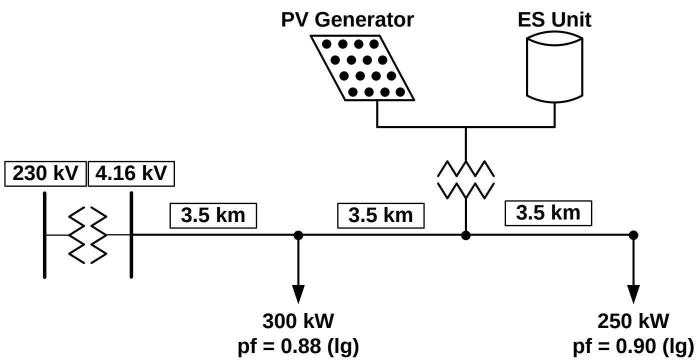


Figure 5. Test system 1—Configuration.

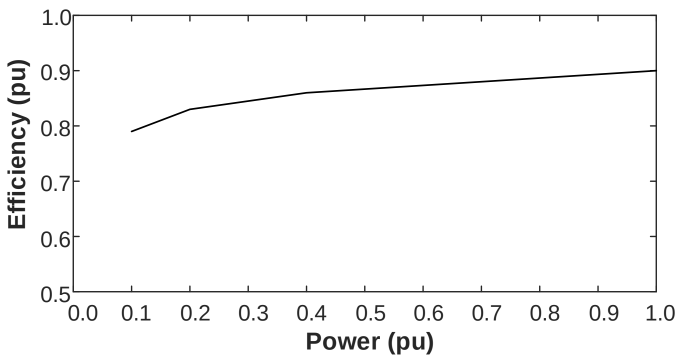


Figure 6. Test system 1—ES unit efficiency curve.

5.2. Simulation results

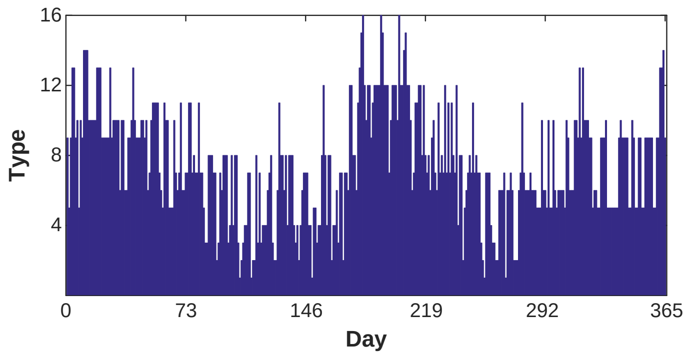
The GPGA procedure was executed for three different scenarios based on the day-classification methods proposed in Section 3 and using the values shown in Table 1. Operational variables without ES and PV generators are as follows:

- Annual energy: 1877440.1 kWh
- Yearly energy losses: 40209.7 kWh
- Maximum yearly power: 427.0 kW
- Yearly power standard deviation: 51.2 kW.

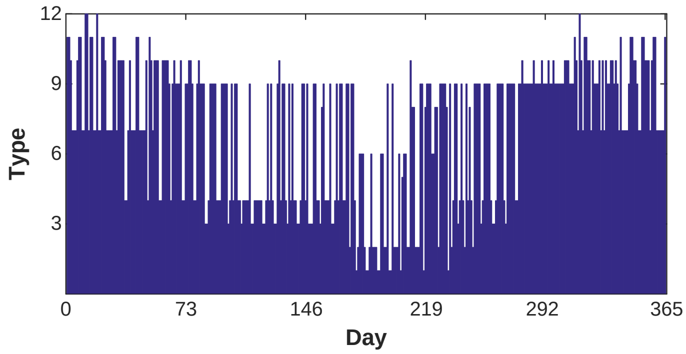
Day-classification was conducted using a 60-kW PV generator (obtained during the initial study). A sample day-classification is shown in Figure 7, where a bar graph is used to show the cluster/group number to which each day of the year is assigned; that is, all days whose bars present the same height have been placed in the same cluster/group and thus, share the same parameters used in the construction of the ES operation curve. As expected, the three classification methods produce different graphs. For the Time-series and Daily-values classification methods the optimization procedure was carried out considering different numbers of clusters. In all scenarios the GPGA seeks to optimize all 4 operational variables (see Section 4).

The results are presented in Table 2. It can be observed that all scenarios converge due to the δ/ε criterion and that best fitness values are consistent throughout the executions. Although operational variables exhibit some variations (based on the scenario under evaluation), they are sufficiently small to consider them negligible and regard the results as acceptable. Except for yearly energy losses (on which the optimization procedure seems to have a very limited effect), there are significant reductions of all operational variables with respect to the case without generation and storage.

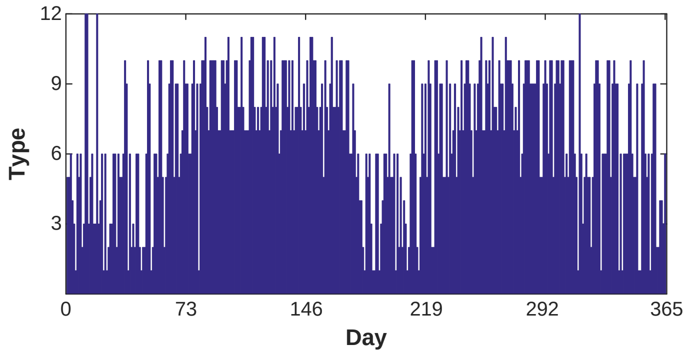
The optimal values for ratings show larger variations than those exhibited by the operational variables; however, in general these values appear to be bound to a specific range, so they can be used as a reference for choosing PV generator and ES unit ratings in real-life applications. The executions based on the Time-series and Daily-values classification methods show a dependency on the number of clusters used for the day-classification; better results are obtained with a smaller number of clusters for both methods. Under these circumstances the CH index could be used to calculate the optimal number of clusters; this approach could prevent the repeated execution of the optimization procedure with different number of clusters for the day-classification.



(a) Daily served energy and PV generation.



(b) Time-series (12 clusters).



(c) Daily-values (12 clusters).

Figure 7. Test system 1—Sample day-classification.

The calculation of the CH index has determined that the optimal number of clusters for the Time-series and Daily-values Clustering methods are 5 and 16 clusters, respectively (the CH index was calculated for a number of clusters ranging from 2 to 24). The results obtained when executing the optimization procedure with the optimal number of clusters for the Time-series and Daily-values classification methods are shown in Table 3. It can be observed that the Time-series method produced a higher value for best fitness with respect to those presented in Table 2 (although energy losses exhibit a slight increment), whereas the results obtained with the Daily-values method show no improvement; in fact, the day-classification with 6 clusters produced better results.

Table 2. Test system 1—Simulation results.

Daily served energy				
Best fitness	2.3803			
Generation count	144			
Simulation time (sec)	23667.8			
PV generator rated power (kW)	70			
ES unit rated power (kW)	55			
ES unit rated energy (kWh)	550			
Interconnection transformer rated power (kVA)	110			
Annual energy (kWh)	1790390.2			
Yearly energy losses (kWh)	40110.5			
Maximum yearly power (kW)	367.9			
Yearly power standard deviation (kW)	44.7			
Time-series classification				
Number of clusters	6	12	18	24
Best fitness	2.4061	2.3854	2.3705	2.3649
Generation count	196	135	91	115
Simulation time (sec)	17146.0	17481.2	15628.0	25119.1
PV generator rated power (kW)	105	70	70	70
ES unit rated power (kW)	65	50	45	40
ES unit rated energy (kWh)	650	500	450	400
Interconnection transformer rated power (kVA)	130	110	100	100
Annual energy (kWh)	1733400.6	1790456.2	1782484.6	1786660.9
Yearly energy losses (kWh)	38820.2	39913.3	39479.6	39739.9
Maximum yearly power (kW)	357.3	373.1	378.5	383.8
Yearly power standard deviation (kW)	46.1	43.7	45.4	44.9
Daily-values classification				
Number of clusters	6	12	18	24
Best fitness	2.4188	2.3698	2.3624	2.3426
Generation count	199	92	92	39
Simulation time (sec)	17829.2	12125.1	16133.6	8502.4
PV generator rated power (kW)	85	45	70	60
ES unit rated power (kW)	70	45	45	50
ES unit rated energy (kWh)	700	450	450	310
Interconnection transformer rated power (kVA)	170	80	110	110
Annual energy (kWh)	1775292.8	1830643.7	1786249.8	1795458.1
Yearly energy losses (kWh)	40515.5	40785.2	39913.0	40125.9
Maximum yearly power (kW)	352.0	378.5	378.5	373.1
Yearly power standard deviation (kW)	42.4	42.8	45.7	48.0

Overall, the highest fitness value was obtained using the Daily-values method with 6 clusters; however, the difference with respect to the highest valued found with the Time-series method is negligible; moreover, the Time-series method required only half of the execution time.

This behavior shows that the Time-series classification method can be used in combination with an optimal number of clusters derived from the calculation of the CH index in order to obtain adequate results from the GPGA. This approach requires the lowest number of clusters, which implies a minimization of the dimension of the optimization problem and a reduction in total execution times.

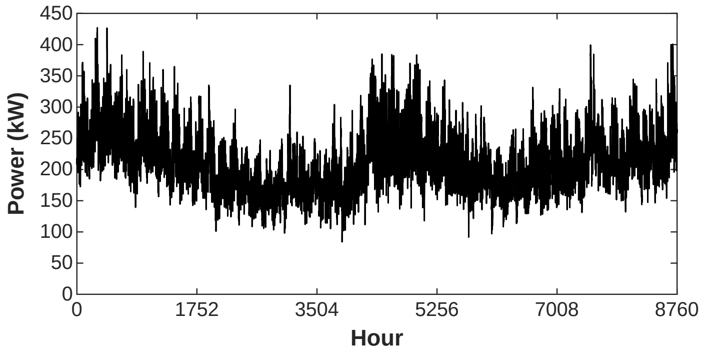
Table 3. Test system 1—Optimal number of clusters.

	Time-series	Daily-values
Number of clusters	5	16
Best fitness	2.4138	2.3682
Generation count	127	197
Simulation time (sec)	9687.6	32236.5
PV generator rated power (kW)	85	75
ES unit rated power (kW)	75	50
ES unit rated energy (kWh)	750	500
Interconnection transformer rated power (kVA)	170	110
Annual energy (kWh)	1771884.2	1778403.4
Yearly energy losses (kWh)	40454.3	39791.2
Maximum yearly power (kW)	346.7	373.2
Yearly power standard deviation (kW)	43.7	46.0

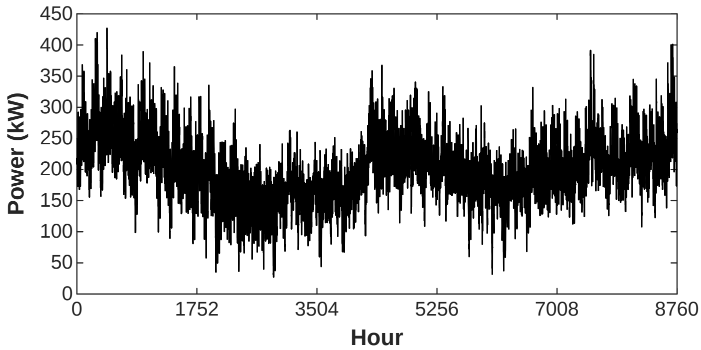
The optimal operation parameters for the Time-series classification method using 5 clusters are shown in Table 4. Figure 8 depicts the behavior of the yearly active power curve obtained following these parameters. It can be observed that PV generation alone is not capable of reducing the yearly maximum power but due to its effect the power curve presents lower minimum values and also exhibits larger variations, see Figure 8b. The introduction of ES manages to reduce the yearly power peak and produce a flatter profile than the curves depicted in Figure 8a and 8b; however, the yearly curve still presents power spikes, see Figure 8c.

Table 4. Test system 1—Optimal operation parameters (Time-series classification-5 Clusters).

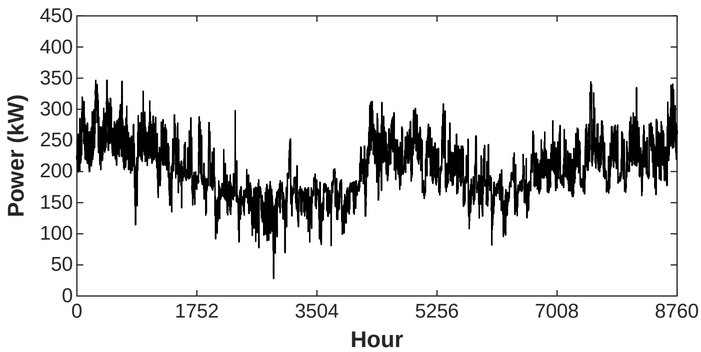
Cluster Number	Charge Limit Factor	Discharge Limit Factor	Charge Correction Factor	Discharge Correction Factor
1	0.68	0.24	2.00	0.46
2	0.20	–0.30	1.19	0.79
3	1.04	0.06	2.00	0.42
4	0.30	1.30	0.55	2.00
5	0.01	0.45	0.94	2.00



(a) Without PV generation.



(b) With PV generation.



(c) With PV generation and ES.

Figure 8. Test system 1—Yearly active power curve (Time-series classification-5 clusters).

5.3. Discussion

One important aspect that should not be ignored is the effect that the fitness function can have on the GPGA performance and how the devices affect the operational variables. PV generation can help to reduce both the annual energy provided by the substation transformer and the yearly energy losses; the peak power occurs at a time when PV generation is zero and the yearly power standard deviation is actually increased due to the presence of PV generation.

As for ES, operation parameters will determine how the operational variables will be affected by ES operation. The operation band defined by the charge and discharge limits will delimit the expected standard deviation by constraining the active power profile to this band; this action will also help to reduce the maximum peak power delivered by the substation transformer. On the other hand, total energy losses will be a result of the interaction between the ES charge (loss increment) and discharge (loss reduction) states. The ES operation curve mirrors the behavior of the active power profile; through this mirrored behavior the ES unit will adjust its power injection/absorption to the system's loading. However, the actual power value that needs to be generated/absorbed in order to produce an overall improvement in system conditions cannot be pre-assumed. Therefore, the initial operation curve defined by the charge and discharge limits must corrected in order to produce the greatest impact on the system: the charge correction factor will ensure that the ES unit has enough energy to meet the pre-dispatched discharge without producing an unacceptable increment in energy losses, whereas the discharge correction factor will help to improve the overall system conditions (standard deviation, peak power, and energy losses) while balancing the use of energy so it is not depleted too quickly. In general, the interactions among the different operation parameters will produce much larger reductions for the yearly maximum power and yearly power standard deviation than for energy losses; therefore, the procedure will drive the optimal operation towards benefiting these operational variables.

Defining a fitness function that accurately reflects the objectives behind the optimization procedure is a critical task. Furthermore, the interconnection transformers represent a new source of losses (i.e., no-load and load losses), which in some cases cannot be compensated by the reduction in conduction losses (i.e., losses in distribution lines) and can limit the connection of larger PV generators and ES units.

The resulting ES operation curves correspond to a system behavior under known load and generator curves; therefore, they cannot be used for actual day-to-day operation. However, the optimal operation parameters found by the GPGA can be used to define a preliminary operation curve based on day-ahead load and generation forecast. For this purpose, the distribution system must be simulated using the forecasted load and generation curves; then the resulting active power profile will be used to construct a preliminary ES operation curve by applying the corresponding operation parameters. The preliminary operation curve can be used as a base to create a set of operation curves (by introducing random variations) that can serve as the initial population of a GPGA aimed at determining the optimal daily operation curve by optimizing every single ES operation point. Using a set of curves based on the preliminary operation curve should speed-up algorithm convergence when compared to a randomly generated initial population.

6. Case study 2: optimization of multiple energy storage units

6.1. Test system configuration

A new test system has been studied in order to test the optimization of multiple ES units; the optimization procedure presented in the previous sections was implemented to be independent of test system size as well as of the number of PV generators and ES units under evaluation.

The new test system is a three-phase 60-Hz overhead system with a simplified representation of the HV transmission system (see Figure 9). ES units and PV generators have the same configuration and characteristics that in the previous test system; see Figures 5 and 6. By default, all distribution loads are connected to the system through MV/LV transformers.

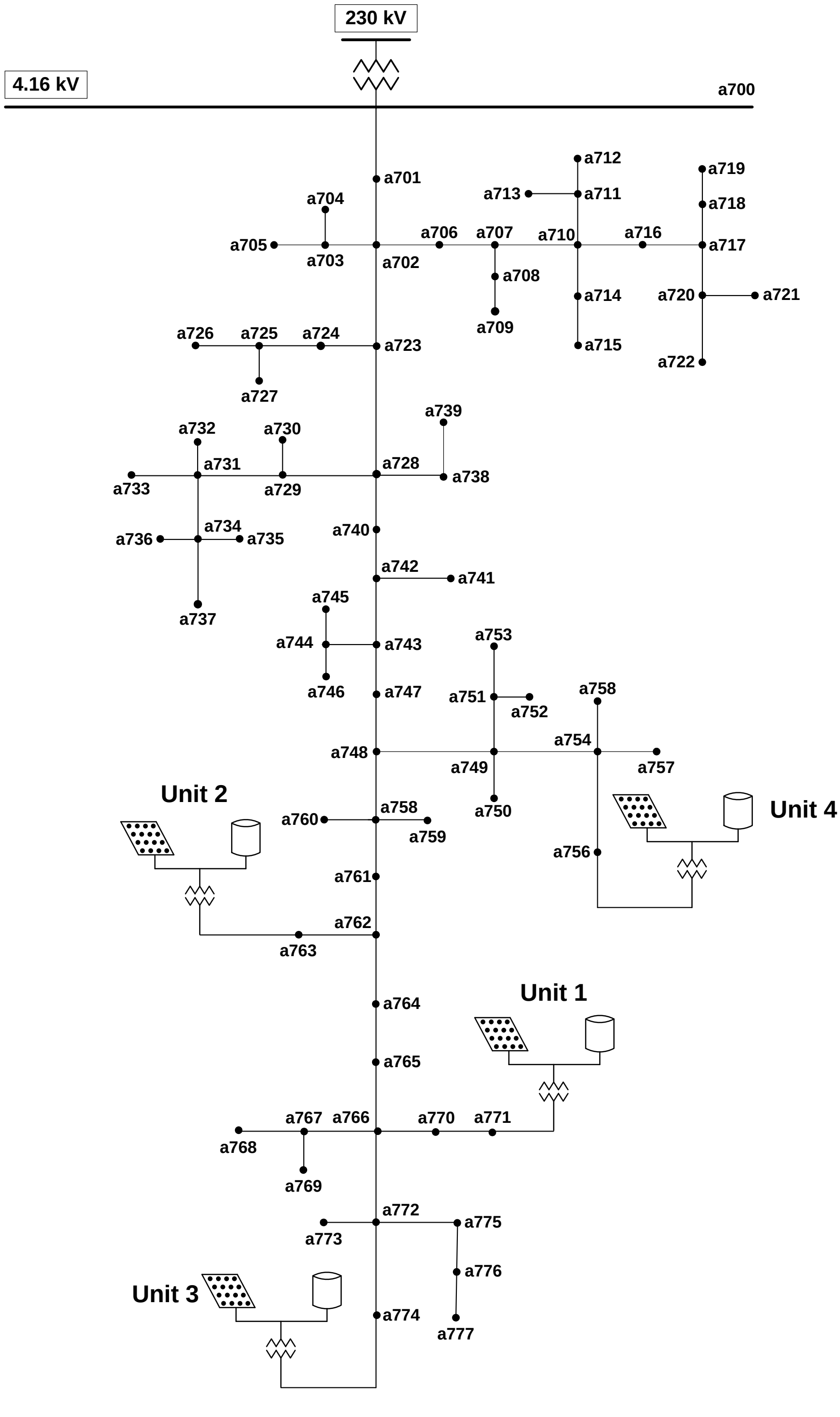


Figure 9. Test system 2—Configuration.

Some important information related to the new test system follows:

- HV rated voltage: 230 kV
- HV actual voltage: 1.03 pu
- MV rated voltage: 4.16 kV
- LV rated voltage: 0.4 kV
- Substation transformer rated power: 2500 kVA
- Total rated active load: 1700 kW.

Operational variables for the new test system (without ES and PV generators) are as follows:

- Annual energy: 7459245.8 kWh
- Yearly energy losses: 199519.6 kWh
- Maximum yearly power: 1473.8 kW
- Yearly standard deviation: 183.9 kW.

As in Test System 1, loads have been modelled as voltage-independent and curve shapes have been derived with the procedure presented in [36]. Day-classification was conducted with the following rated powers for PV generators 1–4: 115, 40, 115, and 270 kW.

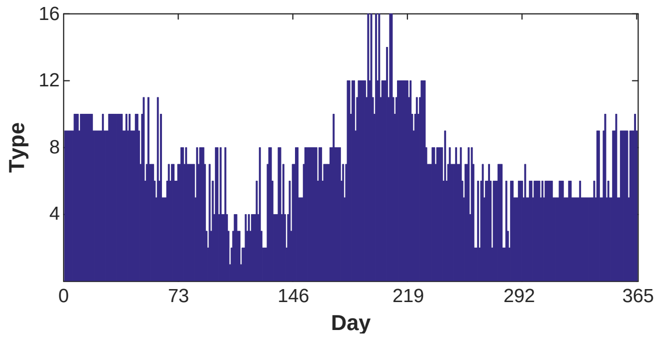
6.2. Simulation results

Simulation results for all three scenarios are presented in Table 5; for the Time-series and Daily-values methods, the optimal number of clusters derived from the CH index (2 and 3 clusters, respectively) were used for the ES operation, see Figure 10.

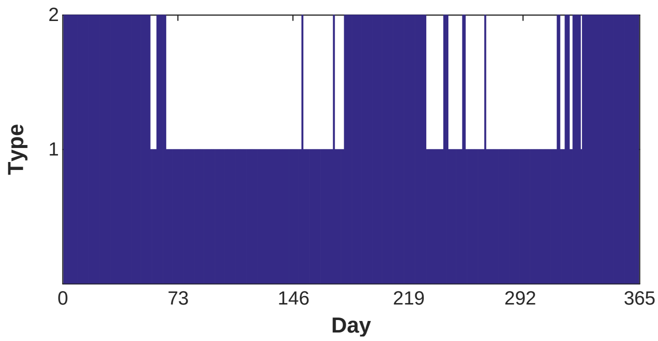
Table 5. Test system 2—Operating conditions with energy storage.

	Day-classification method		
	Daily served energy and PV generation	Time-series	Daily-values
Number of clusters	-	2	3
Best fitness	2.3064	2.4478	2.4201
Generation count	110	128	131
Simulation time (sec)	48688.6	17891.6	22258.4
PV generator rated power (kW)	40-80-15-15	80-220-15-15	105 -105-25-90
ES unit rated power (kW)	80-15-15-40	105-220-15-15	50-40-50-170
ES unit rated energy (kWh)	800-105-150-400	895-2200-150-150	450-400-500-1700
Interconnection transformer rated power (kVA)	160-160-30-50	210-440-30-20	110-210-60-340
Annual energy (kWh)	7294475.4	7149353.3	7136362.4
Yearly energy losses (kWh)	207009.7	211395.4	212202.3
Maximum yearly power (kW)	1403.5	1127.6	1164.5
Yearly power standard deviation (kW)	162.1	142.8	148.1

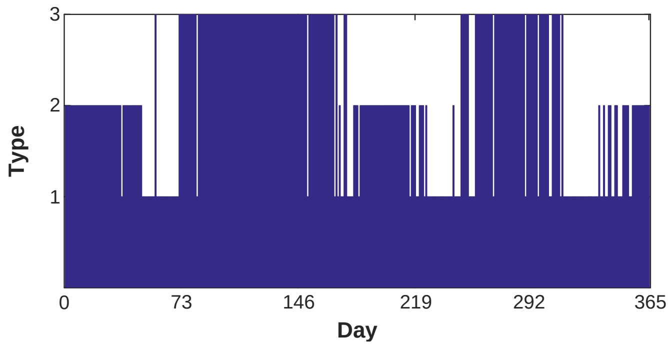
Note that optimal numbers of clusters are much lower than those obtained with the previous test system. This behavior is due to the smoothening effect produced by the aggregation of loads: Test System 1 only served two loads, which caused a more varying power profile, whereas the larger number of loads in the new test system generates a smoother profile. It can be observed that the Time-series classification produces the highest best fitness value and requires the shortest simulation times; these results are in line with those presented in the previous section. With the exception of energy losses, all classification methods were able to produce reductions in the operational variables when compared to the values without ES and PV generation; moreover, the Time-series and Daily-values methods produced better results than the Daily served energy and PV generation method. Consequently, the optimization procedure can be regarded as satisfactory, since the increment in energy losses is clearly outweighed by the improvement in the rest of the operational variables.



(a) Daily served energy and PV generation.



(b) Time-series (2 clusters).



(c) Daily-values (3 clusters).

Figure 10. Test system 2 day—Classification.

The optimal values for PV generators and ES units present large variations for the ratings of individual units; however, the Time-series and Daily-values classification methods present similar penetration factors (i.e., total installed power) for both generation and storage.

The results presented in [51] show that different combinations of rated powers and locations of PV generators can produce similar values of energy losses; therefore, it can also be expected that similar values for the operational variables can be obtained with different ratings of PV generators and ES units.

All GPGA executions converged due to the *δ/ε* criterion; however, since GA is a stochastic optimization procedure, it is possible for it to converge before reaching the optimal solution. If the test system size and number of clusters used in the ES operation mode result in affordable simulation times, the GPGA can be executed without the *δ/ε* criterion, that is, the procedure will stop when it has reached the maximum number of generations.

The optimization procedure was executed for 200, 250, and 300 generations using the Time-series classification method with 2 clusters. Results in Table 6 show an improvement with respect to the values without ES and PV generation; moreover, the best fitness value seems to indicate that 200 generations are enough for obtaining optimal results. Note that operational variables present consistent values among executions. The optimal ratings of PV generators and ES units show clear variations among executions, whereas penetration factors exhibit a more constant trend, a behavior that has remained constant throughout all conducted studies. Consecutive executions of the procedure can be used to produce several combinations of rating values that can be selected in a real-life application or to explore the optimal penetration factor for PV generation and energy storage.

Table 6. Test system 2—Simulation results-time-series classification.

Number of clusters	2	2	2
Best fitness	2.4703	2.4631	2.4675
Generation count	200	250	300
Simulation time (sec)	28645.6	35565.5	43038.9
PV generator rated power (kW)	25-15-145-220	15-50-105-335	15-270-115-15
ES unit rated power (kW)	15-15-115-210	15-50-50-300	15-220-105-15
ES unit rated energy (kWh)	150-150-1150-2100	150-500-500-3000	45-2200-1050-150
Interconnection transformer rated power (kVA)	30-20-160-270	20-60-110-360	20-280-120-30
Annual energy (kWh)	7012252.4	6880414.1	7013129.7
Yearly energy losses (kWh)	204681.5	205104.7	206536.3
Maximum yearly power (kW)	1131.2	1133.8	1123.0
Yearly power standard deviation (kW)	141.7	146.6	142.4

7. Discussion and future work

The main focus of most previous works on the main topics covered in this paper is the optimum sizing and operation of ES units within microgrids or small-size power systems, being the typical optimization period for ES operation 24 hours; larger evaluation periods are generally divided into a reduced number of representative days (usually through clustering techniques) and ES operation is optimized for each representative day. For this work, an evaluation period of one year with a time step of one hour has been considered. Since only through the optimization of every single operation point (i.e., every hour of the year) the global optimum can be reached, the problem dimension becomes very large, and even with the aid of a multi-core installation simulation times are prohibitive. Furthermore, a stochastic optimization method, such as GA, cannot guarantee the optimal solution for an optimization problem with such a large number of input parameters. Therefore, there must exist a tradeoff between optimality and the achievement of a feasible solution.

A comparison between the results derived from the developed methodology and other previous works has not been possible due to the difficulty (or impossibility) to adequate other procedures to the conditions established in the present work. The long term evaluation (i.e., one year) and system solution by means of a power flow simulator appear to be incompatible with the formulation used in other methodologies. On the other hand, assuming the restrictions imposed in other works (such as the evaluation period represented as a group of representative days or the introduction of simplifications into the solution of the electrical system) would defeat the purpose of a procedure that seeks to solve the problem through a detailed representation of the system under study and its behavior during the entire evaluation period.

Future work could be aimed at improving the criteria used in the classification of the evaluation period days, expanding the evaluation period (i.e., several years) and taking into account yearly load variation, refining the proposed operation mode, and defining a fitness function that best balances the improvements in each of the operational variables. Further work could also be based on an optimization procedure that does not rely on dimension reduction techniques and could cope with time steps shorter than one hour (e.g., 15 minutes).

8. Conclusions

This paper has presented the application of a General Parallel Genetic Algorithm to determine the optimal ratings of multiple fixed-location PV generators and ES units when they are operated from the utility's perspective. The evaluation is performed with OpenDSS, which allows for a detailed modeling of the distribution system and an accurate estimation of network conditions; that is, no simplifications are introduced into the solution of the electrical system, the evaluated system can be of any size, and detailed models and advanced control strategies algorithms for system components can be included.

The optimization of ES ratings cannot be done without considering an optimal dispatch strategy; the operation mode proposed in this work attempts to simplify the arduous task of optimizing the complete yearly ES dispatch. The implemented strategy reduces the number of variables to be optimized by using clustering techniques that group the days of the year into different clusters. However, unlike other works, each group of days is not independently optimized. Each possible solution is evaluated by simulating the test system during the entire evaluation period. This approach ensures that the evaluation period is assessed as a whole (not as a decoupled group of representative days) and the ES current state is the result of the complete ES operation history.

The introduction of parallel computing into the optimization procedure has allowed for affordable execution times. Although each individual simulation does not require large simulation times, the repeated evaluation of the system in a sequential manner would result in prohibitive execution times. Therefore, only optimization techniques that are prone to parallelization could be used in conjunction with a power flow simulator and a long term evaluation, since sequential techniques would be unfeasible due to the resulting execution times.

It is important to emphasize that the locations of PV generators and ES units were arbitrarily chosen in this work; however, one must keep in mind that the location of a power source (generation or consumption) can have a non-negligible impact on energy losses and system voltages.

The main features of the implemented methodology may be listed as follows: (i) system conditions are estimated by means of a power flow simulator, (ii) the methodology is independent of system size and the number of units to be optimized, (iii) ES operation is optimized for an evaluation period of one year, (iv) the operation strategy of ES units is based on the daily variation of load and generation curves, (v) hierarchical clustering is used to group daily power curves in order to reduce the dimension of the optimization problem, (vi) time-driven simulation ensures that the evaluation period is assessed as a whole and the ES current state is the result of the ES operation history, and (vii) the parallel approach used in the GA implementation allows for affordable execution times.

Conflict of interest

All authors declare no conflict of interest in this paper.

References

1. New York Independent System Operator (2014) A Review of Distributed Energy Resources.
2. Farret FA, Godoy SM (2006) Integration of Alternative Sources of Energy. John Wiley Press, 301–332.
3. Ackerman T, Andersson G, Söder L (2001) Distributed generation: a definition. *Electr Power Syst Res* 57: 195–204.
4. Willis HL, Scott WG (2000) Distributed power generation: planning and evaluation. Crc Press.
5. Sandia National Laboratories and NRECA (2015) DOE/EPRI Electricity Storage Handbook.
6. Luo F, Meng K, Dong ZY, et al. (2015) Coordinated operational planning for wind farm with battery energy storage system. *IEEE T Sustain Energ* 6: 253–262.
7. International Electrotechnical Commission (2011) Electrical Energy Storage.
8. Farzin H, Fotuhi-Firuzabad M, Moeini-Aghtaie M (2017) A stochastic multi-objective framework for optimal scheduling of energy storage systems in microgrids. *IEEE T Smart Grid* 8: 117–127.
9. Lazaroiu GC, Dumbrava V, Balaban G, et al. (2016) Stochastic optimization of microgrids with renewable and storage energy systems. International Conference on Environment and Electrical Engineering. *IEEE*.
10. Silvestre MLD, Graditi G, Ippolito MG, et al. (2011) Robust multi-objective optimal dispatch of distributed energy resources in micro-grids. PowerTech, 2011 IEEE Trondheim. *IEEE*, 1–5.
11. Agamah SU, Ekonomou L (2016) Peak demand shaving and load-levelling using a combination of bin packing and subset sum algorithms for electrical energy storage system scheduling. *Iet Sci Meas Technol* 10: 477–484.
12. Levron Y, Shmilovitz D (2012) Power systems' optimal peak-shaving applying secondary storage. *Electr Pow Syst Res* 89: 80–84.
13. Jayasekara N, Wolfs P, Masoum MAS (2014) An optimal management strategy for distributed storages in distribution networks with high penetrations of PV. *Electr Pow Syst Res* 116: 147–157.
14. Ippolito MG, Silvestre MLD, Sanseverino ER, et al. (2014) Multi-objective optimized management of electrical energy storage systems in an islanded network with renewable energy sources under different design scenarios. *Energy* 64: 648–662.
15. Rahmani-Andebili M (2017) Stochastic, adaptive, and dynamic control of energy storage systems integrated with renewable energy sources for power loss minimization. *Renew Energ* 113: 1462–1471.
16. Meirinhos JL, Rua DE, Carvalho LM, et al. (2017) Multi-temporal Optimal Power Flow for voltage control in MV networks using Distributed Energy Resources. *Electr Pow Syst Res* 146: 25–32.

17. Hejazi H, Mohsenian-Rad H (2016) Energy storage planning in active distribution grids: a chance-constrained optimization with non-parametric probability functions. *IEEE T Smart Grid*, 1–13.
18. Rahmani-Andebili M, Shen H (2017) Cooperative distributed energy scheduling for smart homes applying stochastic model predictive control. IEEE International Conference on Communications. *IEEE*, 1–6.
19. Kargarian A, Hug G (2016) Optimal sizing of energy storage systems: a combination of hourly and intra-hour time perspectives. *Iet Gener Transm Dis* 10: 594–600.
20. Kerdphol T, Qudaih Y, Mitani Y (2016) Optimum battery energy storage system using PSO considering dynamic demand response for microgrids. *Int J Elec Power* 83: 58–66.
21. Carpinelli G, Mottola F, Proto D (2016) Probabilistic sizing of battery energy storage when time-of-use pricing is applied. *Electr Pow Syst Res* 141: 73–83.
22. Brown PD, Peas Lopes JA, Matos MA (2008) Optimization of pumped storage capacity in an isolated power system with large renewable penetration. *IEEE T Power Syst* 3: 523–531.
23. Wen S, Lan H, Fu Q, et al. (2015) Economic allocation for energy storage system considering wind power distribution. *IEEE T Power Syst* 30: 644–652.
24. Korpaas M, Holen AT, Hildrum R (2003) Operation and sizing of energy storage for wind power plants in a market system. *Int J Elec Power* 25: 599–606.
25. Correia PF, Jesus JMFD, Lemos JM (2014) Sizing of a pumped storage power plant in S. Miguel, Azores, using stochastic optimization. *Electr Pow Syst Res* 112: 20–26.
26. Arabali A, Ghofrani M, Etezadi-Amoli M, et al. (2013) Genetic-algorithm-based optimization approach for energy management. *IEEE T Power Deliver* 28: 162–170.
27. Nick M, Cherkaoui R, Paolone M (2014) Optimal allocation of dispersed energy storage systems in active distribution networks for energy balance and grid support. *IEEE T Power Syst* 29: 2300–2310.
28. Silvestre MLD, Graditi G, Sanseverino ER (2014) A generalized framework for optimal sizing of distributed energy resources in micro-grids using an indicator-based swarm approach. *IEEE T Ind Inform* 10: 152–162.
29. Yang P, Nehorai A (2014) Joint optimization of hybrid energy storage and generation capacity with renewable energy. *IEEE T Smart Grid* 5: 1566–1574.
30. Abedi S, Alimardani A, Gharehpetian GB, et al. (2012) A comprehensive method for optimal power management and design of hybrid RES-based autonomous energy systems. *Renew Sust Energ Rev* 16: 1577–1587.
31. Meng N, Wang P, Wu H, et al. (2015) Optimal sizing of distributed generations in a connected DC micro-grid with hybrid energy storage system. Energy Conversion Congress and Exposition. *IEEE*, 3179–3183.
32. Erdinc O, Paterakis NG, Pappi IN, et al. (2015) A new perspective for sizing of distributed generation and energy storage for smart households under demand response. *Appl Energ* 143: 26–37.
33. Papaefthymiou SV, Papathanassiou SA (2014) Optimum sizing of wind-pumped-storage hybrid power stations in island systems. *Renew Energ* 64: 187–196.
34. Dugan RC (2016) Reference Guide. The Open Distribution System Simulator (OpenDSS). EPRI.
35. Dugan RC, McDermott TE (2011) An open source platform for collaborating on smart grid research. Power and Energy Society General Meeting. *IEEE*, 1–7.

36. Martinez-Velasco JA, Guerra G (2015) Analysis of large distribution networks with distributed energy resources. *Ingeniare* 23: 594–608.
37. Dugan RC, Taylor JA, Montenegro D (2017) Energy storage modeling for distribution planning. *IEEE T Ind Appl* 53: 954–962.
38. Eisen MB, Spellman PT, Brown PO, et al. (1998) Cluster analysis and display of genome-wide expression patterns. *P Natl Acad Sci USA* 95: 14863–14868.
39. Tsekouras GJ, Hatziargyriou ND, Dialynas EN (2007) Two-stage pattern recognition of load curves for classification of electricity customers. *IEEE T Power Syst* 22: 1120–1128.
40. Iglesias F, Kastner W (2013) Analysis of similarity measures in times series clustering for the discovery of building energy patterns. *Energies* 6: 579–597.
41. Chicco G, Napoli R, Postolache P, et al. (2003) Customer characterization for improving the tariff offer. *IEEE T Power Syst* 18: 381–387.
42. Sakoe H, Chiba S (1978) Dynamic programming algorithm optimization for spoken word recognition. *IEEE T Acoustics Speech Signal Process* 26: 43–49.
43. Calinski T, Harabasz J (1974) A dendrite method for cluster analysis. *Commun Stat-Theor M* 3: 1–27.
44. Passino KM (2006) Biomimicry for optimization, control, and automation. *IEEE T Automat Contr* 51: 1406.
45. Michalewicz Z (1996) Genetic Algorithms + Data Structures = Evolution Programs. Springer, 347–348.
46. Yeh EC, Venkata SS, Sumic Z (1995) Improved distribution system planning using computational evolution. *IEEE T Power Syst* 11: 668–674.
47. Mendoza F, Bernal-Agustin JL, Dom ńguez-Navarro JA (2006) NSGA and SPEA applied to multiobjective design of power distribution systems. *IEEE T Power Syst* 21: 1938–1945.
48. Abido MA (2006) Multiobjective evolutionary algorithms for electric power dispatch problem. *IEEE T Evolut Comput* 10: 315–329.
49. Buehren M (2007) MATLAB Library for Parallel Processing on Multiple Cores. Available from: http://www.mathworks.com.
50. Martinez JA, Guerra G (2014) Parallel Monte Carlo approach for distribution reliability assessment. *Iet Gener Transm Dis* 8: 1810–1819.
51. Guerra G, Martinez JA (2016) Optimum allocation of distributed generation in multi-feeder systems using long term evaluation and assuming voltage-dependent loads. *Sust Energ Grid Network* 5: 13–26.