RESEARCH ARTICLE

Embedding quasi-static time series within a genetic algorithm for stochastic optimization: the case of reactive power compensation on distribution systems

Juan M. Lujano-Rojas[1,2,*], Ghassan Zubi[3], Rodolfo Dufo-López[4], José L. Bernal-Agustín[4], José L. Atencio-Guerra[5] and João P. S. Catalão[6]

[1]Department for Management of Science and Technology Development, Ton Duc Thang University, Ho Chi Minh City 758307, Vietnam; [2]Faculty of Electrical and Electronics Engineering, Ton Duc Thang University, Ho Chi Minh City 758307, Vietnam; [3]Triangle Research and Development Center, Kafr Qara 30075, Israel; [4]Department of Electrical Engineering, Universidad de Zaragoza, Calle María de Luna 3, 50018 Zaragoza, Spain; [5]Faculty of Electrical Engineering, Technological University of Panama, Campus Metropolitano Víctor Levi Sasso, 0819-07289 El Dorado, Panama City, Panama and [6]INESC TEC and Faculty of Engineering, University of Porto, R. Dr Roberto Frias, 4200-465 Porto, Portugal

*Corresponding author. E-mail: lujano.juan@tdt.edu.vn

Abstract

This paper presents a methodology for the optimal placement and sizing of reactive power compensation devices in a distribution system (DS) with distributed generation. Quasi-static time series is embedded in an optimization method based on a genetic algorithm to adequately represent the uncertainty introduced by solar photovoltaic generation and electricity demand and its effect on DS operation. From the analysis of a typical DS, the reactive power compensation rating power results in an increment of 24.9% when compared to the classical genetic algorithm model. However, the incorporation of quasi-static time series analysis entails an increase of 26.8% on the computational time required.

Keywords: probabilistic power flow; quasi-static time series; genetic algorithm; reactive power compensation

List of symbols

b :	Index for each bit of genetic algorithm individual ($b = 1, \ldots, B$).
c :	Index for each Monte Carlo simulation ($c = 1, \ldots, C$).
g :	Index for each individual of genetic algorithm ($g = 1, \ldots, G$).
h :	Index for each hour of the day ($h = 1, \ldots, 24$).
n :	Index for each node of distribution system ($n = 1, \ldots, N$).
n_{RPC} :	Node for the installation of reactive power compensation.
t :	Index for each hour of the year ($t = 1, \ldots, 8760$).
A :	Matrix of genetic algorithm population.
$\vec{a}_{(g)}$:	Individual g of genetic algorithm population (A).
$a_{(b, g)}$:	Bit b of individual g of genetic algorithm.
$ACC_{(n)}$:	Capital cost of compensation device at node n (€).
$ACC_{(n)}^{Fxd}$:	Capital cost of fixed compensation device at node n (€).

$ACC_{(n)}^{Var}$: Capital cost of variable compensation device at node n (€).

$ARC_{(n)}$: Replacement cost of compensation device at node n (€).

$ARC_{(n)}^{Fxd}$: Replacement cost of fixed compensation device at node n (€).

$ARC_{(n)}^{Var}$: Replacement cost of variable compensation device at node n (€).

$AMC_{(n)}$: Maintenance cost of compensation device at node n (€).

$AMC_{(n)}^{Fxd}$: Maintenance cost of fixed compensation device at node n (€).

$AMC_{(n)}^{Var}$: Maintenance cost of variable compensation device at node n (€).

$BCBV$: Branch-current to bus-voltage matrix.

$BIBC$: Bus-injection to branch-current matrix.

CRF : Capital recovery factor.

$CAP_{(l, b)}$: Reactive power injected to the system for block l and node b (kVAr).

$d_{(h, n)}$: Load profile at time h and node n.

$D_{(t, n, c)}$: Load time series at time t, node n, and Monte Carlo trial c.

$D_{(t, n)}^{k++}$: Load time series selected by k-means++ algorithm.

$D_{(n)}^{max}$: Capacity of distribution transformer at node n (kVA).

e_{PV} : Electron charge (C).

E_{max} : Maximum value of electricity prices (€/MWh).

$E_{(t)}$: Electricity price at time t (€/MWh).

$E_{LS(n)}$: Cost of energy losses for compensation installed at node n (€).

$f_R(\cdot)$: Function to calculate the heat loss by radiation.

$f_C(\cdot)$: Function to calculate the heat loss by convection.

$f_S(\cdot)$: Function to calculate the heat gained by solar radiation.

$f_{MPPT}(\cdot)$: Implementation of maximum power point tracking algorithm.

$F_{\tilde{Z}}(\cdot)$: Cumulative distribution of time series $\tilde{z}_{(t, n, c)}$.

$F_D(\cdot)$: Cumulative distribution of load-demand time series $D_{(t, n, c)}$.

$FF_{PV(t, n)}^o$: Maximum fill factor at time t for generator n.

$FF_{PV(n)}$: Fill factor of generator n.

$GMD_{(n)}$: Geometric mean distance between phase conductors of branch n (m).

$GMR_{(n)}$: Geometric mean radius of phase conductor of branch n (m).

$G_{(t)}$: Solar radiation at time t (W/m²).

$I_{PV(t, n)}^{SC}$: Short-circuit current at time t for generator n (A).

$I_{PV, STC(n)}^{SC}$: Short-circuit current under standard test conditions for unit n (A).

$I_{PV(n)}^{MAX}$: Current at maximum power production for generator n (A).

$I_{PV(t, n)}$: Photovoltaic cell current at time t for generator n (A).

$\vec{I}_{(t)}$: Vector of branch currents at time t (A).

$I_{(t, n)}$: Current at time t and branch n (A).

$\Delta I_{(t,n)}$: Difference between the maximum current and the actual (A).

$I_{(t, n)}^{max}$: Ampacity at time t of branch n (A).

k_{DS} : Coefficient that depends on the system frequency (Ω/km).

k_{PV} : Parameter of solar transmittance and absorptance.

k_B : Boltzmann constant (J/K).

$k_{CS(n)}$: Auxiliary variable of objective function for node and branch n.

m_{PV} : Photovoltaic cell ideality factor.

$NPC_{(n)}$: Net present cost for compensation equipment at node n (€).

$NPC_{GA(g)}$: Net present cost of individual g (€).

$N_{PV(n)}^{PS}$: Number of panels connected in serial on generator n.

$N_{PV(n)}^{PP}$: Number of panels connected in parallel on generator n.

$N_{PV(n)}^{CS}$: Number of cells connected in serial on generator n.

$N_{PV(n)}^{CP}$: Number of cells connected in parallel on generator n.

$NOCT_{(n)}$: Nominal operating cell temperature of generator n (°C).

$Obj_{GA(g)}$: Value of the objective function for individual g (€).

$p_{I(t, n)}$: Normalized power of inverter at time t for generator n.

$P_{I(n)}$: Rated power of converter of generator n (kW).

$P_{(t, n)}$: Active power consumed at time t and node n (kW).

$P_{R(t, n)}$: Heat loss by radiation at time t for branch n (W/m).

$P_{C(t, n)}$: Heat loss by convection at time t for branch n (W/m).

$P_{S(t, n)}$: Heat gained by solar radiation at time t for branch n (W/m).

$P_{PV(t, n)}$: Photovoltaic cell power at time t for generator n (W).

$P_r\{\cdot\}$: Probability of a determined event.

$Q_{(t, n)}$: Reactive power consumed at time t and node n (kVAr).

Q_C^l : Lowest value of confidence interval of reactive power (kVAr).

Q_C^r : Highest value of confidence interval of reactive power (kVAr).

$Q_{RPC(n)}^{Fxd}$: Rated capacity of fixed compensation at node n (kVAr).

$Q_{RPC(n)}^{Var}$: Rated capacity of variable compensation at node n (kVAr).

$R_{(n)}$: Conductor resistance of branch n (Ω/km).

$R_{PV(t, n)}^S$: Photovoltaic cell resistance at time t for generator n (Ω).

$S_{(t, n)}$: Simulated load series at time t and node n (kVA).

$T_{A(t)}$: Ambient temperature at time t (°C).

$T_{PV(t, n)}$: Photovoltaic cell temperature at time t for generator n (°C).

$u_{PV(t, n)}^{OC}$: Relative photovoltaic cell open-circuit voltage at time t for unit n.

$\vec{U}_{(t)}$: Vector of system voltage at time t (kV).

$U_{(t, n)}$: Voltage at time t and node n (kV).

$\vec{U}_R$: Vector of rated voltage (kV).

$\overline{\Delta U}_{(t)}$: Vector of voltage reduction at time t (kV).

$U_{PV(n)}^{SYS}$: Voltage of photovoltaic generator at node n (V).

U_{min} : Minimum allowed voltage of distribution system (kV).

U_{max} : Maximum allowed voltage of distribution system (kV).

$U_{PV(t, n)}^{OC}$: Photovoltaic cell open-circuit voltage at time t for generator n (V).

$U_{PV, STC(n)}^{OC}$: Open-circuit voltage under standard test conditions of unit n (V).

$U_{PV(t, n)}^e$: Thermal voltage at time t for generator n (V).

$U_{PV(t, n)}$: Photovoltaic cell voltage at time t for generator n (V).

$U_{PV(n)}^{MAX}$:	Voltage at maximum power production for generator n (V).
$V_{(t)}$:	Wind speed at time t (m/s).
$\bar{z}_{(h, n)}$:	Standardized daily load profile at time h for node n.
$\bar{z}_{(t, n, c)}$:	Standardized and normalized series at time t, node n, and trial c.
$\bar{z}_{(t, n)}$:	Standardized yearly load profile at time t for node n.
$Z_{(n)}$:	Conductor impedance of branch n (Ω/km).
$\alpha_{L(t)}$:	Value of white noise at time t.
$\alpha_{PV(n)}$:	Temperature coefficient of generator n (%/$^{\circ}$C).
$\alpha_{I(n)}$:	Parameter of inverter efficiency model for generator n.
$\beta_{I(n)}$:	Parameter of inverter efficiency model for generator n.
δ :	Significance level.
$\eta_{I(t, n)}$:	Inverter efficiency at time t for generator n.
$\eta_{PV(n)}$:	Cell efficiency of photovoltaic generator n.
$\theta_{I(n)}$:	Parameter of inverter efficiency model for generator n.
$\lambda_{(t, n, c)}$:	Random variable at time t, node n, and Monte Carlo trial c.
μ_d :	Mean value of daily load profile.
σ_d :	Standard deviation value of daily load profile.
$\emptyset$:	Autocorrelation coefficient of load time series (%).

1. Introduction

The cultural and technological development of human society is closely related to the idea of management and allocation of natural, human, and economic resources in an appropriate manner. Management of resources and budgets needed for developing social welfare requires feasible strategies, frequently based on optimization techniques. With the increasing complexity of industrial systems and human interactions, optimized management of real-life activities involves the incorporation of stochastic processes. Decision-makers incorporate various sources of uncertainty through probability theory or fuzzy logic (FL) to enhance the effectiveness and reliability of the decision process.

Moreover, the construction of a sustainable society requires an electricity network with smart capabilities to integrate renewable resources and mechanisms to their full exploitation. Due to the random behavior of clean energies, optimal management of energy systems supported by these technologies has become a necessity because of their growing adoption.

The topic under discussion in this paper is essentially a planning problem subject to the uncertain effects of renewable generation. Next, subsection 1.1 briefly presents a general perspective of the state-of-the-art optimization techniques under uncertainty, while subsection 1.2 explores particular issues of the electricity system operating under uncertain conditions. Finally, subsection 1.3 carefully explains the main contributions of the approach developed in this work.

1.1. Optimization under uncertainty

In a general sense, decision-makers can solve optimization problems with uncertain variables using robust optimization (RO), stochastic programming (SP), chance-constrained programming (CCP), and FL, among other techniques.

Concerning RO, Bertsimas, Gupta, and Kallus (2018) developed a robust version of the sample average approximation to overcome the weaknesses related to the limited number of data points used to represent the behavior of uncertain processes. The authors combined features of distributionally RO with hypothesis testing. Similarly, Esfahani and Kuhn (2018) developed a technique to face the problems related to the availability of a limited number of data points by using the Wasserstein metric and distributionally RO. Following a similar trend, Ning and You (2018) modeled the uncertainty using the Dirichlet process mixture combined with a maximum likelihood estimator. Ma (2019) produced a model based on RO that incorporates the risk of life cycle assessment, minimizing the environmental impact of product usage.

RO can be combined with FL and SP to improve optimization results. Farrokh, Azar, Jandaghi, and Ahmadi (2018) developed a hybrid robust fuzzy SP model to consider the influence of different types of uncertainty sources expressed as distributions and fuzzy numbers. In the same way, Ghahremani-Nahr, Kian, and Sabet (2019) developed a model formulated as a robust fuzzy programming problem that employs heuristic techniques. Tsao, Vo-Van, Lu, and Yu (2018) combined two-stage SP and FL for multiobjective optimization. Zou, Ahmed, and Sun (2019) evaluated the weaknesses of multistage stochastic integer programming and proposed an enhanced technique based on stochastic dual dynamic integer programming. The authors reformulated the problem and combined decomposition algorithms with Lagrangian cuts. Baptista, Barbosa-Povoa, Escudero, Gomes, and Pizarro (2019) developed a technique for two-stage multiperiod stochastic optimization that combined CCP and second-order stochastic dominance for risk aversion.

Van Ackooij, Berge, Oliveira, and Sagastizábal (2017) developed a technique to determine the area in which the probabilistic constraints are feasible. Jie, Prashanth, Fu, Marcus, and Szepesvári (2018) developed a technique to solve those problems in which human perspectives are involved. The technique is based on the cumulative prospect theory, where the estimation and optimization are sequentially determined. The approach developed by Buchheim and Pruente (2019) aimed to define a set of feasible solutions instead of a single optimal one. Liu, Lau, and Kananian (2019) proposed a constrained stochastic successive convex approximation methodology to the solution of non-convex problems. The technique solves several convex optimization problems using a surrogate model. Liu, Liu, and Tao (2019) designed a technique to face stochastic composition optimization problems with two expected value functions. The approach embedded an inner objective function into an outer one. The stochastic variance reduction gradient and stochastic average gradient methodologies are combined to estimate the value of the inner objective function, while the duality-free concept is employed to estimate the value of the outer function.

Al-Juboori and Datta (2019) employ non-dominated sorted genetic algorithm-II (NSGA-II) on the optimal design of hydraulic water-retaining structures considering reliability concepts. The formulation considers several stochastic ensemble surrogate models to incorporate the effect of uncertainty.

Regarding CPP, Tan, Gong, Chiclana, and Zhang (2019) proposed a model based on goal programming. The interaction between decision-makers' opinions is represented using uncertain variables, while a combined technique based on the Monte Carlo (MC) simulation and a genetic algorithm (GA) solves the optimization problem. Further on this topic, Kinay, Saldanha-da-Gama, and Kara (2019) created a multicriteria model and analyzed the performance of vectorial optimization and goal programming. Tavana, Shiraz, and Di Caprio (2019) proposed a

model based on random rough mathematical programming formulated using CCP.

On the one hand, the computational tractability of the problem deserves particular attention. On the other hand, the knowledge about the specific application becomes crucial to mitigate the adverse effects of the lack of information. Concerning this reasoning, the next subsection gives an overview of the problem analyzed in this paper.

1.2. Reactive power compensation on distribution networks

Full-scale deployment of renewable generation could produce operating problems on a distribution system (DS), specifically on the voltage profile. A representative condition is when a high amount of renewable power is injected into the DS during periods of low load, resulting in a voltage increase out of the safe limits. Under this panorama, planning engineers analyze the incorporation of reactive power compensation (RPC) to guarantee the effective operation of the smart grid (SG).

RPC can be performed using different devices: a shunt capacitor bank (SCB), a distribution static compensator (DSTATCOM), a unified power quality conditioner (UPQC), or a static VAR compensator (SVC).

Consequently, the sizing and management of these devices have been extensively studied in the literature, resulting in several techniques. Das, Das, and Patra (2017) developed a method that aims to improve the voltage of a node remotely located from another under our control by injecting reactive power. Azzam and Mousa (2010) combined a GA and the ε-dominated definition with the technique for order preference by similarity to the ideal solution (TOPSIS), considering multiple optimization objectives. Pereira, Costa, Contreras, and Mantovani (2016) presented a model for the allocation of distributed generation (DG) and RPC, which uses a tabu search (TS) algorithm and Chu–Beasley GA (CBGA). The TS is used to optimize the installation place of the SCB and DG, while the CBGA is used to perform their power dispatch. Jannat and Savić (2016) developed a model for the optimal allocation and size of an SCB. The model considers the variability of renewable resources, the quality of the voltage profile, and the capacity of reactive power. Thus, MC simulations are used to consider the uncertainty, while an NSGA-II solves the multiobjective optimization problem. Fard and Niknam (2014) developed a methodology that jointly analyzes several reliability indexes and the power loss cost. The authors designed the method as the solution to a multiobjective optimization problem using a firefly algorithm. Su, Masoum, and Wolfs (2016) proposed a sequential approach to maximize the net annual savings, taking into account the voltage limits and the maximum size of the SCB.

Using the literature review presented until this point, the next subsection discusses the main contributions of this work.

1.3. Contributions

Stochastic optimization techniques presented in the technical literature show different viewpoints on how to describe the uncertain processes as well as how to make the resulting formulation computationally tractable. On the other hand, the performance of RPC devices, such as SCBs, DSTATCOMs, UPQCs, and SVCs, depends on the load profile and behavior of the DG so that the operation is affected by the randomness of renewable generation.

Probabilistic constraints related to the problem of installing RPC in a DS supported by renewable energy are directly related to the interactions between the DG and load demand through time. Both of these variables, solar photovoltaic (PV) generation and load demand, are modeled by a time series with a determined correlation degree and daily profile, along with other features. In this paper, quasi-static time series is combined with a GA to determine the optimal size and placement of RPC devices, such as SCB and DSTATCOM.

On the one hand, heuristic optimization approaches, such as GA, GA-TOPSIS, TS-CBGA, NSGA-II, firefly, and PSO, provide flexibility on their formulation and implementation, which allows us to incorporate the non-linear influence of RPC devices on the DS operation. On the other hand, the technical literature considers a limited number of load and DG operating conditions due to the high complexity of the problem under study.

Quasi-static time series consists of evaluating the performance of the DS in terms of voltage and current at each node and branch of the system along the time. The method performs a probabilistic power flow (PPF), obtaining the probability distribution function (PDF) of each voltage and current of the system. These PDFs allow us to verify whether the voltage and ampacity at each node and branch of the network are within the appropriate operating interval.

As the main contribution of this paper, the proposed methodology embedded this general concept into an optimization model based on a GA, using probabilistic constraints to guarantee the feasibility of the solution (sizing and placing of RPC equipment), and minimizing the net present cost (NPC). This contribution fulfills the gap of many works that only assess and optimize the system behavior under a limited number of operating conditions and assumptions.

The paper organization is as follows: Section 2 describes the energy system; Section 3 explains the simulation method; Section 4 presents the proposed methodology, and Section 5 illustrates the proposed method. Finally, Section 6 discusses the main conclusions and findings.

2. Energy System Model

This section describes several essential aspects: the model of environmental resources, the characterization of electricity consumption, the model of the energy conversion systems used for DG, and the behavior of energy prices.

2.1. Environmental variables and renewable resources

Many research centers analyze the meteorological variables such as solar radiation, wind speed, and ambient temperature. The United States National Aeronautics and Space Administration (NASA) provides a general-purpose database of natural resources across the globe (NASA, 2019), which many optimization tools use for the integration of renewable energies.

This work pays special attention to the influence of solar radiation, wind speed, and ambient temperature on the ampacity of DS feeders. The proposed method's simulation model adopts the retrospective reanalysis approach; Pfenninger and Staffell (2016) and Staffell and Pfenninger (2016) implemented this technique for the estimation of the natural resources. The website Renewables.ninja (2019) provides an online implementation of this technique.
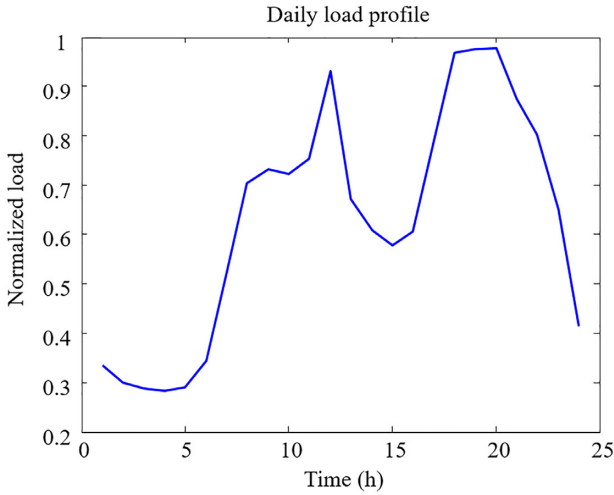
Figure 1: Daily load profile.

Figure 2: CDF of load demand.

2.2. Electricity demand, real-time pricing, and smart distribution system

With the integration of communication equipment and remote sensing (Zaji, Bonakdari, & Gharabaghi, 2019) into the DS, the interest in the behavior of electricity consumption has grown continuously. Many approaches have been proposed in the literature to represent and classify the consumption pattern of electricity. This work uses a statistical approach to numerically create a time series able to represent the daily behavior and PDF of energy consumption. However, other methodologies based on the combination of a time series regression with an artificial neural network (Moeeni & Bonakdari, 2017) or those based on a Markov process (Bonakdari, Zaji, Binns, & Gharabaghi, 2019) are also available.

The simulation approach used in this paper assumes the availability of a consumption profile, its correlation coefficient, and its PDF. Thus, the process described as follows can be used to synthetically generate hourly values in 1 year (Lambert, Gilman, & Lilienthal 2006):

Step 1: Select a determined node (n) classified as a demand node (PQ model for power flow). For those nodes with PV generation, go to subsection 2.3.

Step 2: Set the index of the MC simulation (c) to 1 ($c \leftarrow 1$).

Step 3: Create the time series of correlated random numbers for the corresponding node (n) and MC simulation (c). It is carried out by applying equation (1).

$$\lambda_{(t,n,c)} = \emptyset \left(\lambda_{(t-1,n,c)} \right) + \alpha_{L(t)} \ \forall \ t = 1, \ldots, 8760 \quad (1)$$

The variable $\alpha_{L(t)}$ is modeled as white noise with mean zero and standard deviation equal to $\sqrt{1 - \emptyset^2}$.

Step 4: Calculate the normalized time series of a typical load profile. A typical load profile in 1 day [$d_{(h,n)} \ \forall \ h = 1, \ldots, 24$] is shown in Fig. 1 (Labeeuw & Deconinck, 2013).

Equation (2) defines the normalized daily profile.

$$\bar{z}_{(h,n)} = \frac{d_{(h,n)} - \mu_d}{\sigma_d} \ \forall \ h = 1, \ldots, 24 \quad (2)$$

Step 5: Calculate the yearly time series of a typical load profile. It is carried out by periodically repeating the daily profile throughout the year (365 times). It results in the yearly profile $\bar{z}_{(t,n)}$ built from the series $\bar{z}_{(h,n)}$.
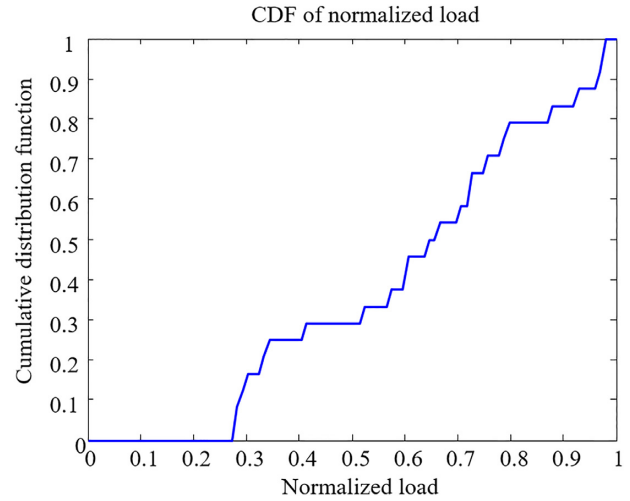
Step 6: Create a yearly time series with a determined correlation degree and with the profile shown in Fig. 1. It is carried out by adding the series $\bar{z}_{(t,n)}$ and $\lambda_{(t,n,c)}$ according to equation (3).

$$\tilde{z}_{(t,n,c)} = \bar{z}_{(t,n)} + \lambda_{(t,n,c)} \ \forall \ t = 1, \ldots, 8760 \quad (3)$$

Step 7: Apply the probability transformation shown in equation (4) to determine the load time series with the required PDF. Figure 2 presents the cumulative density function (CDF) of load time series [$F_D(\cdot)$], which was calculated from the load profile of Fig. 1.

$$D_{(t,n,c)} = F_D^{-1} \left(F_{\bar{Z}} \left(\tilde{z}_{(t,n,c)} \right) \right) \ \forall \ t = 1, \ldots, 8760 \quad (4)$$

Step 8: If $c < C$, set $c \leftarrow c + 1$, and go back to Step 3; else go to Step 9.

Step 9: Select the time series to be used in the simulation process employing the k-means++ clustering algorithm. All of the time series $D_{(t, n, c)}$ are input to the k-means++ algorithm to find a single time series [$D_{(t, n)}^{k++}$] to represent the load at node n.

Step 10: Scale the time series obtained in Step 9 [$D_{(t, n)}^{k++}$] using the capacity of distribution transformer [$D_{(n)}^{max}$], according to equation (5).

$$S_{(t,n)} = \left[D_{(n)}^{max} \right] \left[D_{(t,n)}^{k++} \right] \ \forall \ t = 1, \ldots, 8760 \quad (5)$$

An essential characteristic of SG is its capability to apply the electricity tariff on a real-time basis so that electricity prices can vary dynamically. Regarding this issue, this study assumes that the electricity prices are proportional to the corresponding electricity demand. It means that electricity prices reach their highest value when load demand is maximal, and they reach their minimum value at those hours of low demand. Accordingly, equation (6) estimates the electricity prices.

$$E_{(t)} = E_{max} \left[\sum_{n=1}^{N} \left\{ S_{(t,n)} \ \forall \ t = 1, \ldots, 8760 \right\} \middle/ \max \left(\sum_{n=1}^{N} \left\{ S_{(t,n)} \ \forall \ t = 1, \ldots, 8760 \right\} \right) \right] \quad (6)$$

The average impedance of each branch of the network, considering the conductor temperature in the range between 25 and 75°C, is a frequently used parameter to perform power flow analysis of three-phase balanced systems. Thus, equation (7) shows its definition (Short, 2004).

$$Z_{(n)} = R_{(n)} + \left(\sqrt{-1} \right) (k_{DS}) \log_{10} \left(\frac{GMD_{(n)}}{GMR_{(n)}} \right) \ \forall \ n = 1, \ldots, N \quad (7)$$

2.3. Distributed PV generation

PV generation distributed over the system is a promising manner to incorporate renewable generation in urban areas. Equations (8–21) describe a complete model of a PV generator connected to a determined node (n). The PV cell temperature depends on the ambient temperature and solar radiation, as shown in equation (8) (Lambert, Gilman, & Lilienthal, 2006). The PV cell model, including the non-linear relationship between voltage and current, is presented in equations (9–17) (Lorenzo et al., 1994).

The PV controller employs the golden section search algorithm (Kheldoun, Bradai, Boukenoui, & Mellit, 2016) applied on equation (17). Equation (18) expresses this idea through the function $f_{MPPT}(\cdot)$. The rated capacity of the power converter is presented in equation (19), while equations (20) and (21) describe its efficiency (Rampinelli, Krenzinger, & Romero 2014).

$$T_{PV(t,n)} = \frac{T_{A(t)} + G_{(t)}\left[\frac{NOCT_{(n)} - 20^\circ C}{800 \ W/m^2}\right]\left[1 - \frac{\eta_{PV(n)}\left(1 - \alpha_{PV(n)}25^\circ C\right)}{k_{PV}}\right]}{1 + \left[NOCT_{(n)} - 20^\circ C\right]\left[\frac{G_{(t)}}{800 \ W/m^2}\right]\left[\frac{\alpha_{PV(n)}\eta_{PV(n)}}{k_{PV(n)}}\right]} \tag{8}$$

$$U^e_{PV(t,n)} = \frac{m_{PV}k_B T_{PV(t,n)}}{e_{PV}} \tag{9}$$

$$U^{OC}_{PV(t,n)} = \frac{U^{OC}_{PV,STC(n)}}{N^{CS}_{PV(n)}} + \alpha_{PV(n)}\left[T_{PV(t,n)} - 25^\circ C\right]$$
$$+ U^e_{PV(t,n)}\log\left[\frac{G_{(t)}}{1000 \ W/m^2}\right] \tag{10}$$

$$u^{OC}_{PV(t,n)} = \frac{U^{OC}_{PV(t,n)}}{U^e_{PV(t,n)}} \tag{11}$$

$$FF^o_{PV(t,n)} = \frac{u^{OC}_{PV(t,n)} - \log\left(u^{OC}_{PV(t,n)} + 0.72\right)}{u^{OC}_{PV(t,n)} + 1} \tag{12}$$

$$I^{SC}_{PV(t,n)} = \frac{I^{SC}_{PV,STC(n)}}{N^{CP}_{PV(n)}}\left(\frac{G_{(t)}}{1000 \ W/m^2}\right) \tag{13}$$

$$FF_{PV(n)} = \frac{U^{MAX}_{PV(n)}I^{MAX}_{PV(n)}}{U^{OC}_{PV,STC(n)}I^{SC}_{PV,STC(n)}} \tag{14}$$

$$R^S_{PV(t,n)} = \left(1 - \frac{FF_{PV(n)}}{FF^o_{PV(t,n)}}\right)\left(\frac{U^{OC}_{PV(t,n)}}{I^{SC}_{PV(t,n)}}\right) \tag{15}$$

$$I_{PV(t,n)} = I^{SC}_{PV(t,n)}\left[1 - \exp\left(\frac{U_{PV(t,n)} - U^{OC}_{PV(t,n)} + I_{PV(t,n)}R^S_{PV(t,n)}}{U^e_{PV(t,n)}}\right)\right] \tag{16}$$

$$P_{PV(t,n)} = I_{PV(t,n)}\left[U^{OC}_{PV(t,n)} - I_{PV(t,n)}R^S_{PV(t,n)}\right.$$
$$\left. + U^e_{PV(t,n)}\log\left(1 - \frac{I_{PV(t,n)}}{I^{SC}_{PV(t,n)}}\right)\right] \tag{17}$$

$$P^{MPPT}_{PV(t,n)} = f_{MPPT}\left(P_{PV(t,n)}\right) \tag{18}$$

$$P_{I(n)} = \left[I^{MAX}_{PV(n)}U^{MAX}_{PV(n)}\right]\left[N^{PS}_{PV(n)}N^{PP}_{PV(n)}\right] \tag{19}$$

$$p_{I(t,n)} = \frac{P^{MPPT}_{PV(t,n)}\left(N^{CS}_{PV(n)}N^{CP}_{PV(n)}\right)\left(N^{PS}_{PV(n)}N^{PP}_{PV(n)}\right)}{P_{I(n)}} \tag{20}$$

$$\eta_{I(t,n)} = \frac{p_{I(t,n)}}{p_{I(t,n)} + \alpha_{I(n)} + \beta_{I(n)}\left(p_{I(t,n)}\right) + \theta_{I(n)}\left(p_{I(t,n)}\right)^2} \tag{21}$$

3. Performance of the Energy System

A long-term analysis of the optimal allocation of compensation devices requires a reliable and accurate simulation model. This section describes how the SG is studied, taking into account the influence of DG. Figure 3 summarizes the general methodology implemented for the assessment of the energy system.

According to Fig. 3, information related to the environmental variables (wind speed, ambient temperature, and solar radiation) is required. Additionally, technical data related to the SG, such as DS topology, branch impedance, load demand, and electricity price time series, are also needed. In this sense, the models previously described in Section 2 can be used to approximate some of the required data.

Another piece of critical data is the information related to the set of nodes (candidate nodes) in which a decision-maker would install compensation devices (SCB and DSTATCOM).

Assume that a decision-maker compensates a determined node (n_{RPC}). It means that the voltage at this node should be equal to the rated value of the system (one in the per-unit system). An RPC device consumes or injects reactive power from or to this node, respectively, at each hour of the year. An RPC compensation device consists of a constant (fixed) provision of reactive power and another provision of reactive power with dynamic (variable) capabilities. The constant provision could be an SCB, and the variable provision could be a DSTATCOM. Once the initial input data are available, solar radiation time series [$G_{(t)}$] is used to estimate the corresponding PV generation. Then, the SG and DS information, as well as load and electricity price time series, are incorporated into the PPF, taking into account that the voltage of the compensation node (n_{RPC}) keeps at the nominal value.

The PPF consists of sequentially solving 8760 power flow problems to consider the variations of load demand and DG generation along the time. Time series of the current at each branch, the voltage at each node, energy losses on the DS, and the reactive power consumed or injected from or to the compensation node (n_{RPC}) are results of the PPF analysis. Time series of all of these variables are used to estimate the cost of integrating RPC at the chosen node and the probability of maintaining the voltage profile and branch current within an acceptable interval. The time series of reactive power at the compensation node (n_{RPC}) is analyzed to estimate the rated value of the equipment (SCB and DSTATCOM) to be installed.

Then, the decision-maker builds the PDF of the RPC at this node. The confidence interval associated with the PDF is calculated considering a determined significance level (δ) and later used to estimate the rated capacity of the RPC device. Thus, the rated capacity consists of the constant provision of reactive power [$Q^{Fxd}_{RPC(n_{RPC})}$] and the dynamic provision [$Q^{Var}_{RPC(n_{RPC})}$], specifically.
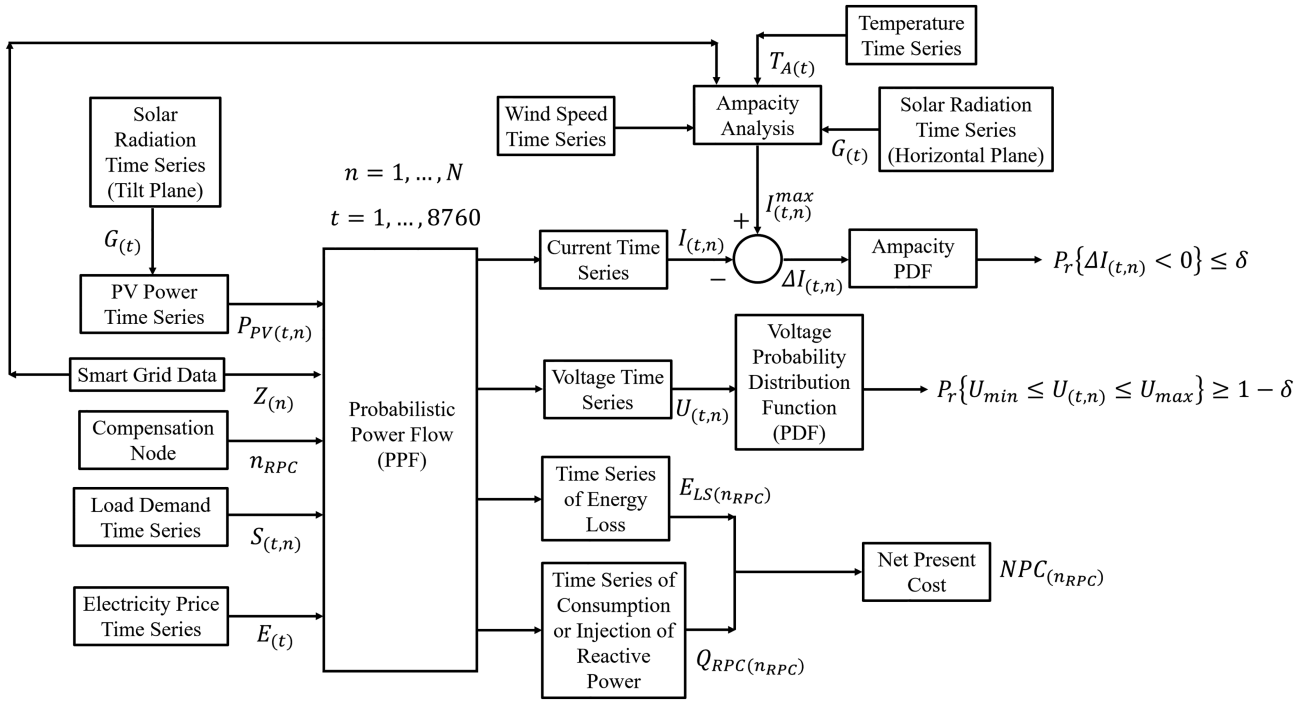
Figure 3: Structure of the simulation model.

Moreover, the time series of energy losses is used to estimate its associated costs [$E_{LS(n_{RPC})}$]. The rated capacity for compensation equipment and the cost related to energy losses (E_{LS}) are combined with other techno-economic parameters to estimate the corresponding NPC, which is the crucial factor to be minimized in this study.

Voltage time series at each node are used to build their corresponding PDFs. These PDFs are later used to evaluate the probability of maintaining the voltage within a determined interval ($P_r\{U_{min} \leq U_{(t, n)} \leq U_{max}\} \geq 1 - \delta$), considering a determined significance level (δ).

Wind speed, solar radiation, and temperature time series, as well as SG data, are used to perform the ampacity analysis. This study obtains the maximum allowed current per branch [$I_{(t, n)}^{max}$]. Then, the model computes the difference between the maximum allowed current and the actual current obtained from the PPF [$\Delta I_{(t, n)} = I_{(t, n)}^{max} - I_{(t, n)}$], to estimate the probability of overloading at each branch [$P_r\{\Delta I_{(t, n)} < 0\} \leq \delta$], considering a determined significance level (δ).

From the perspective of an optimization study, adequacy analysis for the installation of a compensation device is evaluated using the NPC as the main objective, and the probabilities $P_r\{U_{min} \leq U_{(t, n)} \leq U_{max}\} \geq 1 - \delta$ and $P_r\{\Delta I_{(t, n)} < 0\} \leq \delta$ as constraints. The next subsections explain each of these steps on the assessment of the SG.

3.1. Probabilistic power flow analysis

The solution of PPF with the load demands and DGs is an iterative process. As aforementioned, load nodes are represented as a PQ model, while DGs are voltage-controlled nodes (PV model for power flow nomenclature). PV generation is connected to DS at point of common coupling (PCC) through the power converter, which has capabilities to maintain the voltage at the connection node at the rated value (IEEE Standards Coordinating Committee 21, 2018).

For each hour of the year ($t = 1, \ldots, 8760$) and node of the system ($n = 1, \ldots, N$), the corresponding apparent power [$S_{(t, n)}$], active power [$P_{(t, n)}$], and reactive power [$Q_{(t, n)}$] are available in the format shown in equation (22).

$$S_{(t,n)} = P_{(t,n)} + \left(\sqrt{-1}\right) Q_{(t,n)} \ \forall \ t = 1, \ldots, 8760; n = 1, \ldots, N \quad (22)$$

In the case of load demand nodes, the model used here calculates the active and reactive power through equation (22) using the time series developed in subsection 2.2 with a determined power factor. In the case of PV generation, the simulation model estimates the active and reactive power combining the output power of PV generation according to subsection 2.3 and the reactive power required to maintain the voltage at PCC at its rated value. The computational procedure to calculate the reactive power of PV nodes is available in Teng, (2003, 2008) and Lujano-Rojas, Dufo-López, Bernal-Agustín, Domínguez-Navarro, and Catalão (2018).

The optimization model solves 8760 load flow problems, obtaining the vectors of voltage and current [$\vec{U}_{(t)}$ and $\vec{I}_{(t)}$] that define the time series of the voltage of each node and current of each branch of the system. Equation (23) defines the form of such vectors.

$$\vec{U}_{(t)} = \begin{bmatrix} U_{(t,1)} \\ \vdots \\ U_{(t,n)} \\ \vdots \\ U_{(t,N)} \end{bmatrix} ; \ \vec{I}_{(t)} = \begin{bmatrix} I_{(t,1)} \\ \vdots \\ I_{(t,n)} \\ \vdots \\ I_{(t,N)} \end{bmatrix} \ \forall \ t = 1, \ldots, 8760 \quad (23)$$

3.2 Ampacity analysis

The maximum admissible current through a determined DS feeder [$I_{(t, n)}^{max}$] can be approximated using the general heat equation of the conductor of interest under steady-state conditions.

This process results in the model shown in equation (24) (Deb, 2000).

$$I_{(t,n)}^{max} = \sqrt{\frac{P_{R(t,n)} + P_{C(t,n)} - P_{S(t,n)}}{R_{(n)}}} \; \forall \, t = 1, \ldots, 8760; n = 1, \ldots, N \tag{24}$$

In this sense, heat loss by radiation depends on temperature, as suggested in equation (25). Heat loss by convection depends on the temperature and wind speed, as suggested in equation (26), while heat gain depends on the solar radiation, as suggested in equation (27) (Deb, 2000).

$$P_{R(t,n)} = f_R \; \left(T_{A(t)}\right) \; \forall \, t = 1, \ldots, 8760; n = 1, \ldots, N \tag{25}$$

$$P_{C(t,n)} = f_C \; \left(T_{A(t)}, V_{(t)}\right) \; \forall \, t = 1, \ldots, 8760; n = 1, \ldots, N \tag{26}$$

$$P_{S(t,n)} = f_S \; \left(G_{(t)}\right) \; \forall \, t = 1, \ldots, 8760; \; n = 1, \ldots, N \tag{27}$$

Specific details about the estimation model represented by the functions $f_R(\cdot)$, $f_C(\cdot)$, and $f_S(\cdot)$ can be found in Deb (2000).

3.3. Net present cost estimation

Estimating NPC depends on many critical techno-economic factors, such as inflation and discount rates and project lifetime, as well as capital replacement and maintenance costs. Capital and replacement costs are directly related to the capacity of the RPC equipment to be installed, while maintenance costs are related to the costs associated with the energy losses on the DS.

The model calculates the fixed and variable capacities of the RPC device [$Q_{RPC(n)}^{Fxd}$ and $Q_{RPC(n)}^{Var}$] from the PDF of reactive power provision at the node of interest ($n = n_{RPC}$). It means that the reactive power to be provided to this node, to maintain its voltage at the nominal value, has to be calculated for each hour of the year.

Otherwise, the model estimates the effect of the RPC (fixed and variable) on the DS by modeling this device as a voltage-controlled node (PV model for power flow nomenclature) able to provide reactive power only.

The approximation of the required amount of reactive power to be injected or drawn from the system is an iterative process. The main result consists of a vector of reactive power injected or drawn to the installation node [$Q_{(t,\, n_{RPC})} \; \forall \, t = 1, \ldots, 8760$]. Then, the model calculates the rated capacity of the compensation device using the CDF. The probabilistic technique accurately estimates the rated capacity from the most extreme value observed on the confidence interval of the PDF, considering a specific significance level (δ). The technique calculates the confidence interval by evaluating the CDF of $Q_{(t,\, n_{RPC})}$ inversely in the points $\delta/2$ and $1 - \delta/2$, to obtain the lowest and highest values of the interval, respectively.

Figure 4 briefly shows this idea, where the model uses the histogram of frequencies of reactive power $Q_{(t,\, n)}$ at the installation node ($n = n_{RPC}$) to calculate the CDF. The technique estimates the highest value of the confidence interval (Q_C^r) by evaluating the corresponding CDF curve at $1 - \delta/2$ in an inverse manner. Similarly, the approach calculates the lowest value of the confidence interval (Q_C^l) by evaluating the CDF curve at the point $\delta/2$ inversely. Finally, it defines the rated capacity of each RPC device [$Q_{RPC(n_{RPC})}^{Fxd}$ and $Q_{RPC(n_{RPC})}^{Var}$] according to equations (28)
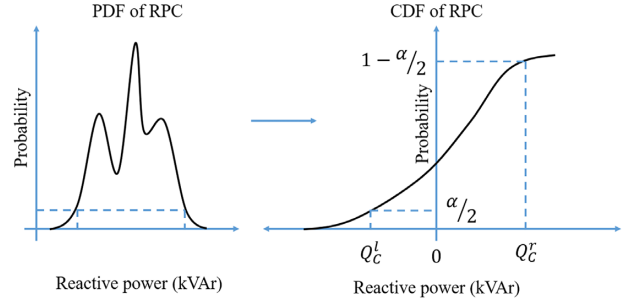

Figure 4: Determining rated capacity of the RPC from the CDF.

and (29).

$$Q_{RPC(n_{RPC})}^{Fxd} = \; min \left(\left|Q_C^l\right|, \left|Q_C^r\right|\right) \tag{28}$$

$$Q_{RPC(n_{RPC})}^{Var} = \; max \left(\left|Q_C^l\right|, \left|Q_C^r\right|\right) - min \left(\left|Q_C^l\right|, \left|Q_C^r\right|\right) \tag{29}$$

This approach allows us to determine the rated capacity of the device for a given node. Thus, exclusively focusing on determining the proper node to install the corresponding device is needed.

The cost of energy losses [$E_{LS(n_{RPC})}$] related to the compensation of the node $n = n_{RPC}$ can be estimated by using the energy losses obtained from the PPF and the electricity price time series [$E_{(t)}$]. Finally, the NPC [$NPC_{(n_{RPC})}$] of reactive power compensation at node $n = n_{RPC}$ is calculated using equations (30–33).

$$ACC_{(n)} = ACC_{(n)}^{Fxd} + ACC_{(n)}^{Var} \tag{30}$$

$$ARC_{(n)} = ARC_{(n)}^{Fxd} + ARC_{(n)}^{Var} \tag{31}$$

$$AMC_{(n)} = AMC_{(n)}^{Fxd} + AMC_{(n)}^{Var} \tag{32}$$

$$NPC_{(n)} = \frac{\sum_{n_{RPC}} \left\{ACC_{(n)} + ARC_{(n)} + AMC_{(n)}\right\} + E_{LS(n)}}{CRF} \tag{33}$$

3.4 Probabilistic voltage profile analysis

As previously mentioned, the optimization model evaluates the quality of the voltage profile by using the probabilistic constraint shown in equation (34).

$$P_r\left\{U_{min} \leq U_{(t,n)} \leq U_{max}\right\} \geq 1 - \delta \; \forall \, n = 1, \ldots, N; \; t = 1, \ldots, 8760 \tag{34}$$

The optimization technique uses the results obtained from the PPF related to the voltage of the DS at each node. Thus, it builds the CDF of voltage at a determined node (n) using $U_{(t,\, n)}$ time series.

Then, the approach computes the value $P_r\{U_{min} \leq U_{(t,\, n)} \leq U_{max}\}$ as the subtraction between $P_r\{U_{(t,\, n)} \leq U_{max}\}$ and $P_r\{U_{(t,\, n)} \leq U_{min}\}$.

3.5 Probabilistic feeder ampacity analysis

The optimization method uses the results from the PPF for the probabilistic assessment of feeders overloading. This evaluation is carried out by verifying equation (35).

$$P_r\left\{\Delta I_{(t,n)} < 0\right\} \leq \delta \; \forall \, n = 1, \ldots, N; \; t = 1, \ldots, 8760 \tag{35}$$

$$\Delta I_{(t,n)} = I_{(t,n)}^{max} - I_{(t,n)} \; \forall \, n = 1, \ldots, N; t = 1, \ldots, 8760 \tag{36}$$

The proposed technique employs the branch current time series $[I_{(t, n)}]$ and the maximum permissible current time series $[I_{(t, n)}^{max}]$ to calculate the margin time series $[\Delta I_{(t, n)}]$, as shown in equation (36). Then, it builds the CDF of the margin $[\Delta I_{(t, n)}]$ and the probability value $P_r\{\Delta I_{(t, n)} < 0\}$ using a linear interpolation algorithm.

4. Proposed Methodology

This section thoroughly explains the proposed methodology. First, in subsection 4.1, the formulation of the optimization problem is carried out in a probabilistic manner so that the random nature of renewable resources is adequately justified. Then, in subsection 4.2, the solution to this problem using a GA is comprehensively described.

4.1. Optimization problem

Optimal allocation and sizing of the RPC devices consists of minimizing equation (37),

$$min\{NPC_{(n)}\} \ \forall n = n_{RPC}, \tag{37}$$

subject to equations (38–42).

$$P_r\{U_{min} \le U_{(t,n)} \le U_{max}\} \ge 1 - \delta \ \forall n = 1, \ldots, N; t = 1, \ldots, 8760 \tag{38}$$

$$P_r\{\Delta I_{(t,n)} < 0\} \le \delta \ \forall n = 1, \ldots, N; t = 1, \ldots, 8760 \tag{39}$$

$$\Delta I_{(t,n)} = I_{(t,n)}^{max} - I_{(t,n)} \ \forall n = 1, \ldots, N; t = 1, \ldots, 8760 \tag{40}$$

$$\overrightarrow{\Delta U}_{(t)} = [BCBV] \ [BIBC] \ \vec{I}_{(t)} \ \forall t = 1, \ldots, 8760 \tag{41}$$

$$\vec{U}_{(t)} = \vec{U}_R - \overrightarrow{\Delta U}_{(t)} \ \forall t = 1, \ldots, 8760 \tag{42}$$

Equation (37) represents the minimization of the NPC of the RPC equipment. Equations (38–40) are the constraints of the voltage profile and feeder ampacity. Constraint equations (41) and (42) represent the energy balance of the system obtained by solving power flow at each time step ($t = 1, \ldots, 8760$).

The next subsection describes the proposed GA implementation.

4.2. GA implementation

Researchers apply heuristic techniques in many engineering fields such as vehicle performance improvement (Morteza & Yildiz, 2016), manufacturing process optimization (Yildiz, 2013), and the cam-roller follower mechanism design (Hamza, Abderazek, Lakhdar, Ferhat, & Yildiz, 2018), among other applications. The optimization method developed in this paper considers all of the available places in the system to the allocation of the RPC equipment. The proposed approach organizes the information in a table similar to that presented in Fig. 5, considering B places for the allocation of RPC devices. The proposed methodology uses a binary-coded GA, where each individual has B bits. This table establishes the relationship between each bit and the DS node. Using the example shown in Fig. 5, if the second bit ($b = 2$) is equal to 1, it means that the installation of a compensation device on node 20 ($n_{RPC} = 20$)

Bit of GA Individual	Candidate Node for Compensation
1	16
2	20
⋮	⋮
b	23
⋮	⋮
B	54

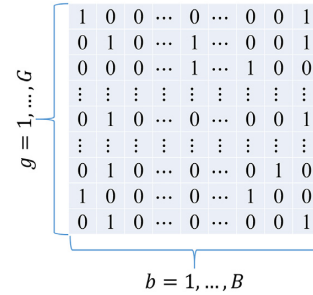**Figure 5:** Example of a table of candidate nodes.

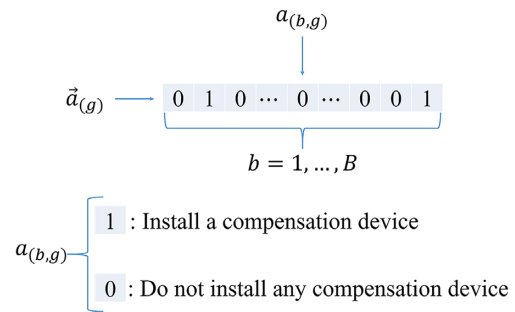**Figure 6:** Scheme of the GA population.

Figure 7: Scheme of the GA individual.

is analyzed. On the contrary, if the second bit is equal to zero, the installation of a compensation device in that node is not considered.

Figure 6 presents an illustration of the GA population (A) with G individuals and B possible places for the installation of compensation equipment. Figure 7 illustrates the structure of the individual g, which considers the compensation of two different places on the system simultaneously.

The optimization algorithm calculates the objective function of a determined individual $[\vec{a}_{(g)}]$ by following the algorithm explained as follows:

Step 1: Considering those bits of the GA individual $[\vec{a}_{(g)}]$ equal to 1 and the corresponding table of candidate nodes (Fig. 5), perform the PPF (subsection 3.1) and ampacity analysis (subsection 3.2), and estimate the capacity of the RPC equipment and the NPC (subsection 3.3). Equation (43) expresses the NPC of the individual under analysis.

$$NPC_{GA(g)} \leftarrow NPC_{(n)} \ \forall n = n_{RPC} \ \forall g = 1, \ldots, G \tag{43}$$

Step 2: Set $n \leftarrow 1$.

Step 3: The CDF of the system voltage and current at each node and branch (n) are built to evaluate the inequality equations (38) and (39). If the constraint equation (38) is false or the constraint equation (39) is false, the respective element of the
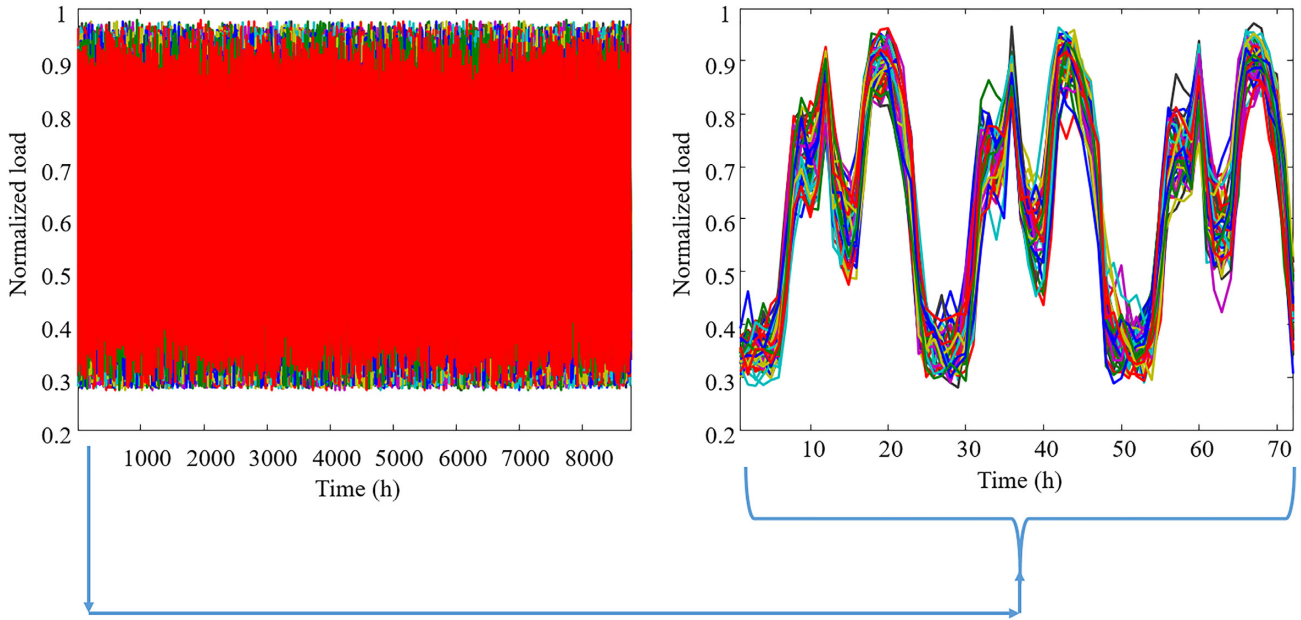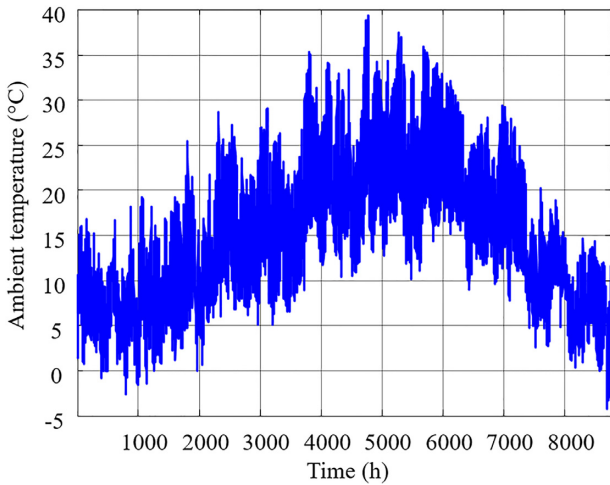
Figure 8: Load demand time series.

Figure 9: Temperature time series.

cedures) appear in the technical literature; more information is available in Goldberg (1989).

5. Case Study

The technique proposed in this work has been analyzed using an extensive simulation process considering a hypothetical DS. The next subsections thoroughly describe the testing conditions as well as the corresponding results.

5.1. Case description

A DS to be located in Zaragoza (Spain) illustrates the proposed methodology in this paper. The simulation model generated the load demand and electricity price time series using the method previously described in subsection 2.2. For a better understanding, Fig. 8 shows the yearly time series of load demand $[D_{(t, n)}^{k++}]$ on the left side and the first 72 hours of the year on the right side. It is possible to observe how the generated time series have the hourly trend of the profile previously shown in Fig. 1.

The typical meteorological year (TMY) for solar radiation, ambient temperature, and wind speed were found on the website Renewables.ninja (2019) for the area of interest (latitude equal to 41.84° and longitude equal to −1.1695°).

Figures 9 and 10 present the TMY for ambient temperature and wind speed, respectively. Figure 11 shows the TMY of solar radiation for 0° and 36° tilt. The TMY of ambient temperature, wind speed, and solar radiation with 0° tilt are used to evaluate the ampacity of each distribution feeder (subsection 3.2). The TMY of solar radiation for 36° tilt is used to evaluate the PV power production; according to Jacobson and Jadhav (2018), this is the optimal tilt angle for the location under study.

Figure 12 shows the structure of the DS, while Table 1 presents the corresponding information related to the conductor type, feeder length, and the location of the DG. The rated voltage was assumed to be 12.47 kV operating at 60 Hz; the tolerance of PPF analysis was 0.001, the maximum number of iterations was 16, and the general power factor was 90%.

constraint vector becomes 1 $[k_{CS(n)} \leftarrow 1]$. On the contrary case, the respective element becomes zero $[k_{CS(n)} \leftarrow 0]$.

Step 4: If $(n < N)$, set $n \leftarrow n + 1$, and go to Step 3; else go to Step 5.

Step 5: Calculate the value of the objective function $[Obj_{GA(n)}]$ of individual g using equation (44).

$$Obj_{GA(g)} = NPC_{GA(g)} \left(1 + \sum_{n=1}^{N} k_{CS(n)}\right) \forall g = 1, \ldots, G \quad (44)$$

The objective function equation (44) becomes higher as the operational feasibility of the solution reduces. In other words, when the constraint equations (38) and (39) are not valid in some nodes or branches, the value of the objective function in equation (44) increases, reducing the fitness of the individual.

The rest of the steps related to the optimization with a GA using binary coding (reproduction, crossing, and mutation pro-
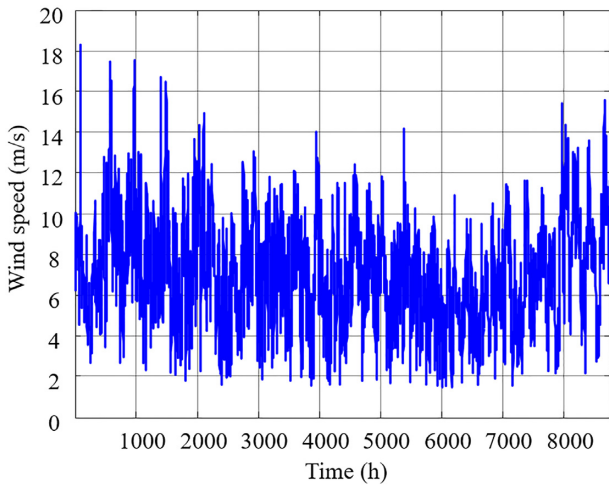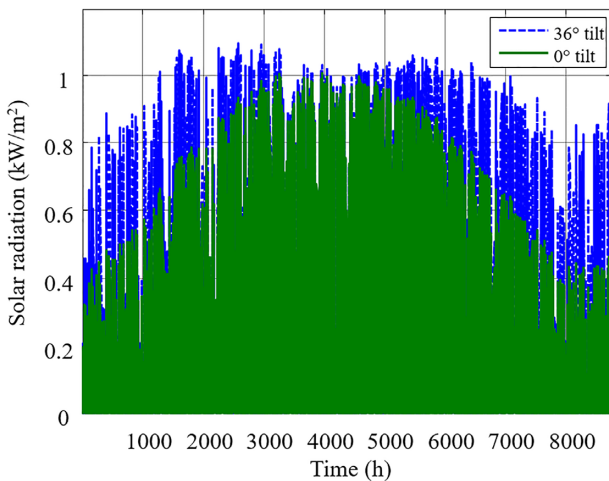
Figure 10: Wind speed time series.

Figure 11: Solar radiation time series.

The substation capacity was 4000 kVA. All of the conductors of the DS were assumed to be an all-aluminum conductor type, while the structure of each feeder was assumed to be built considering $GMD_{(n)} = 1.26$ m. The types of nodes considered were load, generation, and compensation. The load nodes are power consumption with constant power factor, while the compensation nodes are places where the decision-maker would install RPC equipment (candidate nodes). Compensation and generation nodes are voltage-controlled nodes to maintain the voltage at PCC to 12.47 kV.

Table 2 presents the data and characteristics of the PV generators used to estimate the power production of the farms connected to nodes 18 and 49.

Table 3 shows the table of candidate nodes (Fig. 5) for the system under analysis. Table 1 was used to build this table, considering the node type. This table presents the relationship between GA individuals and the candidate nodes available for the installation of the RPC devices. According to the number of rows in this table, the number of bits of each individual is set to 20 ($B = 20$), obtaining $2^{20} = 1048576$ possible combinations for the installation of the RPC.

The binary-coded GA with the objective function described in subsection 4.2 was implemented considering 100 generations,

50 individuals in the population ($G = 50$), a crossover rate of 95%, and a mutation rate of 5%. The maximum value of electricity prices considered is 110€/MWh ($E_{max} = 110$€/MWh). The calculated NPC assumes a discount rate of 5% and an inflation rate of 3.5%.

The example considered a project lifetime of 25 years. Capital costs for fixed and variable RPC equipment were assumed to be 50€/kVA and 500€/kVA, respectively, and their lifetime was 15 years. The simulation has been performed in MATLAB® using a computer with an i7-3630QM CPU at 2.40 GHz with 8 GB of memory and 64-bit operating system.

5.2. Optimization results

Table 4 presents the rated values of the fixed [$Q_{RPC(n)}^{Fxd}$] and variable [$Q_{RPC(n)}^{Var}$] compensation devices. These values were estimated considering a significance level of 5% ($\delta = 0.05$).

The computational time spent was 58 hours, approximately. The iterative process needed to estimate the reactive power to be injected or drawn at each compensation node and time step requires a considerable amount of computational resources. Nevertheless, the proposed approach can obtain a reliable solution with a probabilistic basis.

Figure 13 describes the behavior of the optimization algorithm, which converges to the values previously shown in Table 4. Figures 14 and 15 present the PDF and CDF of the RPC, respectively. The values of reactive power are negative, which means that reactive power is injected into the system most of the time. As shown in Fig. 14, a continuous provision of reactive power is required. This amount of reactive power (463.39 kVAr) is supposed to be supplied by a device similar to an SCB. On the other hand, the hourly fluctuations of the reactive power, mainly related to the diurnal variations of load demand, are supposed to be provided by a device with dynamic capabilities, similar to a DSTATCOM of 1464.69 kVAr.

Table 5 presents the probability constraint equations (38) and (39) considering a range of ±5% of the rated voltage of the system ($U_{min} = 95\%$ and $U_{max} = 105\%$). As can be observed, all of the probabilities of fulfilling the constraint equation (38) are higher than 95% (with a significance level of 5%) when the RPC devices are installed. Initially, the DS has problems at those nodes located far from the substation. Nodes 62, 64, and 65 have low probabilities of fulfilling the required operating conditions. These nodes have probabilities of 94.908, 90.325, and 87.594%, respectively (i.e. all of them are lower than 95%, with a significance level of 5%).

However, when the decision-maker evaluates the installation of RPC in node 8, the corresponding condition becomes valid for all the nodes of the system. On the other hand, the probability of overloading remains negligible because the wind speed is higher than zero most of the time, which increases the ampacity of each distribution feeder.

5.3. Proposed versus traditional GA implementation

This subsection presents a comparison between the proposed and the traditional GA implementation. Complete details related to the implementation of a traditional GA are given in Appendix A. Table 6 presents the rated capacity of the RPC at each node of the system, while Table 7 shows a comparative analysis between the total RPC to be added to the system according to the method implemented.

From Tables 4 and 6, the proposed methodology only suggests the integration of the RPC in specific nodes, while the traditional methodology suggests the integration of compensation
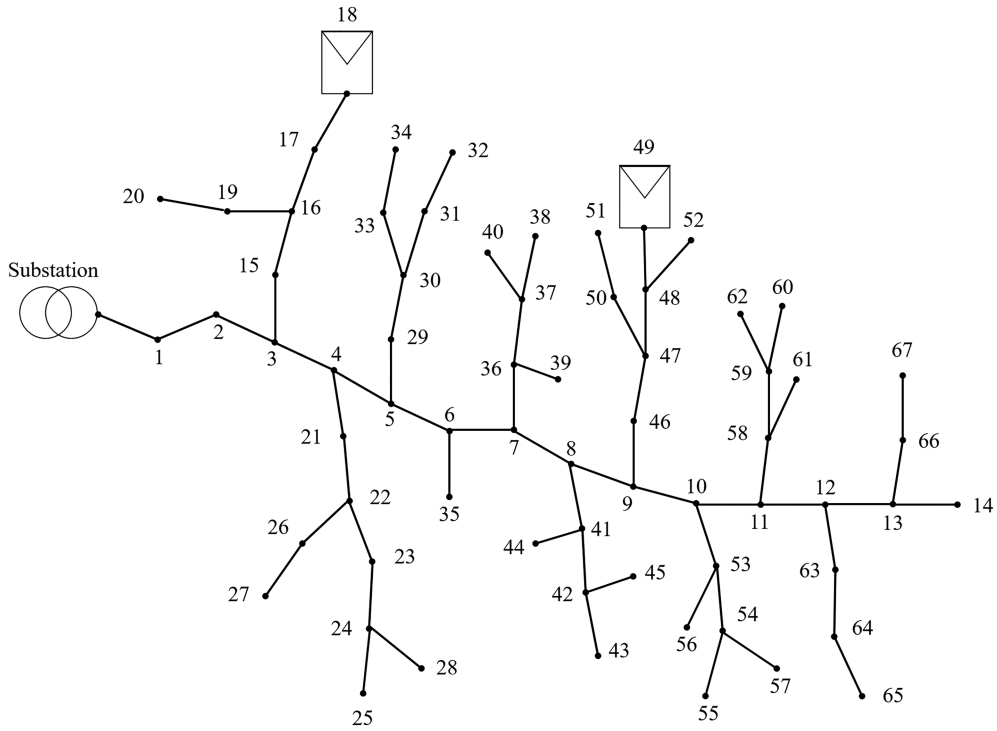
Figure 12: Distribution system under study.

Table 1: Distribution system data.

Sending node	Receiving node	Node type	Mathematical model	Size (kcmil)	Strands	Length (km)	Rating voltage (kV)	Transformer capacity (kVA)
Substation	1	Load	PQ	300	19	2.583		75
1	2	Load	PQ	300	19	1.461		75
2	3	Load	PQ	300	19	2.731		75
3	4	Load	PQ	266.8	7	2.842		75
4	5	Compensation	PV	250	7	2.139	12.47	
5	6	Load	PQ	250	7	1.163		50
6	7	Load	PQ	250	7	1.381		75
7	8	Compensation	PV	250	7	2.721	12.47	
8	9	Load	PQ	250	7	1.203		75
9	10	Compensation	PV	83.69	7	1.903	12.47	
10	11	Load	PQ	41.74	7	1.496		75
11	12	Compensation	PV	26.24	7	2.328	12.47	
12	13	Load	PQ	26.24	7	1.081		75
13	14	Compensation	PV	26.24	7	2.215	12.47	
3	15	Load	PQ	26.24	7	2.710		75
15	16	Compensation	PV	26.24	7	2.434	12.47	
16	17	Load	PQ	26.24	7	2.129		50
17	18	Generation	PV	26.24	7	1.644	12.47	
16	19	Load	PQ	26.24	7	2.322		75
19	20	Load	PQ	26.24	7	2.948		50
4	21	Load	PQ	26.24	7	1.936		75
21	22	Load	PQ	26.24	7	2.713		75
22	23	Load	PQ	26.24	7	1.710		50
23	24	Load	PQ	26.24	7	1.125		25
24	25	Compensation	PV	26.24	7	1.631	12.47	
22	26	Compensation	PV	26.24	7	2.306	12.47	
26	27	Load	PQ	26.24	7	3.126		75
24	28	Load	PQ	26.24	7	1.802		25
5	29	Load	PQ	26.24	7	2.033		15
29	30	Compensation	PV	26.24	7	1.489	12.47	

Table 1: Continued

Sending node	Receiving node	Node type	Mathematical model	Size (kcmil)	Strands	Length (km)	Rating voltage (kV)	Transformer capacity (kVA)
30	31	Load	PQ	26.24	7	1.359		75
31	32	Load	PQ	26.24	7	2.496		50
30	33	Load	PQ	26.24	7	1.411		75
33	34	Compensation	PV	26.24	7	1.851	12.47	
6	35	Compensation	PV	26.24	7	3.046	12.47	
7	36	Load	PQ	26.24	7	1.601		75
36	37	Load	PQ	26.24	7	2.205		75
37	38	Compensation	PV	26.24	7	1.866	12.47	
36	39	Load	PQ	26.24	7	2.326		75
37	40	Load	PQ	26.24	7	1.914		75
8	41	Compensation	PV	26.24	7	2.082	12.47	
41	42	Load	PQ	26.24	7	2.147		25
42	43	Load	PQ	26.24	7	3.011		75
41	44	Load	PQ	26.24	7	3.115		75
42	45	Load	PQ	26.24	7	3.070		25
9	46	Compensation	PV	26.24	7	1.529	12.47	
46	47	Load	PQ	26.24	7	1.471		50
47	48	Load	PQ	26.24	7	1.536		75
48	49	Generation	PV	26.24	7	2.518	12.47	
47	50	Compensation	PV	26.24	7	2.612	12.47	
50	51	Load	PQ	26.24	7	1.869		75
48	52	Load	PQ	26.24	7	1.243		25
10	53	Load	PQ	26.24	7	1.702		50
53	54	Load	PQ	26.24	7	1.483		75
54	55	Load	PQ	26.24	7	2.801		75
53	56	Compensation	PV	26.24	7	2.359	12.47	
54	57	Load	PQ	26.24	7	2.256		75
11	58	Compensation	PV	26.24	7	1.438	12.47	
58	59	Load	PQ	26.24	7	1.892		50
59	60	Load	PQ	26.24	7	2.425		75
58	61	Compensation	PV	26.24	7	1.794	12.47	
59	62	Load	PQ	26.24	7	2.770		75
12	63	Compensation	PV	26.24	7	2.656	12.47	
63	64	Load	PQ	26.24	7	2.520		50
64	65	Load	PQ	26.24	7	2.139		75
13	66	Load	PQ	26.24	7	1.973		15
66	67	Compensation	PV	26.24	7	1.612	12.47	

Table 2: Characteristics of PV panels.

Parameter	$n = 18$	$n = 49$
$NOCT_{(n)}$ (°C)	47.9	47.9
m_{PV}	1	1
$U^{OC}_{PV,STC(n)}$ (V)	30.3	30.3
$I^{SC}_{PV,STC(n)}$ (A)	8.06	8.06
$\alpha_{PV(n)}$ (%/°C)	−0.377	−0.377
$\eta_{PV(n)}$ (%)	20	20
$U^{MAX}_{PV(n)}$ (V)	27.79	27.79
$I^{MAX}_{PV(n)}$ (A)	9.3	9.3
$N^{CS}_{PV(n)}$	48	48
$N^{CP}_{PV(n)}$	1	1
$N^{PP}_{PV(n)}$	40	60
$U^{SYS}_{PV(n)}$ (V)	600	600

devices spread over the system. On the other hand, the traditional method suggests the incorporation of less compensation capacity than the proposed technique.

At this point, the most important differences between both techniques influence the results. In the proposed methodology, adding or subtracting reactive power fixes the voltage of a determined node of the system at the rated value. Thus, this action impacts the behavior of the rest of the system. In the traditional approach, the addition or subtraction of reactive power reduces the energy losses of the system, maintaining the voltages within a determined operating interval, but not necessarily at the rated value all the time.

On the one hand, the proposed approach considers the local effects of reactive power injection of DG at the PCC, and the interaction between the load demand and DG. On the other hand, the traditional method considers the DG's effects in a centralized manner, using the estimation of the net load. In a general sense, the estimation of the net load results in the analysis of a DS with lower demand than in the case of the proposed approach, resulting in a lower estimation of the RPC and a difference of 24.9%.

Nevertheless, the formulation of both methods has necessary implications for the structure of the optimization problem. The proposed technique depends only on the number of candidate nodes, while the traditional formulation depends on several aspects; these aspects are the number of blocks of load duration curve (LDC), the number of candidate nodes, and the integer magnitudes of reactive power.

Table 3: GA individuals and candidate nodes.

Bit of GA individual	Candidate node for compensation
1	5
2	8
3	10
4	12
5	14
6	16
7	25
8	26
9	30
10	34
11	35
12	38
13	41
14	46
15	50
16	56
17	58
18	61
19	63
20	67

Table 4: RPC planning.

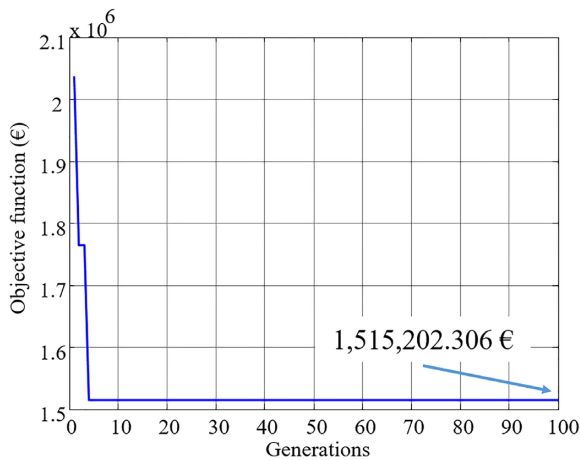| $n = n_{RPC}$	$Q_{RPC(n)}^{Fxd}$ (kVAr)	$Q_{RPC(n)}^{Var}$ (kVAr)	$NPC_{(n)}$ (€)	Time (min)
8	463.39	1464.69	1 515 202.306	3501.22

Figure 13: Convergence of a GA.

The reactive power injected to the system per load block and node [$CAP_{(l, b)} \forall l = 1, \ldots, 12; b = 1, \ldots 20$] was represented as an integer in the interval $[-114, 0]$ using 115 steps. Under this assumption, the number of possible combinations is $115^{(O=BL)} = 115^{240}$. From this reasoning, the proposed methodology considerably reduces the number of possible combinations of the optimization problem and its computational complexity (the proposed method has only 2^{20} possible combinations). However, the traditional implementation requires only the solution of $L = 12$ power flow problems for a single individual, while the proposed method requires the solution of 8760, increasing it by 26.8%.
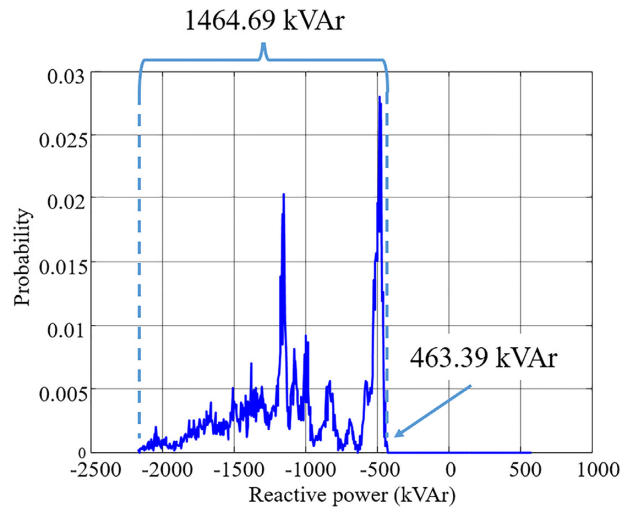

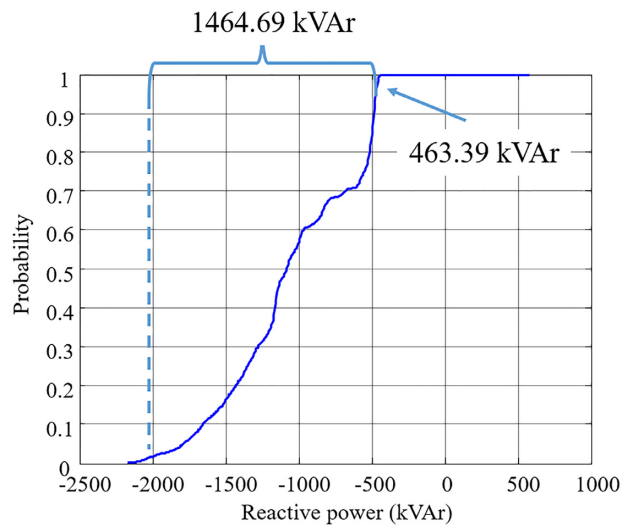**Figure 14:** The PDF of the RPC at node 8.

Figure 15: The CDF of the RPC at node 8.

6. Conclusions

In this paper, a methodology for the optimal placement and sizing of RPC devices in a DS with PV generation has been presented and tested through an extensive computational analysis.

The proposed methodology reduces the computational complexity when compared to the traditional GA implementation based on the discretization of the LDC, because it considers only binary variables. However, the evaluation of each GA individual of the proposed approach requires the solution of PPF, which means solving 8760 deterministic power flow calculations. Conversely, the evaluation of a GA individual of the traditional implementation depends on the number of discretization intervals used to build the LDC. Consequently, the proposed approach requires 26.8% more computational time than the classic version, according to the analyzed case. This issue could be overcome by implementing parallel computing solutions or by reducing the number of data points of input time series using a clustering algorithm. Another significant problem of the proposed approach is its scalability because the computational time also increases with the number of DS nodes.

Table 5: Analysis of voltage and current limits.

Receiving node or branch	Voltage limits		Current limits	
	Without compensation	With compensation	Without compensation	With compensation
1	1	1	0	0
2	1	1	0	0
3	1	1	0	0
4	1	1	0	0
5	1	1	0	0
6	1	1	0	0
7	1	1	0	0
8	1	1	0	0
9	1	1	0	0
10	1	1	0	0
11	1	1	0	0
12	0.97642694	1	0	0
13	0.97223174	1	0	0
14	0.97223174	1	0	0
15	1	1	0	0
16	1	1	0	0
17	1	1	0	0
18	1	1	0	0
19	1	1	0	0
20	1	1	0	0
21	1	1	0	0
22	1	1	0	0
23	1	1	0	0
24	1	1	0	0
25	1	1	0	0
26	1	1	0	0
27	1	1	0	0
28	1	1	0	0
29	1	1	0	0
30	1	1	0	0
31	1	1	0	0
32	1	1	0	0
33	1	1	0	0
34	1	1	0	0
35	1	1	0	0
36	1	1	0	0
37	1	1	0	0
38	1	1	0	0
39	1	1	0	0
40	1	1	0	0
41	1	1	0	0
42	1	1	0	0
43	1	1	0	0
44	1	1	0	0
45	1	1	0	0
46	1	1	0	0
47	1	1	0	0
48	1	1	0	0
49	1	1	0	0
50	1	1	0	0
51	1	1	0	0
52	1	1	0	0
53	1	1	0	0
54	0.99292237	1	0	0
55	0.97514269	1	0	0
56	1	1	0	0
57	0.97751142	1	0	0
58	0.99531963	1	0	0
59	0.96906393	1	0	0

Table 5: Continued

Receiving node or branch	Voltage limits		Current limits	
	Without compensation	With compensation	Without compensation	With compensation
60	0.9524258	1	0	0
61	0.99531963	1	0	0
62	**0.94908676**	**1**	0	0
63	0.95316781	1	0	0
64	**0.90325342**	**1**	0	0
65	**0.87594178**	**1**	0	0
66	0.97106164	1	0	0
67	0.97106164	1	0	0

Table 6: RPC sizing and placing analysis (traditional method).

Candidate node	Rated RPC (kVAr)
5	66
8	93
10	81
12	55
14	105
16	71
25	70
26	53
30	97
34	69
35	82
38	88
41	34
46	78
50	106
56	64
58	36
61	79
63	83
67	38

Table 7: Comparison of total RPC.

Method	Rated RPC (kVAr)	Time (min)
Proposed	1928.099	3501.224
Traditional	1448	2561.801
Difference (%)	24.9	26.8

Although the proposed technique requires considerable computational resources, it can provide a reliable solution with a probabilistic basis. In other words, traditional GA implementation offers a general approximation of a required RPC; this method calculates it from the discretized LDC (Table A1 of the appendix). Relevant characteristics, such as autocorrelation and daily profile behavior, are not considered. Conversely, the proposed approach considers the hourly interaction between the load demand (Fig. 8) and renewable resources (Figs 9–11), resulting in a 24.9% increase of the RPC's capacity, when compared with the classical approach. Additionally, PPF considers the connection node of the RPC on the DS in each evaluation of GA individuals. These features offer higher accuracy and a solution

with high robustness because the PDF of the RPC is calculated (Figs 14 and 15).

Another disadvantage relies on the quality of the time series simulation method of environmental variables and electricity demand because the proposed approach is highly dependent on this data. In this regard, recently developed approaches related to big data analytics applied to the SG can be useful to estimate the time series of the load and the environmental variables of interest. Besides, a more in-depth study on the relationship between the behavior of load demand and electricity prices (equation 6) should be performed to increase the feasibility of the NPC estimation.

Supplementary Data

Supplementary data are available at JCDENG online.

Acknowledgement

Conflict of interest statement

Declarations of interest: none.

References

Al-Juboori, M., & Datta, B. (2019). Optimum design of hydraulic water retaining structures incorporating uncertainty in estimating heterogeneous hydraulic conductivity utilizing stochastic ensemble surrogate models within a multi-objective multi-realisation optimization model. *Journal of Computational Design and Engineering*, 6, 296–315.

Azzam, M., & Mousa, A. A. (2010). Using genetic algorithm and TOPSIS technique for multiobjective reactive power compensation. *Electric Power Systems Research*, 80, 675–681.

Baptista, S., Barbosa-Povoa, A. P., Escudero, L. E., Gomes, M. I., & Pizarro, C. (2019). On risk management of a two-stage stochastic 0–1 model for the closed-loop supply chain design problem. *European Journal of Operational Research*, 274(1), 91–107.

Bertsimas, D., Gupta, V., & Kallus, N. (2018). Robust sample average approximation. *Mathematical Programming*, 171(1–2), 217–282.

Bonakdari, H., Zaji, A. H., Binns, A. D., & Gharabaghi, B. (2019). Integrated Markov chains and uncertainty analysis techniques to more accurately forecast floods using satellite signals. *Journal of Hydrology*, 572, 75–95.

Buchheim, C., & Pruente, J. (2019). *K*-adaptability in stochastic combinatorial optimization under objective uncertainty. *European Journal of Operational Research*, 277, 953–963.

Das, D. (2002). Reactive power compensation for radial distribution networks using genetic algorithm. *International Journal of Electrical Power and Energy Systems*, 24(7), 573–581.

Das, S., Das, D., & Patra, A. (2017). Reconfiguration of distribution networks with optimal placement of distributed generations in the presence of remote voltage controlled bus. *Renewable and Sustainable Energy Reviews*, 73, 772–781.

Deb, A. K. (2000). *Power line ampacity system: Theory, modeling, and applications*. (1st ed.). Boca Raton, Florida: CRC Press LLC.

Esfahani, P. M., & Kuhn, D. (2018). Data-driven distributionally robust optimization using the Wasserstein metric: Perfor-

mance guarantees and tractable reformulations. *Mathematical Programming*, 171(1–2), 115–166.

Fard, A. K., & Niknam, T. (2014). Optimal stochastic capacitor placement problem from the reliability and cost views using firefly algorithm. *IET Science, Measurement and Technology*, 8(5), 260–269.

Farrokh, M., Azar, A., Jandaghi, G., & Ahmadi, E. (2018). A novel robust fuzzy stochastic programming for closed loop supply chain network design under hybrid uncertainty. *Fuzzy Sets and Systems*, 341, 69–91.

Ghahremani-Nahr, J., Kian, R., & Sabet, E. (2019). A robust fuzzy mathematical programming model for the closed-loop supply chain network design and a whale optimization solution algorithm. *Expert Systems with Applications*, 116, 454–471.

Goldberg, D. E. (1989). *Genetic algorithms in search, optimization, and machine learning*. (1st ed.). Boston, Massachusetts: Addison-Wesley Publishing Company, Inc.

Hamza, F., Abderazek, H., Lakhdar, S., Ferhat, D., & Yildiz, A. R. (2018). Optimum design of cam-roller follower mechanism using a new evolutionary algorithm. *The International Journal of Advanced Manufacturing Technology*, 99, 1267–1282.

IEEE Standards Coordinating Committee 21(2018). *IEEE Standard for Interconnection and Interoperability of Distributed Energy Resources with Associated Electric Power Systems Interfaces*. IEEE Std 1547™-2018. IEEE Standards Association.

Jacobson, M. Z., & Jadhav, V. (2018). World estimates of PV optimal tilt angles and ratios of sunlight incident upon tilted and tracked PV panels relative to horizontal panels. *Solar Energy*, 169, 55–66.

Jannat, M. B., & Savić, A. S. (2016). Optimal capacitor placement in distribution networks regarding uncertainty in active power load and distributed generation units production. *IET Generation, Transmission and Distribution*, 10(2), 3060–3067.

Jie, C, Prashanth, L. A., Fu, M., Marcus, S., & Szepesvári, C. (2018). Stochastic optimization in a cumulative prospect theory framework. *IEEE Transactions on Automatic Control*, 63, 2867–2882.

Kheldoun, A., Bradai, R., Boukenoui, R., & Mellit, A. (2016). A new golden section method-based maximum power point tracking algorithm for photovoltaic systems. *Energy Conversion and Management*, 111, 125–136.

Kinay, O. B., Saldanha-da-Gama, F., & Kara, B. Y. (2019). On multicriteria chance-constrained capacitated single-source discrete facility location problems. *Omega*, 83, 107–122.

Labeeuw, W., & Deconinck, G. (2013). Residential electrical load model based on mixture model clustering and Markov models. *IEEE Transactions on Industrial Informatics*, 9(3), 1561–1569.

Lambert, T., Gilman, P., & Lilienthal, P. (2006). Micropower system modeling with HOMER. In: A. Felix, B. E. Farret, & G. Simões (Eds.), *Integration of alternative sources of energy*(pp. 379–418), Hoboken, New Jersey: John Wiley & Sons.

Liu, A., Lau, V. K. N., & Kananian, B. (2019). Stochastic successive convex approximation for non-convex constrained stochastic optimization. *IEEE Transactions on Signal Processing*, 67, 4189–4203.

Liu, L., Liu, J., & Tao, D. (2019). Dualityfree methods for stochastic composition optimization. *IEEE Transactions on Neural Networks and Learning Systems*, 30, 1205–1217.

Lorenzo, E., Araujo, G, Cuevas, A., Egido, M., Minano, J., & Zilles, R. (1994). *Solar electricity: Engineering of photovoltaic systems*. Progensa.

Lujano-Rojas, J. M., Dufo-López, R., Bernal-Agustín, J. L., Domínguez-Navarro, J. A., & Catalão, J. P. S. (2018). Probabilistic methodology for estimating the optimal photovoltaic ca-

pacity in distribution systems to avoid power flow reversals. *IET Renewable Power Generation*, 12(9), 1045–1064.

Ma, J. (2019). Robust optimal usage modeling of product systems for environmental sustainability. *Journal of Computational Design and Engineering*, 6, 429–435.

Moeeni, H., & Bonakdari, H. (2017). Forecasting monthly inflow with extreme seasonal variation using the hybrid SARIMA-ANN model. *Stochastic Environmental Research and Risk Assessment*, 31, 1997–2010.

Morteza, K., & Yildiz, A. R. (2016). A comparative study of non-traditional methods for vehicle crashworthiness and NVH Optimization. *Archives of Computational Methods in Engineering*, 23, 723–736.

NASA https://power.larc.nasa.gov/ Accessed 11 November 2019.

Ning, C., & You, F. (2018). Data-driven stochastic robust optimization: general computational framework and algorithm leveraging machine learning for optimization under uncertainty in the big data era. *Computers and Chemical Engineering*, 111, 115–133.

Pereira, B. R., Jr, Costa, G. R. M. D., Contreras, J., & Mantovani, J. R. S. (2016). Optimal distributed generation and reactive power allocation in electrical distribution systems. *IEEE Transactions on Sustainable Energy*, 7(3), 975–984.

Pfenninger, S., & Staffell, I. (2016). Long-term patterns of European PV output using 30 years of validated hour reanalysis and satellite data. *Energy*, 114, 1251–1265.

Rampinelli, G. A., Krenzinger, A., & Romero, F. C. (2014). Mathematical models for efficiency of inverters used in grid connected photovoltaic systems. *Renewable and Sustainable Energy Reviews*, 34, 578–587.

Renewables.ninja. https://www.renewables.ninja/ Accessed 11 November 2019.

Short, T. A. (2004). *Electric power distribution handbook*. Boca Raton, Florida: CRC Press LLC.

Staffell, I., & Pfenninger, S. (2016). Using bias-corrected reanalysis to simulate current and future wind power output. *Energy*, 114, 1224–1239.

Su, X., Masoum, M. A. S., & Wolfs, P. J. (2016). PSO and improved BSFS based sequential comprehensive placement and real-time multi-objective control of delta-connected switched capacitors in unbalanced radial MV distribution networks. *IEEE Transactions on Power Systems*, 31(1), 612–622.

Tan, X., Gong, Z., Chiclana, F., & Zhang, N. (2019). Consensus modeling with cost chance constraint under uncertainty opinions. *Applied Soft Computing*, 67, 721–727.

Tavana, M., Shiraz, R. K., & Di Caprio, D. (2019). A chance-constrained portfolio selection model with random-rough variables. *Neural Computing and Applications*, 31(2), 931–945.

Teng, J.-H. (2003). A direct approach for distribution system load flow solutions. *IEEE Transactions on Power Delivery*, 18(3), 882–887.

Teng, J.-H. (2008). Modelling distributed generations in three-phase distribution load flow. *IET Generation, Transmission and Distribution*, 2(3), 330–340.

Tsao, Y.-C., Vo-Van, T., Lu, J.-C., & Yu, V. (2018). Designing sustainable supply chain networks under uncertain environments: Fuzzy multi-objective programming. *Journal of Cleaner Production*, 174, 1550–1565.

Van Ackooij, W., Berge, V., Oliveira, W. D., & Sagastizábal, C. (2017). Probabilistic optimization via approximate p-efficient points and bundle methods. *Computers and Operations Research*, 77, 177–193.

Yildiz, A. R. (2013). A new hybrid artificial bee colony algorithm for robust optimal design and manufacturing. *Applied Soft Computing*, 13, 2906–2912.

Zaji, A. H., Bonakdari, H., & Gharabaghi, B. (2019). Applying upstream satellite signals and a 2-D error minimization algorithm to advance early warning and management of flood water levels and river discharge. *IEEE Transactions on Geoscience and Remote Sensing*, 57, 902–910.

Zou, J., Ahmed, S., & Sun, X. A. (2019). Stochastic dual dynamic integer programming. *Mathematical Programming*, 175(1–2), 461–502.

Appendix A

This appendix presents the most relevant results obtained from the implementation of a traditional GA for optimal allocation and sizing of DSTATCOM.

A.1 Traditional GA implementation

The technique proposed by Das (2002) was considered a benchmark for comparison. Traditional implementation consists of building the system's LDC, as shown in Fig. A1. Then, the traditional method discretizes this curve into $l = 1, \ldots, L$ blocks considering the corresponding relationship between the load demand [$\Delta P_{(l)}$] and time [$\Delta t_{(l)}$].

Calculations have been carried out by considering the net load of the system, while the power injection at each generation node was zero.
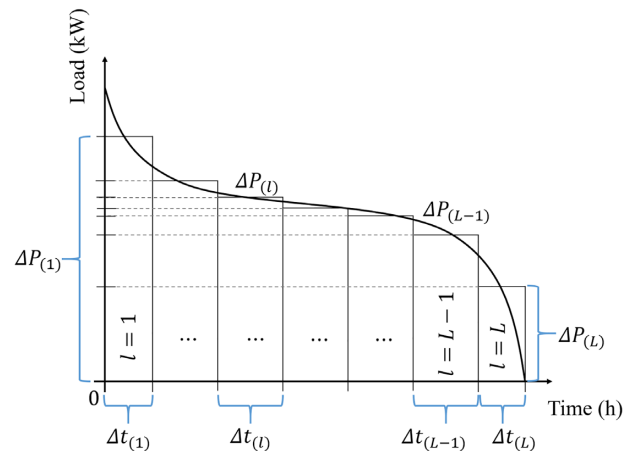

Figure A1: Discretization of the net load demand curve.

The traditional methodology estimates rated compensation capacity [$Q_{RPC(n)}$] to minimize the objective shown in equation (A.1) subject to the constraint presented in equation (A.2). The
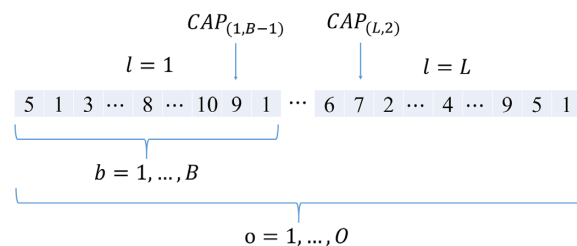

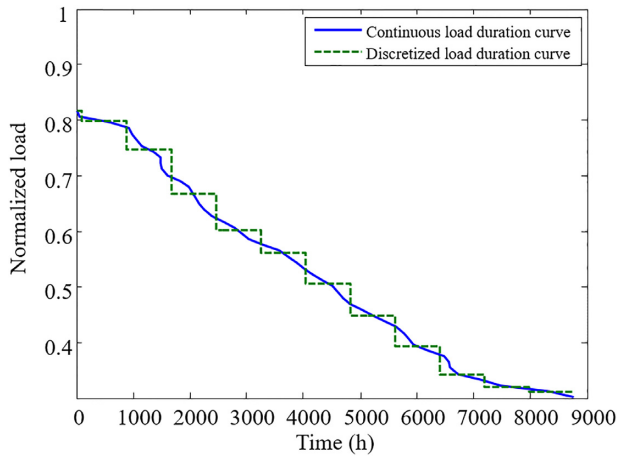**Figure A2:** Structure of an integer-coded GA individual.
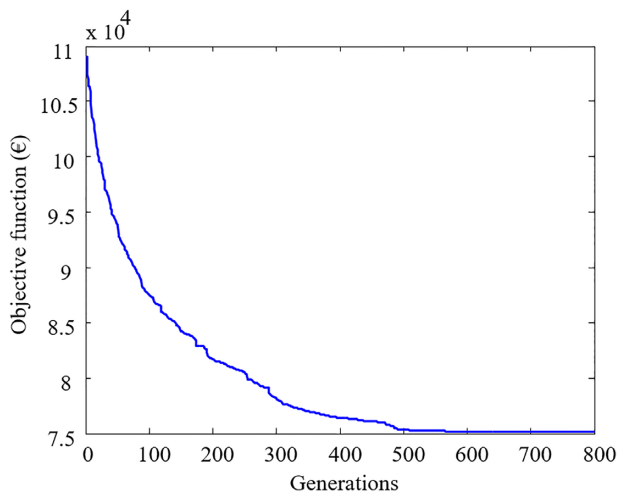
Figure A3: Discretization of the LDC.

Figure A4: Convergence of the GA (traditional method).

objective (equation A.1) is related to the energy losses on the DS and the investment in RPC devices, while the constraint (equation A.2) is related to the voltage profile improvements for each operating condition ($l = 1, \ldots, L$). The energy cost (K_e) is the mean value of the price signal [$E_{(t)}$], while the acquisition cost (K_c) is the value mentioned above 500€/kVA.

$$min\left\{ K_e \sum_{l=1}^{L} \Delta t_{(l)} \Delta P_{LS(l)} + K_c \sum_{n=1}^{N} Q_{RPC(n)} \right\} \qquad (A.1)$$

$$U_{min} \leq U_{(n,l)} \leq U_{max} \; \forall \; n = 1, \ldots, N; \; l = 1, \ldots, L \qquad (A.2)$$

Figure A2 shows the structure of a typical individual of this implementation, which uses integers to specify a determined type and size of RPC. The variable $C AP_{(l, b)}$ characterizes each element of the GA individual; this variable depends on the block number and the installation node. $C AP_{(l, b)}$ represents the magnitude of the reactive power injected or drawn from the system during the analysis of the load condition, which corresponds to the block l at the installation node b.

Figure A3 presents the LDC and the corresponding discretization.

The traditional GA was implemented considering 800 generations, a population size of 7500 individuals, a crossover rate of 95%, and a mutation rate of 5%. The addition or subtraction of reactive power was evaluated using integer magnitudes. It means that the sizes associated with the variable $C AP_{(l, b)} \; \forall \; l = 1, \ldots, 12; \; b = 1, \ldots 20$ were injections or extractions of reactive power using integers in the step of 1 kVAr. For our case study, $C AP_{(l, b)} \in [-114, 0]$. Figure A4 shows the convergence of the traditional GA for the case under study.

Table A1 shows the results of the optimization process. It reports the amount of reactive power to be injected per candidate node ($b = 1, \ldots, 20$) and discretization block ($l = 1, \ldots, 12$). As expected, the highest injection of reactive power occurs at the highest load demand, which defines the capacity of the equipment according to this method.

Table A1: Injection of reactive power per candidate node and load block (kVAr).

b/l	1	2	3	4	5	6	7	8	9	10	11	12
1	−65	−66	−66	−65	−57	−19	−22	−43	−7	−37	−53	−8
2	−93	−93	−93	−75	−85	−67	−5	−59	−21	−23	−12	−7
3	−79	−81	−81	−75	−70	−64	−26	−56	0	−13	−36	−38
4	−55	−55	−55	−51	−10	−50	−45	−9	−39	−16	−2	−43
5	−105	−105	−90	−56	−25	−77	−68	−43	−46	−47	−50	−21
6	−71	−71	−71	−70	−71	−71	−71	−71	−71	−71	−71	−69
7	−69	−70	−70	−65	−67	−66	−70	−70	−54	−49	−32	−40
8	−53	−53	−53	−52	−53	−51	−53	−53	−48	−44	−13	−33
9	−97	−97	−97	−97	−80	−85	−58	−44	−35	−9	−68	−31
10	−69	−69	−69	−61	−55	−53	−8	−18	−53	−5	−5	−49
11	−82	−82	−81	−67	−77	−82	−75	−67	−16	−29	−40	−4
12	−88	−88	−88	−68	−87	−86	−56	−40	−72	−69	−18	−57
13	−34	−34	−34	−34	−9	−19	−30	−13	−12	−29	−12	−3
14	−78	−78	−78	−67	−58	−28	−39	−40	−12	−35	−16	−31
15	−106	−106	−106	−106	−72	−37	−106	−18	−54	0	−17	−22
16	−63	−64	−64	−45	−62	−63	−41	−55	−60	−45	−64	−44
17	−35	−36	−36	−30	−36	−14	−34	−36	−14	−13	−18	−25
18	−79	−79	−79	−79	−51	−79	−79	−11	−78	−72	−32	−34
19	−83	−83	−82	−80	−83	−21	−31	−60	−3	−2	−29	−1
20	−38	−38	−36	−31	−28	−20	−25	−22	−28	−28	−8	−17