An Interdisciplinary Approach on Efficient Virtual Microgrid to Virtual Microgrid Energy Balancing Incorporating Data Preprocessing Techniques

Paraskevas Koukaras[1,2] 0000-0002-1183-9878, Christos Tjortjis*[1,2] 0000-0001-8263-9024, Paschalis Gkaidatzis[1] 0000-0003-4444-5370, Napoleon Bezas[1], Dimosthenis Ioannidis[1] 0000-0002-5747-2186, Dimitrios Tzovaras[1] 0000-0001-6915-6722

[1]Information Technologies Institute, Centre of Research & Technology - Hellas
6th km Harilaou - Thermi, 57001, Thessaloniki, Greece

[2]School of Science and Technology, International Hellenic University, 14th km Thessaloniki – N. Moudania, 57001 Thermi, Thessaloniki, Greece

[1]{p.koukaras, c.tjortjis, pgkaidat, bezas, djoannid, dimitrios.tzovaras}@iti.gr

Abstract

A way to improve energy management is to perform balancing both at the Peer-to-peer (P2P) level and then at the Virtual Microgrid-to-Virtual Microgrid (VMG2VMG) level, while considering the intermittency of available Renewable Energy Source (RES). This paper proposes an interdisciplinary analytics-based approach for the formation of VMGs addressing energy balancing. Our approach incorporates Computer Science methods to address an Energy sector problem, utilizing data preprocessing techniques and Machine Learning concepts. It features P2P balancing, where each peer is a prosumer perceived as an individual entity, and Virtual Microgrids (VMGs) as clusters of peers. We conducted several simulations utilizing clustering and binning algorithms for preprocessing energy data. Our approach offers options for generating VMGs of prosumers, prior to using a customized Exhaustive brute-force Balancing Algorithm (EBA). EBA performs balancing at the cluster-to-cluster level, perceived as VMG2VMG balancing. To that end, the study simulates on data from 94 prosumers, and reports outcomes, biases, and prospects for scaling up and expanding this work. Finally, this paper outlines potential ideal usages for the approach, either standalone or integrated with other toolkits and technologies.

Keywords

Machine Learning, Preprocessing, Data Analysis, Virtual Microgrid, Energy Balancing

Abbreviations

Demand Response (DR), Peer-to-peer (P2P), Photovoltaic (PV), Renewable Energy Source (RES), Distribution System Operator (DSO), Distributed Energy Resource (DER), Demand Response (DR), Microgrid (MG), Virtual Microgrid (VMG), VMG to VMG (VMG2VMG), Virtual Power Plants (VPP), Exhaustive Balancing Algorithm (EBA), Peer-to-Grid (P2G), Smart Grid (SG), Artificial Intelligence (AI) and Machine Learning (ML), Time of Use (ToU), Real Time Pricing (RTP), Critical Peak Pricing (CPP), Key Performance Indicator (KPI), nearly Zero Energy Building (nZEB), Total Active Energy (TAE)

Declarations

Funding The project eDREAM has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement N. 774478.

Conflicts of interest/Competing interests Researchers in Information Technologies Institute, Centre of Research & Technology or School of Science and Technology, International Hellenic University.

Availability of data and material (data transparency): Data partly available in the appendices due to confidentiality clauses for the ongoing eDREAM project.

Code availability (software application or custom code): N/A due to confidentiality clauses for the ongoing eDREAM project.

Acknowledgments

This work has been conducted within the eDREAM project Grant number 774478, co-funded by the European Commission as part of the H2020 Framework Programme (H2020-LCE-2017-SGS).

1. Introduction

Enhancing Microgrid (MG) management has recently drawn significant research attention [1]. MGs offer great opportunities for improving energy distribution through balancing, CO_2 emissions reduction, energy production, cost reduction etc. [2].

The main challenges that the energy sector faces refer to securing seamless and reliable power system operation, the best way possible [3]. For example, Renewable Energy Source (RES) integration could aid reducing the fluctuations on daily energy loads paired with improvements in energy storage. However, RESs are intermittent in nature. Therefore, novel decision support systems are required to manage their smooth integration. Moreover, it is important to utilize them in various ways for minimizing power losses, as well as optimal energy management regarding peak demands [4].

We need ways to improve energy management in the presence of RES. Energy balancing at the P2P and VMG2VMG level can act as a tool for that. This is important since it could cause: i) potential energy cost reduction in energy transactions between VMGs and Aggregators, as well as Aggregators and the Distribution System Operator (DSO), ii) main grid tariff costs reduction, iii) better energy distribution by enabling power tradeoffs in-between prosumers and iv) prosumer participation to energy-sharing programs. Energy can be transferred either uni- or bi-laterally. Opportunities for Demand Response (DR) power transfer at the P2P level arise. These points offer novel advancements in the energy sharing research domain.

Reviewing the literature, as detailed in section 2, shows that energy balancing draws the attention of many researchers from various fields. This can be attributed to the wide range of applications and domains required for developing the next generation of power systems, ie. The Smart Grid (SG). These requirements attempt to meet the increasing needs for technological implementations brought to the energy sector by the rapid spread of RES integration. Energy distribution/balancing is considered a challenging domain, as it involves multiple factors, such as RES intermittency, peak hour demands, DR signals, etc. Historical data can offer an insight into the times that energy demand is at its peak. Stochastic indicators can be retrieved, about the times during the day when grid functionality relaxes, or energy demands are nearly even.

This paper elaborates on forming *clusters* or *bins* of prosumers to achieve improved energy management for Virtual Microgrids (VMGs). It adopts an interdisciplinary approach, incorporating concepts from the domains of Computer Science and Energy. Energy management is an important task as confirmed by the review of the literature. Balancing energy inputs/outputs to the grid, the main aim of this work, is equally important. This paper provides an analysis offering various functionalities for VMG balancing, that can be incorporated to a stand-alone toolkit. These functionalities can be combined with external components. Alternatively, they can act as baseline architectural system components, such as Blockchain, Energy Business Intelligence. This way they can address other energy related problems, like minimizing costs/kWh and CO_2 emissions. Thus, they can help to conceive a complete energy management system.

The dataset we used for simulations refers to 94 prosumers. Our goal is balancing at the Peer-to-peer (P2P) and VMG2VMG level. We verify its achievement by multiple simulations and visual presentations. The simulations incorporate clustering and binning algorithms to handle energy prosumer data, as reported in section 5. An Exhaustive brute-force Balancing Algorithm (EBA) handles VMG2VMG energy balancing. Simulation outputs can be combined with the Use Cases described in section 6.1. For example, i) cost minimization or profit maximization, during the energy transactions between VMGs and Aggregators and Aggregator and DSO, ii) opportunities arising from VMG2VMG energy transactions, when RES are used at a large scale and iii) VMGs capability to utilize P2P energy transactions, i.e. energy trade among prosumers.

This work presents an interdisciplinary approach investigating the energy balancing problem. Its contributions can be summarized as follows:

1. A heuristic approach utilizing unsupervised learning to form and balance VMGs.
2. A high-level, generic approach that addresses P2P and VMG2VMG balancing. The presentation of simulations combining clustering/binning for preprocessing is detailed. This approach can act as a standalone toolkit for baseline estimation, or as part of a comprehensive energy management system.
3. A potential solution to the energy balancing problem which incorporates a Computer Science centric approach for an Energy sector problem. The investigation conducted emphasizes the need for multidisciplinarity for such problems, due to the increased complexity of fast evolving energy management systems.

The remainder of this paper is structured as follows: section 2 sets the context with the presentation of key background concepts and a review of the state of the art; section 3 states the problem; the core concepts of the proposed VMG formation and balancing methodology are presented in section 4; the simulations conducted are detailed in section 5; the paper concludes in sections 6 and 7 by presenting a research summary, elaborating on final thoughts, implications, biases and directions for future work.

2. Background & Literature review

This section sets the context for our research, introducing key background concepts and reviewing the state of the art.

2.1 Background

2.1.1 *Peer-to-peer energy sharing*

Peer-to-peer (P2P) energy sharing provides a modern way to exchange energy, in a distributed manner that is more consistent with the SG concept. It allows MGs or prosumers, possibly consisting of various Distributed Energy Resources (DER), to exchange surplus energy. This is more appealing than conventional Peer-to-Grid (P2G) trading, since it is handled in a distributed manner, among peers, rather than in a centralized one, as has been the case in P2G. However, for P2P to function, a detailed IT infrastructure is needed, including sensors, energy meters and communication systems, called the Internet of Energy (IoE) [5], to establish a management tool that effectively offers the necessary interaction within the energy market [6]. At the same time, great improvements to cost effectiveness of P2P mechanisms are required, so that they can offer financial advantages [7].

2.1.2 MGs & VMGs

MGs have been tested and evaluated over the past decades through lab work, pilot demonstrations and expos, ultimately debuting in energy markets. Their operation and implementation have been radically improved, offering technological novelties, cost optimizations and other benefits. Such benefits include better electricity infrastructure, reliability and better green energy resource integration, such as Photovoltaic (PV), wind turbine etc. [8], or even underwater kite systems [9] through improved local and more immediate control and coordination among them. At the same time, they have contributed to reducing CO_2 emissions via optimally utilizing RESs. Moreover, they aided electricity provision to isolated areas, where a connection to the rest of the grid was previously considered either impossible, or unaffordable.

VMGs emerge profoundly as they can expand MG capabilities, by releasing them from their physical form. Research attempts aim at expanding capabilities through novel hybrid designs for improving the desired control of MGs [10]. They conceptually target solving the problem of optimized DR, scheduling within limited geographical areas with integrated RES (PV installations, Wind systems, etc.), enhancing the current functionality of MGs.

2.1.3 Consumer/Prosumer/Prosumagers

The electricity sector is being transformed as DER integration rises [11]. At the same time, price of energy storage drops daily, enabling prosumers to be transformed into prosumagers [12].

Prosumagers can produce, consume, and store their own energy using an effective management system. Each prosumager can be considered a MG. New opportunity arise regarding P2P energy transactions, DR optimization and the concept of Virtual Power Plants (VPPs). VPP is another form of VMGs, where loads, generation and energy storage are managed optimally and efficiently, regardless of physical proximity.

2.1.4 Power Distribution

Regarding power distribution, current directives from the European Commission promote the transit from a conventional power system model, as shown in Figure 1, to the SG one, shown in Figure 2. To that end, the generation/load balance needs to be improved [13]. Especially under high-RES integration, or when there are optimization requirements with regards to DR signals.

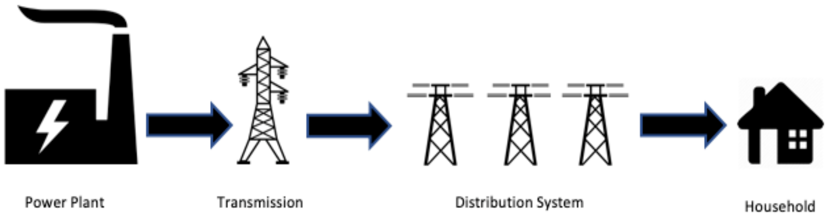


Figure 1: Conventional power system structure

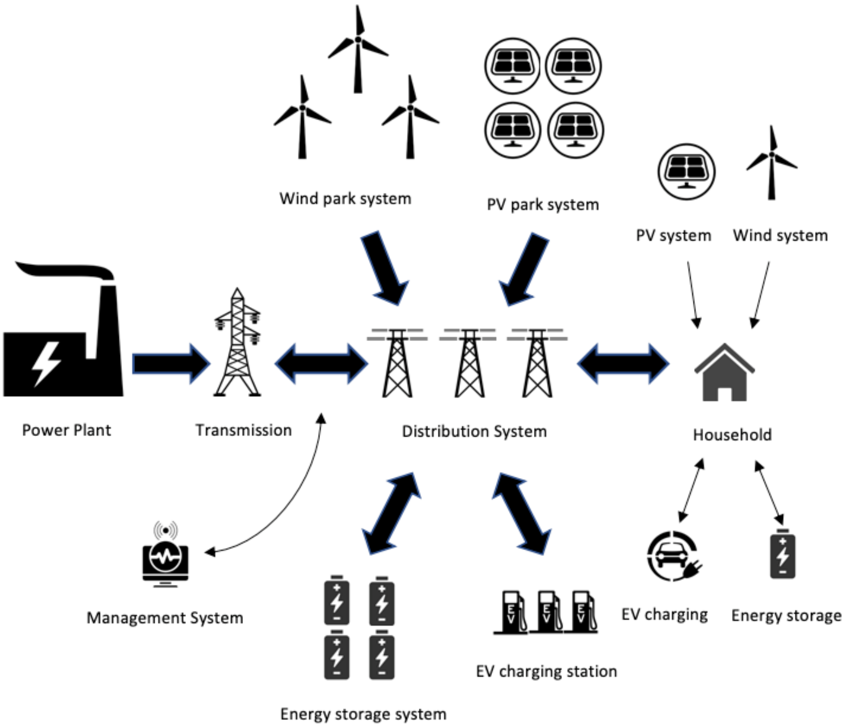


Figure 2: Smart Grid structure

Therefore, the energy management system needs to incorporate a variety of Computer Science related technologies. For example, Artificial Intelligence (AI) and Machine Learning (ML) techniques have proved useful for forecasting techniques. These are in turn required for various optimized schedules of DERs either in long-term basis (i.e., a day) or short-term basis (i.e., a few hours or minutes ahead) [14-15]; Internet of Things (IoT), Blockchain and Big Data have been useful for the P2P energy transactions [16]. Practical examples of the aforementioned are presented in sections 2.1.5 and 2.2 that follow.

2.1.5 Demand Response (DR) programs

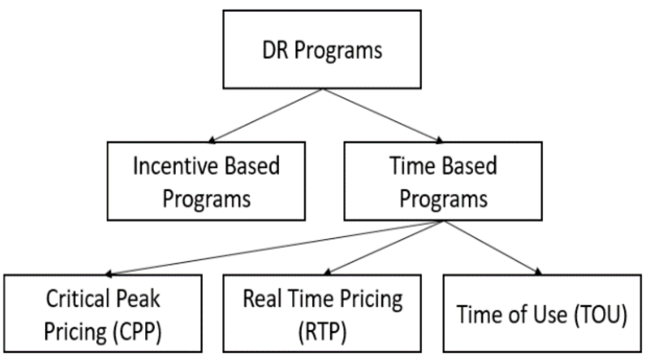


Figure 3: Demand Response categories

For DR programs, there are two main categories related with this work: incentive- and time-based programs [17]. There are time-based programs for kWh pricing (Figure 3*).* This study distinguishes the Time of Use (ToU), Real Time Pricing (RTP) and Critical Peak Pricing (CPP), where the price per kWh depends on the supply cost during different periods.

According to ToU a period of a day is divided into various pricing periods that correspond to different electricity prices, e.g., low, medium and high.

According to RTP, the electricity prices in the retail market, follow mostly the price fluctuations of the wholesale electricity market. Therefore, prices are adjusted dynamically. In both cases the fluctuation of the prices, either hourly in RTP, or in a period of hours in ToU, reflect the fluctuations of the load, that is the demand of electric energy.

2.2 Literature review

This section is devoted to a review of the research relevant to our approach, incorporating concepts from the research fields of Computer Science and Energy. More specifically, it showcases innovative attempts regarding MGs while presenting state of the art implementations in DR programs. The aforementioned concepts are widely used in this work.

Work on energy management strategies attempts to improve the MG energy exchange mechanism by utilizing their DR and energy storage capacity [18]. It focuses on measuring three KPIs: a) load or energy consumption patterns, b) the energy by distributed generation resources and c) an electricity cost reduction system.

The work in [19] elaborates on energy sharing mechanisms inside a MG that comprises P2P prosumers. It takes into consideration the willingness of prosumers to shift their loads to the grid. The cost per kWh is calculated after a consensus is reached, among all the prosumers involved in this energy sharing. They also propose a day- or an hour-ahead pricing model. This is achieved utilizing a distributed iterative algorithm. For testing their framework, they use real data and compare it to using the traditional trade via feed-in-tariff, while improving the local consumption of PV generated energy.

Another general evaluation methodology targets P2P cost-efficiency energy sharing models [20]. It identifies the values, estimates billing, and presents the performance index value of P2P sharing models. It comes up with the result that the Supply and Demand Ratio (SDR) model’s economic performance is better than that of Mid-Market Rate (MMR), whilst both are much better than Bill Sharing (BS) models.

The work in [21] addresses the intermittence of renewable sources generation. There are various options regarding the utilization of energy sources. Also, more than one energy sources can be exploited simultaneously. This way, greater prospects are provided, since there is no unique dependency on renewables, but rather on other sources as well. In order to develop such a multi-objective system, the authors use historical data to forecast the renewables’ complementary energy output. DR and Demand-Side Management (DSM) are necessary to calculate the need of intermittent capacity and to schedule production variations. They optimize the RES mix, as well as the minimum peak load share and minimum global cost, by calculating RES complementarity.

The study in [22] elaborates on industrial MGs. They incorporate and examine the concepts of grid connectivity, dispatchable and non-dispatchable energy sources and storage control systems. They highlight the benefits of operating DERs in MGs, such as CO_2 emissions and cost reduction during times of peak loads on the grid. Another study deals with an energy management system based on high renewable energy usage on MGs [23]. They deal with the intermittent nature of RES generation considering various attributes, such as loads, DR/generation costs, worst-case transaction costs etc. They use a dual decomposition method to break the problem into smaller chunks of problems, outputting the optimal solution.

In [24], Quiggin et al. use a linear programming approach in order to model a MG as a mix of RESs, storage and DR system. Their work resulted in an improved DR with balanced fluctuations. Similarly, in [25] Ding et al. propose a mixed-integer-linear-programming management system for industrial DERs. They optimize DER functionality by using day-ahead electricity prices to manipulate peak demands lowering the energy costs of the industrial facility. Hawkes and Leach proposed a linear programming cost minimization model for designing MG with DERs [26]. They concluded that grid-connected MGs can be more economical than isolated, or islanded ones.

A P2P energy trading in a MG is presented in [27]. The authors consider small scale DERs for local energy trading between prosumers and consumers. A hierarchical system architectural model is proposed to identify the key elements and technologies required for P2P energy trading. Tests show that P2P energy trading can improve the local balancing, regarding energy generation and consumption, especially when low voltage MGs can be utilized. Another approach to energy P2P energy trading utilizing VMGs is developed with a game theoretic approach [28]. The results show that by utilizing VMGs, improvements regarding CO2 reduction along with energy cost reduction can be achieved.

Vergados et al. address the issue of generating clusters of prosumers, thus forming VMGs, in order to participate to the energy market as a single entity, reducing total energy cost and forecasting inaccuracies [29]. They used a dataset of 33 prosumers, and experimented with different clustering algorithms (spectral, genetic and an adaptive algorithm). The results show that a cost reduction can be achieved by grouping prosumers to VMGs given that VMG aggregators emerge.

A scalable DR framework featuring a Blockchain based near real time DR validation scheme is presented in [30]. The proposed scheme allows the involved prosumers to verify the authenticity and integrity of all the DR events. Ongoing progress regarding a decentralized energy flexibility system is presented focusing on the requirements and the use cases were developed using ledger technologies [31]. The usage of decentralized Blockchain mechanisms is evaluated in [32], as well as their capabilities to deliver reliable, timely and secure energy flexibility in energy demand profiles of Prosumers. These will be utilized by stakeholders of the energy flexibility markets, such as DSO, Aggregators etc.

The state of the art regarding energy forecasting requirements evaluates when prosumers are to be integrated with SGs. Energy demand is assessed individually (per prosumer) and in an aggregated way (sum of prosumers) forecasting the energy demand in order to inform the DSO about possible grid imbalances [33]. A platform called FUSE, integrates multi-purpose functionalities regarding the energy sector. This platform is designed based on the state of the art and the most recent real-life requirements and expectations from energy distribution systems [34].

Another paper envisions the usage of 2nd life Li-ion batteries for enhancing distribution network operations. Batteries are to be utilized as “Storage as a service” for two use cases. In the first case the DSO manages the battery storage system aiming to increase power quality and improve Low Voltage management efficiency. In the second use case the batteries are managed by a District Energy Manager (DEM) performing peak saving and power smoothing according to a profile that is requested by the DSO. Results show that the usage of 2nd life Li-ion batteries looks promising for the power distribution network. Experimentation on DEM is ongoing [35].

In [36] researchers, address the issues regarding slow adoption of Blockchain technologies in the energy sector. Such issues are related to the low scalability and high processing overhead when dealing with near real time data from smart energy meters. A scalable 2nd tier solution is proposed, combining Blockchain ledger with distributed queuing systems and NoSQL databases. Thus, energy transactions are registered less frequently on the system, while retaining the tamper-evident benefits of Blockchain technology. Also, researchers propose a tamper-evident registration methodology for energy data transactions. Results show that the prototype looks promising regarding scalability and tampering of energy data.

The problem of inefficient and unsustainable operation of data centers with regards to their optimal fusion with the local energy grids is addressed in [37]. A solution is proposed by dealing with matters related to the active participation in DR programs. Another study reviews the most recent state of the art regarding Blockchain initiative incorporating a 3-tier architecture enabling the classification of all technological solutions based on decentralized applications in the energy sector [38].

In [39] a particular interconnected MG system is proposed, with the aim to optimally coordinate the operation of several MGs within a market environment. It considers the technical constraints of the distribution network. In [40] authors propose a game theoretic non-cooperative distributed coordination control scheme. It addresses the multi-operator energy trading and facilitates a powerful control structure for MMGs. The difficulty of such a task is identified, when multiple MGs are contemplated and especially when multiple beneficiaries coexist.

In this paper, the proposed approach tries to facilitate such a process by identifying VMGs with similar needs. We use clustering and binning techniques attempting to generate a methodology that exchanges energy more effectively, leading to improved energy sharing and balancing.

3. Problem Statement

Energy balancing in a portfolio of residential, commercial and industrial assets can be quite challenging, due to the great intermittency and stochasticity involved i.e., diverse energy consumption behavior, weather conditions, the intermittency of RES and more. Also, there are local ecosystem constraints, such as power loads, a variety of RES, costs, grid topology or local fees that ensure the fairness and stability of such an energy system [41]. Often the available assets form VMGs conceptualized as clusters, which have specific parameters i.e., Key Performance Indicators (KPIs), such as energy consumption or generation (*Figure 4*).

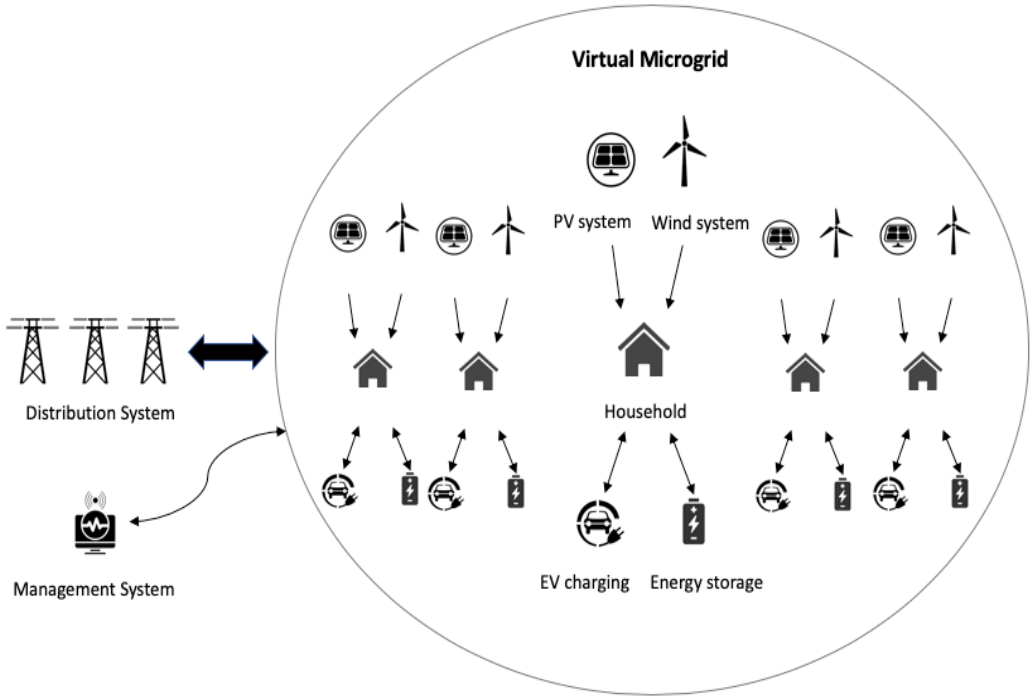


Figure 4: Generic VMG structure

In order to achieve better energy management, especially when extensive usage of RES can be exploited, we need ways for Energy balancing at the P2P and VMG2VMG level. The DSO and Aggregators need tools that aid accuracy regarding DER power operation information, while exposing new possibilities regarding portfolio management. The information available should be presented in an aggregated form, enhancing monitoring capabilities of both DSO and aggregator; the former in terms of managing the grid, the latter in terms of managing its portfolio. Most of the approaches like in [27-29] either attempt to generate VMGs to reduce the cost of energy or address the issue of energy balancing at the P2P level.

Developing a tool addressing the points mentioned above may offer a long-term improvement on applied balancing strategies. At some point, these can be addressed by introducing the idea of local or global optimal balancing. Such balancing can expose *diagnostic* and *prescriptive* capabilities in data analytics, similar with other identified data science domains [42]. This exposes new possibilities, since prosumer portfolios can be virtually clustered and balanced. Local balancing is perceived through balancing at the P2P level, while global balancing is perceived after performing balancing at the VMG2VMG level.

Implications include the combination of such a tool with a cost minimization process that involves transferring energy from individual VMGs to the VMG portfolio of an aggregator or the DSO. Before energy transfer to the latter commences, a global VMG2VMG balancing could be performed. Further investigation, associated with VMG2VMG energy transfer during excessive usage of RES, could take place. VMGs enable tradeoffs of energy in-between VMGs, offering optimal energy distribution in close proximity, and maximizing savings from main grid tariffs. Such a tool should be provided as software, hardware or data as a service. For deployment and operation, cloud computing technologies may offer a complete information technology solution [43].

4. Methodology

This section presents a high-level overview of the methodology, including the techniques and the dataset used. Results from simulations along with outputs are presented in the next section.

4.1 Overview

This study elaborates on an interdisciplinary approach for energy balancing, based on Energy and Computer Science introducing a methodology and testing the validity of its results. The approach utilizes ML preprocessing concepts, such as clustering, binning, and a customized Exhaustive Balancing Algorithm (EBA). It exploits clusters and bins of energy consumption to address various use cases, presented in 6.1.

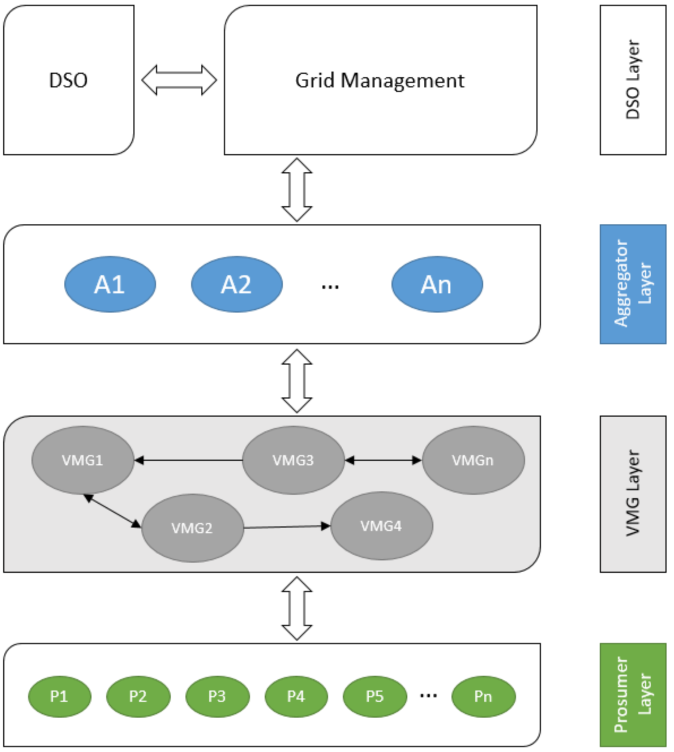


Figure 5: Layers of balancing

Figure 5 demonstrates the conceptual layers of the architectural interaction between VMG2VMG, VMG to Aggregator/s, and Aggregators to DSO. The approach focuses on the VMG and Prosumer layers circulated in red in the figure, running multiple simulations to support a methodology for local and global balancing.

By showcasing a solution in the form of a VMG formation problem, this study also envisions to reduce costs or increase profits. To that end, the proposed methodology enables a possible DR strategy accompanied with specific constraints, detailed in section 2.1.5. The DR strategy can be exploited given that VMG aggregators emerge in the energy market/ecosystem. This tackles the issue of energy balancing, based on unsupervised learning. Our goal is to improve portfolio management, using a visualization-clustering/binning approach offered as a service, while enhancing DR functionality and efficiency for energy stakeholders i.e., Aggregators and DSOs [44].

Our methodology addresses local P2P Balancing and global VMG2VMG Balancing with four simulations that test and validate our approach. These simulations, detailed in section 5, involve data preprocessing, rule enforcement and decision making, using an EBA. Simulations exploit data from a full day of energy usage, splitting load-demand measurements, as well as grid inputs/outputs, in the 24-hour scale. Thus, we use a "timestamp" variable, ranging between 1 and 24.

4.2 Dataset

The dataset to test the proposed approach contains records corresponding to prosumers with various attributes, as listed in Table 1. These data were generated by capturing various measurements from CERTH/ITI nearly Zero Energy Building (nZEB) Smart Home [45]. This nZEB incorporates a wide variety of novel technologies. It is a rapidly evolving prototyping infrastructure equipped with IoT devices and technologies that allow to measure and validate many experimental conditions attempting to imitate real life living conditions.

The data for experimentation were retrieved from the extended nZEB API. That way access is granted to measurements from hundreds of devices within the Smart Home. The energy measurements (power generation and consumption to and from the grid, respectively) were collected for the period between 2017 and 2020. The rationale for choosing this data source for simulations resides in that nZEB excels at mitigating residential housing energy profiles by incorporating distinct housing units and rooms. That way distinct measurements can be retrieved, being either energy generation or consumption [46]. Therefore, according to preliminary results from ongoing research, this study distinguishes around 100 possible residential energy profiles along the observation of four years of energy generation and consumption patterns based on past research results [47]. This admission forms the baseline approach for considering these energy residential profiles to mimic individualized prosumers to be used for the simulations. To that end, a conceptual, yet very close to real life conditions, prosumer dataset emerges.

Furthermore, python's pandas library [48] is utilized for enforcing various data preprocessing techniques, like removing duplicates and missing values as well as data transformation and reduction needed to normalize and clean the dataset. The raw dataset is a 980 MB .csv file with attributes listed in Table 1.

Table 1: Initial (raw) dataset description

Attribute	Type of value
A	Meter Id
B	Quarter-hour progressive number in the day
C	Latitude of prosumer and longitude
D	Total Active Energy (TAE) delivered from the prosumer, injected into the grid
E	TAE absorbed by the prosumer from the grid
F	Fixed value (15): indicating that data gathering is carried out every 15 minutes
G	Datetime

The preprocessed dataset is a 33.781 MB .csv file that contains 815.208 rows of data, which represents entries for 94 conceptual prosumers with attributes listed in Table 2.

The raw dataset time resolution was in quarters of an hour. Yet, it is aggregated to match the 1-hour resolution, for enhancing the observation clarity for the presentation of the results.

Table 2: Preprocessed dataset description

Attribute	Description
A (Id)	Prosumer Id
B (Datetime)	Datetime in format (“%Y-%m-%d %H:%M:%S”)
C (coordinates)	Latitude of prosumer and longitude in the format
D (Ea)	TAE absorbed by the prosumer from the grid
E (E0)	TAE delivered from the prosumer to the grid

The dataset is transformed based on the description shown in Table 3 to match the requirements of our approach. Measurement units are also added.

Table 3: Dataset used for simulations

Attribute	Description
A (ID)	Prosumer Id
B (Datetime)	Datetime in format (“%Y-%m-%d %H:%M:%S”)
C (Ea)	TAE absorbed by the prosumer from the grid in Watts.
D (E0)	TAE delivered from the prosumer to the grid in Watts.

As an extra preprocessing step, coordinates attributes are discarded since our approach uses data from various prosumers in an assumed close vicinity .

Finally, columns C (E0a) with D (E0) are merged into column C (E0a) (as shown in Table 4).

Table 4: Dataset after merging columns

Attribute	Description
A (ID)	Prosumer Id
B (Datetime)	Datetime in format (“%Y-%m-%d %H:%M:%S”)
C (E0a)	TAE absorbed (in Watts) by the prosumer from the grid is signed as negative, while the TAE delivered from the prosumer to the grid is signed as positive.

4.3 Methods and algorithms

In order to achieve P2P balancing towards the envisioned VMG2VMG balancing, this study utilizes a hybrid approach performing clustering through rule enforcement, seeking to attain and verify achievements. A detailed presentation of this mixture of techniques is presented in section 5.

To formulate the proposed approach of VMG generation, we conceive rules that create classes of prosumers. That way we enable an initial categorization of the prosumers in the dataset. Therefore, three district classes of prosumers are generated, each class conforming to Rules#1-3. Each rule, when satisfied assigns a prosumer to a specific class. These classes of prosumers are:

a) The ones whose energy generation is less than their consumption, i.e. they draw energy from the grid (class1).

b) The ones who balance their generation and consumption levels, i.e. a minimal interaction with the grid (class2).

c) The ones whose energy generation is greater than their consumption, i.e. they inject energy to the grid (class3).

Therefore, we distinguish three classes that associate with any entries indexed by timestamp and other related values within the dataset. In that way, the retrieval of the aggregated values of generation/production for each class is straightforward. The dataset description utilized for enforcing the rules is presented in Table 4. Figure 6 shows the energy data from all prosumers at a specific timestamp, before enforcing the rules. Figure 7-9 depict the formation of each class on a specific timestamp.

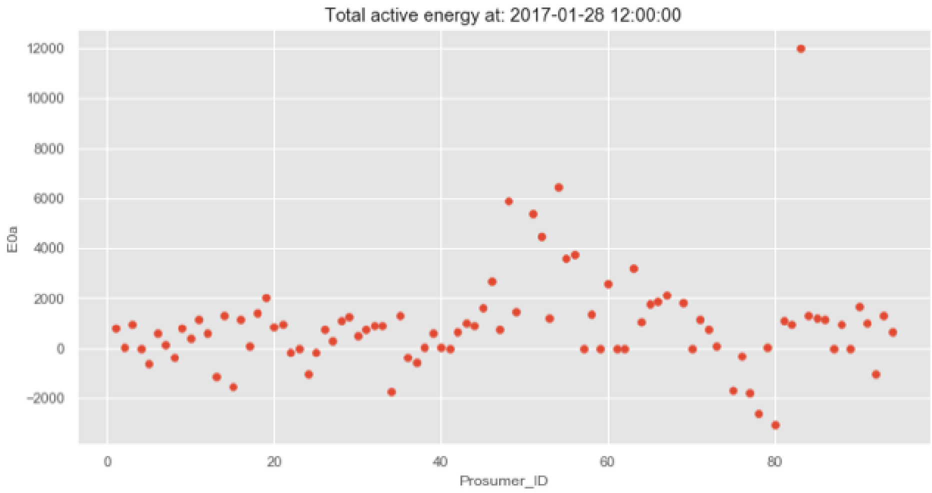


Figure 6: TAE per prosumer on a specific timestamp

Applying Rule#1, we generate a class of prosumers (class1) who draw power from the grid. More precisely, all prosumers with E0a < -1 (Watt) enroll to this class for the specific timestamp, as shown in *Figure 7*.

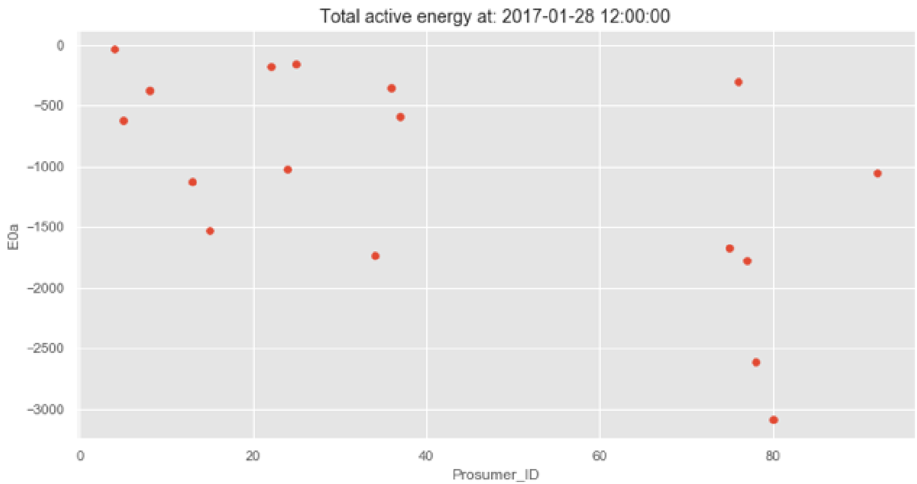


Figure 7: TAE per prosumer abiding with Rule#1

Applying Rule#2, we create generate a class of prosumers (class2) who virtually neither draw from nor inject power to the grid. To be precise, all prosumers with:

$$-1 \leq E0_a \leq +1 \; (Watt) \; (1)$$

enroll to this class for the specific timestamp (*Figure 8*). These value boundaries for Rule#2 are set because injection to or draws from the grid between -1 and +1 Watt can be considered negligible.

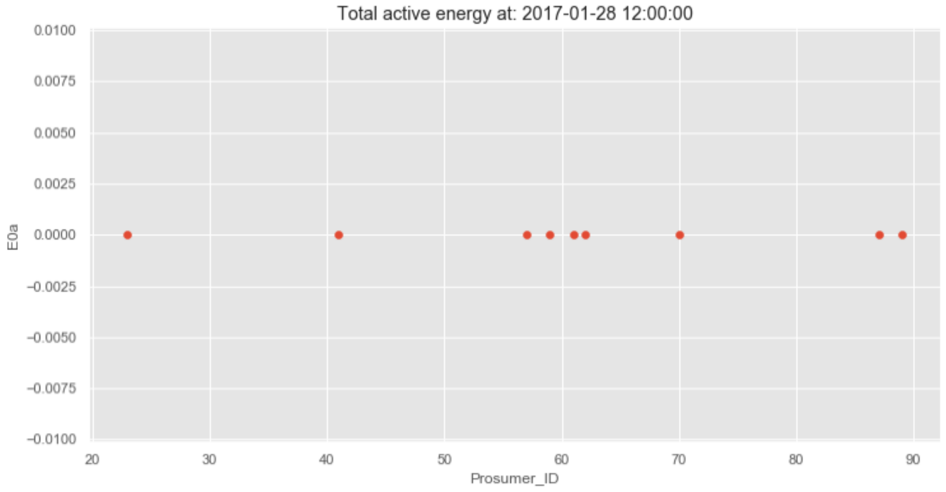


Figure 8: TAE per prosumer abiding with Rule#2

Applying Rule#3, we generate a class of prosumers (class3) who inject power into the grid. To be precise, all prosumers with E0a > +1 (Watt) are enrolled to this class, for that specific timestamp, shown in Figure 9. At this point, various measurements can be utilized regarding the information contained in each class at a specified timestamp, such as the exact sum/max/mean/std (standard deviation) draw from/injection to the grid, the number of prosumers etc. Next, this information is used for the process of VMG formation.

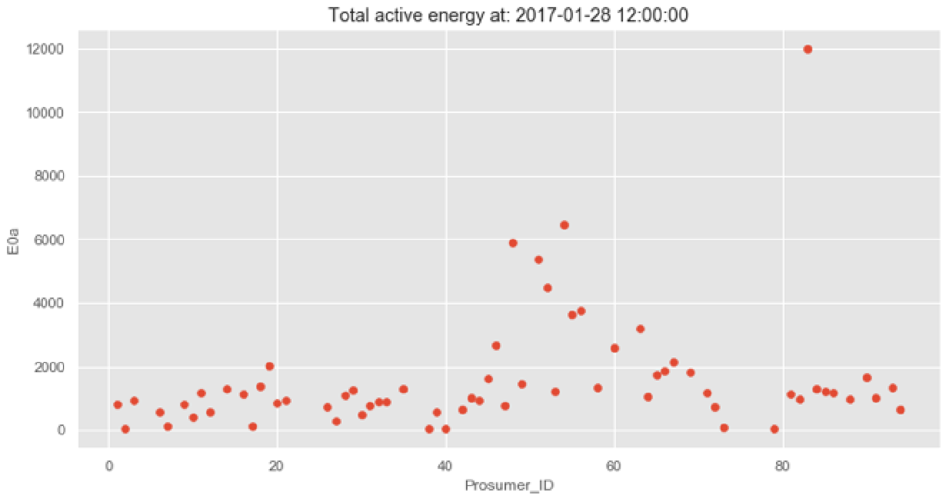


Figure 9: TAE per prosumer abiding with Rule#3

K-means [49] CUT and QCUT [48] cluster or bin prosumers, respectively. QCUT is based on the enforcement of the quantile-based discretization function. It is used on the raw data to discretize the consumptions/productions into equal-sized bins (buckets), based on ranks or based on the dataset entries' quantiles. For example, 100 production/consumption values for five quantiles produce a categorical object indicating quantile membership for each of the consumption values.

CUT is based on the enforcement of bin values into discrete intervals. This method is used on the raw data attempting to segment and sort the consumptions/ productions into bins. It performs better when the aim is to transpose from continuous variables to categorical. For example, 100 production/ consumption values for five bins produce groups of production/ consumption value ranges.

The EBA algorithm balances individual, or aggregated consumption/generation energy values resulting from prosumer clusters or bins. Also, Within-Cluster Sum of Squares (WCSS) [50] and the silhouette score [50] determine the best value k for clusters, as an input to the k-means algorithm. There are many methods for defining the number of clusters. For example, the silhouette score calculates the average similarity of objects in a cluster and the distance from objects in other clusters. As a method for defining number of clusters, it involves more computations, but it is more informative [51]. The elbow method iterates on an algorithm that attempts to measure a point where a clustering score starts to decrease the most. In practice when a sharp elbow appears it may cause misinterpretations [52]. For the purposes of this paper, although we also experimented with elbow method and WCSS, we selected the silhouette score since it is more thorough, and less prone to wrong inspection of results.

5. Simulations and Results

To test the proposed approach, tests are carried out (Simulations #1-4) on data regarding hourly timestamps for a specific day. The prosumers retrieved from the utilized dataset have PVs for their generation. Therefore, for improved prosumer profile the time horizon of (08:00-17.00) for a specific day has been chosen; that is the time horizon of PV generation. Between these hours there is sunlight in the nZEB region. Appendix A depicts how the initial data (described in Table 4) are clustered by utilizing WCSS and silhouette score for defining the number of k clusters and k-means [49] algorithm for clustering. The reason for preferring the silhouette score is explained in section 4.3. An example of k-means clustering for a specific timestamp, using the elbow method with WCSS, as well as the silhouette score is presented in section 5.1. The same process was repeated for all timestamps when clustering was involved.

As described in section 4.2 various data preprocessing techniques were performed (removing duplicates, dealing with missing values, data transformation and reduction) to prepare the initial dataset for the simulations. Then, to get a glimpse of the initial state of the data, clustering is performed, before conducting the simulations. The reason for initiating this process is to ensure that this analysis is not impaired by any disparities. At this point it is clarified that the clustering is performed on each timestamp to create clusters of the different values of the C attribute (described in Table 4).

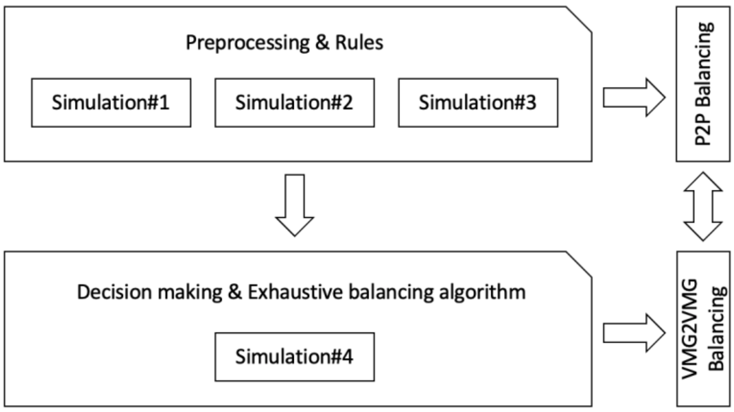


Figure 10: Simulations Overview

Rules detailed in 4.3, classify prosumers according to their energy profile. For example, some inject energy to the grid, some draw energy and some are virtually idle. It is assumed that all prosumers can generate energy through PV panels. The energy for each timestamp is the absolute value of generation minus consumption.

In Simulation#1, the rules aid P2P level balancing by linking prosumers with maximum generation with prosumers with maximum consumption through an iterative process.

In Simulation#2 and Simulation#3 this approach expands its analytical capabilities by forming VMGs for the stakeholders. Therefore, binning on the prosumer dataset is performed with a view to present the data in a quantitative, as well as a qualitative way. The last simulation essentially presents the final stage of this approach, supporting decision making for balancing the formed clusters. Simulations 1-3 produce clusters and Simulation 4 inputs aggregated energy values (resulting from consumption or injection of energy to the grid) per cluster and utilizes EBA with user defined target values, as shown in Figure 10.

5.1 Simulation#1

For the first simulation, the three rules are utilized to perform balancing for a specific timestamp, linking the prosumer with the maximum production value to the prosumer with the maximum consumption value. The methodology for this simulation comprises four steps depicted in Figure 11 and explained below:

Step 1: Using the rules to transform and define the sum (injection or draw from grid) of intermittent resources for each formed class per requested timestamp (based on historical data or live feed data[1]).

Step 2: Once precise mapping of intermittent resources injection/draw of energy is finalized, perform balancing of injection/draw by connecting the prosumer with maximum injection value to the prosumer with the maximum draw value, aiming to balance the highest energy generation with the highest energy consumption. Iterate as long as there are loads (generation/consumption) to balance.

[1]In the simulations, historical data are utilized to have a realistic measure of outcomes and balancing goals.

Step 3: Generate a new dataset that contains all the data resulting from Step 2. Then cluster the results to form VMGs that can be characterized by their levels of positive or negative energy injected or drawn from the grid. These VMGs can interact with the VMG layer to exchange energy with the rest of the VMG layer structure.

Step 4: Visualize and elaborate on the results. At this point plots are generated to offer an analytical explanation of the experimentation performed.

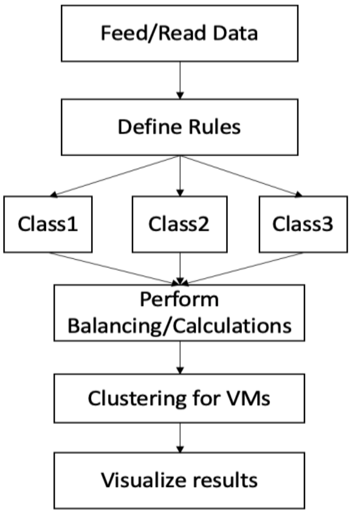


Figure 11: Simulation#1 methodology

Balancing is performed with an iterative procedure. Subtracting from the prosumers with the maximum injection values to the grid (class3) from prosumers with the maximum draw values from the grid (class1), essentially enforcing a virtual link between them. Since rules are set, the balancing process takes into consideration just class1 and class3 prosumers because class2 prosumers are already balanced. Table 5 shows an example of this process by linking Prosumer's 83 injection to Prosumer's 80 draw resulting in a balanced value. The same applies for Prosumer 48 and Prosumer 77, as well as Prosumer 54 and Prosumer 78, respectively.

Once the balanced dataset is formed, preprocessing is performed to standardize [50] the dataset to be fed into two algorithms for deciding the best possible choice for the number k of clusters.

The algorithms used for comparison reasons is WCSS [50] and the silhouette score [50].

Table 5: Example after balancing for timestamp: '2017-01-28 12:00:00'

ID	ID2	E0aclass3	E0aclass1	balancedE0a
83	80	11998	-3089	8909
48	77	5906	-1774	4132
54	78	6450	-2605	3845

The data are standardized so that all input values are to be centered on the value zero. This would indicate a perfect balance since it means that the energy generated and the energy consumed, are exactly equal. This is achieved by enforcing the following formula:

$$y = \frac{x_i - k}{s} \quad (2)$$

Where y is the data which are rescaled in a way that k=0 and s=1.

WCSS is used to measure the sum of distances of the available observations from the cluster centroid as stated in (3):

$$WCSS = \sum_{i \in n} (X_i - Y_i)^2 \quad (3)$$

Where Y_i is the centroid for observation X_i.

The silhouette coefficient for all data points I in a partition of k clusters is given by:

$$\bar{s}_k = \frac{1}{|I|} \sum_{i=1}^{|I|} s(i) \quad (4)$$

The higher the value of $\bar{s}_k$ the better the quality of clustering.

After performing WCSS to the data, we observe that the ideal number of clusters is three or four, as shown in Figure 12.

To conclude on the number k of clusters required by the clustering method, the silhouette score is also utilized. It results in four as the ideal number of clusters, as shown in Table 6, where the maximum value for the silhouette coefficient is 0.5183 for four clusters.

Next, k-means clustering algorithm [49] is used to form clusters that will conceptually form the P2P balanced VMGs abiding with the data feed to the approach of this simulation.

k-means partitions n objects into k clusters in a way that each object belongs to the cluster with the nearest mean.

$$J = \sum_{J=1}^{k} \sum_{i=1}^{n} \left\| x_i^{(j)} - C_j \right\|^2 \quad (5)$$

where k is the number of clusters, n the number of cases, x_i case i, and C_j the centroid for cluster j.

The distance function is calculated as: $\left\| x_i^{(j)} - C_j \right\|^2$.

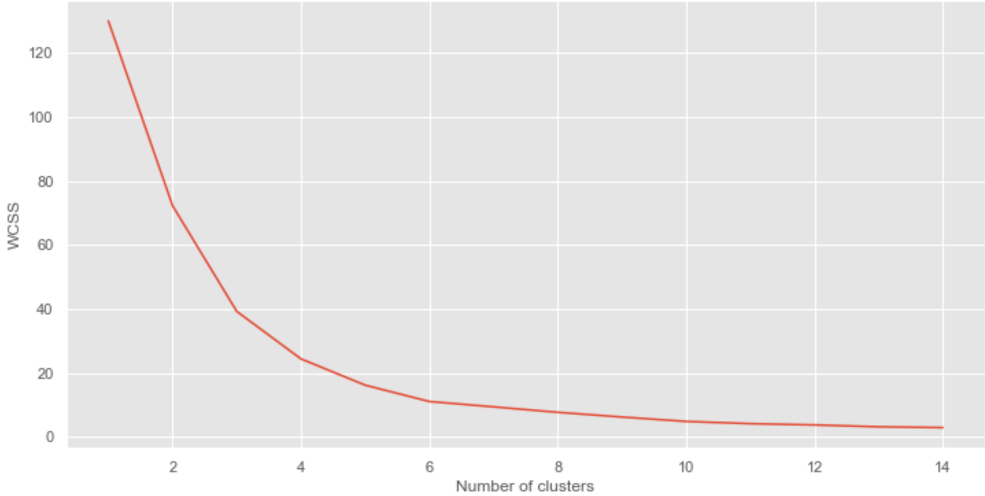


Figure 12: WCSS for defining number of clusters after balancing for timestamp: '2017-01-28 12:00:00'

Table 6: Silhouette coefficients for defining the number of clusters after balancing for timestamp: '2017-01-28 12:00:00'

No_clusters	Silhouette coefficient
2	0.47220146897061493
3	0.4934618404423085
4	0.5183793358157461
5	0.4514708099396544
6	0.4360998186060542
7	0.4134683471816738
8	0.42552728693807457
9	0.4189035658276343
10	0.4516080362887115
11	0.44350308931017823
12	0.4384334292996846
13	0.4551119965543649
14	0.43860146210831386

Figure 13 and Figure 14 depict the output of k-means for k=3 and k=4 respectively.

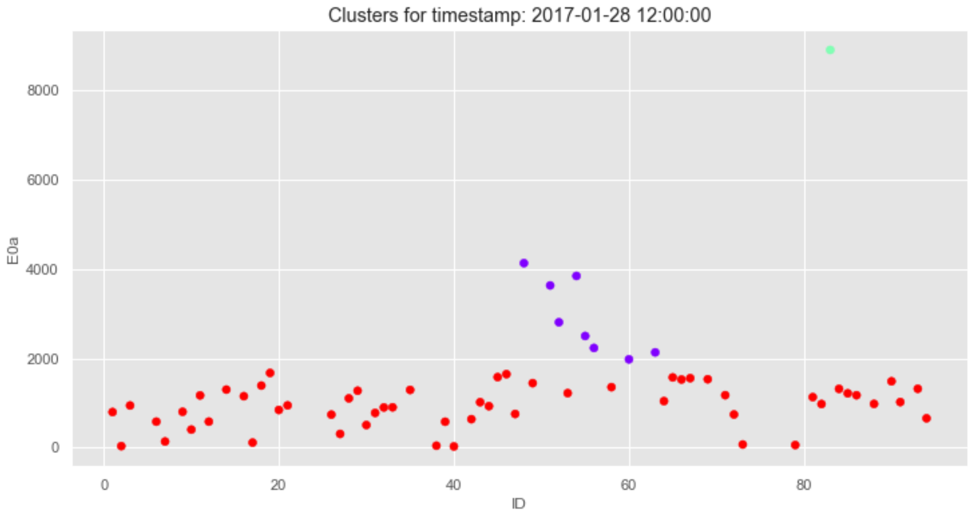


Figure 13: VMGs formed for k=3

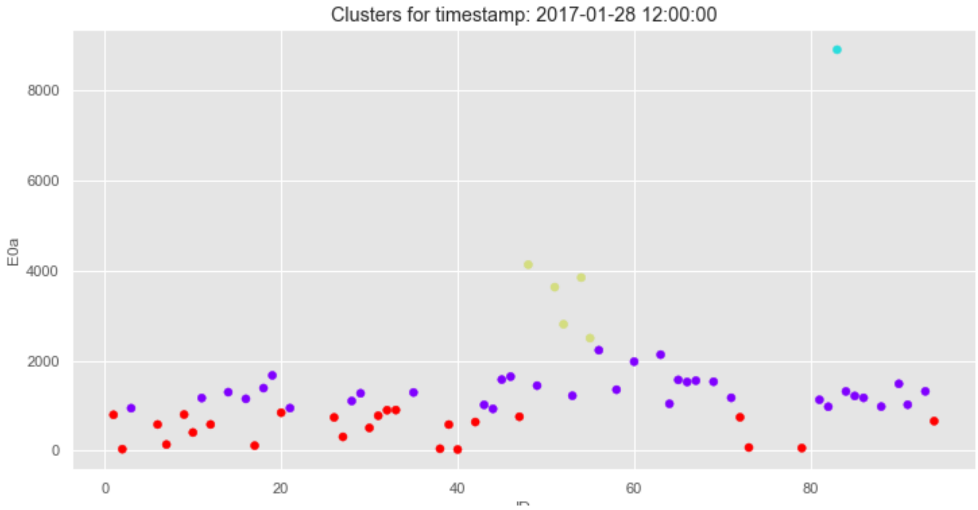


Figure 14: VMGs formed for k=4

In Figure 13 and Figure 14 it is evident that at this specific timestamp, even after balancing, there is an outlier due to the PV production from prosumer #83, shown in Table 5.

The figures in Appendix B show how clustering is performed after linking for each timestamp the prosumer with the max value with the consumer with the min value. This simulation is labeled *"V1 of energy balancing"*.

In the next 2 simulations, binning is performed on the raw dataset, according to two distinct binning methods QCUT and CUT, respectively [48]. The output of these methods is depicted in Appendix C. It shows the overview of prosumer allocation after using binning with both techniques. The reason for doing so, is to enhance the visualization and analytical capabilities of the proposed approach, offering more options for the stakeholders when considering forming VMGs.

5.2 Simulation#2

This simulation describes binning as it is performed incorporating QCUT [48] on raw data. This method attempts to create bins with almost the same number of consumptions/productions entries on each of the bins (Figure 15).

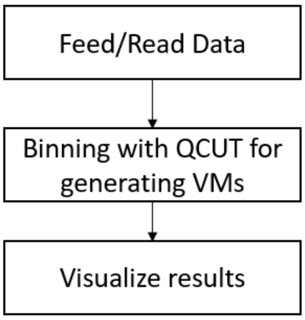


Figure 15: Overview of methodology Simulation#2

Step 1: Given a specific timestamp, the QCUT method is performed on data, resulting in five bins of prosumers.

Step 2: The aggregated consumption/generation for each bin is calculated. As stated earlier, the values signed as negative represent energy drawn from the grid, while values signed as positive are values that inject energy to the grid.

Step 3: Visualize and elaborate on the results. At this point plots are generated to offer an analytical explanation of the conducted experimentation. Figure 16 showcases the process and output of Step 1 (each bin contains nodes with the same color), while Table 7 showcases the process and output of Step 2.

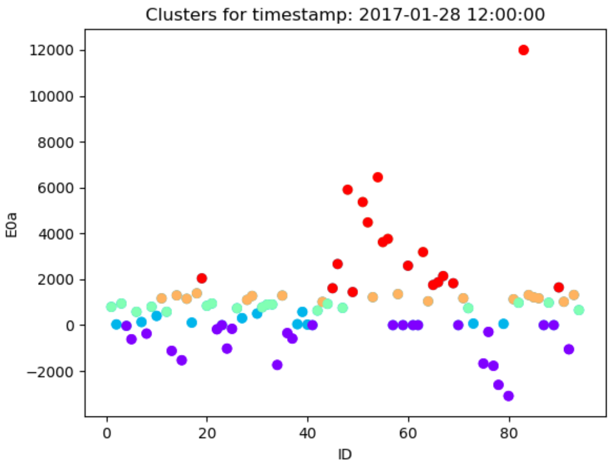


Figure 16: QCUT on raw data

Table 7: Prosumer QCUT binning with sums of consumption/production per bin.

BinLabel	E0a Count	Sum
1	26	-18197
2	11	2255
3	18	14458
4	18	21635
5	18	64359

Table 7 refers to Figure 16 as it describes the count of Prosumers that were placed on each of the bins along with the overall consumption/production.

Figures on the right side in Appendix D depict QCUT method output, showing how QCUT binning performs on specific timestamps within a single day (sun hours), where each table refers to the figure to its right. This simulation is labeled *"V2 of energy balancing"*.

5.3 Simulation#3

This simulation describes binning with the CUT [48] method on raw data. This method tries to create bins that contain consumptions/ productions entries within similar ranges i.e., belonging to the same value group, but with different bin frequency (Figure 17).

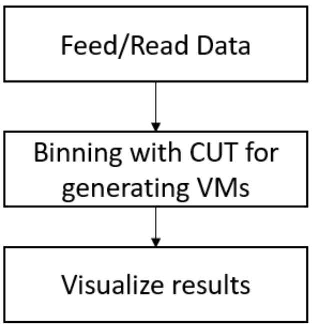


Figure 17: Overview of methodology Simulation#3

Step 1: Similarly, with the previous simulation, given a specific timestamp, CUT is performed on data resulting in five bins of prosumers.

Step 2: The aggregated consumption/ production is calculated for each bin. As stated earlier, the values signed as negative represent energy drawn from the grid while values signed as positive inject energy to the grid.

Step 3: The results are visualized and elaborated on. At this point plots are generated to offer an analytical explanation for this simulation. Figure 18, explains the process and output of Step 1 (each bin contains nodes with the same color), while Table 8 showcases the process and output of Step 2.

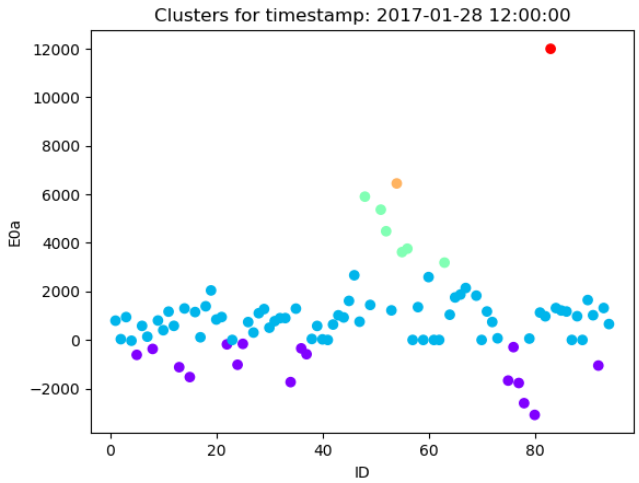


Figure 18: CUT on raw data

Table 8: Prosumer CUT binning with sums of consumption/production per bin.

BinLabel	**E0a Count**	**Sum**
1	16	-18165
2	67	57902
3	6	26325
4	1	6450
5	1	11998

Table 8 refers to Figure 18, as it describes the number of Prosumers placed in each bin along with the overall consumption/production.

Appendix D shows how CUT binning performs on specific timestamps within a single day (sun hours). On the left side figures depict CUT method output, whilst each table refers to the figure to its right. This simulation is labeled *"V3 of energy balancing"*.

5.4 Simulation#4

The last simulation outlines the final stage of the approach (Figure 19); the final definition of the VMGs, incorporating decision making and EBA. This involves two steps.

Step 1: At this point, a decision on which clustering, or binning process is more suitable for balancing (balanced VMG formation), needs to be taken. In case a naivenaïve P2P balancing and clustering of prosumer VMGs is requested to act as a baseline estimator for an energy balancing operation, the output of Simulation#1 is utilized. In case the formation of VMGs based on a relative measure of consumptions/ productions is requested, the CUT method for binning is employed i.e., the output of Simulation#2. Finally, in case the absolute measure of consumptions/productions is requested to form VMGs, the QCUT method for binning is employed i.e., the output of Simulation#3. These options are offered aiming to increase the number of possible outcomes of this simulation, whilst acting as individual and distinct sub-results that feed the main EBA algorithm, expanding the analytical capabilities of the proposed approach, for evaluation purposes.

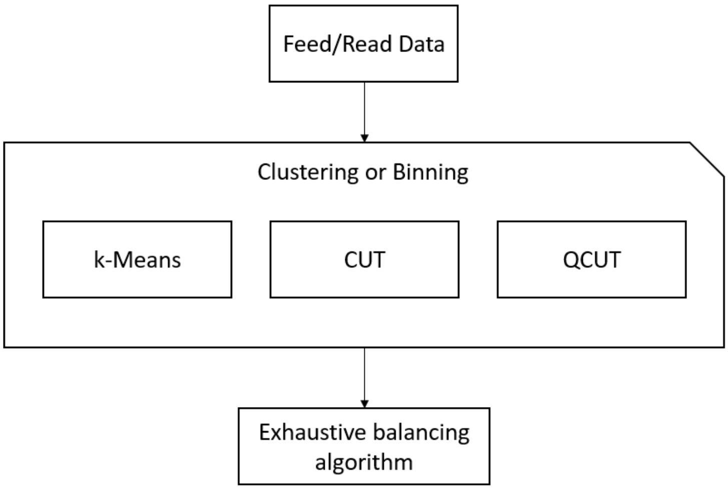


Figure 19: Overview of methodology Simulation#4

Step 2: Once the binning measure or clustering technique is determined, balancing between the available bins or clusters (VMGs) that are formed is performed. This is achieved by utilizing a method that effectively decides which bins or clusters need to be linked with one or more additional bins/clusters. A customized exhaustive (brute-force) searching algorithm (referred as EBA) is implemented, that creates lists of the best matching values (consumptions/ productions) it receives as input, which are closer to user defined numeric target values.

The following example uses as input the output of one of the binning approaches (Simulations#2 or #3) and utilizes EBA, described in Step 2, to perform balancing on a VMG level (VMG2VMG balancing).

Consider the following arrays:

Console Input:

Numbers = [-12277, -14027, +33929, -33000, +12630, -21612, +520, +24000, +14000]

Targets = [+500, +1500, -1000]

where each entry on the *Numbers* array refers to the aggregated values of either draw (signed with "-") or injection (signed with "+"), and each entry of the *Targets* array refers to the requested by the stakeholder (DSO, Aggregator) energy target value for the VMG (or cluster of prosumers).

Executing the balancing algorithm, inputting as arguments the two arrays,

search_closest_to_target (*Numbers*, *Targets*), the EBA output is the following:

Console Output:

Arithmetic value to approach for balancing: 500, List of combined values forming a VMG: (-14027, 520, 14000)

Arithmetic value to approach for balancing: 1500, List of combined values forming a VMG: (33929, -12277, 12630, -33000)

Arithmetic value to approach for balancing: -1000, List of combined values forming a VMG: (24000, -21612)

6. Discussion

Our approach incorporates various alternatives for VMG formation giving more options for testing the validity of VMG2VMG balancing output. These options are implemented throughout Simulation#1-3 respectively, while their association and connection with Simulation#4 was explained in section 5. The VMG2VMG energy balancing problem is addressed by incorporating EBA, an exhaustive search and balancing algorithm described in section 5.4, while the P2P energy balancing problem is solved utilizing the rules' approach. Sub-section 6.1 discusses perceived implications of the conducted simulations and their output.

6.1 Implications

The approach presented can be implemented as a standalone toolkit offering the functionalities described in this study. Based on the current state of ongoing research on this domain, possible use cases are identified that can act as an expansion, or can be incorporated as external components, utilizing or being dependent on the output of the conducted research.

Use Case #1

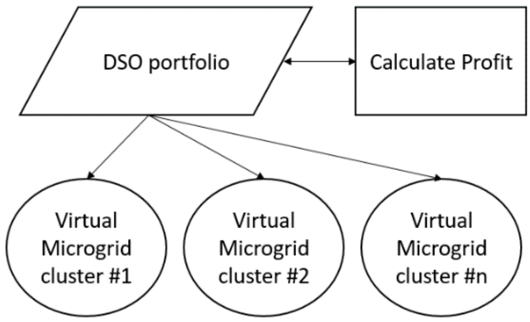
The first use case envisions potential cost minimization or profit maximization when transferring energy from VMGs handled by aggregators to the virtual portfolio of the DSO (Figure 20). These grids are conceptualized as clusters, which have specific parameters regarding consumptions, productions, KPIs etc. Then the DSO can be more accurate regarding DER load distribution for calculating profits or minimizing functional costs.

Such an objective function given a ToU DR can be:

$$\mathrm{Cost} = \sum_{h=1}^{24} [\mathrm{ToU}(h) * \mathrm{pgrid}(h)] \quad (5)$$

Table 9: Use Case #1 Nomenclature

Nomenclature	Description
H	time step ∈ {1…24}
ToU(h)	price of grid kW on a specific timestamp
pgrid(h)	kWh injected/drawn on a specific timestamp



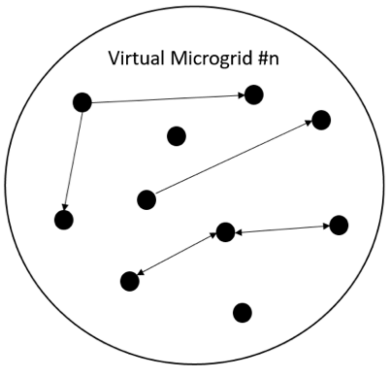
According to Table 9 and (5), it is considered that the value of ToU(h) is known for h {1…24} from a day-ahead price forecasting algorithm such as in [53]. The variable pgrid(h) can have both positive and negative values. When pgrid(h) is positive, the prosumer injects power to the grid at the hour h, while pgrid(h) is negative the prosumer draws power from the grid at the hour h.

Figure 20: DSO to VMG management example

Use Case #2

The second use case addresses the opportunities arising from VMG2VMG transfer of energy, when RES can be utilized in a large scale. VMGs are conceived while enabling tradeoffs of energy in-between prosumers, offering optimal energy distribution in close proximity. They are also considered for maximizing savings from main grid tariffs and reducing green gas emissions. Local ecosystem VMGs constraints can be enforced to ensure fairness and stability of such an energy management system.

Use Case #3



The third use case refers to VMGs capability to utilize P2P transfer, meaning that energy can be traded from prosumer to prosumer a.k.a., energy trading to cover DR needs of prosumers. Figure 21 shows a VMG perceived as a cluster #n while black dots indicate prosumers. There are directed edges that inform the direction of the DR energy transfer on P2P level between various nodes. Each node represents a prosumer that agreed to participate to an energy-sharing program. Energy can be exchanged uni- or bilaterally.

Figure 21: DR Distribution on P2P level, optimized single VMG schema

6.2 Biases

The main identified biases of this interdisciplinary energy balancing approach refer to:

A. Our study poses as a Computer Science centric approach attempting to investigate an Energy sector problem, prescribing the requirements for a multidisciplinary approach on a problem that requires merging concepts from multiple research fields. This statement imposes that certain levels of domain details are disregarded. For example, the study does not consider constraints, such as power transmission loss or heating loss for calculating the VMG formulation. The values for PV production on the dataset represent the final product that interacts with the grid, whether injecting to, or drawing energy from it.

B. Observing the outputs after running the VMG balancing process multiple times, it is evident that despite any balancing approach on VMGs, every attempt is still biased by the nature of RES. That is to say, a high production is expected during peak hours while the sun is up, and zero production after the sun sets. To resolve this issue, possible solutions could be either i) sufficient energy storage means or capacity for saving the excess produced energy, or ii) the implementation and enforcement of a globally accepted VMG scheme that allows energy transfer from geographical areas with sunlight to areas without sunlight during the same timestamps.

7. Conclusions

7.1 Research Summary

This paper proposes an approach which implements high-level energy balancing at the VMG and P2P level, while reviewing some state-of-the-art computer science and energy attempts, highlighting the necessity for interdisciplinarity in this problem domain. The approach utilizes heuristics through rules and unsupervised learning to form and balance VMGs. Biases in the proposed methodology are stated in section 6.2, and assumptions in section 3. These biases and assumptions inferably generate directions for improvements. The methodology and its connection with result outline are depicted in Figure 22.

A step-by-step approach is followed, incorporating experimentation and elaboration to produce results that aid organizing our next research steps described in section 7.2. Each of the simulations, results in a version of energy balancing that can act as stand-alone component, able to feed the final simulation that deals with VMG2VMG balancing, the main research aim of this paper.

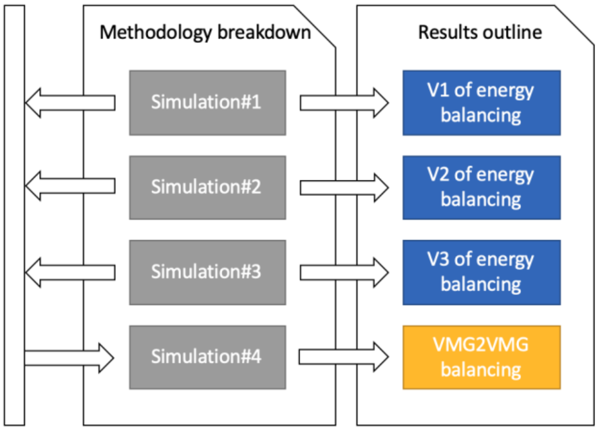


Figure 22: Research summary and outcomes

Simulation#1 enforces the rules defined (section 4.3) for balancing at the P2P level. Simulation#2 utilizes a quantile-based discretization for appointing prosumers to bins. Simulation#3 transposes prosumer energy values from continuous variables to categorical, appointing the prosumer entries to equally sized bins. Finally, Simulation#4 performs VMG2VMG balancing by utilizing EBA, an exhaustive search and balancing algorithm that is activated by the results from either of the previous simulations (#1-3).

Appendices A-D are provided for a more detailed presentation of the experimentation process. These aim at further clarifying our approach by presenting the process steps analytically.

7.2 Future work

This study can be extended by addressing the following points.

A. Continue monitoring of the evolution of research novelties on RES energy distribution balancing.

B. Improve the methodology by tackling biases presented in 6.2 and enhancing the perception of balancing in a more energy-centric manner.

C. Extend ongoing work on distinct energy profile generation (as described in section 4.2) and update the dataset entries to form lists of P2P energy transfer, aiming to generate closed communities within VMGs. This approach would aid further improvement of energy balancing and allow further simulations with more datasets in order to calibrate the approach.

D. Improve the model by further automating our methodology to act as a stand-alone component or toolkit. An extended reference to such tools was made in section 6.1 and specifically to attempts incorporating Blockchain technology, as in [30-32]. Blockchain can offer the infrastructure for a secure [54] and distributed energy balancing ecosystem [38] through smart contracts. Our architectural approach can thus be complemented at the application level [36], whilst exposing Business Intelligence opportunities [55].

E. Regarding DR programs, this study conceptualizes that our approach can be expanded to use the Time of Use (ToU) category or any other DR program described in section 2.1.5. We initially considered utilizing ToU, since currently the approach solely relies on PV energy production for VMG energy output. This seems more appropriate as the dataset contains data that represent prosumer energy generation originated from RES, such as PV. Considering these points, a minimization on operating costs and pollution emissions in an individualized way (per VMG), or in a super-cluster of multiple VMGs can be achieved. Having presented future prospects of this work, we envision that it can assist academics and practitioners alike when referring to DERs, VMG2VMG, P2P energy transfer models and energy balancing utilizing clustering, and exhaustive balancing techniques merged with distributed energy transactions using technologies such as Blockchain i.e. [34] and [53].

References

[1] Cagnano, A., De Tuglie, E., & Mancarella, P. (2020). Microgrids: Overview and guidelines for practical implementations and operation. Applied Energy, 258, 114039.

[2] Tabatabaei, N. M., Kabalci, E., & Bizon, N. (Eds.). (2019). Microgrid architectures, control and protection methods. Springer.

[3] Asmus, P. (2010). Microgrids, virtual power plants and our distributed energy future. The Electricity Journal, 23(10), 72-82.

[4] Badami, M., Fambri, G., Mancò, S., Martino, M., Damousis, I. G., Agtzidis, D., & Tzovaras, D. (2020). A Decision Support System Tool to Manage the Flexibility in Renewable Energy-Based Power Systems. Energies, 13(1), 153.

[5] Kafle, Y. R., Mahmud, K., Morsalin, S., & Town, G. E. (2016, September). Towards an internet of energy. In 2016 IEEE International Conference on Power System Technology (POWERCON) (pp. 1-6). IEEE.

[6] Tzovaras, D., Damousis, I., Papanikolaou, A., Pitsiladis, G., & Barbagelata, G. (2019). DRIMPAC—Unified Demand Response Interoperability Framework Enabling Market Participation of Active Energy Consumers. In Multidisciplinary Digital Publishing Institute Proceedings (Vol. 20, No. 1, p. 15).

[7] Long, C., Wu, J., Zhou, Y., & Jenkins, N. (2018). Peer-to-peer energy sharing through a two-stage aggregated battery control in a community Microgrid. Applied energy, 226, 261-276.

[8] Hirsch, A., Parag, Y., & Guerrero, J. (2018). Microgrids: A review of technologies, key drivers, and outstanding issues. Renewable and sustainable Energy reviews, 90, 402-411.

[9] Paiva, L. T., & Fontes, F. A. (2018). Optimal electric power generation with underwater kite systems. Computing, 100(11), 1137-1153.

[10] Bintoudi, A. D., Zyglakis, L., Apostolos, T., Ioannidis, D., Al-Agtash, S., Martinez-Ramos, J. L., ... & Demoulias, C. (2017, November). Novel hybrid design for microgrid control. In 2017 IEEE PES Asia-Pacific Power and Energy Engineering Conf. (APPEEC) (pp. 1-6). IEEE.

[11] Balijepalli, V. M., Pradhan, V., Khaparde, S. A., & Shereef, R. M. (2011, December). Review of demand response under smart grid paradigm. In ISGT2011-India (pp. 236-243). IEEE.

[12] Sioshansi, F. P. (Ed.). (2019). Consumer, Prosumer, Prosumager: How Service Innovations Will Disrupt the Utility Business Model. Academic Press.

[13] Charalambous, A., Hadjidemetriou, L., Zacharia, L., Bintoudi, A. D., Tsolakis, A. C., Tzovaras, D., & Kyriakides, E. (2019). Phase Balancing and Reactive Power Support Services for Microgrids. Applied Sciences, 9(23), 5067.

[14] Ellahi, M., Abbas, G., Khan, I., Koola, P. M., Nasir, M., Raza, A., & Farooq, U. (2019). Recent approaches of forecasting and optimal economic dispatch to overcome intermittency of wind and photovoltaic (PV) systems: A review. Energies, 12(22), 4392.

[15] Khuntia, S. R., Rueda, J. L., & van der Meijden, M. A. (2016). Forecasting the load of electrical power systems in mid-and long-term horizons: a review. IET Generation, Transmission & Distribution, 10(16), 3971-3977.

[16] Livingston, D., Sivaram, V., Freeman, M., & Fiege, M. (2018). Applying blockchain technology to electric power systems. Council on Foreign Relations.

[17] Nosratabadi, S. M., Hooshmand, R. A., & Gholipour, E. (2017). A comprehensive review on microgrid and virtual power plant concepts employed for distributed energy resources scheduling in power systems. Renewable and Sustainable Energy Reviews, 67, 341-363.

[18] Akbary, P., Ghiasi, M., Pourkheranjani, M. R. R., Alipour, H., & Ghadimi, N. (2019). Extracting appropriate nodal marginal prices for all types of committed reserve. Computational Economics, 53(1), 1-26.

[19] Liu, N., Yu, X., Wang, C., Li, C., Ma, L., & Lei, J. (2017). Energy-sharing model with price-based demand response for microgrids of peer-to-peer prosumers. IEEE Transactions on Power Systems, 32(5), 3569-3583.

[20] Zhou, Y., Wu, J., Long, C., Cheng, M., & Zhang, C. (2017). Performance evaluation of peer-to-peer energy sharing models. Energy Procedia, 143, 817-822.

[21] Moura, P. S., & de Almeida, A. T. (2010). Multi-objective optimization of a mixed renewable system with demand-side management. Renewable and Sustainable Energy Reviews, 14(5), 1461-1468.

[22] Blake, S. T., & O'Sullivan, D. T. (2018). Optimization of distributed energy resources in an industrial microgrid. Procedia CIRP, 67, 104-109.
[23] Zhang, Y., Gatsis, N., & Giannakis, G. B. (2013). Robust energy management for microgrids with high-penetration renewables. IEEE Transactions on Sustainable Energy, 4(4), 944-953.
[24] Quiggin, D., Cornell, S., Tierney, M., & Buswell, R. (2012). A simulation and optimisation study: Towards a decentralised microgrid, using real world fluctuation data. Energy, 41(1), 549-559.
[25] Ding, Y. M., Hong, S. H., & Li, X. H. (2014). A demand response energy management scheme for industrial facilities in smart grid. IEEE Transactions on Industrial Informatics, 10(4), 2257-2269.
[26] Hawkes, A. D., & Leach, M. A. (2009). Modelling high level system design and unit commitment for a microgrid. Applied energy, 86(7-8), 1253-1265.
[27] Zhang, C., Wu, J., Zhou, Y., Cheng, M., & Long, C. (2018). Peer-to-Peer energy trading in a Microgrid. Applied Energy, 220, 1-12.
[28] Anoh, K., Maharjan, S., Ikpehai, A., Zhang, Y., & Adebisi, B. (2019). Energy peer-to-peer trading in virtual microgrids in smart grids: a game-theoretic approach. IEEE Transactions on Smart Grid.
[29] Vergados, D. J., Mamounakis, I., Makris, P., & Varvarigos, E. (2016). Prosumer clustering into virtual microgrids for cost reduction in renewable energy trading markets. Sustainable Energy, Grids and Networks, 7, 90-103.
[30] Cioara, T., Anghel, I., Pop, C., Bertoncini, M., Croce, V., Ioannidis, D. & D'Oriano, L. (2018, June). Enabling new technologies for demand response decentralized validation using blockchain. In 2018 IEEE Int'l Conf. on Environment and Electrical Engineering and 2018 IEEE Industrial and Commercial Power Systems Europe (EEEIC/I&CPS Europe) (pp. 1-4). IEEE.
[31] D'Oriano, L., Mastandrea, G., Rana, G., Raveduto, G., Croce, V., Verber, M., & Bertoncini, M. (2018, November). Decentralized blockchain flexibility system for Smart Grids: Requirements engineering and use cases. In 2018 Int'l IEEE Conf. and Workshop in Obuda on Electrical and Power Engineering (CANDO-EPE) (pp. 39-44). IEEE.
[32] Pop, C., Cioara, T., Antal, M., Anghel, I., Salomie, I., & Bertoncini, M. (2018). Blockchain based decentralized management of demand response programs in smart energy grids. Sensors, 18(1), 162.
[33] Petrican, T., Vesa, A. V., Antal, M., Pop, C., Cioara, T., Anghel, I., & Salomie, I. (2018, September). Evaluating forecasting techniques for integrating household energy prosumers into smart grids. In 2018 IEEE 14th Int'l Conf. on Intelligent Computer Communication and Processing (ICCP) (pp. 79-85). IEEE.
[34] Stecchi, U., Gomez, J., Gallego, L., Santori, F., & Rico, J. (2019, May). A multipurpose ICT platform for supporting energy transition: first results in flexibility profiling. In 2019 1st Int'l Conf. Energy Transition in the Mediterranean Area (SyNERGY MED) (pp. 1-6). IEEE.
[35] Bragatto, T., Cresta, M., Croce, V., Paulucci, M., Santori, F., & Ziu, D. (2019, June). A real-life experience on 2 nd life batteries services for Distribution System Operator. In 2019 IEEE Int'l Conf. Environment and Electrical Engineering and 2019 IEEE Industrial and Commercial Power Systems Europe (EEEIC/I&CPS Europe) (pp. 1-6). IEEE.
[36] Pop, C., Antal, M., Cioara, T., Anghel, I., Sera, D., Salomie, I. & Bertoncini, M. (2019). Blockchain-based scalable and tamper-evident solution for registering energy data. Sensors, 19(14), 3033.
[37] Vesa, A. V., Cioara, T., Anghel, I., Antal, M., Pop, C., Iancu, B., & Salomie, I. (2020). Energy Flexibility Prediction for Data Center Engagement in Demand Response Programs. Sustainability, 12(4), 1417.
[38] Pop, C., Cioara, T., Anghel, I., Antal, M., & Salomie, I. (2020). Blockchain based Decentralized Applications: Technology Review and Development Guidelines. arXiv preprint arXiv:2003.07131.
[39] Kong, X., Liu, D., Wang, C., Sun, F., & Li, S. (2020). Optimal operation strategy for interconnected microgrids in market environment considering uncertainty. *Applied Energy*, *275*, 115336.
[40] Liu, W., Gu, W., Wang, J., Yu, W., & Xi, X. (2018). Game theoretic non-cooperative distributed coordination control for multi-microgrids. *IEEE Transactions on Smart Grid*, *9*(6), 6986-6997.
[41] Cardoso, G., Brouhard, T., DeForest, N., Wang, D., Heleno, M., & Kotzur, L. (2018). Battery aging in multi-energy microgrid design using mixed integer linear programming. Applied energy, 231, 1059-1069.
[42] Koukaras, P., & Tjortjis, C. (2019). Social media analytics, types and methodology. In Machine Learning Paradigms (pp. 401-427). Springer, Cham.

[43] Dikaiakos, M. D., Katsaros, D., Mehra, P., Pallis, G., & Vakali, A. (2009). Cloud computing: Distributed internet computing for IT and scientific research. IEEE Internet computing, 13(5), 10-13.
[44] Papazoglou, M. P., & Georgakopoulos, D. (2003). Introduction: Service-oriented computing. Communications of the ACM, 46(10), 24-28.
[45] "Digital Innovation Hubs - Smart Specialisation Platform.", http://s3platform.jrc.ec.europa.eu/digital-innovation-hubs-tool/-/dih/1521/view.
[46] Vafeiadis, T., Zikos, S., Stavropoulos, G., Ioannidis, D., Krinidis, S., Tzovaras, D., & Moustakas, K. (2017, October). Machine learning based occupancy detection via the use of smart meters. In 2017 Int'l Symposium on Computer Science and Intelligent Controls (ISCSIC) (pp. 6-12). IEEE.
[47] Adamopoulou, A. A., Tryferidis, A. M., & Tzovaras, D. K. (2016). A context-aware method for building occupancy prediction. Energy and Buildings, 110, 229-244.
[48] McKinney, W. (2010, June). Data structures for statistical computing in python. In Proc. 9th Python in Science Conf. (Vol. 445, pp. 51-56).
[49] Arthur, D., & Vassilvitskii, S. (2006). k-means++: The advantages of careful seeding. Stanford.
[50] Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O. & Vanderplas, J. (2011). Scikit-learn: Machine learning in Python. Journal of machine learning research, 12(Oct), 2825-2830.
[51] Rousseeuw, P. J. (1987). Silhouettes: a graphical aid to the interpretation and validation of cluster analysis. Journal of computational and applied mathematics, 20, 53-65.
[52] Ketchen, D. J., & Shook, C. L. (1996). The application of cluster analysis in strategic management research: an analysis and critique. Strategic management journal, 17(6), 441-458.
[53] Conejo, A. J., Contreras, J., Espínola, R., & Plazas, M. A. (2005). Forecasting electricity prices for a day-ahead pool-based electric energy market. Int'l journal of forecasting, 21(3), 435-462.
[54] Tsolakis, A. C., Moschos, I., Votis, K., Ioannidis, D., Tzovaras, D., Pandey, P., ... & García-Castro, R. (2018, July). A Secured and Trusted Demand Response system based on Blockchain technologies. In 2018 Innovations in Intelligent Systems and Applications (INISTA) (pp. 1-6). IEEE.
[55] Mendling, J., Weber, I., Aalst, W. V. D., Brocke, J. V., Cabanillas, C., Daniel, F., ... & Gal, A. (2018). Blockchains for business process management-challenges and opportunities. ACM Transactions on Management Information Systems (TMIS), 9(1), 1-16.