

Article

Data Analytics-Based Multi-Objective Particle Swarm Optimization for Determination of Congestion Thresholds in LV Networks

Javier Leiva ^{1,2,*}, Rubén Carmona Pardo ¹ and José A. Aguado ²

¹ Endesa, 28042 Madrid, Spain; ruben.carmona@enel.com

² Department of Electrical Engineering, University of Málaga, 29016 Málaga, Spain; jaguado@uma.es

* Correspondence: javier.leiva@enel.com

Received: 26 February 2019; Accepted: 2 April 2019; Published: 4 April 2019



Abstract: A growing presence of distributed energy resources (DER) and the increasingly diverse nature of end users at low-voltage (LV) networks make the operation of these grids more and more challenging. Particularly, congestion and voltage management strategies for LV grids have usually been limited to some elemental criteria based on human experience, asset oversizing, or grid reinforcement. However, with the current massive deployment of sensors in modern LV grids, new approaches are feasible for distribution network assets operation. This article proposes a multi-objective particle swarm optimization (MOPSO) approach, combined with data analytics through affinity propagation clustering, for congestion threshold determination in LV grids. A real case study from the smart grid of Smartcity Malaga Living Lab is used to illustrate the proposed approach. Within this approach, distribution system operators (DSOs) can take decisions in order to prevent situations of risk or potential failure at LV grids.

Keywords: congestion management; low-voltage networks; multi-objective particle swarm optimization; affinity propagation clustering; optimal congestion threshold

1. Introduction

Conventional development of the distribution networks has been driven by a worst-case fit-and-forget principle to tackle most of medium-voltage (MV) and low-voltage (LV) congestion challenges due to the long lifetime of their assets and planning horizon [1–4]. This led to limited distributed energy resources (DER) capacity allowed to be connected and costly network reinforcements, although peak loads generally occur only for a few hours in a year [1]. It is definitely recommendable to stop operating the distribution networks as static black boxes [5,6].

Despite promising examples of voltage and reactive power regulation that have been implemented in Germany and Italy according to national standards VDE 4105 and CEI 0-2 respectively [7], congestion management is normally approached dynamically only in transmission networks [8], for example by rescheduling and readjusting set points in generation units [9,10]. Nonetheless, the threats posed by the intermittent and dynamic nature of distributed generators, electric vehicles (EVs), or energy storage systems, to cite some of the most relevant DER technologies [11,12], require new means to help distribution system operators (DSOs) in their key function of congestion management in those complex scenarios [10]. This is the case of control algorithms proposed for operations such as curtailment of MV wind generators [13] or EV smart charging [14]. Therefore, flexibility in distribution networks is an active research question, including DER units and end users in market models that exploit their capabilities to relieve congestion [15] and improve security of supply [16].

However, in the case of LV networks, other means that can be considered as an alternative to those commonly widespread, such as asset oversizing or reinforcement, are scarce [17]. In addition

to this, congestion limits are not clearly defined for LV distribution network assets, such as MV/LV power transformers or LV feeders, apart from conventional criteria initially designed for static network planning [18,19] or the maximum admissible power determined by their nominal characteristics.

Determining what we have defined here as optimal congestion threshold may help DSOs in their dynamic operation of the LV networks. These thresholds provide them with technical references for the value of current measured individually in a certain distribution network element. Thereby, DSOs may enable preventive actions, by means of their own controllable network assets or with the help of other participating agents such as demand aggregators, before congestion definitively takes place and potential risk of failure may severely increase. Consequently, those thresholds must be optimized, as detailed in Section 3, considering not only the highest values of current experienced in the distribution network elements and their admissible congestion limits, but also the time duration and repetition of those congested or close to congestion situations over time.

Congestion management is addressed by classical methods by means of market mechanisms, as in the case of demand side management models [20], or by means of network planning-based methodologies, as in hosting capacity models [21], to cite some relevant examples, but limitations may arise at the time of their implementation on LV networks. Indeed, these models may be adapted to this new paradigm, with the help of present innovative solutions such as LV network-monitoring technologies and MV/LV state estimation algorithms [22]. Particularly, hosting capacity approaches are strongly conditioned by the nominal admissible limits on the network element of study [21] or the number and characteristics of scenarios considered [23], while demand side management requires the development of models oriented to facilitate end users' decision-making [24] and read locational market signals in order to boost participation in LV networks [20].

The methodology proposed here is especially designed for LV networks and seeks a compromise between the subjective human-based criteria and the precise but limited, as explained before, classical models. A data analysis is proposed here to address this challenge by means of clustering a significant group of distribution network elements. The aim is to provide insight into underlying patterns of data, to accelerate knowledge discovery, elements' classification, and subsequent computational efforts [25,26]. In addition to this, an optimization problem is formulated in order to help make decisions for preventing and managing congestion in heterogeneous massive sets of distribution network elements [27]. This provides DSOs with enriched objective criteria for the proactive preventive operation and maintenance of their assets, unlike the mentioned precedents based on their subjective, conservative, and generalized experience. In particular, particle swarm optimization (PSO) is proposed here thanks to its applicability to solve multi-objective optimization problems, and to its capability to avoid finding solutions biased by predetermined human initial decisions [28].

Authors in Reference [9] surveyed some applications of PSO on power systems, emphasizing its capabilities in facing issues such as uncertainty in load demand incorporating distributed generation, EV charging management, or economical dispatch to determine generation operating conditions while network constraints are met. Other relevant applications are those oriented to optimal power flow calculation and congestion management [10]. PSO-based algorithms can also deal with non-smooth functions, especially those related to frequency regulation and voltage constraints [10]. Further approaches are outlined in Reference [9], such as balancing loads between feeders or deciding on the optimal configuration, size, and topology of the distribution system.

PSO bears similarities to other advanced optimization algorithms, such as evolutionary computation techniques, where genetic and ant colony algorithms stand out for their popularity. Nevertheless, those techniques are more complex and present some serious restrictions to be applied in our methodology. Genetic algorithms may be faster than PSO and restrict the reproduction of weak solutions, but their crossover and mutation operations result as incompatible with our optimization problem [29], since it is formed by load duration curves, with individuals defined by non-independent characteristics [30]. Ant colony algorithms, despite of being based on swarm behavior [31] as in the case of PSO, assure convergence in problems where source and destination are predefined and

specific [32], unlike in the congestion threshold determination problem proposed here. The objective of this article is threefold: (1) To determine optimal technical thresholds to prevent congestion in distribution network assets, such as MV/LV power transformers and LV feeders; (2) to contribute with an optimization methodology not based on subjective previous experience and replicable in any kind of network, by employing clustering and multi objective particle swarm optimization (MOPSO); and (3) to apply the methodology to a real dataset obtained from Smartcity Malaga Living Lab, an area with more than 15,000 real end users, where 750 sensors installed in 56 MV/LV secondary substations are measuring current, voltage, power, and energy every 5 min [33].

This article is organized as follows. Section 2 addresses the congestion management problem in LV distribution networks, presenting the limitations of the DSOs to tackle saturation problems. In Section 3, the methodology for optimal congestion threshold determination is presented, detailing the data analytics-based and optimization algorithm developed, as well as its objectives, formulation, constraints, and pseudocode. The results of this methodology are discussed in Section 4 thanks to its application on power transformers and LV feeders from a real case of study from Smartcity Malaga Living Lab. Finally, Section 5 presents the conclusions and the fundamental ideas discussed in the article.

2. Congestion in LV Distribution Networks

Distribution networks have their end users spread, distributed in any form, independently if regarding a LV feeder and its phases, or among different feeders connected to the same power transformer in a secondary substation [4,34,35]. Moreover, each end user, whether consumer, prosumer, or generator, has its own arbitrary pattern, even more diverse with the proliferation of DER technologies.

A clear example of this highly heterogeneous distribution of clients of the LV networks can be seen in Figure 1 for two MV/LV (medium voltage/low voltage) secondary substations of Smartcity Malaga Living Lab [36], despite being located in the same geographical area, having the same number of LV feeders and nominal rating in their power transformers, and also the same number of predominantly residential single-phase end users, which is slightly above 500 in each case.

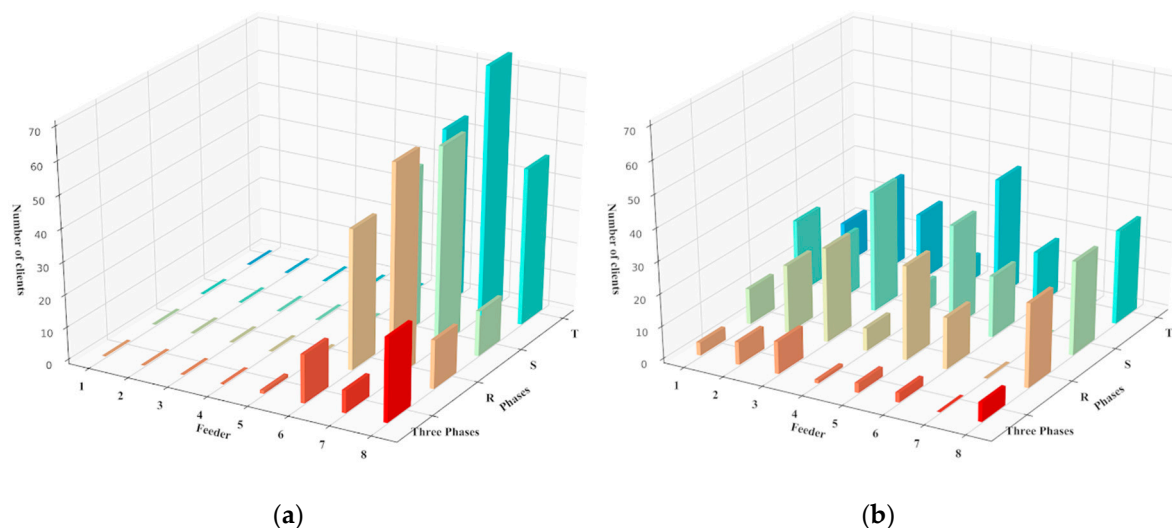


Figure 1. An example of single-phase (in orange, green, and blue) and three-phase (in red) client distribution per low-voltage (LV) feeder in two secondary substations (a,b) of Smartcity Malaga Living Lab.

This heterogeneity influences severe congestion and voltage in the network. In order to compensate for the voltage drop and to avoid subsequent undervoltage at the end of the feeder, the voltage at the LV side of a MV/LV power transformer has been traditionally set over 1 p.u. [1] with

a conventional off-line fixed setting [37]. Meanwhile, conventional options to tackle congestion go predominantly through network oversizing and extension, as in the case of rating upgrade of power transformer or construction of new secondary substations to overcome critical situations that take place in short periods of time [38,39], as can be seen in the load duration curves of a set of Smartcity Malaga power transformers displayed in Figure 2.

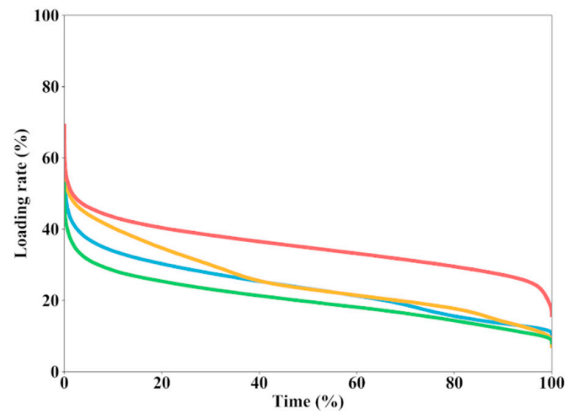


Figure 2. Load duration curves in power transformers in four secondary substations over a period of a year of Smartcity Malaga Living Lab.

Furthermore, congestion variations experienced in adjacent LV feeders and their phases, as displayed in Figure 3, evidence a totally decoupled unbalanced performance, leading to highly diverse characteristic values. Therefore, this means having situations close to congestion along a LV feeder, or particular phase, at the same moment the neighbor ones may not.

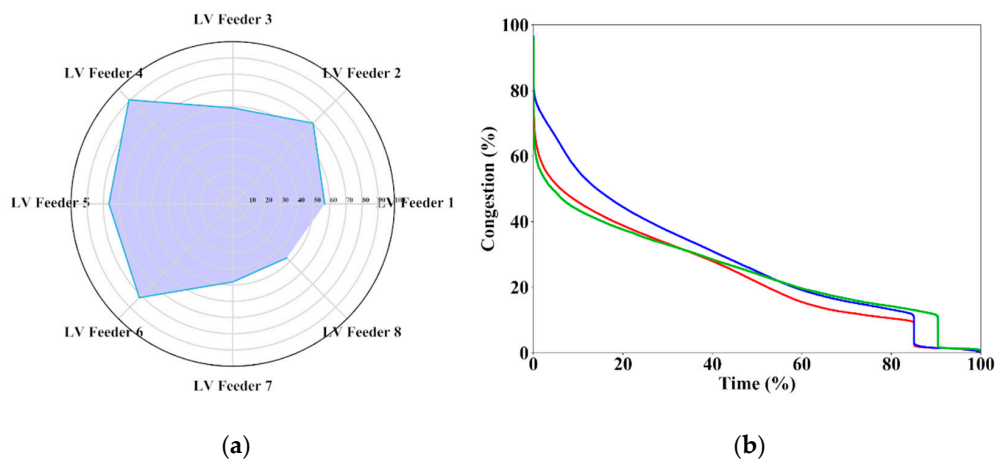


Figure 3. (a) Maximum current measured in the LV feeders with respect to their nominal admissible currents, and (b) load duration curves of the three phases in its LV feeder 5, of a power transformer of Smartcity Malaga Living Lab over a period of a year.

Unlike voltage, network congestion limits are not clearly defined by international grid codes, such as in EN50160 [40], which states that voltage on MV and LV distribution network nodes must remain within $\pm 10\%$ for 95% of the week, and between $+10\%$ and -15% for all time. Apart from the maximum limit determined by the nominal admissible power of each asset of the network, other values commonly used to set maximum congestion thresholds are 95% by the transmission system operator in Spain for their cables [18], or 75% by the regulator of the electric sector in Peru for both cables and power transformers [19], to cite some examples. However, these values are based on their local experience.

Power transformers are critical expensive assets in distribution networks, hence special attention must be given to their operation and maintenance [41,42]. Nonetheless, detailed information is scarce at the LV network level, such as the particular phase where a single-end user or DER unit is connected [43], despite successful smart grid pilot projects such as Smartcity Malaga [44], INTEGRIS [5], IDE4L [45], or MONICA [46] have dedicated special efforts in distribution network digitalization and state estimation. Therefore, technical solutions are available today to address congestion in singular elements such as LV feeder phases, avoiding bottlenecks that may arise even earlier in the case of significant DER penetration or demand growth [47,48], since technical constraint violations may occur at an earlier stage [49].

Therefore, DSOs may operate the distribution networks by means of a permanent dynamic supervision of the congestion level, taking into account the optimal congestion threshold assigned by this methodology to any particular network element. As pointed out in Section 1, those thresholds provide DSOs with technical references that prevent them before congestion definitively takes place in a particular distribution network element and, consequently, potential risk of failure increases. This lets them leave behind corrective processes and act before an incident occurs, therefore optimizing budget allocation and the quality of service provided to end users [50] by means of preventive actions, such as adjusting on-load tap changing (OLTC) MV/LV power transformers [51] or interacting with DER units so that they adjust their operating regime to the existing network conditions [7].

3. Proposed Approach for LV Congestion Determination

The proposed methodology consists of three stages, as displayed in Figure 4. Firstly, data acquisition allows for the creation of the search space, which is formed in this problem by load duration curves of distribution network elements. They may represent a numerous diverse dataset for modern DSOs, hence, data analysis techniques must be applied. On the one hand, this allows for the classification of assets by grouping them regarding their similarities. On the other hand, the most representative elements can be identified, so deeper studies can be performed on them, and later be extrapolated to the rest of similar assets for the sake of the efficiency of the analysis.

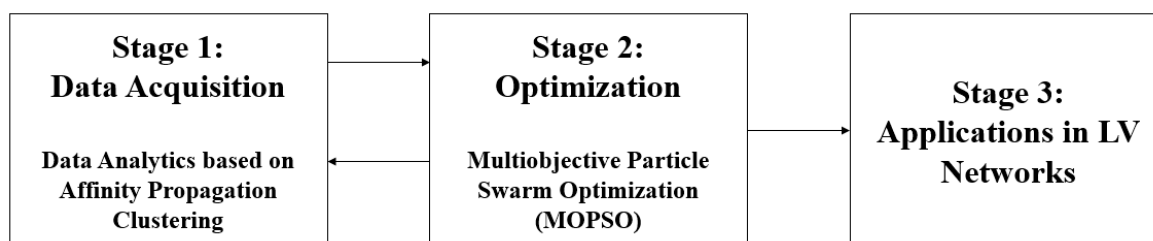


Figure 4. Stages of the methodology proposed.

Data analytics is approached here by means of an affinity propagation clustering technique [25]. Apart from initially considering any of those distribution network elements as potential cluster centers, this technique is characterized by not requiring any previous specification about the number of clusters to be formed, in addition to exchanging messages iteratively between elements until a high-quality set of exemplars and clusters emerges, doing so in less than one hundredth the time and at a lower error rate than other exemplar-based methods [25,52]. Consequently, only a basic identification and characterization of those assets to be clustered is needed to initiate the analytics.

The second stage focuses on determining the optimal congestion thresholds. For this, a particle swarm optimization (PSO) problem is formulated, which is a stochastic-based artificial intelligence search technique inspired on natural life [28]. A population of particles, called *swarm*, formed by points, *flows* through the search space taking into consideration the historical best position for each particle itself and the rest as a whole, naturally orienting the search towards an optimal or near-optimal solution [9,17].

Multi-objective PSO (MOPSO) is proposed here due to its capacity for simultaneous resolution of multiple conflicting objectives in congestion threshold determination [17]. In addition to this, MOPSO presents a low computational cost to provide a set of solutions, that, unlike classical multi-objective methods, are diverse and spread enough [17]. This allows for finding the best, non-trivial trade off among more than one objective, since no single solution may simultaneously optimize the whole search objective [53].

The affinity propagation clustering previously executed allows for the application of MOPSO only on the most representative element of each cluster. Then, the third stage addresses the extrapolation of the result obtained in the optimization process to the rest of assets.

3.1. Data Analytics based on Affinity Propagation Clustering

The affinity propagation algorithm [25] applied here is set with a dumping factor of 0.9 in order to avoid high numerical oscillations and favor the convergence of the clustering process, and is trained with a dataset formed by the characteristics listed on Table 1.

Table 1. Characteristic values of each network element of the dataset.

Criteria	Characteristic
Loading	Maximum congestion value measured
Repetition	Maximum congestion value measured after the 3 most critical days in a year
	Maximum congestion value measured after the 10 most critical days in a year
	Maximum congestion value measured after the 15 most critical days in a year
Time duration	Maximum congestion value measured after the 0.01% most critical scenarios in a year
	Maximum congestion value measured after the 0.1% most critical scenarios in a year
	Maximum congestion value measured after the 1% most critical scenarios in a year
Geographical	Number of points of delivery
	Total length of the element ¹
End users	Number of clients
	Percentage of domestic clients
Power	Nominal admissible power
	Total power contracted

¹ Total length of the element: 0 in the case of power transformers; including any segment downstream associated in the case of power lines.

Standardization in the clustering process is important, since the values of those characteristics are often measured in different units, meaning that some of them could be dominating other ones, hence influencing the course of the cluster analysis [54]. To overcome this, the characteristic values considered in this methodology are re-scaled by using min-max standardization.

In addition to the nominal admissible power or the maximum congestion value for each asset, such as power transformers or LV feeder, the time duration of those congested situations experienced has to be carefully considered, as in the case of Spain, where the annual maximum demand is denoted by the congestion level reached around the 100 most critical hours [55], about 0.01% of the year. Furthermore, here it is proposed to consider the number of different days where the maximum congestion levels took place in order to take into consideration the degree of repetition of those risky situations.

Other key aspects, such as the number and type of end users connected to every distribution network element, its length (which is zero in the case of power transformers, but includes any segment downstream associated in the case of LV feeder phases), or the total power contracted, are also considered. Hence, the clustering process provides a data analysis based not only on graphical features of each load duration curve, but also based on physical parameters of the corresponding asset.

3.2. Multi-Objective Particle Swarm Optimization

The determination of the optimal congestion threshold is a constrained optimization problem where a specific multi-objective function is minimized. The aim is to find the optimal characteristic value of congestion in the load duration curve of a distribution network element considered, valid for the longest time scenarios of operation and being the closest to the maximum demand level experienced on it.

The multi-objective function is formed by three subfunctions. The first objective subfunction F_{obj_1} aims to minimize the standard deviation of the congestion characteristic values ϕ of the load duration curve, leading to select groups of points that are more concentrated around a certain congestion level. This diminishes the influence of abrupt changes in congestion characteristic values, exceptionally unconnected with neighboring values, and typically corresponding to the highest demand scenarios.

The second objective subfunction F_{obj_2} tackles the percentage of time scenario characteristic values ρ of the load duration curve. Having the maximum standard deviation here aims to find a solution far from unrepresentative short time variations. This subfunction is characterized by the inverse of the standard deviation in order to fit it in the minimization of the multi-objective problem.

Finally, the third objective subfunction F_{obj_3} addresses the inclination angle φ by scoring linearly between 0 and 10 points of the linear regression of any group of points forming a particle of the load duration curve. An inclination angle of 45° presents the best compromise between congestion measurements and time scenarios, namely, between its vertical and horizontal projections. However, with the aim of not losing the influence of the highest demand experienced, a favorable inclination angle, called φ_{fav} , must be set above 45° , helping the optimization problem converge towards more stabilized congestion characteristic values but without being too close to 90° and then falling into unrepresentative, extremely brief, and highest-congestion scenarios. A band of the most favorable angles, φ_{band} , is also set in order to help define diverse punctuation areas, as shown in Figure 5. This objective subfunction is characterized by the inverse, as expressed in Equation (7), in order to fit it in the minimization of the multi-objective problem, as in the case of objective subfunction 2.

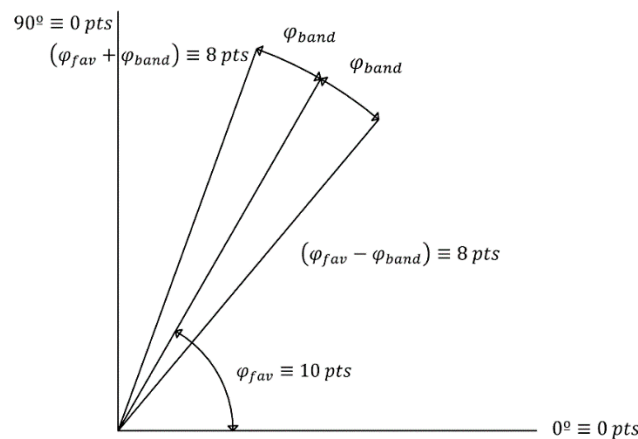


Figure 5. Representation of value assignment according to the favorable inclination angle set.

Therefore, the formulation of this multi-objective is stated as follows:

$$\begin{aligned} \text{Minimize : } & F_{obj}(\phi, \rho, \varphi) = f\left(F_{obj_1}(\phi), F_{obj_2}(\rho), F_{obj_3}(\varphi)\right) \\ \text{Subject to : } & \phi \geq \phi_{limit} \end{aligned} \quad (1)$$

The problem constraint $\phi \geq \phi_{limit}$ refers to all those congestion characteristic values above a particular value, which is considered to represent a realistic, considerable, potential risk of saturation of the distribution network elements. It is recommended to consider a wider range of congestion states

than in traditional criteria, as mentioned in Section 2 for 95% [18] and 75% [19], in order to cover a broader set of representative scenarios and then to avoid extremely singular transient states.

By checking compliance with this restriction, MOPSO is not applied in the case of those clusters whose elements are not satisfying the problem constraints for none of the points of their load duration curves, since they are not considered to be above a minimum level of potential risk.

In our methodology, and according to MOPSO algorithm, the *swarm* is formed by N particles, each of them composed by n points, as displayed in Figure 6, that move along the search space, that is to say, in a certain load duration curve. Consequently, each particle j is characterized by the congestion ϕ_j and the corresponding percentage of time scenarios ρ_j of its center, and its position $x_j(t)$ in a moment of time t along the load duration curve depends on its previous position $x_j(t - 1)$ and its velocity $v_j(t)$ [27]:

$$x_j(t) = x_j(t - 1) + v_j(t). \quad (2)$$

This velocity $v_j(t)$ is composed by terms related to the inertia ω_j and the velocity of the particle in the previous state $v_j(t - 1)$ to represent its tendency to continue in the same direction, and the attraction towards the best position $p_{j_{best}}$ ever found by the particle and by any of the particles of the swarm $p_{global\ best}$ [27]:

$$v_j(t) = \omega_j \cdot v_j(t - 1) + \psi_1 \cdot rand_1 \cdot (p_{j_{best}} - x_j(t - 1)) + \psi_2 \cdot rand_2 \cdot (p_{global\ best} - x_j(t - 1)). \quad (3)$$

Authors in Reference [27] pointed that the sum of constants ψ_1 and ψ_2 should be 4.0, meaning $\psi_1 = \psi_2 = 2$ to give the same weight to the individual and group experience. $rand_1$ and $rand_2$ are random values between 0 and 1 [27], and the inertia of the particle ω_j is given by:

$$\omega_j = \omega_{max} - \frac{\omega_{max} - \omega_{min}}{iter_{max}} \cdot iter_i \quad (4)$$

where ω_{max} and ω_{min} are also constant values set at 1 and 0 respectively, in order to reach an initial high value near to 0.9 as pointed in Reference [27], hence moving fast towards the global optimum.

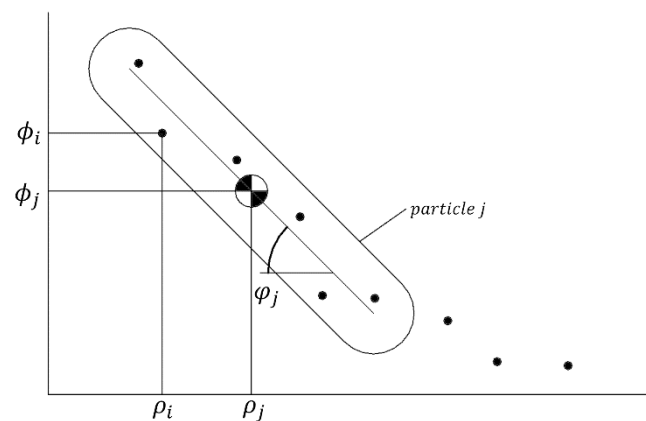


Figure 6. Representation of congestion, time scenarios, and inclination characteristic values of a generic particle formed by a series of points.

Therefore, in each iteration, every ϕ_j and ρ_j corresponding to the center of each particle, in addition to ϕ_i , ρ_i and the inclination angle of the linear regression ϕ_j of their n points, will be evaluated in the following subfunctions of the multi-objective function, leading to a solution:

$$F_{obj1}(\phi_j) = \sqrt{\frac{1}{n-1} \sum_{i=1}^n (\phi_i - \bar{\phi}_j)^2} \quad (5)$$

$$F_{obj_2}(\rho_j) = \frac{1}{\sqrt{\frac{1}{n-1} \sum_{i=1}^n (\rho_i - \bar{\rho}_j)^2}} \quad (6)$$

$$\frac{1}{F_{obj_3}(\varphi_j)} = \begin{cases} 8 \cdot \frac{\varphi_j}{(\varphi_{fav} - \varphi_{band})} & \text{if } \varphi_j \leq \varphi_{fav} - \varphi_{band} \\ 8 + 2 \cdot \frac{\varphi_j - (\varphi_{fav} - \varphi_{band})}{\varphi_{band}} & \text{if } \varphi_j \in (\varphi_{fav} - \varphi_{band}, \varphi_{fav}) \\ 10 & \text{if } \varphi_j = \varphi_{fav} \\ 8 + 2 \cdot \frac{(\varphi_{fav} + \varphi_{band}) - \varphi_j}{\varphi_{band}} & \text{if } \varphi_j \in (\varphi_{fav}, \varphi_{fav} + \varphi_{band}) \\ 8 \cdot \frac{90 - \varphi_j}{90 - (\varphi_{fav} + \varphi_{band})} & \text{if } \varphi_j \geq \varphi_{fav} + \varphi_{band} \end{cases} \quad (7)$$

where: $\bar{\varphi}_j$ is the mean congestion value in the particle j ; $\bar{\rho}_j$ is the mean percentage time scenario value in the particle j .

Pareto dominance must be checked every time the position of a swarm is updated, storing those non-dominated solutions among the obtained for the N particles after their evaluation on the mentioned subfunctions of the multi-objective function, in order to approximate the Pareto front [9].

In relation to this, the first stopping criterion is determined for the algorithm developed here: The global optimum results from the best particle of the swarm after a certain number of consecutive iterations without any progress in the Pareto front. In addition to this, the second stopping criterion consists of a maximum number of iterations.

As pointed out in Section 1, other advanced optimization algorithms, such as genetic and ant colony algorithms, apart from being more complex than this PSO-based approach, present serious restrictions to be applied in our methodology due to the impossibility to cross and mutate points defined by non-independent congestion and time characteristics [29,30], and to the absence of a clear, predefined optimal solution area [32].

3.3. Applications in LV Networks

This methodology can be applied on wide groups of distribution network elements by means of the ratio between the optimal congestion threshold $\phi_k^{opt_calc}$ determined by the MOPSO algorithm and the maximum congestion value measured ϕ_k^{max} for each cluster center k . Therefore, this so-called *threshold ratio_k* can be employed individually on any of the members m belonging to that cluster k in order to determine its particular optimal congestion threshold ϕ_m^{opt} , where:

$$threshold\ ratio_k(\%) = \frac{\phi_k^{opt_calc}}{\phi_k^{max}} \cdot 100 \quad (8)$$

$$\phi_m^{opt} = threshold\ ratio_k \cdot \phi_m^{max} \quad (9)$$

Moreover, the use of the threshold ratio reflects the existing differences among the congestion experienced in each element, instead of simply imposing the optimal of the cluster center. Thus, the application of the methodology follows the flowchart described in Figure 7.

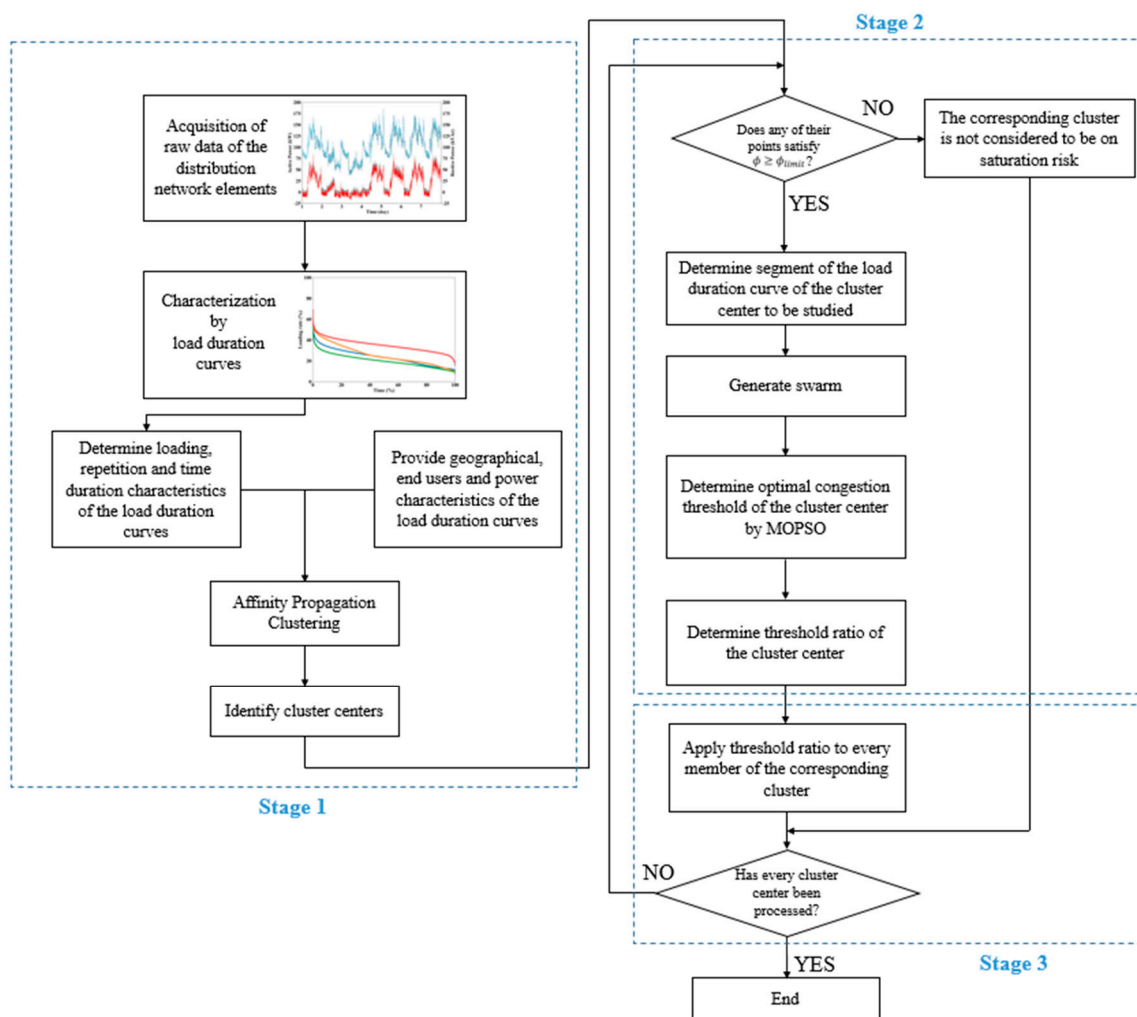


Figure 7. General flowchart of the methodology.

4. Results and Discussion

This methodology has been applied on real data obtained from Smartcity Malaga Living Lab. In particular, data from five MV/LV secondary substations are considered here, including a total of six power transformers and their 18 corresponding LV side phases. In addition to this, as displayed in Table 2, their corresponding 45 LV feeder lines, with three phases in each case, are also here considered, therefore taking into account their highly dispersed and heterogeneous distribution of clients, as shown in Figure 1. Consequently, this methodology has been implemented on load duration curves from 153 distribution network elements, a group of elements linked to around 5000 out of the 15,000 real end users of the Living Lab.

Table 2. Distribution network elements of the group considered in the case study.

Secondary Substation	Power Transformers	LV Feeders	Distribution Network Elements
A	1	7	24
B	1	7	24
C	1	7	24
D	2	12	42
E	1	12	39

4.1. Data Analytics for Data Pre-Processing

Every distribution network element has been characterized by 5-min data acquisition during a whole year period, meaning more than 109,000 current measurements in each case. Thus, this group formed by a total of 153 elements involves handling around 16,000,000 data. Consequently, the size of this dataset provides a wide, varied dataset to reflect the reality of the operation and maintenance of the distribution networks, not only because of the number of assets considered, but also because of the time interval, including summertime, wintertime, working days, and holidays.

According to the methodology detailed in Section 3, the first stage is data analytics. Applying affinity propagation to cluster those 153 elements results in the composition of four different clusters, as shown in Figures 8–10, and the identification of their corresponding cluster centers listed in Table 3.

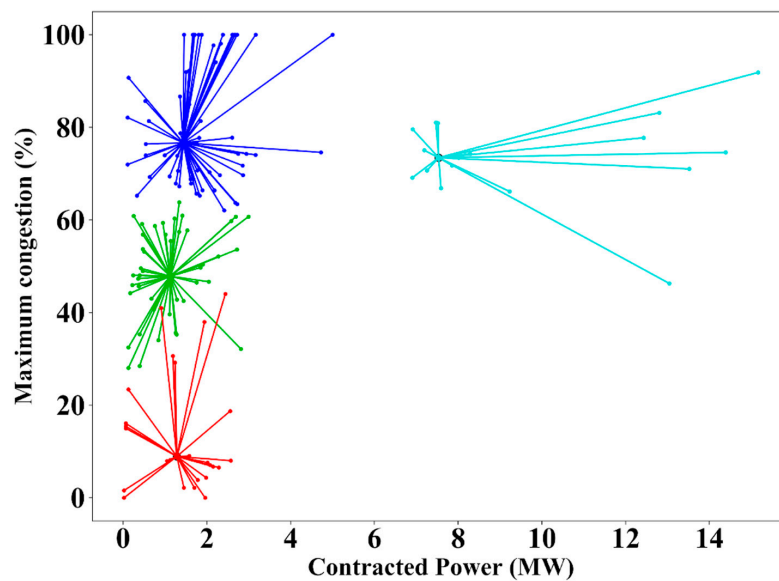


Figure 8. Maximum congestion in relation to the total contracted power in the case study.

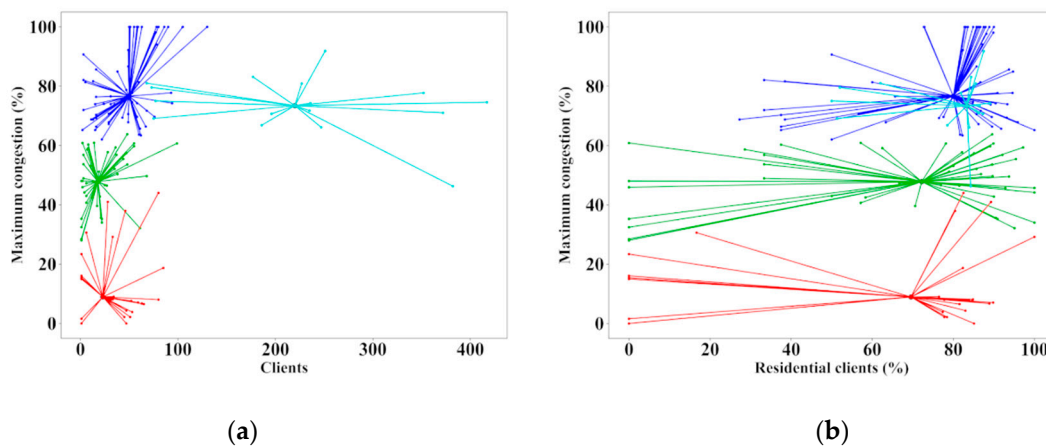


Figure 9. (a) Maximum congestion in relation to the number of clients and (b) in relation to the proportion of residential clients in the case study.

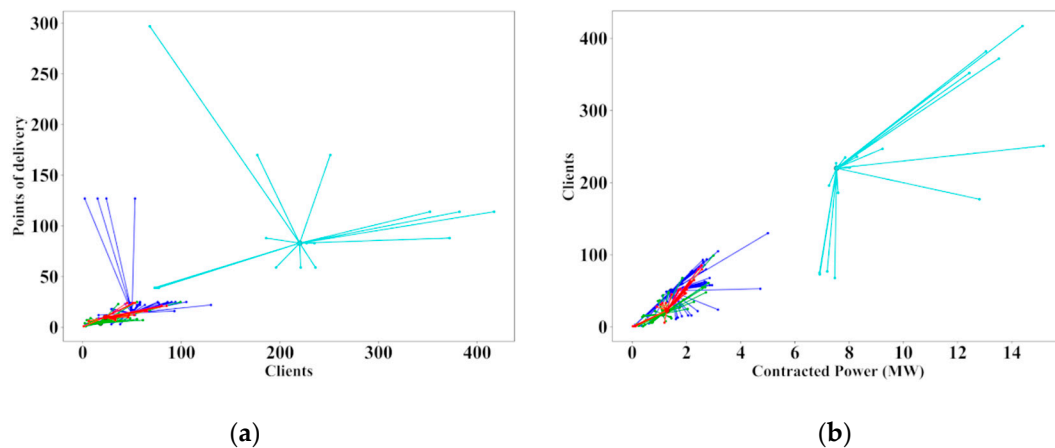


Figure 10. (a) Number of clients in relation to the number of points of delivery and (b) in relation to the contracted power in the case study.

Table 3. Composition of the clusters.

Cluster	Cluster Center	Color	Cluster Members
1	Secondary Substation D LV Feeder 7 Phase T	Dark blue	61
2	Secondary Substation D LV Feeder 14 Phase S	Green	47
3	Secondary Substation D Transformer 2 Phase S	Light blue	18
4	Secondary Substation E LV Feeder 9 Phase T	Red	27

The lowest values obtained at the maximum congestion characteristic can be found in the members of cluster 4. Although all the power transformers of the group belong to cluster 3, which has the least number of members, clusters 1, 2, and 4 are highly heterogeneous since they have LV feeder phases belonging to every secondary substation of the group. In addition to this, three out of the four cluster centers belong to the same secondary substation, which evidences the diversity among elements that can be found even in the same asset of the distribution network.

Regarding maximum congestion, 16 distribution network elements have measured values over 95% of their maximum admissible power [18], and 14 out of them reached their nominal maximum congestion of 100%. Regarding the conventional congestion threshold of 75% [19], 40 elements experienced congestion over that limit, and only 10 out of them remain over it at 1% of time scenarios. At the same time, 29 elements present a maximum congestion value measured below 35%, representing almost the fifth part of the group, and 17 out of them remained below 10%.

The highest number of clients in a secondary substation, by grouping those connected to the three LV sides of its power transformer, is around 1300 clients, and the lowest is 225, while the average per LV feeder phase is 37 clients, ranging from 1 to 130. Despite the area being predominantly residential, representing 50% or more of the clients in 100 LV feeder phases, only 15 LV feeder phases present 100% industrial clients.

Geographically, more than 600 points of delivery are included in the group of elements, ranging from 1 to 25 per LV feeder phase and with an average of 14 per each of them. The average length per LV feeder phase is 200 m. The clustering executed here makes it possible to identify differences that are not easily discernible a priori, as can be seen in Figure 10, where most of the elements from clusters 1, 2, and 4 are coincident for the represented characteristics.

4.2. Optimization

The second stage of the methodology is optimization. To evaluate the different cluster centers obtained, a particular $\phi_{limit} = 35\%$ is set in order to focus the analysis only on potentially risky congestion states, according to the reality observed in the real electrical network of study of Smartcity

Malaga Living Lab. In addition to this, ϕ_{limit} adapts the methodology to the search space without rejecting any predetermined proportion or number of assets.

Once ϕ_{limit} restrictions have been checked, it is observed that the cluster center 4, in red in Figures 8–10, does not satisfy it for any of its points. Therefore, this cluster is out of the following optimization process because of not being in a situation of risk minimally considerable. At the same time, applying ϕ_{limit} to the cluster centers 1, 2, and 3 allow for the focus of the optimization problem in a limited set of points of their load duration curves, listed in Table 4, corresponding to the most saturated scenarios.

Table 4. Number of points considered for the execution of the multi-objective particle swarm optimization (MOPSO) algorithm.

Cluster Center	Points in Raw Load Duration Curve	Points above ϕ_{limit}
1	1196	600
2	1772	117
3	597	195

MOPSO is then applied on cluster centers 1, 2, and 3, where a favorable inclination angle $\varphi_{fav} = 75^\circ$, a tolerance band $\varphi_{band} = 5^\circ$, a maximum number of iterations without any progress in Pareto front $iter_{pareto\ front} = 10$, and a maximum number of iterations $iter_{max} = 100$ have been set. The initialization of the population is carried out by means of a random stochastic process that distributes them arbitrarily along the search space. The optimization algorithm that has been executed for the series of particles and their corresponding points is shown in Table 5.

Table 5. Populations considered for the execution in the MOPSO algorithm.

Case	Particles	Points per Particle
1	15	3
2	20	3
3	30	3
4	30	5
5	20	5
6	15	5

The results obtained from the MOPSO algorithm developed are displayed in Table 6 showing a good robustness of the method, with a low standard deviation despite being a stochastic-based method. Therefore, the methodology can be considered to be slightly dependent on the size of the population.

Table 6. Optimal congestion thresholds of the cluster centers.

Cluster Center	Maximum Congestion Measured	Optimal Congestion Threshold						Standard Deviation
		Case 1	Case 2	Case 3	Case 4	Case 5	Case 6	
1	76.65%	68.00%	66.62%	66.00%	67.15%	66.00%	66.00%	0.01
2	47.84%	34.97%	35.55%	35.90%	35.65%	39.62%	37.49%	0.04
3	73.44%	60.97%	64.30%	55.15%	64.57%	64.57%	65.68%	0.05

A representation of the evolution in the values obtained in the three objective subfunctions is displayed in Figure 11. Particularly, those are the values obtained by the best particle in each iteration of the algorithm in the Case 1, which is formed by a swarm of 15 particles with 3 points each, applied to cluster center 1. It must be noted that the role of the best particle can be played by different particles of the swarm along the different iterations, according to Pareto compliance.

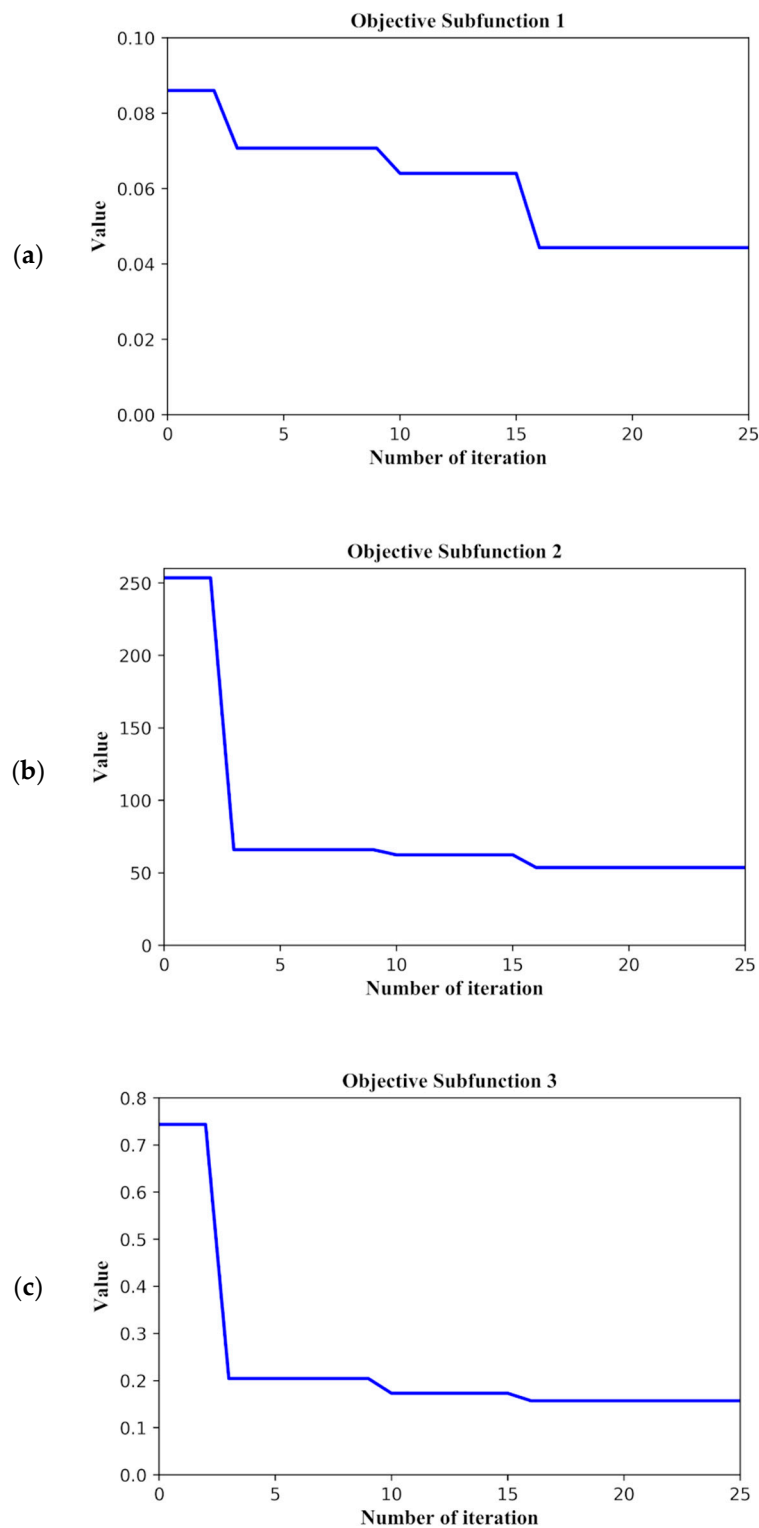


Figure 11. Evolution in the values obtained, for the best particle of each iteration, in case 1 applied to cluster 1, in the objective subfunctions of (a) congestion characteristic values, (b) time scenario characteristic values, and (c) inclination angle.

The example displayed in Figure 11 shows how values obtained for objective subfunctions 1, 2, and 3 are decreasing progressively, but not continuously, towards the minimum, leading to identifying a particle with the corresponding optimal congestion characteristic value of 68.00% displayed in Table 6. The global optimum finally results after 25 iterations, meaning that the stopping criterion of

10 consecutive iterations without any progress in the Pareto front has been applied; hence, the other stopping criterion of a maximum number of iterations of 100 has not been reached.

Apart from considering a favorable inclination angle of $75^\circ \pm 5^\circ$, the MOPSO algorithm has also been executed for $70^\circ \pm 5^\circ$ and $65^\circ \pm 5^\circ$. As shown in the results displayed in Figure 11 for cluster center 2, the algorithm presents a good stability in the results obtained and therefore a low dependence on this parameter, as evidenced in Figure 12, despite the varied characteristics of the swarms employed. In addition to this, it evidences that higher values of favorable inclination angles result in higher optimal values due to the greater influence of the maximum congestion values measured in the element.

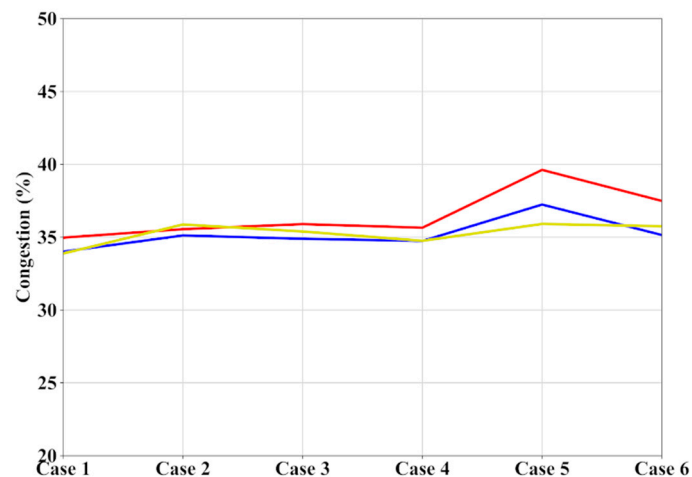


Figure 12. Optimal congestion thresholds for cluster center 2 considering favorable inclination angles of 75° (in red), 70° (in blue), and 65° (in yellow).

4.3. Application on the Distribution Network of the Case Study

The third and last stage of the methodology is its application in LV networks. To do so, once the congestion thresholds have been determined for each cluster center, the mean threshold ratio is calculated for each cluster center of the case study, as shown in Table 7.

Table 7. Threshold ratios calculated for the cluster centers.

Cluster Center	Threshold Ratio						Mean
	Case 1	Case 2	Case 3	Case 4	Case 5	Case 6	
1	88.71%	86.91%	86.11%	87.61%	86.11%	86.11%	86.93%
2	73.10%	74.31%	75.04%	74.52%	82.82%	78.37%	76.36%
3	83.02%	87.55%	75.10%	87.92%	87.92%	89.43%	85.16%

Therefore, by applying the corresponding mean threshold ratio, an individual, particular optimal congestion threshold is set for 135 out of the 153 distribution network elements of Smartcity Malaga Living Lab belonging to clusters 1, 2, and 3. To cite some relevant figures, 22 distribution network elements have been set with optimal thresholds over the conventional 75% [19] of their maximum admissible power, 63 elements over 60%, and 78 over 50%.

The numerical results here obtained, as displayed in Figure 13, provide the network operators with an objective, individual determination of the optimal congestion threshold in order to establish supervision and control strategies such as those mentioned in Sections 1 and 2. In addition to this, as shown in Table 8 in the case of the cluster centers, this methodology provides the time scenarios corresponding to those optimal congestion thresholds.

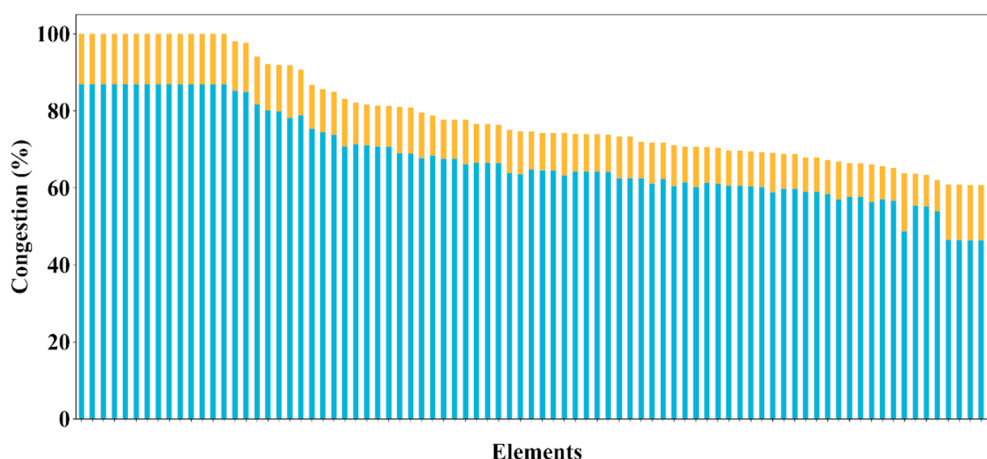


Figure 13. Maximum congestion measured (in orange) and optimal congestion threshold assigned (in blue) in the most saturated elements in the case study.

Table 8. Threshold ratios and time characteristics of the cluster centers.

Cluster Center	Maximum Congestion Measured	Optimal Congestion Threshold	Threshold Ratio	Optimal Percentage of Time Scenario	Optimal Number of Time Scenarios	Different Days of Occurrence
1	76.65%	66.63%	86.93%	0.025%	27	13
2	47.84%	36.53%	76.36%	0.038%	40	22
3	73.44%	62.54%	85.16%	0.050%	53	28

For instance, in the case of cluster center 1, the optimal congestion threshold determined at 66.63% potentially means intervening in 0.025% of the year of time considered in this dataset. Even more, 27 different scenarios experienced congestion above that threshold, each one of a 5-min duration, taking place over 13 different days. In this way, the distribution network elements of the group can be ranked according to their optimal congestion threshold but also considering the number of congested time scenarios, as displayed in Figure 14, given the useful additional information provided.

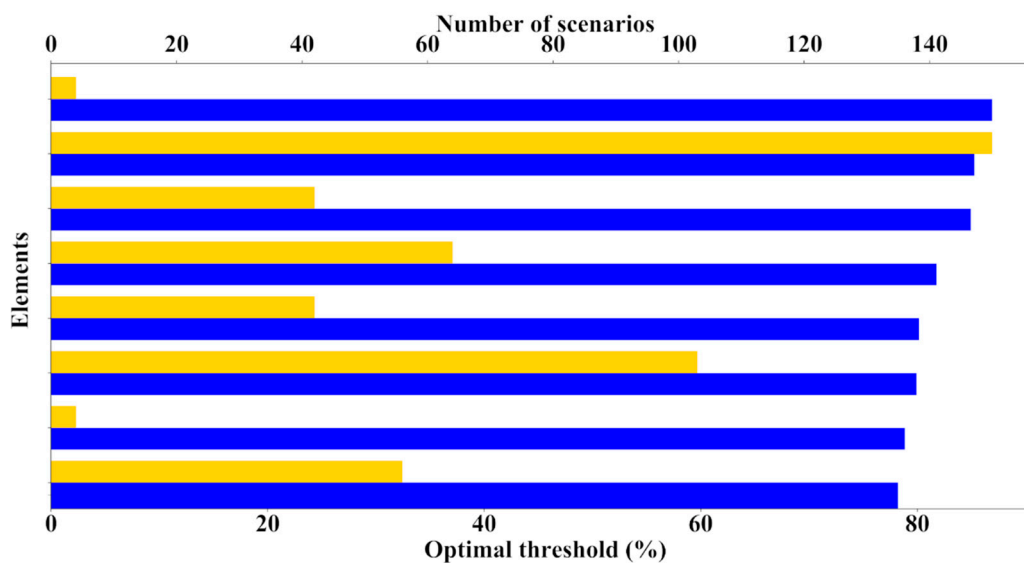


Figure 14. Optimal congestion threshold (in blue) and number of time scenarios (in orange) in a set of elements in the case study.

Moreover, the previous set of distribution network elements can also be ranked according to their optimal congestion threshold but considering now the number of different days presenting congestion, as displayed in Figure 15; hence, a totally different ranking can be composed, since there is not a direct relation between them.

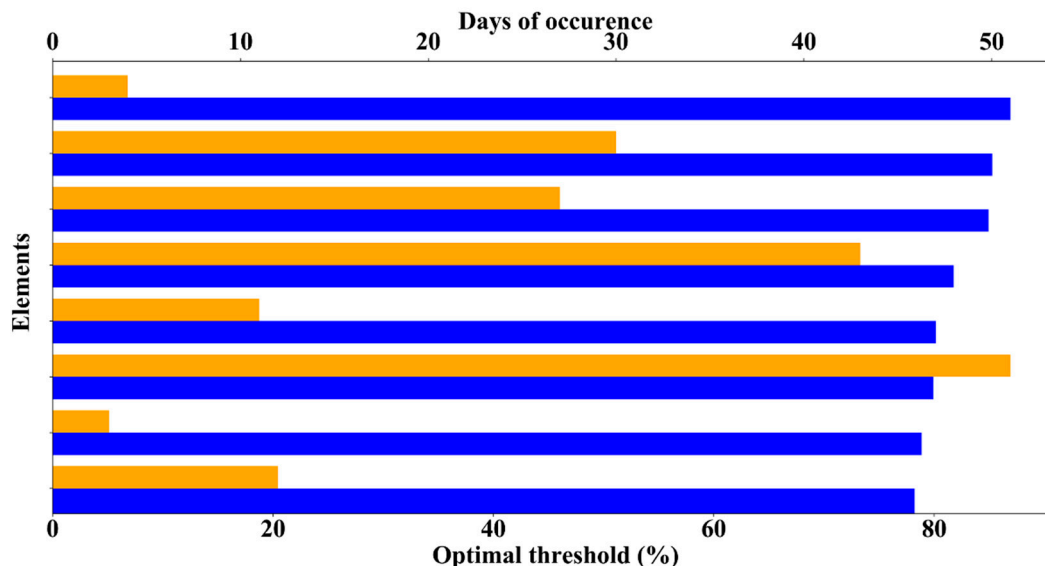


Figure 15. Optimal congestion threshold (in blue) and days of occurrence (in orange) in a set of elements in the case study.

Consequently, this methodology allows DSOs to tackle potential congestion states in their assets by prioritizing the operation of those distribution network elements which are closer to saturation according to both the optimal threshold determined and the expected congested situations.

5. Conclusions

This methodology provides a tool that supports DSOs in making decisions to manage congestion in any of the elements of the distribution network, such as MV/LV power transformers or LV feeders and their corresponding phases. It allows them to operate and maintain their assets in a preventive way beyond the classic planning criteria traditionally used, hence bringing an alternative to oversizing and reinforcements, and not based on previous, subjective precedents. Moreover, the methodology has been validated with real data obtained from Smartcity Malaga Living Lab, under real operating conditions.

Due to the means available today for the digitalization of the network, the large volumes of data to be handled far outweigh human experience, as in the case study presented here, based on a dataset formed by more than 16,000,000 data for a single year from a network related to around 5000 real end users. The pre-processing performed by means of affinity propagation clustering results in a more efficient computational analysis. Thus, the 153 distribution network elements considered here were divided into four clusters, with three of them meeting the restrictions of the optimization problem. In the case of the fourth cluster, formed by 18% of the distribution network elements, their congestion levels were not considered to be over a minimum level of risk.

The optimization process carried out provides enriched criteria with respect to traditional methods, such as considering a fixed threshold of 75%, which implies only 26% of the elements were considered to be over a minimum risky saturation level, and even less in the case of 95% fixed threshold. Meanwhile, our MOPSO algorithm not only takes into account congestion, but also time scenarios, in order to determine an optimal congestion threshold to 82% of the elements of the group of study, characterizing them individually and under an objective basis. Furthermore, 41% of the elements of the case study were provided with optimal congestion thresholds over 60% of their nominal admissible power,

evidencing that using this methodology gives an expanded view of congestion situations that are not negligible under preventive operation and maintenance standards.

The methodology can be applied to online processing. Nonetheless, an adaptation of the methodology may be necessary. The computational time required, and the granularity and size of the dataset considered, must be carefully assessed, having in mind that, as in the case of the clustering process considered here, data analytics techniques may help in handling those massive datasets. In addition to this, the availability of online data is a major issue, since proper data acquisition, processing, and provision to the methodology must be designed, for example by means of big data techniques [56], since, as the network conditions change, the optimization solution and clusters might change as well.

Author Contributions: Conceptualization, J.L. and J.A.A.; methodology, J.L. and R.C.P.; software, J.L. and R.C.P.; supervision, J.A.A.; validation, J.L.; writing—original draft, J.L.; writing—review and editing, R.C.P. and J.A.A.

Funding: This work has been partially funded by the project ENE2016-80638-R granted by *Ministerio de Economía y Competitividad* of the Government of Spain.

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

References

- Hague, A.N.M.M.; Nguyen, P.H.; Vo, T.H.; Blik, F.W. Agent-based unified approach for thermal and voltage constraint management in LV distribution network. *Electr. Power Syst. Res.* **2017**, *143*, 462–473. [[CrossRef](#)]
- You, S.; Segerberg, H. Integration of 100% micro-distributed energy resources in the low voltage distribution network: A Danish case study. *Appl. Therm. Eng.* **2014**, *71*, 797–808. [[CrossRef](#)]
- Chua, K.H.; Wong, J.; Lim, Y.S.; Taylor, P.; Morris, E.; Morris, S. Mitigation of Voltage Unbalance in Low Voltage Distribution Network with High Level of Photovoltaic System. *Energy Procedia* **2011**, *12*, 495–501. [[CrossRef](#)]
- Kadam, S.; Bletterie, B.; Gawlik, W. A Large Scale Grid Data Analysis Platform for DSOs. *Energies* **2017**, *10*, 1099. [[CrossRef](#)]
- Lu, S.; Repo, S.; Della, G.; Della Gustina, D.; Álvarez-Cuevas, F.; Löf, A.; Pikkarainen, M. Real-Time Low Voltage Network Monitoring-ICT Architecture and Field Test Experience. *IEEE Trans. Smart Grid* **2015**, *6*, 2002–2012. [[CrossRef](#)]
- Li, R.; Gu, C.; Li, F.; Shaddick, G.; Dale, M. Development of Low Voltage Network Templates-Part I: Substation Clustering and Classification. *IEEE Trans. Power Syst.* **2015**, *30*, 3036–3044. [[CrossRef](#)]
- Caldon, R.; Coppo, M.; Turri, R. Distributed voltage control strategy for LV networks with inverter-interfaced generators. *Electr. Power Syst. Res.* **2014**, *107*, 85–92. [[CrossRef](#)]
- Håberg, M.; Bood, H.; Doorman, G. Preventing Internal Congestion in an Integrated European Balancing Activation Optimization. *Energies* **2019**, *12*, 490. [[CrossRef](#)]
- Zhang, Y.; Wang, S.; Ji, G. A Comprehensive Survey on Particle Swarm Optimization Algorithm and Its Applications. *Math. Probl. Eng.* **2015**, *2015*, 931256. [[CrossRef](#)]
- Hazra, J.; Sinha, A.K. Congestion Management Using Multiobjective Particle Swarm Optimization. *IEEE Trans. Power Syst.* **2007**, *22*, 1726–1734. [[CrossRef](#)]
- Reihani, E.; Motalleb, M.; Ghorbani, R.; Saoud, L.S. Load peak shaving and power smoothing of a distribution grid with high renewable energy penetration. *Renew. Energy* **2016**, *86*, 1372–1379. [[CrossRef](#)]
- Tuballa, M.L.; Abundo, M.L. A review of the development of Smart Grid technologies. *Renew. Sustain. Energy Rev.* **2016**, *59*, 710–725. [[CrossRef](#)]
- Vandoorn, T.L.; de Vyver, J.V.; Gevaert, L.; Degroote, L.; Vandeveld, L. Congestion Control Algorithm in Distribution Feeders: Integration in a Distribution Management System. *Energies* **2015**, *8*, 6013–6032. [[CrossRef](#)]
- Cao, C.; Wang, L.; Chen, B. Mitigation of the Impact of High Plug-in Electric Vehicle Penetration on Residential Distribution Grid Using Smart Charging Strategies. *Energies* **2016**, *9*, 1024. [[CrossRef](#)]
- Romero-Ruiz, J.; Pérez-Ruiz, J.; Martín, S.; Aguado, J.A.; de la Torre, S. Probabilistic Congestion Management using EVs in a Smart Grid with intermittent Renewable Generation. *Electr. Power Syst. Res.* **2016**, *137*, 155–162. [[CrossRef](#)]

16. Esmat, A.; Usaola, J.; Moreno, M.A. Distribution-Level Flexibility Market for Congestion Management. *Energies* **2018**, *11*, 1056. [CrossRef]
17. Lalwani, S.; Singhal, S.; Kumar, R.; Gupta, N. A comprehensive survey: Applications of multi-objective particle swarm optimization (MOPSO) algorithm. *Trans. Comb.* **2013**, *2*, 39–101. [CrossRef]
18. Spanish Electricity System 2017 Report. Red Eléctrica de España. Available online: <https://www.ree.es/en/statistical-data-of-spanish-electrical-system/annual-report/spanish-electricity-system-2017-report> (accessed on 25 February 2019).
19. Instalaciones en Alerta. Organismo Supervisor de la Inversión en Energía y Minería de Perú. Available online: <http://www.osinergminorienta.gob.pe/web/rcc/catalogos> (accessed on 25 February 2019).
20. Siano, P.; Sarno, D. Assessing the benefits of residential demand response in a real time distribution energy market. *Appl. Energy* **2016**, *161*, 533–551. [CrossRef]
21. Sherif, I.; Shady, A.A.; Almoataz, A.; Ahmed, Z. State-of-the-art of hosting capacity in modern power systems with distributed generation. *Renew. Energy* **2019**, *130*, 1002–1020. [CrossRef]
22. Abdel-Majeed, A.; Braun, M. Low Voltage System State Estimation Using Smart Meters. In Proceedings of the 47th International Universities Power Engineering Conference (UPEC), London, UK, 4–7 September 2012; pp. 1–6. [CrossRef]
23. Niederhuemer, W.; Schwalbe, R. Increasing PV hosting capacity in LV grids with a probabilistic planning approach. In Proceedings of the 2015 International Symposium on Smart Electric Distribution Systems and Technologies (EDST), Vienna, Austria, 8–11 September 2015; pp. 537–540.
24. Roldán-Blay, C.; Escrivá- Escrivá, G.; Roldán-Porta, C. Improving the benefits of demand response participation in facilities with distributed energy resources. *Energies* **2019**, *169*, 710–718. [CrossRef]
25. Frey, B.J.; Dueck, D. Clustering by passing messages between data points. *Science* **2007**, *315*, 972–976. [CrossRef] [PubMed]
26. Zhang, X.; Furtlehner, C.; Germain-Renaud, C.; Sebag, M. Data Stream Clustering With Affinity Propagation. *IEEE Trans. Knowl. Data Eng.* **2014**, *26*, 1644–1656. [CrossRef]
27. del Valle, Y.; Venayagamoorthy, G.K.; Mohagheghi, S.; Hernandez, J.C.; Harley, R.G. Particle Swarm Optimization: Basic Concepts, Variants and Applications in Power Systems. *IEEE Trans. Evol. Comput.* **2008**, *12*, 171–195. [CrossRef]
28. Kennedy, J.; Eberhart, R. Particle swarm optimization. In Proceedings of the ICNN'95—International Conference on Neural Networks, Perth, Australia, 27 November–1 December 1995; pp. 1942–1948. [CrossRef]
29. Jones, K.O. Comparison of Genetic Algorithm and Particle Swarm Optimisation. In Proceedings of the International Conference on Computer Systems and Technologies, Varna, Bulgaria, 16–17 June 2005; Section IIIA.
30. Bai, Q. Analysis of Particle Swarm Optimization Algorithm. *Comput. Inf. Sci.* **2010**, *3*, 180–184. [CrossRef]
31. Chen, R.M.; Shen, Y.M.; Wang, C.T. Ant Colony Optimization Inspired Swarm Optimization for Grid Task Scheduling. In Proceedings of the International Symposium on Computer, Consumer and Control (IS3C), Xian, China, 4–6 July 2016; pp. 461–464. [CrossRef]
32. Selvi, V.; Umarani, R. Comparative Analysis of Ant Colony and Particle Swarm Optimization Techniques. *Int. J. Comput. Appl.* **2010**, *5*, 1–6. [CrossRef]
33. Endesa desarrolla un sistema para controlar en tiempo real la red de distribución. Available online: <https://www.endesa.com/es/prensa/news/d201811-endesa-desarrolla-un-sistema-para-controlar-en-tiempo-real-la-red-de-distribucion.html> (accessed on 25 February 2019).
34. Sathiskumar, M.; Kumar, A.N.; Lakshminarasimman, L.; Thiruvenkadam, S. A self adaptive hybrid differential evolution algorithm for phase balancing of unbalanced distribution system. *Int. J. Electr. Power Energy Syst.* **2012**, *42*, 91–97. [CrossRef]
35. Velez, V.M.; Hincapie, R.A.; Gallego, R.A. Low voltage distribution system planning using diversified demand curves. *Int. J. Electr. Power Energy Syst.* **2014**, *61*, 691–700. [CrossRef]
36. Smartcity Málaga Living Lab: A laboratory to create the city of the future. Available online: <https://www.endesa.com/en/projects/a201801-living-lab-malaga-city-future.html> (accessed on 25 February 2019).
37. Calderaro, V.; Conio, G.; Galdi, V.; Piccolo, A. Reactive power control for improving voltage profiles: A comparison between two decentralized approaches. *Electr. Power Syst. Res.* **2012**, *83*, 247–254. [CrossRef]

38. Soyster, A.L.; Eynon, R.T. The Conceptual Basis of the Electric Utility Sub-model of Project Independence Evaluation System. In Proceedings of the Seminar of the United Nations Economics Commission for Europe, Washington, DC, USA, 24–28 March 1980; Energy Modelling Studies and Conservation, 1982; pp. 417–429. [[CrossRef](#)]
39. Wu, A.; Ni, B. *Line Loss Analysis and Calculation of Electric Power Systems*, 2nd ed.; Wiley: Hoboken, NJ, USA, 2016. [[CrossRef](#)]
40. AENOR. *Standard UNE-EN 50160:2011/A1:2015*; AENOR: Madrid, Spain, 2015.
41. Zhao, Z.Y.; Tang, C.; Zhou, Q. Identification of Power Transformer Winding Mechanical Fault Types Based on Online IFRA by Support Vector Machine. *Energies* **2017**, *10*, 2022. [[CrossRef](#)]
42. Tenbohlen, S.; Coenen, S.; Djamali, M.; Müller, A.; Samimi, M.H.; Siegel, M. Diagnostic measurements for power transformers. *Energies* **2016**, *9*, 347. [[CrossRef](#)]
43. Wang, K.; Skiena, S.; Robertazzi, T.G. Phase balancing algorithms. *Electr. Power Syst. Res.* **2013**, *96*, 218–224. [[CrossRef](#)]
44. Carillo-Aparicio, S.; Heredia-Larrubia, J.R.; Perez-Hidalgo, F.M. Smart City Málaga, a real-living lab and its adaptation to electric vehicles in cities. *Energy Policy* **2013**, *62*, 774–779. [[CrossRef](#)]
45. Barbato, A.; Dede, A.; Della Giustina, D.; Massa, G.; Angioni, A.; Lipari, G.; Ponci, F.; Repo, S. Lessons learnt from real-time monitoring of the low voltage distribution network. *Sustain. Energy Grids Netw.* **2018**, *15*, 76–85. [[CrossRef](#)]
46. Gomez-Exposito, A.; Arcos-Vargas, A.; Maza-Ortega, J.M.; Rosendo-Macías, J.A.; Alvarez-Cordero, G.; Carillo-Aparicio, S.; González-Lara, J.; Morales-Wagner, D.; González-García, T. City-Friendly Smart Network Technologies and Infrastructures: The Spanish Experience. *Proc. IEEE* **2018**, *106*, 626–660. [[CrossRef](#)]
47. Navarro-Espinosa, A.; Mancarella, P. Probabilistic modeling and assessment of the impact of electric heat pumps on low voltage distribution networks. *Appl. Energy* **2014**, *127*, 249–256. [[CrossRef](#)]
48. Nijhuis, M.; Gibescu, M.; Cobben, S. Gaussian Mixture Based Probabilistic Load Flow for LV-Network Planning. *IEEE Trans. Power Syst.* **2017**, *32*, 2878–2886. [[CrossRef](#)]
49. Rodriguez-Calvo, A.; Cossent, R.; Frias, P. Integration of PV and EVs in unbalanced residential LV networks and implications for the smart grid and advanced metering infrastructure deployment. *Int. J. Electr. Power Energy Syst.* **2017**, *91*, 121–134. [[CrossRef](#)]
50. Tehrani, M.K.; Fereidunian, A.; Lesani, H. Financial planning for the preventive maintenance of power distribution systems via fuzzy AHP. *Complexity* **2016**, *21*, 36–46. [[CrossRef](#)]
51. Long, C.; Ochoa, L.F. Voltage Control of PV-Rich LV Networks: OLTC-Fitted Transformer and Capacitor Banks. *IEEE Trans. Power Syst.* **2016**, *31*, 4016–4025. [[CrossRef](#)]
52. Qian, Y.; Yao, F.; Jia, S. Band selection for hyperspectral imagery using affinity propagation. *IET Comput. Vis.* **2009**, *3*, 213–222. [[CrossRef](#)]
53. Das, S.; Mullicka, S.S.; Suganthan, P.N. Recent advances in differential evolution—An updated survey. *Swarm Evol. Comput.* **2016**, *27*, 1–30. [[CrossRef](#)]
54. Trebuña, P.; Halčinová, J.; Fil’o, M.; Markovič, J. The importance of normalization and standardization in the process of clustering. In Proceedings of the 12th International Symposium on Applied Machine Intelligence and Informatics (SAMII), Herľany, Slovakia, 23–25 January 2014; pp. 381–385. [[CrossRef](#)]
55. Elías, X.; Bordas, S. *Energía, Agua, Medioambiente, Territorialidad y Sostenibilidad*, 1st ed.; Díaz de Santos: Madrid, Spain, 2011.
56. Jiang, Z.; Liu, K. Real Time Interpretation and Optimization of Time Series Data Stream in Big Data. In Proceedings of the 3rd IEEE International Conference on Cloud Computing and Big Data Analysis (ICCCBDA), Chengdu, China, 20–22 April 2018; pp. 243–247.

