UNIVERSITY OF THESSALY

DEPARTMENT OF ELECTRICAL AND COMPUTER ENGINEERING

Next Generation Energy Markets and Applications

Author:
Antonia NASIAKOU

Supervisor:
Dr. Manolis VAVALIS

A dissertation submitted in fulfillment of the requirements for the degree of Doctor of Philosophy

in the

Department of Electrical and Computer Engineering

April 2018

Declaration of Authorship

I, Antonia Nasiakou, declare that this dissertation titled, "Next Generation Energy Markets and Applications" and the work presented in it are my own. I, Antonia Nasiakou, confirm that:

- This work was done wholly or mainly while in candidature for a research degree at this University.
- Where any part of this dissertation has previously been submitted for a degree or any other qualification at this University or any other institution, this has been clearly stated.
- Where I have consulted the published work of others, this is always clearly attributed.
- Where I have quoted from the work of others, the source is always given. With the exception of such quotations, this dissertation is entirely my own work.
- I have acknowledged all main sources of help.
- Where the dissertation is based on work done by myself jointly with others, I have made clear exactly what was done by others and what I have contributed myself.

Signed: ______________________________

Date: ______________________________

The dissertation of Antonia Nasiakou is approved by:

- Prof. Manolis Vavalis, University of Thessaly, Department of Electrical and Computer Engineering
- Prof. Lefteris Tsoukalas, University of Thessaly, Department of Electrical and Computer Engineering
- Prof. George Papavassilopoulos, National Technical University of Athens, School of Electrical and Computer Engineering
- Associate Prof. Dimitrios Bargiotas, University of Thessaly, Department of Electrical and Computer Engineering
- Assistant Prof. Dimitrios Katsaros, University of Thessaly, Department of Electrical and Computer Engineering
- Prof. George Liberopoulos, University of Thessaly, Department of Mechanical Engineering
- Associate Prof. Iordanis Koutsopoulos, Athens University of Economics and Business Department of Informatics

University of Thessaly

Abstract

Department of Electrical and Computer Engineering

Doctor of Philosophy

Next Generation Energy
Markets and Applications

Antonia Nasiakou

The transition from the conventional to the smart operation of the grid has emerged a lot of opportunities for improving the efficiency and the safety in the power grid. The power grid is in stable state only when all its components (power plants, Renewable Energy Sources (RES), consumers (residential, commercial, industrial) transmission and distribution lines, transformers, phasor measurement units, regulators) are properly configured. The last decades, researchers have focused on improving the efficiency of the power grid using Demand Response (DR) programs including the consumers in the operation of the power grid and more specifically in the operation of the energy market giving to them in that way an active role. Additionally, the integration of the RES can further increase the energy savings by reducing the need of the energy storage. However, the impact of the stochasticity at the renewable energy generation side has to be considered. In this dissertation, applications based on energy market models that focusing on i) energy savings by shifting the demand on off-peak hours from on peak hours, ii) adjusting appropriate energy tariffs at each consumer, iii) RES energy savings by selling its energy in lower prices when it is needed, iv) integration of RES to cover the energy need in the agricultural area, and v) partitioning methodologies for improving the operation of the distribution grid in terms of energy savings are presented. Specifically, our ultimate goal is to design and implement advanced energy systems simulation and validation models, able to integrate more and more the consumers and the RES in the operation of the power grid. These models are integrated into simulation engines. In the presented applications in this dissertation, a new power distribution system simulation and analysis tool, the GridLAB-D, game theory models, fuzzy logic theory, clustering algorithms and DR programs are utilized.

Due to the liberalization of electricity market, the vertical integrated, the monopolistic markets are replaced by several components with different responsibilities. The power producers are independent utility companies, either conventional or RES, that produce and sell power in specific prices in the deregulated energy markets. The consumers are the buyers in the market, which may be residential, commercial or industrial. The market operator has to follow the rules of transporting the power on time to the consumers. The energy cannot be stored in large scale and there is excess amount of energy production that does not contribute to anything to the network.

So, the market operator tries to sell as more energy as it is possible. In this context, the different types of electricity markets are studied and electricity market models introducing competition in those markets through the active participation of both consumers and producers in a bid-based, security-constrained electricity market are proposed. In particular, first, the different types of electricity markets by focusing in the market's characteristics that Hogan (1996) proposes are introduced. The development of the new electricity market models addresses the problem of the high energy consumption during the on-peak hours for both residential and agricultural load. The policy for the solution of this problem considers various features such as the priority at the operation of specific appliances and the willingness of the consumer to shift his energy consumption.

The DR programs have vital role for the steady state operation of the power grid. Their primary goal is to smooth the energy consumption by offering better prices to the consumers or by providing them with incentives to reduce the consumption at specific time intervals during the day. Many researchers have aimed to Demand-side Management (DSM) techniques paying more attention to the consumers' side.

For the steady state, reliable and efficient operation of both transmission and distribution networks various techniques can be considered. Moreover, an anomaly or disruption in the operation of the electricity network could be mitigated through advances in robust decision making mechanisms. Those techniques aim to efficient reaction of the electricity network under various types of disturbances. For instance, a huge disturbance in electricity network could be handled by partitioning it, and subsequently targeting at isolating the problem to prevent diffusion in the whole grid. Further, the delivery of electricity has to be non-stop, so that the isolation of the disturbance, leaded by partitioning methodologies, could be conducted in an efficient way.

Πανεπιστήμιο Θεσσαλίας

Περίληψη

Τμήμα Ηλεκτρολόγων Μηχανικών και Μηχανικών Υπολογιστών

Διδακτορικό Δίπλωμα

Σύγχρονες Αγορές Ενέργειας και Εφαρμογές

Αντωνία Νασιάκου

Η μετάβαση από τη παραδοσιακή στην έξυπνη λειτουργία του δικτύου έχει αναδεύσει πολλές ευκαιρίες για τη βελτίωση της απόδοσης και της ασφάλειας του ηλεκτρικού δικτύου. Το δίκτυο ηλεκτρικής ενέργειας βρίσκεται σε σταθερή κατάσταση μόνο όταν όλες οι συνιστώσες που το απαρτίζουν (σταθμοί παραγωγής ηλεκτρικής ενέργειας, γεννήτριες ανανεώσιμων πηγών ενέργειας (ΑΠΕ), καταναλωτές (οικιακοί, εμπορικοί, βιομηχανικοί), σταθμοί μετασχηματιστών, μονάδες μέτρησης φάσης, ρυθμιστές) βρίσκονται σε ισορροπία. Τις τελευταίες δεκαετίες, οι ερευνητές μέσα από ένα μεγάλο αριθμό δημοσιεύσεων επικεντρώθηκαν στη βελτίωση της αποδοτικότητας του ηλεκτρικού δικτύου με τη χρήση τεχνικών απόκρισης ζήτησης, με την άμεση συμμετοχή των καταναλωτών στη λειτουργία του ηλεκτρικού δικτύου και πιο συγκεκριμένα στην λειτουργία της αγοράς ενέργειας. Με αυτόν τον τρόπο οι καταναλωτές έχουν ενεργό ρόλο στην διαμόρφωση της τιμής της αγοράς. Επιπλέον, η ενσωμάτωση των ΑΠΕ μπορεί να αυξήσει περαιτέρω την εξοικονόμηση ενέργειας αυξάνοντας έτσι την ανάγκη αποθήκευσης ενέργειας. Ωστόσο, πρέπει να ληφθεί υπόψη ο αντίκτυπος της στοχαστικότητας στην πλευρά της παραγωγής της ενέργειας. Στην παρούσα διατριβή, παρουσιάζουμε μερικές εφαρμογές που βασίζονται σε μοντέλα της αγοράς που εστιάζουν στην 1) εξοικονόμηση ενέργειας μεταβάλλοντας τη ζήτηση εκτός ωρών αιχμής από τις ώρες αιχμής, 2) προσαρμογή κατάλληλων τιμολογίων ενέργειας σε κάθε καταναλωτή, 3) πώληση ενέργειας σε χαμηλότερες τιμές, 4) ενσωμάτωση των ΑΠΕ στον δίκτυο για την κάλυψη των αναγκών σε ενέργεια στην αγροτική περιοχή, και 5) μεθοδολογίες διαχωρισμού του δικτύου για τη βελτίωση της λειτουργίας του δικτύου διανομής όσον αφορά την εξοικονόμηση ενέργειας. Συγκεκριμένα, ο απώτερος στόχος μας είναι να σχεδιάσουμε και να εφαρμόσουμε μοντέλα προσομοίωσης και επαλήθευσης προηγμένων ενεργειακών συστημάτων, ικανά να ενσωματώσουν περισσότερο τους καταναλωτές και τις ΑΠΕ στην λειτουργία του ηλεκτρικού δικτύου. Στις εφαρμογές που παρουσιάζονται και αναλύονται στην παρούσα διατριβή χρησιμοποιείται το GridLAB-D , ένα νέο εργαλείο προσομοίωσης και ανάλυσης του ηλεκτρικού δικτύου στο επίπεδο της διανομής, μοντέλα θεωρίας παιγνίων, η θεωρία ασαφούς λογικής, τεχνικές ομαδοποίησης των καταναλωτών και τεχνικές απόκρισης της ζήτησης.

Λόγω της ελευθέρωσης της αγοράς ηλεκτρικής ενέργειας, η συμβατική αγορά και οι μονοπωλιακές αγορές αντικαθίστανται από διάφορες συνιστώσες με διαφορετικές αρμοδιότητες. Αυτές οι συνιστώσες είναι οι παραγωγοί ενέργειας, οι καταναλωτές και ο διαχειριστής της αγοράς. Οι παραγωγοί ηλεκτρικής ενέργειας είναι ανεξάρτητες επιχειρήσεις, είτε συμβατικές είτε ΑΠΕ, που παράγουν και πωλούν ενέργεια σε συγκεκριμένες τιμές. Οι καταναλωτές είναι οι πελάτες, οι οποίοι μπορεί να είναι οικιακοί, εμπορικοί ή βιομηχανικοί. Ο διαχειριστής της αγοράς είναι υπεύθυνος για την ομαλή λειτουργία της αγοράς στοχεύοντας στην διατήρηση της ισορροπίας μεταξύ της ζήτησης και της προσφοράς.

Η ενέργεια δεν μπορεί να αποθηκευτεί σε μεγάλη κλίμακα. Έτσι, ο διαχειριστής της αγοράς προσπαθεί να πουλήσει όσο περισσότερη ενέργεια είναι δυνατόν. Σε αυτό το πλαίσιο, μελετάμε τους διάφορους τύπους αγορών ηλεκτρικής ενέργειας και προτείνουμε μοντέλα για την αγορά της ηλεκτρικής ενέργειας που εισάγουν ανταγωνισμό μέσω της ενεργούς συμμετοχής τόσο των καταναλωτών όσο και των παραγωγών.

Συγκεκριμένα, πρώτα, εισάγουμε διάφορους τύπους αγορών ηλεκτρικής ενέργειας επικεντρώνοντας στα χαρακτηριστικά της αγοράς που ακολουθείται από τις Ηνωμένες Πολιτείες της Αμερικής. Η ανάπτυξη των νέων μοντέλων της αγοράς ηλεκτρικής ενέργειας αντιμετωπίζει το πρόβλημα της υψηλής κατανάλωσης ενέργειας κατά τη διάρκεια των ωρών αιχμής τόσο για την οικιακή όσο και για την γεωργική ζώνη. Η πολιτική επίλυσης αυτού του προβλήματος εξετάζει διάφορα χαρακτηριστικά όπως η προτεραιότητα στη λειτουργία συγκεκριμένων συσκευών και η προθυμία του καταναλωτή να μετατοπίσει την κατανάλωση ενέργειας σε διαφορετικές στιγμές μέσα σε ένα συγκεκριμένο χρονικό διάστημα.

Οι τεχνικές απόκρισης ζήτησης διαδραματίζουν σημαντικό ρόλο στην ασφαλή και σταθερή λειτουργία του ηλεκτρικού δικτύου. Πρωταρχικός στόχος τους είναι να εξομαλύνουν την κατανάλωση ενέργειας προσφέροντας καλύτερες τιμές στους καταναλωτές ή παρέχοντάς τους κίνητρα για να μειώσουν την κατανάλωση σε συγκεκριμένα χρονικά διαστήματα κατά τη διάρκεια της ημέρας. Πολλοί ερευνητές έχουν ως στόχο τις τεχνικές διαχείρισης της ζήτησης που δίνουν μεγαλύτερη έμφαση στην πλευρά των καταναλωτών.

Για σταθερή, αξιόπιστη και αποδοτική λειτουργία τόσο των δικτύων μεταφοράς όσο και των δικτύων διανομής, μπορούν να ληφθούν υπόψη διάφορες τεχνικές. Επιπλέον, μια ανωμαλία ή διακοπή της λειτουργίας του δικτύου ηλεκτρικής ενέργειας θα μπορούσε να μετριαστεί με την εφαρμογή διάφορων μηχανισμών λήψης αποφάσεων. Οι τεχνικές αυτές στοχεύουν στην αποτελεσματική αντίδραση του ηλεκτρικού δικτύου υπό διάφορες μορφές διαταραχών. Για παράδειγμα, μια τεράστια διαταραχή στο δίκτυο ηλεκτρικής ενέργειας θα μπορούσε να αντιμετωπιστεί με την απομόνωση του προβλήματος για την αποτροπή της διάχυσης του σε ολόκληρο το δίκτυο. Περαιτέρω, η διανομή ηλεκτρικής ενέργειας πρέπει να είναι συνεχής, έτσι ώστε η απομόνωση τυχόν προβλημάτων μπορεί να αποτραπεί με μεθόδους ομαδοποίησης.

Acknowledgements

This dissertation represents the research that was conducted towards my PhD degree from the Department of Electrical and Computer Engineering, University of Thessaly, Greece.

First, I would like to express my special appreciation and thanks to Prof. Manoli Vavali for being a great mentor and teacher and for giving me the opportunity to work with him. I would also like to thank him for taking me as an undergraduate and Master's student and for encouraging my research and for allowing me to grow as a research scientist. His advice on both research as well as on my career have been priceless. I would like to express my thanks and appreciation to the members of the dissertation committee; Prof. Lefteri Tsoukala, Prof. George Papavassilopoulos, Associate Prof. Dimitrios Bargiotas, Assistant Prof. Dimitrios Katsaros, Prof. George Limperopoulos and Associate Prof. Iordanis Koutsopoulos.

Finally, special thanks go to my parents, my siblings and my fiancée for their love, patience and their support in writing.

Publications

The results of this dissertation are described in the following publications.

Articles in Journals

1. A. Nasiakou, M. Alamaniotis, and L. H. Tsoukalas. Extending the k-means clustering algorithm to improve the compactness of the clusters. *Journal of Pattern Recognition Research*, 11(1):61-73, November 2016.

2. A. Nasiakou, M. Vavalis, and D. Zimeris. Smart energy for smart irrigation. *Computers and Electronics in Agriculture*, 129:74-83, 2016.

3. M. Foti, A. Nasiakou, L. Vasilaki, and M.Vavalis. On Visualizing Distribution Systems for Next Generation Power Distribution Grids. *International Journal of Computational & Neural Engineering (IJCNE)*,3(1): 16-27, June 2016.

4. R. Fainti, A. Nasiakou, L.H Tsoukalas, and M. Vavalis. Design and early simulations of next generation intelligent energy systems. *Int. J. Monit. Surveill. Technol. Res.*, 2(2):58-82, April 2014.

Articles in Books or Conference Proceedings

1. A. Nasiakou, M. Alamaniotis, L. Tsoukalas, and M. Vavalis. Dynamic Data Driven Partitioning of Smart Grid Using Learning Methods. *Handbook of Dynamic Data Driven Applications Systems, Springer International Publishing*, 2018.

2. E. Houstis, A. Nasiakou, M. Vavalis, and D. Zimeris. Linking Smart Energy and Smart Irrigation: Integration, System Architecture, Prototype Implementation and Experimentation,*3rd International Congress on Energy Efficiency and Energy Related Materials (ENEFM2015), Springer Proceedings in Energy. Springer, Cham*, pages 143-149, 2017.

3. A. Nasiakou, M. Alamaniotis, and L. H. Tsoukalas. Matgridgui; a toolbox for gridlab-d simulation platform. *In 2016 7th International Conference on Information, Intelligence, Systems Applications (IISA)*, pages 1-5, July 2016.

4. A. Nasiakou, M. Alamaniotis, and L. H. Tsoukalas. Power distribution network partitioning in big data environment using k-means and fuzzy logic. *IET Conference Proceedings*, pages 89-97, January 2016.

5. A. Nasiakou, M. Vavalis, and D. Bargiotas. Simulating active and reactive energy markets.*In 2015 6th International Conference on Information, Intelligence, Systems and Applications (IISA)*, pages 1-6, July 2015.

6. E. Tsoukalas, M. Vavalis, A. Nasiakou, R. Fainti, E. Houstis, G. Papavasilopoulos, E. Sarri, C. Nikolaou, and G. Koutras. Towards next generation intelligent energy systems: Design and simulations engines. *In IISA 2014, The 5th International Conference on Information, Intelligence, Systems and Applications*, pages 412-418, July 2014.

7. R. Fainti, A. Nasiakou, E. Tsoukalas, and M. Vavalis. Large scale simulations for electric energy markets. *In IISA 2014, The 5th International Conference on Information, Intelligence, Systems and Applications*, pages 160-165, July 2014.

Other Publications in Conferences and Journals

The concepts and the technologies (fuzzy logic theory, GridLAB-D) used in the following publications helped me to obtain a deeper knowledge how to use them in order to design the models presented in this dissertation.

1. M. Alamaniotis, A. Nasiakou, R. Fainti, and L. H. Tsoukalas. Leaky bucket approach implementing anticipatory control for nodal power ow management in smart energy systems. *In 2016 IEEE PES Innovative Smart Grid Technologies Conference Europe (ISGT-Europe)*, pages 1-6, Oct 2016.

2. P. L. Lagari, A. Nasiakou, R. Fainti, K. Mao, L. H. Tsoukalas, R. Bean, and M. Alamaniotis. Evaluation of human machine interface (hmi) in nuclear power plants with fuzzy logic method. *In 2016 7th International Conference on Information, Intelligence, Systems Applications (IISA)*, pages 1-6, July 2016.

3. A. Nasiakou, R. Bean, and M. Alamaniotis. Development of Human Machine Interface (HMI) for Digital Control Rooms in Nuclear Power Plants.*10th International Topical Meeting on Nuclear Plant Instrumentation, Control and Human Machine Interface Technologies*,pages 1-6, June 2017.

4. R. Fainti, A. Nasiakou, E. Tsoukalas, and M. Vavalis. Simulations of all electric vehicles on fixed routes. *International Journal of Electric and Hybrid Vehicles*,5(4), pages 296–316, 2013.

Contents

List of Figures

List of Tables

List of Abbreviations

AC	**A**lternating **C**urrent
AI	**A**rtificial **I**ntelligence
AMI	**A**dvanced **M**etering **I**nfastructure
BRPs	**B**alance **R**esponsible **P**arties
CDI	**C**luster **D**ispersion **I**ndicator
CH	**C**alinski **H**arabasz
CPP	**C**ritical **P**eak **P**ricing
DAM	**D**ay-**A**head **M**arket
DB	**D**avies **B**ouldin
DDDAS	**D**ynamic **D**ata **D**riven **A**pplication **S**ystems
DES	**D**istributed **E**nergy **S**ources
DI	**D**unn **I**ndex
DLC	**D**irect **L**oad **C**ontrol
DoD	**D**epth **o**f **D**ischarge
DOE	**D**epartment **o**f **E**nergy
DR	**D**emand **R**esponse
DRPs	**D**emand **R**esponse **P**roviders
DSM	**D**emand **Side** **M**anagement
DPWS	**D**evices **P**rofile for **W**eb **S**ervices
EDR	**E**lectrical **D**ata **R**ecorder
FLS	**F**uzzy **L**ogic **S**ystems
GUI	**G**raphical **U**ser **I**nterface
HANs	**H**ome **A**rea **N**etworks
HVAC	**H**eating **V**entilation and **A**ir **C**onditioning
IDC	**I**nternational **D**ata **C**orporation
IoT	**I**nternet **of** **T**hings
ISO	**I**ndependent **S**ystem **O**perator
LMP	**L**ocational **M**arginal **P**rice
MO	**M**arket **O**perator
MVC	**M**odel **V**iew **C**ontroller
NEM	**A**ustralian **N**ational **E**lectricity **M**arket
PCUs	**P**ower **C**onditioning **U**nits
PNNL	**P**acific **N**orthwest **N**ational **L**aboratory
PV	**P**hotovoltaic
RES	**R**enewable **E**nergy **S**ources
RMSSTD	**R**oot **M**ean **S**quare **S**tandard **D**eviation
RS	**R** **S**quared
RTP	**R**eal **T**ime **P**ricing
S	**S**ilhouette
SMI	**S**imilarity **M**atrix **I**ndicator
SO	**S**ystem **O**perator
TOU	**T**ime **of** **U**se
TSO	**T**ransmission **S**ystem **O**perator

UPHS	**U**nderground **P**ump **H**ydroelectric **S**torage
WAN	**W**ide **A**rea **N**etwork
WoT	**W**eb **of** **T**hings

Dedicated to my family and my fiancée

Chapter 1

Introduction

1.1 Motivation

Concerns about global climate change, economic stability and the need for consumers' participation in the operation of the power grid lead to the transition from the conventional power grid to the smart grid. In the smart grids, renewable energy generation technologies, smart metering infrastructures, and energy storage technologies are modernizing the electricity infrastructure. The smart grid through the RES introduces stochasticity on the generation besides the stochasticity on the demand side.

Demand Side Management (DSM) techniques for both commercial and residential consumers characterize the liberalized electricity markets. To assess such an overall impact, occasionally, a lot of theoretical studies and a set of case studies have been conducted and published (Wogrin, Barquín, and Centeno, 2013), (Li, Shi, and Qu, 2011b), (Grimm et al., 2016), (Koliou et al., 2014). Not many associated simulation studies have been performed in large scale using real data. So far, in the current simulation studies, the data are not as realistic as the ones that are used in this dissertation; the data used in this dissertation concerns information regarding the distribution network.

The power grid remains in a healthy and stable state only when all its components (conventional generators, RES, customer-loads, kilometers of transmission and distribution lines etc.) are properly configured. The solution of such a problem is difficult to be found. Some of the widely used tools for addressing such complex problems are machine learning (Michalski, Carbonell, and Mitchell, 2013) and the multi-agent approaches. Both the aforementioned tools utilized for addressing the objective of this dissertation. An agent-based simulation engine with integrated modules is developed for integrating the proposed models of this dissertation. Using these models, the system can be characterized by a very desirable feature; self-healing which means that whenever the system operates at its optimal operating point.

The primary objective of this dissertation is inspired by the observation that there is a big amount of wasted produced energy due to the inefficient operation of the power grid. Moreover, the observation that there is a lot of space (grey area in Figure 1.1) for a whole day in which the energy consumption can be shifted is a challenge for developing models that can fill these gaps. So, the objective of this dissertation is the design and the analysis of models that focus on the optimal operation of the power grid with less energy losses and the more integration of the RES. In Figure 1.2, the energy savings in the cases where DSM techniques are applied are shown. In both figures, energy demand is shown with yellow and energy waste in grey color. More specifically, the power pattern for an entire working week (five peaks) is depicted. It is clear that there are significant differences in the demand during the day and the night. Ideally, the generated power must be in balance with the demand and it has to be approximately 10% of the peak demand for security reasons. As a result, a significant amount of energy (grey area) it does not actually consumed to cover any demand. The energy that is not produced for that mid-size smart grid is approximately 40% of the energy actually used.

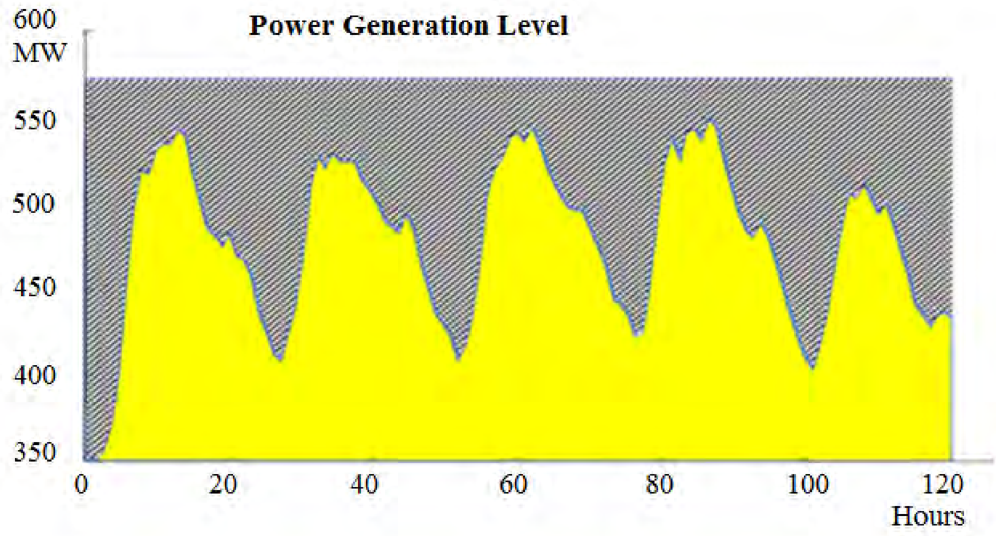

FIGURE 1.1: Demand on a mid-size smart grid of max power in the range of 600 MW (Tsoukalas et al., 2014).

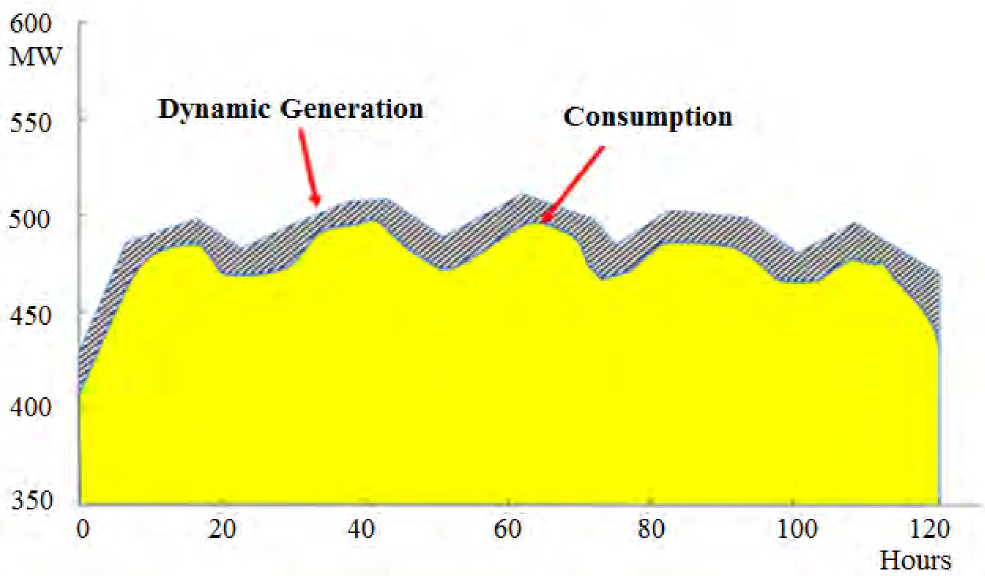

FIGURE 1.2: Expected energy savings using demand side management techniques (Tsoukalas et al., 2014).

More precisely, this dissertation focuses on the presentation of the design, the implementation and the analysis of the simulation results of models that

- pay attention to the decentralized markets focusing on the competitiveness. Producers and consumers participate in this type of market through auctions for determining the price and the quantity of the energy that is going to be dispatched.

- involves both the producers and the consumers in the operation of the power grid and it provides support for both the transmission and the distribution networks.
- introduce the RES to the operation of the power grid.
- combine several characteristics such as accuracy, efficiency, robustness and reliability in the operation of the next generation energy systems. Additionally, the basic physical, technical and societal principles are also involved.
- use artificial intelligence techniques to develop. optimal market strategies, partitioning methods that focus on the optimal, reliable and efficient operation of the power grid.

In this dissertation, the GridLAB-D simulation platform is used for all the demonstrations presented.

1.2 Contribution of this Dissertation

In this section, the contributions of this dissertation which are analytically presented are described as follows.

In Chapter 2, basic concepts used in this dissertation are described and analyzed in detail. First, an analysis how the power grid works focusing on the operation of the smart grid has been made. The Big Data is related to the emergence of the smart grid, so a brief definition of this concept as well is declared. A general background about Demand Response (DR) programs is also presented. Additionally, this chapter provides background details about the game theory, the fuzzy sets, the clustering algorithms, the Renewable Energy Sources (RES), the batteries and the distribution systems. Finally, some information about the power theory is also presented.

In Chapter 3, a detailed description of the GridLAB-D simulation platform that is used to test, design and develop the proposed models and analyze the results is presented. Moreover, due to the fact that the GridLAB-D provides only a text-based interface difficult to be handled by any user, a friendly Graphical User Interface (GUI), called MatGridGUI, is proposed. In particular, it provides to the users a simple way to run and create a simulation model from scratch without having any knowledge about how the psychical components of a power grid have to be connected. Moreover, a visualization tool, called GridLAB-Dvisor, is also presented. It uses the results of a GridLAB-D simulation and it visualizes them using various technologies.

In Chapter 4, market models for both the active and reactive power are presented in detail. The need of reducing the "wasted" energy leads to the development of models that use DR programs and distributed energy resources for shifting the demand from the on peak hours to the off peak hours. Moreover, the importance of the reactive power for the safe and reliable operation of the power grid offers the challenge of modeling the operation of the reactive power markets. Moreover, the effect on the price, that the consumers have to pay if there is more need for reactive power than the one that is provided is also a big challenge.

In Chapter 5, an extension of the standard k-means algorithm which improves its performance concerning the selection of the initial cluster centers is presented. The different way the initial cluster centers are selected, based on the test cases presented in this dissertation, leads to better clustering results concerning the compactness and the well-separation of the clusters. This algorithm besides with other partitioning algorithms are utilized to design the models and methodologies regarding the partitioning of the distribution grid which focus on

- selling energy at lower prices to a specific partition which is chosen through fuzzy logic decision making methods
- grouping the residential consumers to groups with similar energy consumption patterns.

In Chapter 6, a very promising model which integrates the RES in the market that has as primary goal to satisfy the demand of the agricultural sector is presented. The results of two different scenarios concerning the irrigation system show that the RES can provide enough energy to satisfy the demand of both the agricultural and residential sector.

In Chapter 7, a summary of the contribution of this dissertation is presented and a small discussion about some future plans is also analyzed.

Chapter 2

Basic Concepts and Enabling Technologies

2.1 How the Grid Works

The power grid is the most complex system which delivers electricity from producers to consumers. It is 100 years old and it must cover the demand of consumers for energy. In short, the electricity is delivered from energy providers to consumers through the power grid — from generation through the transmission and the distribution network to the consumers either residential, commercial or industrial.

2.2 Smart Grid

The U.S. Department of Energy (Energy, 2009) describes the transformation of the grid into the smart grid as *"a transformation from a centralized, producer-controlled network to one that is less centralized and more consumer-interactive. A smarter grid makes this transformation possible by bringing the philosophies, concepts, and technologies that enabled the Internet to the utility and the electric grid. More importantly, it enables the industry's best ideas for grid modernization to achieve their full potential. Many of these ideas are already in operation. Yet it is only when they are empowered by means of the two-way digital communication and plug-and-play capabilities that exemplify a smarter grid that genuine breakthroughs begin to multiply. Adoption of the Smart Grid will enhance every facet of the electric delivery system, including generation, transmission, distribution, and consumption."*

The smart grid is the evolution of the traditional electrical power grid that integrates more technologies related to the monitoring, controlling and communication to ensure the optimized usage of its capacity (Fang et al., 2012). Smart grid technologies facilitate grid-connected intermittent generation from RES such as the wind, the solar, by using the energy produced by these sources during peak usage hours and replenishing it during off peak hours. Additionally, these technologies give the ability to read any available data from the meters, record power consumption habits and create load patterns necessary for forecasting and clustering purposes, provision of real-time power energy costs, any kind of remotely support, and auto-healing any disturbances in the grid by isolating the problem. The smart grid encourages the active participation of both producers and consumers using smart metering devices and bi-directional communications. An important technology for this interaction is the Advanced Metering Infrastructure (AMI) (Balaji, Ram, and Nair, 2017) which enables the two-way communication between the producers and the consumers. AMI, as an integral part of a smart grid, has been paid more attention the last decades due to the ability to offer the energy in a more efficient way using electric metering and communications; the smart meters are equipped with technologies which provide two-way communications. AMI provides a lot of capabilities to both consumers and producers. Real time consumption monitoring and controlling, automated billing for the side of the consumers, distribution automation and DR programs methods for the producers' side are available through AMI. The architecture of an AMI system is depicted in Figure 2.1 (Balaji, Ram, and Nair, 2016). According to this architecture each consumer is equipped with a smart meter. The data of the meter through properly configured frameworks and networks are available to the energy providers and market operators. More specifically, Home Area Networks (HANs) connect the smart meters, the smart devices within the residency, the generation and the energy storage devices. The Wide Area Network (WAN) provides connections between the devices to request specific data, to initiate tasks. The smart meters of an AMI can capture electricity usage data on a periodic basis and transmit that data to the energy providers/producers. This energy data can be driven many times per second. This facilitates the consumers to change their consumption according to the energy price any particular time instance.

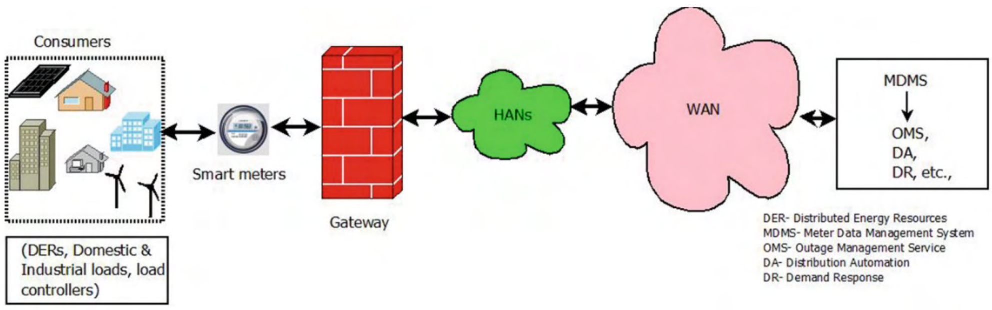

FIGURE 2.1: Architecture of an AMI system (Jayanth Balaji, Harish Ram, and Nair, 2016).

The energy of the traditional power grid is one way flow; it starts from the producers and through the grid, it ends to the consumers. As it is observed, it has just one direction, from producers via the power grid transmission lines to consumers via distribution lines. The traditional power grid consists of four main components: (a) generators, (b) transmission network, (c) distribution network and (d) consumers. The generators and the transmission network belong to the supply-side system while the distribution network and the consumers to the demand-side system. Generators are located far away from consumers and both the transmission and distribution network have to deliver the energy to long distances which lead to energy losses (Yoldas et al., 2017), (Tuballa and Abundo, 2016).

The aforementioned drawback, the smart grid is called to resolve. A key element of the smart grid besides the auto-healing is the distributed generation. Using distributed generation, the generation is close to the end-users. The smart grid tries to eliminate the distance between the consumption and the energy production, transforming in that way the traditional power grid into a decentralized power system (Li, Shi, and Qu, 2011a), (Hejazi et al., 2013). This leads to the integration of the RES, which can be characterized as small-scale power generators. Part of produced energy by the RES can be stored in batteries for future use. The consumers act both as buyers and sellers, commonly called prosumers, establishing in that way a bi-directional communication.

The energy in the smart grid flows in a loop in which the energy is exchanged between the participants. Besides the energy flow, another basic difference between the smart grid and the traditional power grid is the Data Flow (Jarrah et al., 2015). In the traditional power grid, there is no feedback from consumers to producers. The producers manually or remotely read the aggregated energy usage from the consumers' meter, and the consumers as a "response" take the monthly billing invoice. Consumers do not have an active role in this procedure and they do not have the ability to identify energy consumption patterns in real-time. On the other hand, in the smart grid, the energy data are available to all the energy providers and consumers. Consumers are able to make decisions when they want to use the electricity, to choose low-cost consumption periods, and reserve electricity for emergency use. The smart grid technologies are used to provide information about energy consumption patterns, real time electricity prices and energy use suggestions.

A concept close to the smart grid is the microgrid. The emergence of the microgrid is a consequence of the distributed generation. The microgrid is characterized as the "building block of smart grid" (Hatziargyriou, 2014). It consists of generators, storage devices, and end-users/consumers that have the ability to disconnect from the rest of the electrical network. This mode is called "islanding mode", and it is used to isolate any problem occurs in a part of transmission/distribution grid. In such a case the microgrid's production from the Distributed Energy Sources (DES) is required to meet the demand for energy. Using the "islanding mode" the reliable and the safe operation of the power grid is achieved. Another mode of operation is the grid-connected mode. In this kind of operation mode, power can be either received or injected from the main grid.

2.3 Big Data

The rapid development of sensor technology, wireless transmission technology, cloud computing, network communication technology, and of smart mobile devices lead to the accumulation of large amounts of data. The volume of data is also growing rapidly with complex forms. In Figure 2.2, the integration of the smart grid with the big data is illustrated. Actually, the "connection" of those two concepts is achieved through a pool of big data which is the intermediate between the smart grid and the data. It is estimated that by 2020 the volume of the data will be increased 50 times of the size that the International Data Corporation (IDC) (Gantz and Reinsel, 2011) pointed out (1.8 ZB in 2011). The era of the Big Data has arrived (Zhou, Fu, and Yang, 2016). The Big Data concept is used in the applications presented in Chapter 5.

2.3.1 Characteristics of Big Data in Energy Sector

In the energy sector, the concept of big data is introduced and studied by developing smart energy management techniques. Energy big data have provided the users with a plethora of opportunities and challenges. Some of the challenges are the following:

- collection, store, and management of the energy big data using efficient ways
- efficient analysis and energy big data mining
- mining of information and values from the energy big data
- prevention of risks using effective ways

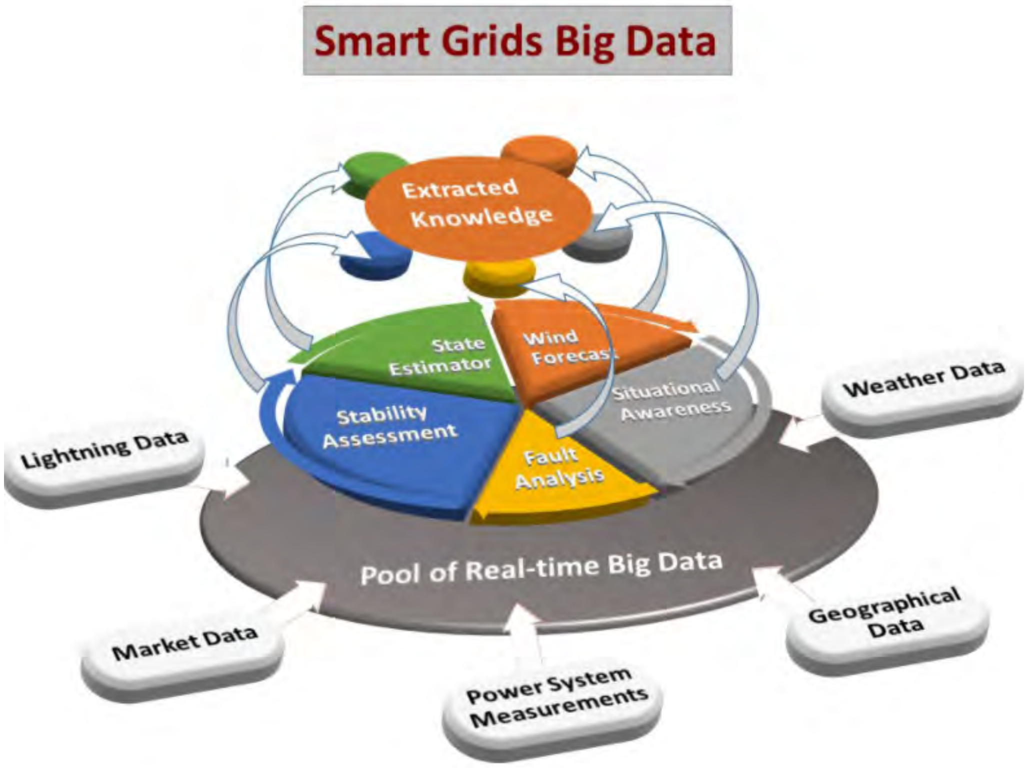

FIGURE 2.2: Smart Grid and Big Data.

Source: http://smartgridcenter.tamu.edu/sgc/web/?page_id=4313

- protection of users' privacy
- support of decision-making methods

Generally, Big Data has the characteristics of "4V"; **volume**, **velocity**, **variety** and **value**. For the energy big data, these characteristics follow the aspects described below (Zhou, Fu, and Yang, 2016), (Hashem et al., 2015).

- **Volume**: It refers to the amount of energy data generated every particular time instance. This big amount of data has to be stored. With the big data technologies and the help of the distributed systems, the data can be stored in different locations and they can be driven together each time.
- **Velocity**: It refers to the speed for the process, collection and the use of the data. Generally, in smart energy systems, the speed ranges from 5 minutes to 1 second.
- **Variety**: It refers to the complexity of the data. In smart energy systems, the data are related with weather data, data associated with the consumers' energy consumption, and the energy produced by both RES and the conventional generators. Actually, the energy big data consists of structured, semi-structured and unstructured data.
- **Value**: The value of the data related with the smart energy systems is an index indicated the information about the data. Information regarding the consumers' energy consumption, how to develop marketing strategies, how to

improve the system's reliability and performance can be some of the possible values of the data.

Besides the "4V", Zhou, Fu, and Yang (2016) declared also the "3E" for the big data related with the energy sector. The "3E" reflects to the **energy, exchange, and empathy**. The **energy** refers to the energy savings taking through big data analytics methods. The **exchange** refers to the exchange of the energy big data from other sources and finally the **empathy** refers to energy services and the consumers' needs. In Figure 2.3, the characteristics of energy big data are depicted.

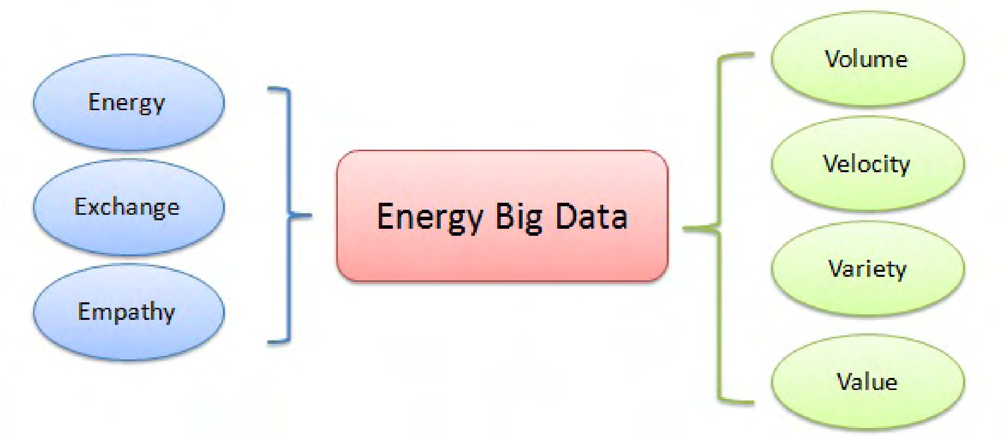

FIGURE 2.3: 3E and 4V characteristics of energy big data (Zhou, Fu, and Yang, 2016).

2.4 Demand Response and Energy Markets Models

2.4.1 Demand Response Applications

The purpose of the traditional power grid is to manage and control the supply from the energy utilities (or producers) for satisfying the energy demand. On the other hand, the emergence of the smart grid provides the active role of the consumers in the operation of the power grid through DR programs. According to Department of Energy in the United States (Energy, 2006), DR is "*tariff or program established to motivate changes in electric use by end-use customers in response to changes in the price of electricity over time, or to give incentive payments designed to induce lower electricity use at times of high market prices or when grid reliability is jeopardized*". DR programs and their technologies (pricing schemes) are used in this dissertation (Chapter 4, 5, and 6) for monitoring and controlling the behavior of the consumers to the changes of the price. It is worth mentioning at that point, that the energy efficiency and the DR are two totally different concepts. The energy efficiency stands with actions that are taken during the operation of the power grid. Some of these actions are listed below (Sebastian and Margaret, 2016).

1. reduction of energy consumption
2. use of more efficient appliances
3. turning on/off the appliances

On the other hand, the DR programs primarily focus on energy use shifting; from the on peak to the off peak hours for energy and money savings (Faria and Vale, 2011). More specifically, the response in electricity usage is achieved through DR programs which are designed to coordinate electricity use with power system operation. Some of the most common applications of the DR programs are listed below.

1. DR provides elasticity in the electricity price. It can decrease the high prices and their volatility. When the power system does not operate under DR programs, the demand is inelastic. Moreover, when the generators work close to their max capacity, the supply is inelastic as well. So, in such a case, the electricity price is high and especially in times that the demand is too high.

2. DR can reduce the number of the generators that work in their max capacity.

3. DR can improve the reliability and the performance of the power grid. Additionally, it can provide flexibility to both the consumers and producers.

4. DR can integrate more easily the RES into the power grid by providing energy balance.

Price elasticity is a concept which is strongly connected with DR programs. Besides the price elasticity of supply as aforementioned, there is price elasticity of demand as well. The price elasticity of supply measures how the supplied energy reacts to a change in its price. The formula for calculated the price elasticity of the supply is the following:

$$\lambda_{supply} = \frac{\%change\_inquantity\_supplied}{\%change\_in\_price} \tag{2.1}$$

On the other hand, the price elasticity of demand is a measure to quantify the change in the quantity demanded with respect to a change in its price. The price elasticity of demand is often used when the price sensitivity needs to be studied and analyzed. The formula for calculated the price elasticity of the demand is the following:

$$\lambda_{demand} = \frac{\%change\_in\_quantity\_demanded}{\%change\_in\_price} \tag{2.2}$$

Both elasticity for demand and supply can be described as *elastic*, *unit elastic*, or *inelastic* (Thimmapuram and Kim, 2013). An elastic demand or elastic supply is one in which the elasticity is greater than one, indicating a high responsiveness to changes in price; any high or even small increase in the price drops the demand to zero. Therefore, when the price increases, the total revenue drops to zero. In inelastic demand or inelastic supply, the price elasticity is less than one, indicating in that way low responsiveness to price changes. In other words, changes in the price do not affect the demand or the supply. Unitary elasticities note proportional responsiveness. In that case, the elasticity is zero.

In Figure 2.4, the behavior of the price's elasticity value of the demand is depicted. The x-axis corresponds to the demand and the y-axis to the price. Getting into more details, it can be observed that the price elasticity is increased when the demand is increased and the slope tends to be bigger. This happens due to the fact that when the price is decreased, more consumers focus on buying with that price, so the price elasticity is increased in order the ratio of the equation 2.2 to be steady.

DR is introduced in the power grid through different ways such as distributed generation, dispatchable load, storage and other resources that may contribute to

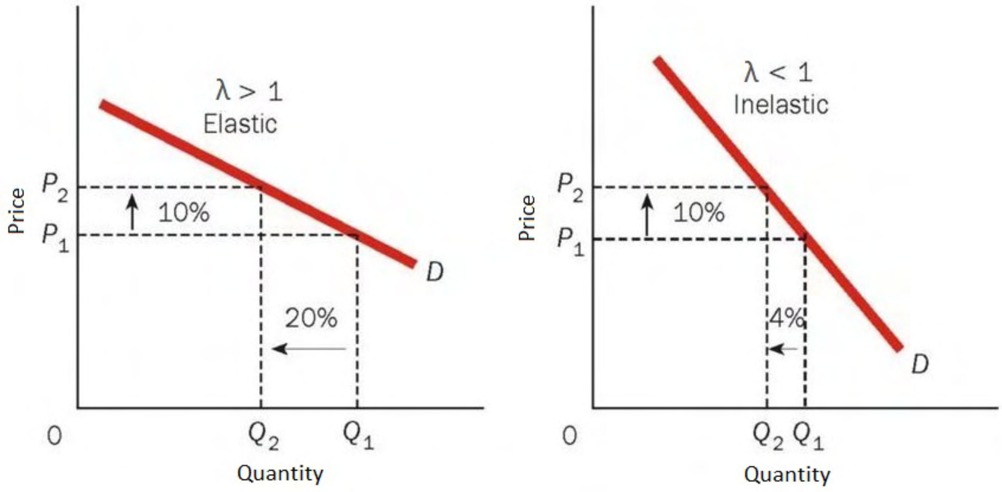

FIGURE 2.4: Elastic and Inelastic Demand.

Source: http://www.economicsguide.me/?page_id=1990

modify the power supplied by the main grid. It is worth noting here that there are three possible ways as following which can be used in order the end-users to comply with the challenges of using the electricity as the System Operator (SO) suggests (Connell et al., 2014):

1. Reduction of the electricity use through suggested strategies
2. Shift energy consumption to off peak-hours or periods with low demand
3. Use of standby generated energy

The strategies that the end-user can follow, usually called load curtailment strategies, can be achieved by changing the set-points of the heating ventilation and air conditioning (HVAC) systems, dimming the lighting system, etc. On the other hand, the shifting of the energy consumption can be attained by pre-heating or pre-cooling of the buildings.

As aforementioned, DR refers to the changes in demand in response to changes in the price or to incentive payments to force lower use of electricity when the price is high or the reliability of the system is in danger. So, DR programs can be categorized into two groups; incentive (Zhong, Xie, and Xia, 2013) and price-based DR (Moghaddam, Abdollahi, and Rashidinejad, 2011). Through the price-based programs, the end-users can participate in wholesale electricity markets. In demand bidding programs and capacity bidding programs, the end-users offer demand reductions via price/quantity bids into electricity markets. If their bids are accepted from the market, they have to provide demand reductions at specified times. The types of price-based programs are presented in the next lines.

Types of DR Programs

The electricity price depends both on the available supply and on the demand, so its volatility may be increased with an increased electricity production. DR programs use techniques to induce end-users to reduce demand. On the other hand, they also could force the end-users to use electricity during periods of low demand and high production for reducing in that way the wasted energy. In the next lines, some types of DR programs (Siano, 2014) are summarized.

Time of use Tariff (TOU) - Common program used in residential consumers; this type of DR program offers various electricity prices (most commonly three levels) per day. Therefore, the consumers have the challenge to shift their energy consumption in low prices' periods. In a traditional power grid where the consumption and production are predictable, this tariff could be effective. However, the structure of the tariff is too static to monitor the state in a power system with a large share of RES.

Critical Peak Pricing (CPP) - This program offers to the end-users decreased electricity prices when the power grid operates in steady state mode. On the other hand, when the power grid is congested, the prices are increased to avoid any disturbances and collapses. However, CPP can be applied for a limited period during a year, so it cannot be used to improve the performance of the power grid.

Direct Load Control (DLC) - This program compensates the end-users if they prefer in a request of the SO to turn on some of their appliances. The energy demand can be increased if an appliance is in operation for a long time of period. Its implementation is simplified.

Power tariffs - It provides kind of incentives for the consumers to reduce the peak demand by increasing the network tariff as their peak demand increases. Traditionally, power tariffs are based on the monthly peak demand but they could be based on shorter time periods as well. This tariff would generally result in an overall peak shaving in the power system but it does not provide any incentives to schedule the demand in order to follow the electricity production. Moreover, the peak demand of a single customer does not necessarily increase the total peak demand in the power system, since the customer's peak may occur at a different time compared to the total peak in the power system.

Real Time Pricing (RTP) - The main idea of an RTP program is to provide the customers with an electricity price that reflects the actual conditions of the power grid. One important aspect, regarding the design of an RTP scheme, is the time difference between the announcement of the price to the consumers and the actual consumption. A long time, e.g. using day-ahead price, which is widely used, would result in a price that less accurately reflects the power system conditions. It may also result in increased need for balancing power if the scheduled demand is not considered in the day-ahead market. A shorter time lag, e.g. based on the intraday market, would result in better reflection of the demand/supply but it is more difficult for the consumers to plan their electricity consumption. The RTP can be calculated based on the area price, as in the Nordic power market, or on the locational marginal price (LMP), as in the New Zealand power market and some markets in the USA, although the RTP is not offered to all customers in all markets (Hiko, 2011).

2.4.2 Overview of Electricity Markets

In the 1990s, the deregulation of many power systems in the U.S. had been started. The conversion from the vertical integration to the wholesale electricity markets was a fact. However, the crisis in California at 2000-2001 prevailed that the active participation of the consumers in the operation of the power grid is necessary. Electricity is a commodity with high impact on the social and economic developments around the world.

Electricity cannot be stored in large quantities and its storage in batteries depending on the batteries' features and size is very expensive. Hence, the supply and the demand has to be in balance. Moreover, the energy losses on the supply side due to the transportation over long distances has to be included. Otherwise, disturbances and collapses may occur. Therefore, the design and the development of electricity markets are necessary for the safe and efficient operation of the power grid (Wang et al., 2015). Different types of electricity markets are arranged in a sequential turn, starting time before and ending time after the actual delivery of the electricity to the end-users.

Agents participating in the electricity market are depicted in Figure 2.5. These agents are the suppliers or producers or generators or utilities or energy providers, the consumers or end-users or customers, the SO (Independent or Transmission or Market) and the retailers or Demand Response Providers (DRPs) or intermediaries. Producers are responsible for producing electricity and selling it to the market. A generator may sell energy to the electricity market and/or directly to consumers through bilateral contracts.

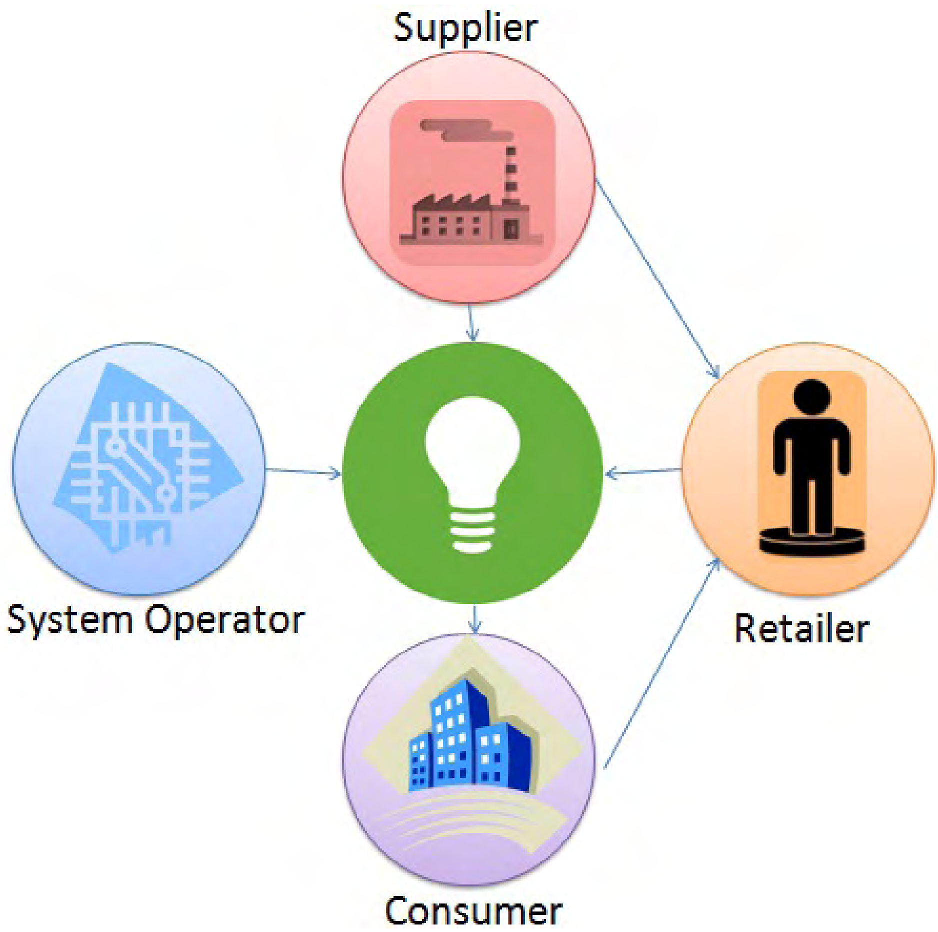

FIGURE 2.5: Components of electricity market.

The balance between the supply and the demand and the security of the power system is controlled by the SO. More specifically, it maintains the physical network. The SO may be owned by the Transmission System Operator (TSO) or may be fully independent called Independent System Operator (ISO). Besides the SO, the Market Operator (MO) is also a significant part of the electricity market. It is responsible for the economic management of the market, to provide and design the market rules and determines prices and quantities of energy traded in the market. In various markets, such as the Australian National Electricity Market (NEM), PJM and New England

ISO markets, the ISO and the MO operate as a single entity (Conejo, Carrión, and Morales, 2010).

End-users can participate in DR programs directly with the supplier, or through a retailer. In the organized markets, such as wholesale electricity markets, the aggregated energy of all the end-users is presented in the market through the retailer. On the other hand, retailers can also be aggregated from the utilities in order to provide to them contracts. The retailers buy energy from the utilities at a specific price and then they sell the energy to the end-users to a different price through contracts or at a specific price per MWh.

Types of electricity markets

Before describing the different types of markets, it is worth describing here the exact meaning of the market clearing process. It is actually, the procedure from which the price, called clearing price, and the quantity, called clearing quantity, are available to the participants of a market. The market clearing process ends when a specific period, called market period, expires; after that, a new market period starts.

Based on the way that the trade between the generators and the end-users is arranged, there are two types of markets; mediated and bilateral (Prabavathi and Gnanadass, 2015). A mediated market can be a dealer, a power exchange or a power pool. A dealer can fit an amount of demand and supply by selling it at a price most usually higher than the one he purchased it. At an exchange, the market is organized as an auction. This type of energy market is used in the applications of this dissertation. In the auction-based market, both consumers and producers submit their bids and offers respectively for a market period. The offers are sorted by the lowest price to the highest, while the bids from the highest to the lowest price. The intersection of supply and demand curves that are created denotes the market clearing price and quantity. In the pool market, the SO buys the whole generation from producers at one uniform price. More specifically, it collects all the bids from the suppliers and it sorts them from the cheapest to the most expensive. This can be achieved by collecting bids from suppliers (i.e., the price for a certain generation capacity) and aggregating them from cheapest to most expensive. The intersection with the (usually inelastic) demand curve provides the market price (Hogan, 1996).

As an alternative to trading at an exchange (mediated), is the bilateral market, usually called OTC (over-the-counter) trading. Bilateral markets are much more flexible concerning the contracts and there is a capability for direct trade between the participants. Usually, bilateral trading stands for end-users that have already signed contracts. In general, bilateral trades can be organized in two ways: by fitting the bids of participants or by direct communication between the utilities. The most significant distinction of the bilateral markets is the continuous trading, with prices unique to each transaction.

Besides the types of the markets based on the way of the trading, the trading is organized in different time frames (Panagiotakopoulou, 2015), hence four different types of markets are formulated as described below. In Figure 2.6, the sequence of these markets is illustrated.

1. *Forward-Future Market*: Forward or Future Markets operate from a year or more ahead down until the clearing period expires. Then, the SO takes his role. Forwards and futures are contracts for delivery of electricity at a predefined time in the future for a price agreed upon today. Forward or Future markets can be organized as bilateral or mediated market. Forwards are traded on a

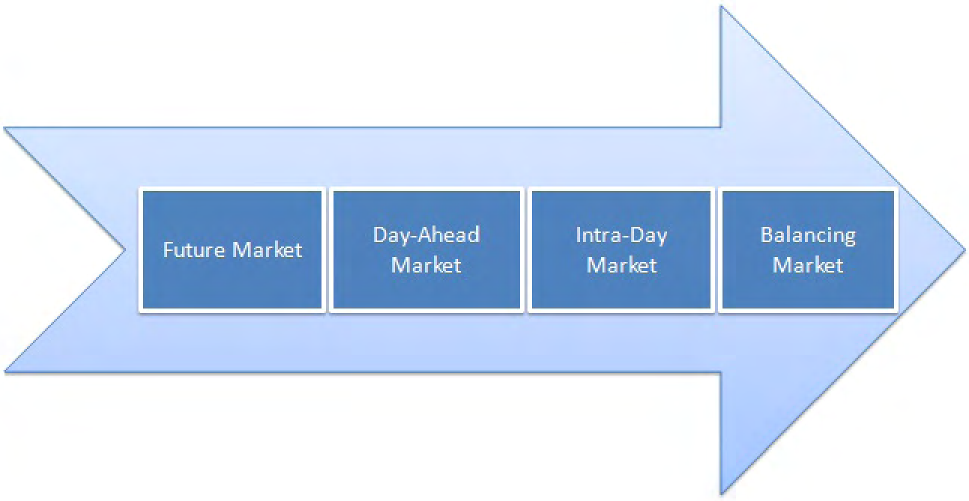

FIGURE 2.6: Market Time Frames.

Source: KU Leuven Energy Institute

bilateral while Futures are traded on an exchange. These types of contracts allow both the producers and consumers to lock in a price and avoid the price volatility.

2. *Day-Ahead Market*: The Day-Ahead Market (DAM) is the most widely used market with the highest traded volumes of energy. The participants of the market sell or buy energy as the day-ahead prices for the following day. The DAM market is based on the bids and offers data submitted to the DAM. DAM is used in conjunction with real-time (spot) markets. This aims to balance the deviation of the participants from the day-ahead strategies.

3. *Intra-Day Market*: In the Intra-Day Market, electricity is traded on the same day. This type of market allows the market participants to optimize their strategies based on the forecasts, power plant outages, etc.

4. *Balancing Market*: This type of market is not a real market where consumers and producers can participate or have contracts. It actually refers to the process via which the SO maintains the physical network. However, it is a single-buyer market where the SO is the only buyer.

Besides the energy market, the last decades the reactive power market (Ganger et al., 2013) has been paid attention. However, it is not yet priced in most of the reactive market models. Many researchers have worked in presented simulations and in modeling the objective function of the reactive power market operation as a multi-objective optimization problem (Saini and Saraswat, 2012). The main characteristics of a reactive power market are based on the facts that:

- Large quantity of reactive power into the transmission system over long distances is costly, so the local production of reactive power is desired.

- A reactive power market is usually monopsony – only one buyer, the SO that actually takes the charges from the customers.

- The market operates on long-term contracts, so any disturbances in the active price do not affect the reactive power price.
- The market is organized as a uniform auction; all the providers receive the highest price offered by all the providers.

2.5 Game theory

A branch of applied mathematics and sciences is the game theory. It has been utilized in economics, computer and social science, engineering and more specifically, smart grid. During the past decade, a plethora of activities uses game theory to model and analyze a variety of problems such as energy management problems (Zhu et al., 2012). Game theory is used in this dissertation for the maximization of the producers' profit in a hierarchy of two energy markets. More details can be found in Chapter 4. The terms commonly used in the study of game theory are the following (Tadelis, 2013):

- **Players**: A decision maker based on the rules of the game.
- **Strategy**: A complete set of actions that a player can take.
- **Payoff**: The profit of a player that he receives after the end of the game.
- **Equilibrium**: The point in a game where all players have made their decisions and the game is over.

Note that the most widely known equilibrium is the Nash Equilibrium which is one of the foundational concepts in game theory. The Nash equilibrium of a game involving two or more players in which each player is assumed to have knowledge of the strategies of the remaining players is when none of the players are going to gain more if he changes his own strategy. More specifically, if a player decides for a strategy and none of the players can gain more by changing his own strategies, the set of the players' strategies with the associated payoffs corresponds to a Nash equilibrium.

The classic game theory problem is the "Prisoner's Dilemma". The Prisoner's dilemma pertains to analyzing the decision-making process when two criminals are arrested but there are not enough evidences to convict them. Thus, the two prisoners are isolated and they got an offer as follows: if one convicts the other, he will get a reduced sentence or go free. However, none of the prisoners have information about the decision of each other. The payoff, if they both say nothing, is good. On the other hand, if one of them says nothing and the other one betrays, then the latter has benefits. If they both confess, they both get a reduced sentence. The dilemma is the choice between the two decisions. More precisely, the Prisoners' Dilemma explains the interaction between any players of a game and that the game theory offers interdependent decision-making. For competitive players, as the users in smart grid are, the communication between the players is small and producers and/or consumers are self-interested, trying to optimize their own profits. Therefore, noncooperative models, as explained next, are the most commonly used in the smart grid.

2.5.1 Noncooperative Games

In noncooperative games (Mangiatordi, Pallotti, and Vecchio, 2013), the players have conflicts (partial or total) regarding the interest over the outcome of a decision making process. The noncooperative term does not mean that the players of the game do not cooperate. For instance, energy providers offer competition over pricing strategies, market control. Such a game provides competitiveness where each player needs to make a decision independent of the other players. The player can have information about the possible choices of the other players and their effect on the player's profits. Note that any cooperation of a noncooperative game is self-enforcing without any communication among the players of the game. For instance, a smart meter can show the real-time electricity price in the SO. The demand-side management decision has to be made without asking "neighbors" whether they turn on/off each of their appliances.

2.6 Fuzzy Logic Preliminaries

In this section of the current chapter, the basic concepts of the fuzzy Logic theory which its technologies are used for the development of a fuzzy logic decision-making method as it is described in Chapter 5 are presented. It is known that several situations in the real world are fuzzy. Zadeh (1965) introduced the Fuzzy Logic Set Theory to overcome and handle problems with uncertainty.

Fuzzy Logic focuses on describing and analyzing reasoning based on fuzzy set theory. Fuzzy logic theory can be considered as an extension of the classical set theory inheriting most of its rules. Generally, fuzzy systems are the systems which are based on fuzzy set theory and/or fuzzy logic. More specifically, a fuzzy system consists of a base relied on linguistic rules and an inference engine based on fuzzy logic.

Fuzzy Systems

In this dissertation, the fuzzy logic toolbox and the Simulink are used to design, implement and test the proposed Fuzzy Logic Systems (FLS). The Fuzzy systems are also called fuzzy rule-based systems or fuzzy inference systems. These systems take the outputs using linguistic rule based and fuzzy inference processes. In Figure 2.7, the structure of an FLS is depicted.

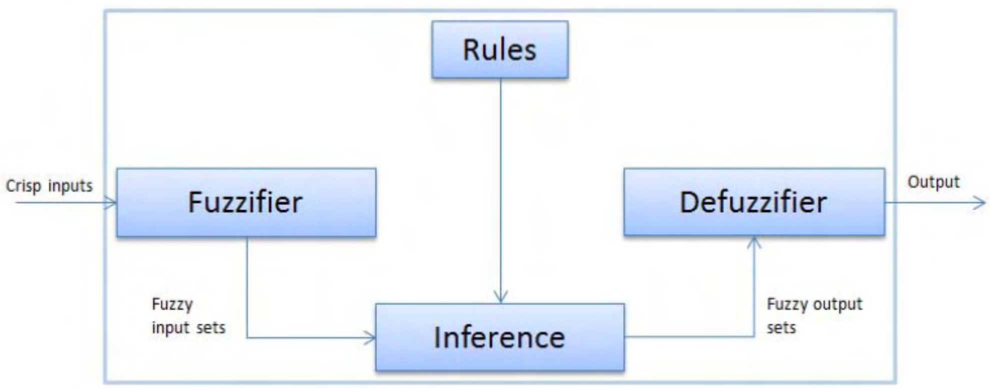

FIGURE 2.7: Fuzzy Logic Inference System Structure (Lagari et al., 2016).

An FLS consists of four main components:

- Fuzzifier: it transforms the crisp input sets into fuzzy.

- Rules: an IF-THEN based set of rules correlates the inputs.
- Inference: it uses the inputs and the IF-THEN rules to make an inference.
- Defuzzifier: it transforms a fuzzy set into a crisp output value.

There are two types of fuzzy systems concerning the type of the fuzzy rules, the *Mamdani* fuzzy systems, and the *Takagi-Sugeno-Kang* fuzzy systems. In Mamdani fuzzy systems both antecedent and consequent are fuzzy propositions. The rules of this type of systems have the following format:

if x_1 *is* A_1 *and* x_2 *is* A_2 *and .... and* x_n *is* A_n ***then*** y_1 *is* B_1

In *Takagi-Sugeno-Kang* systems, the antecedent is a fuzzy proposition while the consequent is a crisp function of the variables in the antecedent. The rules of this type of systems have the following format:

if x_1 *is* A_1 *and* x_2 *is* A_2 *and .... and* x_n *is* A_n ***then*** y_1 *is f(*x_1*)*

where $f()$ is a function which combines the inputs of the fuzzy system.

Fuzzy Logic

As it aforementioned, Fuzzy logic can be considered as an extension of the classical logic. In the next lines, the concept of *linguistic variable*, *Negation*, *Conjunction*, *Disjunction* and *fuzzy proposition* are presented.

A *linguistic variable*, is defined as a tuple (x, T(X), U, G, M) where:

- X is the name of the base variable
- T(X) is the set of linguistic terms of X (e.g., T(X) = high, medium)
- U is the universe of discourse of the base variable
- G is the rule for generating the linguistic terms based on the primary linguistic terms (e.g., very low, not high);
- M is the rule for giving semantic meaning to each linguistic term

A *fuzzy proposition* is a logical proposition for each of the fuzzy sets. If a proposition P is assigned to the fuzzy set A the degree of truth of P is defined as follows:

$$T(P) = \mu_A(x), x \in U \tag{2.3}$$

For the next lines, a *fuzzy proposition* $P : x \in A$ for a fuzzy set A and a *fuzzy proposition* $R : y \in B$ for a fuzzy set B is considered. The *Negation* is given by the formula $T(-P) = \mu_{\overline{A}(x)} = 1 - \mu_A(x)$. The *Conjunction* and the *Disjunction* are defined as follows:

$$T(P \wedge Q) = min(\mu_A(x), \mu_B(x)) \tag{2.4}$$

$$T(P \vee Q) = max(\mu_A(x), \mu_B(x)) \tag{2.5}$$

Fuzzy Set Theory

The uniqueness of a fuzzy set is that an object can belong to two different sets with a different degree, called membership value. The membership value is calculated through a membership function. The results of a membership function range from 0 (complete exclusion) to 1 (complete membership) instead of 0, 1 in the classical set theory.

A fuzzy set A can be uniquely defined by a membership function of the form:

$$\mu_A : U \rightarrow [0, 1] \tag{2.6}$$

where U is the universe of discourse and the μ_A declares the membership degree of x in the fuzzy set A. Any kind of function can be used to define the membership function. It is obvious that the fuzzy set theory extends the classical set theory in which an object has two options; to belong to the set A or not. Some commonly used in the literature membership functions are the as piece-wise linear functions, Gaussian distribution function, sigmoid curve, quadratic or cubic polynomial curves. In this dissertation, in most of the cases, the triangular form is used due to the symmetry that provides.

In the next lines, the properties and the operations of the fuzzy sets are described in detail. Beginning with some of the properties the *support* $S(A)$ of a fuzzy set A is a crisp set with all elements that have a membership value greater than zero and belongs to fuzzy set A. More formally:

$$S(A) = \{x \in U | \mu(x) > 0\} \tag{2.7}$$

The *height* hgt(A) of a fuzzy set A is the least upper bound of the membership values and it is defined as

$$hgt(A) = sup_{x \in U} \mu_A(x) \tag{2.8}$$

The fuzzy set is called *normal* if the height of the fuzzy set is equal to 1, otherwise it is called *subnormal*.

The α-cut of a fuzzy set is the most important property. The α belongs to the range [0,1] and the α-cut of a fuzzy set A, denoted by A_α , is the crisp set defined as follows:

$$A_\alpha = \{x \in U | \mu_A(x) \geq \alpha\} \tag{2.9}$$

As in classical set theory, in fuzzy set theory operations such as *intersection*, *complement* and *union* exist.

The *intersection* and the *union* of two fuzzy sets A and B is defined as follows:

$$\mu_c(x) = f(\mu_A(x), \mu_B(x)) \tag{2.10}$$

The main difference is that the *intersection* follows the operation $C = A \cap B$ and $f(\alpha, 1) = \alpha$ while the *union* follows the $C = A \cup B$ and $f(\alpha, 0) = \alpha$

The *complement* of a fuzzy set A is given by the following equation:

$$\mu_{\overline{A}(x)} = f(\mu_A(x)) \tag{2.11}$$

The operators used to compute the *intersection*, the *union* and the *complement* of fuzzy sets are $\mu_c(x) = min(\mu_A(x), \mu_B(x))$, $\mu_c(x) = max(1 - \mu_A(x), \mu_B(x))$ and $\mu_{\overline{A}(x)} = 1 - \mu_A(x)$ respectively.

2.7 Overview of Clustering Algorithms

As described in sub chapter 2.3, the last decades the amount of data is getting larger and larger. Data repositories can store an enormous number of records associated with information from various sectors such as economics, science, energy. One way to deal with this big amount of data in order to analyze them and mining the information is to classify them into a set of different classes with meanings that vary. The clustering algorithms belong to the wide area of machine learning and it is a standard procedure in data analysis. In fact, it is a branch of Artificial Intelligence (AI) that uses algorithms to find patterns in data. There exists two types of machine learning; supervised and unsupervised learning. The clustering is designed to group objects where objects in the same cluster are as similar as possible and objects in different clusters are as dissimilar as possible. Clustering methods are applied in many fields, such as psychology, energy, economics and pattern recognition. The features that have to be determined in a clustering procedure are the following:

- the number of clusters
- the size and the shape of clusters
- the density of the clusters
- the positions of the clusters

The data in clustering procedure is called instances and each instance consists of attributes. In supervised learning, each instance can be represented as the tuple (x, y), where x is a set of attributes and y is the target attribute. The basic difference between the supervised and the unsupervised learning is that in the latter there is no need of target attribute or external classification. Since there is no a prior knowledge, its aim is to find similarities among the instances within the set of data. A simple example of a clustering for a set of 2-D data points is depicted in Figure 2.8. The number of clusters/ groups depends on biases related to the data instances. Theoretically, the number of clusters can b as many as the data instances, although that cancels the primary aim of clustering (Aggarwal and Reddy, 2014).

In this dissertation, there is no need of any external information for the clustering of the energy data, so in the next lines, an overview of the most popular unsupervised machine learning clustering algorithms is presented. The choice of a clustering algorithm determines the setting of the parameters. Based on the parameters the algorithms are divided into (a) hierarchical clustering algorithms and (b) partitioning clustering algorithms. For each type, in this dissertation only the ones that are utilized are presented.

2.7.1 Partitioning Algorithms

Partitioning clustering methods partition a set of data points into clusters where every pair of the data points is either distinct or not. Partitioning clustering begins with an initial clustering which is iteratively improved until an error criterion function is minimized. The initial clusters can be either random or the clusters from a pre-processing clustering. In the next lines, the algorithms and a list of some clustering validity measures utilized in the applications presented in Chapter 5 are described.

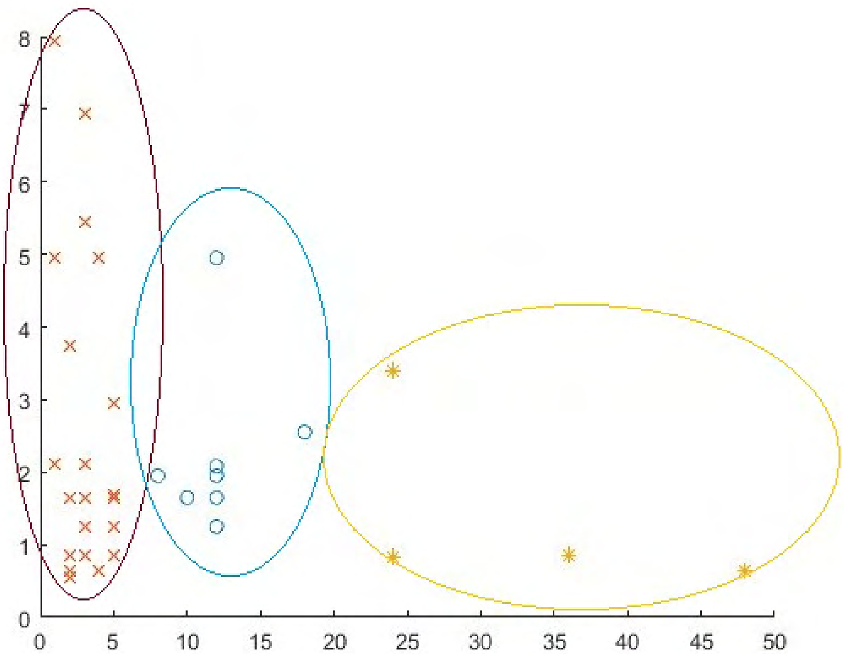

FIGURE 2.8: Example of clustering a set of 2-D data points.

k-means

k-means type algorithms (Celebi, Kingravi, and Vela, 2013) are the most used algorithms in data mining applications. The main principle of a partitioning algorithm is to partition a data set into several subsets, i.e., data clusters, in order to achieve compactness within each cluster and low similarity between different clusters. Regarding the k-means algorithm, it consists of two phases: the first one is associated with the definition of the k initial centroids; each initial centroid is selected randomly. The second phase is to assign each data point of the data set to the nearest centroid. The Euclidean distance is utilized to calculate the distance between the centroid of each cluster and the data points. Actually, the closer the distance of one data point to a centroid, the larger possibility this data point to be assigned in that cluster has. An error criterion function is used in order to evaluate the accuracy and the performance of the clustering. The error criterion function has to be minimized and it is expressed as follows:

$$E = \sum_{i=1}^{k} \sum_{p \in X_i} \parallel p - c_i \parallel^2 \tag{2.12}$$

where p is a data point from the data set and c_i is a cluster centroid. The steps of the k-means algorithm are the following:

1. Choose randomly k data points from the data set as initial centroids.
2. Repeat the following steps
   (a) Assign each data point to the closest centroid based on the Euclidean distance.
   (b) Recompute the new centroids of each cluster until the convergence criterion is met. Recalculation of the centroid is based on the formula below:

$$c_i = \frac{1}{C_i \sum_{j=1}^{C_i} x_j} \tag{2.13}$$

where c_i is the new centroid, C_i is the number of data points in cluster i and x_j is a data point that belongs to cluster i.

The algorithm runs iteratively until the criterion function in equation 2.12 converges to a single value. In other words, the method converges when there are no more differences between the centroids between two sequentially iterations. The major drawback of the k-means algorithm is the random selection of the initial centroids. Therefore, for different runs, it produces different clustering results and the clusters may be unbalanced. Moreover, the computation time is proportional to the number of data points, the number of clusters and the number of iterations. Actually, its computation time is O(N*k*n); N is the number of data points, k is the number of clusters and n is the number of the algorithm's iterations. So, the number of clusters which is defined by the user, and the initial centroids impact both the efficiency of the k-means and the accuracy of the clustering results.

k-means++

The k-means++ algorithm (Arthur and Vassilvitskii, 2007) is an extension of the standard k-means algorithm which tries to optimize the selection of the initial centroids. The initial centroids are chosen based on a probability that combines the distance of the selected data point as a center. More specifically, the steps for selecting the initial cluster centers are the following:

1. Uniformly random selection of a cluster centroid among the data points
2. For each data point x, compute the distance $d(x)$ between x and the nearest center that has been selected in the previous iteration
3. Selection of a new data point as the next centroid with probability proportional to $d(x)^2$
4. Repeat steps 2 and 3 until k cluster centroids have been selected

Once the above procedure ends, the k cluster centroids are used as initial centroids for the standard k-means algorithm. A drawback of k-means++ algorithm is that the selection of the next cluster centroid depends on the current set of centroids.

Fuzzy c-means

The Fuzzy c-means clustering is very similar to the k-means algorithm. The main difference between these two algorithms is that using the fuzzy c-means algorithm each data point may belong to more than one clusters with a different degree. Fuzzy c-means is usually used in pattern recognition applications (Sanchez-Gomez et al., 2016).

2.7.2 Hierarchical Algorithm

The hierarchical clustering algorithm (Sadaaki, 2012) clusters the data at the same time creating in that way a cluster tree. Figure 2.9 shows such a tree. The tree is actually a hierarchy where the clusters are created by joining clusters at each level.

The advantage of hierarchical clustering algorithm over the other clustering algorithms is that the data is kept without any changes. The hierarchical clustering runs as follows:

1. In the beginning, each data point corresponds to a cluster
2. The most similar pair of clusters is merged into one cluster based on the similarity criterion
3. Find the similarities between the new cluster and the old ones
4. Repeat step 2 and 3 until the cutting criterion is met

The grouping of the data points is based on linkage criteria such as average distance, shortage distance, centroid distance and Ward distance.

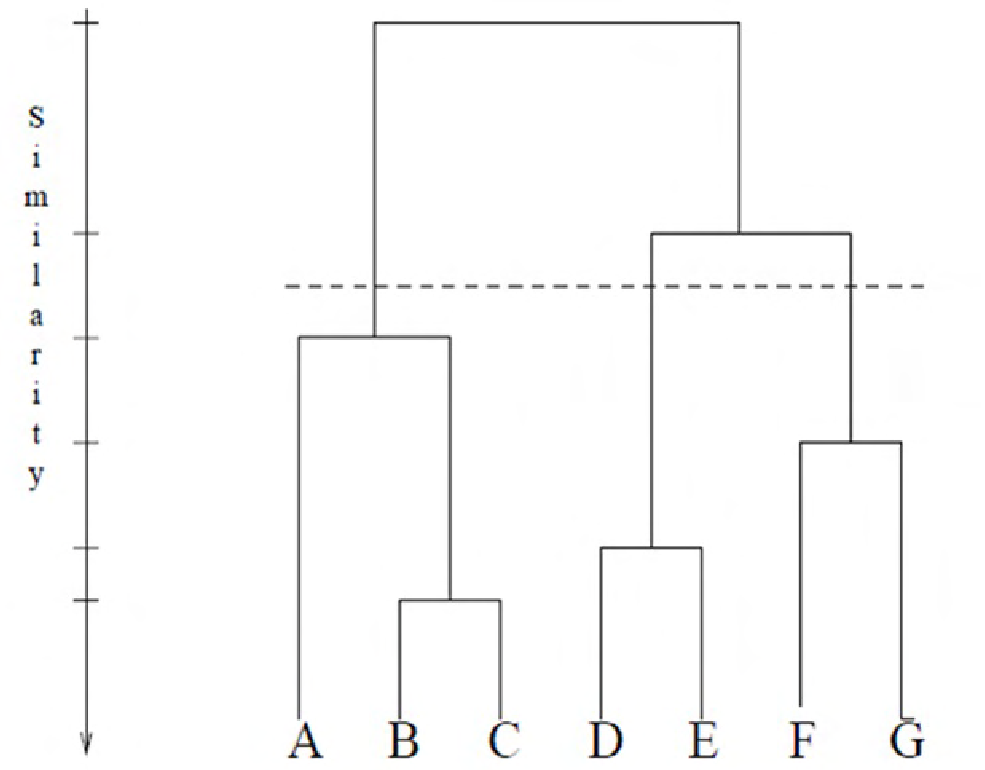

FIGURE 2.9: A tree obtained using Hierarchical Algorithm (Jain, Murty, and Flynn, 1999).

2.7.3 Clustering Evaluation Measures

One of the most significant issues in proving the validity of a clustering algorithm is the use of clustering evaluation measures, which evaluate the degree to which the clustering results are acceptable or not. There are two types of clustering evaluation measures: the external and the internal. The main difference between those two types is the use of external information for the validation of the clustering results. On one hand, the external validation measures are used to define the right number of clusters and to choose the more efficient clustering algorithm for the available data set. On the other hand, internal validation measures are also applied to distinguishing the best clustering algorithm as well as defining the best number of clusters; no additional other than the available information is needed (Wong, 2015).

In this dissertation, two internal validation measures, namely, the **compactness** and **separation** (Rendon et al., 2011) are studied. With regard to the compactness of the cluster, there are a few metrics such as the variance, and the distance of a data point from a center. In fact, these metrics measure how much related the data points in the same cluster are. Furthermore, the separation criterion measures how different-distinct the clusters are, comparing each cluster with another one considering either the distance between the cluster centers or the minimum distance between the data points of the different clusters. Some of the validation measures consider both the aforementioned criteria are presented in the next lines (Rendon et al., 2011), (Hassani and Seidl, 2017).

The **Root-mean-square standard deviation** (RMSSTD) checks how much homogeneous the data points on the clusters are by using the square root of the variance of the data points' attributes. **R-squared** (RS) measures the difference between the clusters and its value is equal to the ratio of the sum of squares between the clusters to the total sum of squares between all the data points belonging to the data set. The **Calinski-Harabasz** (CH) index is calculated based on the sum of squared errors between clusters, i.e. the squared difference of all the data points and each of the cluster centers.

The **Silhouette** (S) index validates the performance of the clustering algorithm. First, the average distance of data points from a center in the same cluster is calculated and then is compared to the average distance between the data points of the other clusters. This factor can be used in order to determine how compact the clusters are. In fact, the S index measures the relative distance of each data point in each cluster to the data points in the other clusters. This has been recognized as a way to determine the optimal number of clusters for a specific data set.

Further, the **Davies-Bouldin** (DB) index is used to identify the clustering which provides the best clustering results. The lower the value of the DB index, the better clustering results the algorithm provides. The DB index is determined as follows: for each cluster, the maximum value of a factor associated with the similarities between this cluster and the other ones is selected. The value of the DB index is the average of the aforementioned factor of all the clusters. The DB gives the average of each cluster's similarity measurement with its most similar cluster (see equation 2.14).

$$DBI = \frac{1}{K}\sum_{k=1}^{K} max\left\{\frac{d\left(x_i, C_k\right) + d(x_j, C_k)}{d(R)}\right\} \tag{2.14}$$

2.8 Overview of Distributed Energy Resources

The objective of this section is to provide an overview of the DES that is widely used the last decades for satisfying the energy demand. Traditional energy technologies have not analyzed in this dissertation since they are already studied technologies. More specifically, the technologies and the features of the solar panels, the wind turbines, and the batteries; DES, which are used in the applications, are presented in Chapters 4, 5, 6.

2.8.1 Solar Energy

Solar energy based technologies provide the ability to provide clean, reliable and efficient energy compared to other forms of energy such as nuclear and biomass.

Some of the solar energy applications (Kalogirou, 2013) are the following:

- heat and cool buildings (both active and passive),
- heat water for domestic and industrial uses,
- heat swimming pools,
- power refrigerators,
- operate engines and irrigation pumps,
- desalinate water for drinking purposes,
- generate electricity,
- for chemistry applications

The solar energy can be applied to three different applications for supplying electricity at the residencies. The solar energy can be converted into different forms, such as heat and electricity. In Figure 2.10, the three different ways of supplying energy are depicted.

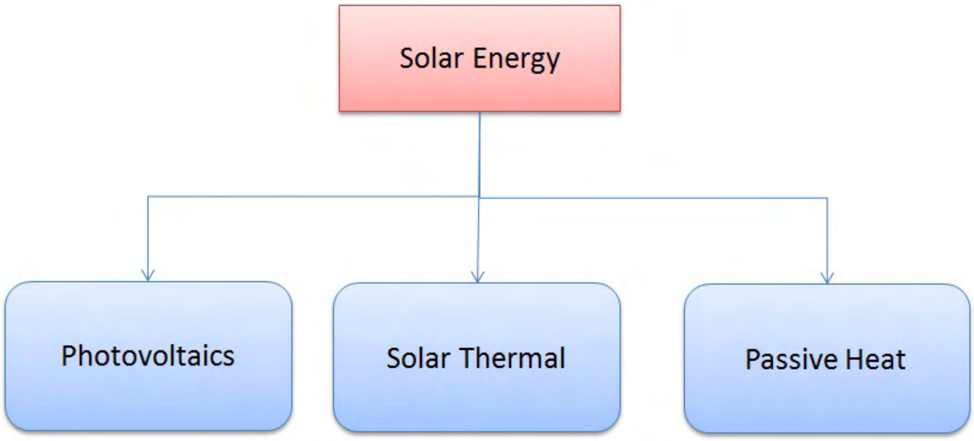

FIGURE 2.10: Types of solar energy use.

More specifically, the photovoltaic panels can generate electricity that can be utilized from domestic appliances. The solar thermal type of solar energy provides heat for hot water and generally for the applications that make use of heating systems. Finally, the passive heat is used to provide heat via the design of a building.

Photovoltaic Energy

The photovoltaic energy is the conversion of the sunlight into electricity. A photovoltaic (PV) cell, called a solar cell, is the technology that is used to convert solar energy into electricity. A photovoltaic cell is usually made from silicon alloys. A set of photovoltaic cells provides a photovoltaic system. It is commonly accepted that PV is one of the future sources of energy concerning the cost reduction of electricity generation. An individual photovoltaic cell can have different sizes from about 1 cm to 10 cm across. It is worth mentioning at that point that just one cell can produce only 1 or two watts of power which are not enough for the most of the applications.

So, for increasing the power output of each photovoltaic system, the panels are connected into a packaged weather-tight module. Modules are PV panels connected in series or in series to form a string. The strings are connected in parallel and a PV system is formed. The number of the modules provides information about the amount of power output (Kannan and Vakeesan, 2016).

Depending on the method of fabrication and the used material, there are different types of cells; **Crystalline** (Monocrystalline and Polycrystalline), **Amorphous Silicon** and other Thin Film cells e.g. Copper Indium di Selenide (CIS) and the second generation thin film technology, **Crystalline Silicon on Glass (CSG)** (Razykov et al., 2011). According to the IEEE standard 929-2000, PV systems are also divided into three categories considering the power output: 1) systems rated at 10 kW or less, 2) systems rated between 10 kW and 500 kW, and 3) systems rated above 500 kW. Moreover, based on the way the PV system is connected, the PV systems are classified into two more categories: grid-connected systems and islanded or stand-alone systems (Cupertino et al., 2012). In figures 2.11 and 2.12 the diagrams of a grid-connected PV system and a stand-alone PV system are depicted respectively.

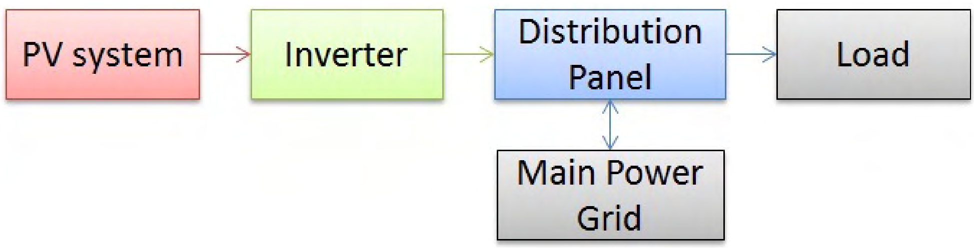

FIGURE 2.11: Diagram of grid-connected PV system (Cupertino et al., 2012)

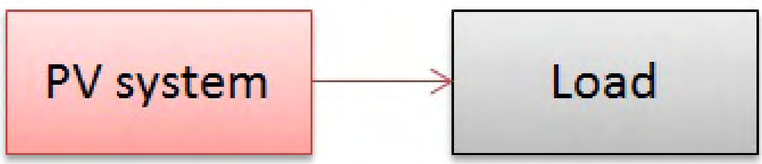

FIGURE 2.12: Diagram of stand-alone PV system (Cupertino et al., 2012).

Climate conditions such as cloud, fogs and very high temperature have a significant effect on the solar energy and, as a consequence, the performance of PV panels. The efficiency of the PV panels ranges from 10 to 20 %.

For converting the DC power produced by the PV systems, Power Conditioning Units (PCUs) are necessary. The PCUs are utilized to control and convert the DC power to high-quality AC power before being injected into the power grid. There are two types of PV systems considering the number of stages of the power processing procedure. In single-stage PV systems, all the control tasks are performed by an inverter. On the other hand, in a two-stage PV system, an inverter follows a DC-DC converter and both of them are responsible for the control tasks. The two-stage systems overwhelm the one-stage systems due to their higher flexibility (Bahaj and James, 2016).

The different way the PV modules and the PCUs are connected provides different topologies to the PV systems. The most common topologies are the (a) Centralized,

(b) Master-slave, (c) String, (d) Team concept, (e) Multi String and (f) Modular (Bahaj and James, 2016).

The **Centralized** topology is one of the well-established topologies and it is the one that is utilized for PV systems. A single inverter is used in the PV system and it is usually used for large PV systems with a high power output of up to several megawatts. One of the basic advantages of this topology is its low cost and the ease maintenance of the single inverter. On the other hand, the low reliability of this topology is one of its main disadvantages.

The **Master Slave** topology focuses on improving the low reliability provided by the Centralized topology. To come up with that, it uses a set of parallel inverters such that if there is any problem with one of the inverters, the other inverters provide the power to the grid. This technique by setting a number of inverters to improve the reliability increases the performance of this topology but the cost is higher compared to the Centralized topology. In similar with the Centralized topology, the power loss and the partial shading remain a problem without an efficient solution. In the **string** topology, each string is connected to one inverter; hence, the reliability of the system is increased. Moreover, the flexibility is a new and basic feature of this topology since new strings can be added any time to the PV system increasing in that way the power output of the system.

The **Team concept** technology is a combination of the **String** and the **Master-Slave** technology. In the **Multi-string** topology, a DC-DC converter, and a String are directly connected and all the DC-DC converters are connected to a single inverter via a DC bus. In that way, the maximum power point and voltage amplification are tracked. In particular, this topology combines the features of both the string and centralized topologies. However, the reliability of PV system is decreased compared to the reliability that the string topology provides. Moreover, the losses are accumulated to the losses of the system.

The **Module** is the most recent topology. Some of the main advantages are the fewer power losses, the flexibility and the improved monitoring at each module failure. On the other hand, it has applications only to small PV systems.

The intermittent and the stochastic nature of the output power of the PV systems may impose special challenges on the reliable and safe operation of the power grid. Grid-connected PV systems, as the ones used in this dissertation, focus on enhancing the performance of the power grid by improving the voltage levels of the power grid and reducing the power losses. However, the integration of the PV systems in the power grid may lead to various problems, such as power and voltage disturbances, overloading of the feeder and even underloading.

2.8.2 Wind Energy

Besides the energy that can be produced by the sun, the wind belongs to the RESs as well. Occasionally, it is utilized to pump water and move ships. With the evolution in the electricity sector and the power grid, wind energy can be applied in new applications by providing electric power using a clean and efficient way in large scale and capacities. Nowadays, there is a big variety of wind-powered generation systems beginning from small residential to utility sizes (Kaplan, 2015). The wind is transformed into a renewable and free source of energy. Wind turbines are a source of energy that exploits this renewable wind power to produce electricity.

Wind Turbines

Wind turbines are used by the humans over the last thousand years. Their mechanism is very simple; the wind turns the blades that are connected to a generator, and in turn, the generator produces electricity. More specifically, wind turbines can produce electrical power by converting the kinetic energy of wind into mechanical power with one more conversion into electrical power by means of an electrical generator. The conversion of the voltage and the frequency provided by the different types of wind turbines are made by the power electronic converters to the level that the various utilities desire. After that, the power is transferred to a properly configured transformer which converts the voltage to the desired level (Blaabjerg and Ma, 2013). In Figure 2.13, the block diagram of a typical wind turbine and the way that it is connected to the grid is depicted.

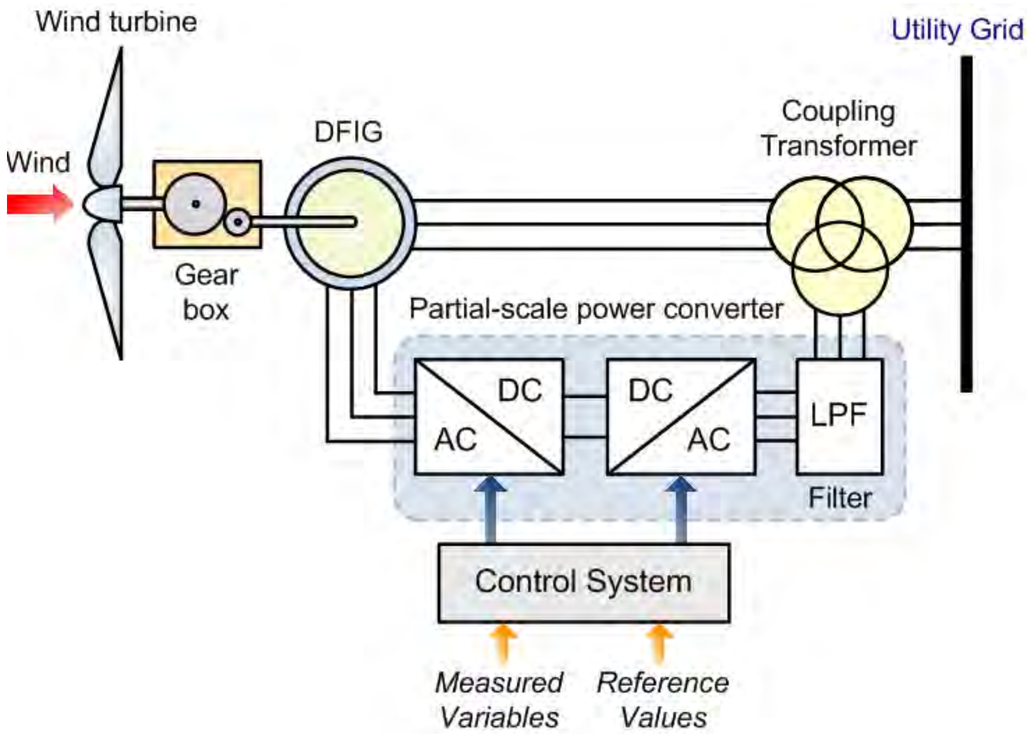

FIGURE 2.13: Diagram of a wind turbine (Gustavo and Gimenez, 2011).

Wind turbines can be separated into two types determined by which way the turbine spins; the horizontal axis wind turbine and the vertical axis wind turbine. The latter are not mature enough to be used in utility scale applications. The main difference between these two technologies is the way the turbine blades rotate by having as origin the ground. In the former, the axis of the turbines' blades rotation is parallel to the ground. On the latter, the axis rotation is vertical to the ground.

On the other hand, based on the different technologies that are used, the wind turbines are classified into four basic technologies as follow (Rajveer Mittal and D.K.Jain, 2010):

1. Fixed speed wind turbines
2. Full converter wind turbines
3. Variable-slip wind turbines

4. Doubly-fed induction generator

The majority of installed wind turbines are connected to the bulk power system. They are rarely installed in a single way, instead, they are installed in sets of tens or hundreds of similar turbines, commonly called wind farms.

It is worth mentioning at that point that, the wind turbines and the wind farms as well may affect the stability of the power system. This depends on the voltage at which the connection system is interconnected with the power system which in turn it depends on the size of the wind turbine and the location. Moreover, one of the main concerns besides the voltage control is the reactive power compensation. However, the wind turbines should be located not far away from load centers. On the other hand, wind turbines can be used as distributed generators, close to the load centers, but in such a case quality is important. In this dissertation, the wind turbines are used as distributed energy providers.

2.8.3 Batteries

Energy storage devices are used to improve the availability of the power generated by the RESs. So, energy storage technologies are needed for the successful incorporation and integration of RES in the smart grid. Among them, the battery is the most commonly used energy storage device due to its efficiency and the direct connection with most of the RES. Moreover, batteries provide and maintain the stability and the flexibility in the operation of the smart grid. The battery is a non-polluted device, its installation is very easy and quick. Generally speaking, it provides an easy way to store energy anytime. The use of the batteries into the smart grid convert the non-dispatchable sources into a dispatchable energy source by introducing new applications of electricity storage (Xiao et al., 2014).

Batteries and other energy storage devices can convert electrical energy from a power network into a form that can be stored for converting back to electrical energy when it is needed. Then, this type of process can offer an opportunity for electricity to be stored at times of either low demand, low generation cost. Moreover, they can be utilized from intermittent energy sources at times of high demand, high generation cost or when no other generation means is available (Bahramirad and Daneshi, 2012), (Xiao et al., 2014).

Technologies

Based on the different chemicals and characteristics that are used in order to produce electrical energy, different types of batteries have been developed. In general, a battery consists of one or more electrochemical cells. At the same time, each cell consists of solid electrolyte together with a positive electrode (anode) and a negative electrode (cathode) (Zheng et al., 2014). Nowadays, there exist a plethora of batteries with their advantages and their disadvantages. Among them, there are specific types that fit better in the smart grid applications and for the smoother integration of them into the power grid. Some of the commonly used are presented in detail below. In the applications of this dissertation, the Lithium-ion batteries are utilized.

1. **Lithium-ion Batteries**: Lithium-ion or Li-ion batteries, as one of the most advanced rechargeable batteries, have been attracting much attention in the last decades. They are commonly used for electronic devices, more usually in cell phones and laptop computers. Li-ion batteries are highly advanced comparing to the other batteries with respect to gravimetric and volumetric energy.

In Figure 2.14, the comparison of various rechargeable batteries is depicted. It is notable that the Li-ion batteries overwhelm over the other batteries. Li-ion batteries are flexible, which means that can be formulated to a variety of shapes and sizes, so as to efficiently fit the electronic devices. A major advantage of Li-ion batteries is its scalability; it can be adapted to any voltage, power, and energy requirement. On one hand, due to its high quality of electronics that requires, it makes its technology complex but on the other hand, it provides high-quality management, control, and analysis in smart grid applications. Some of the basic characteristics of Li-ion batteries are the following:

- High efficiency
- High energy density
- Long cycle life
- Versatility: electrodes can be optimized for energy patterns

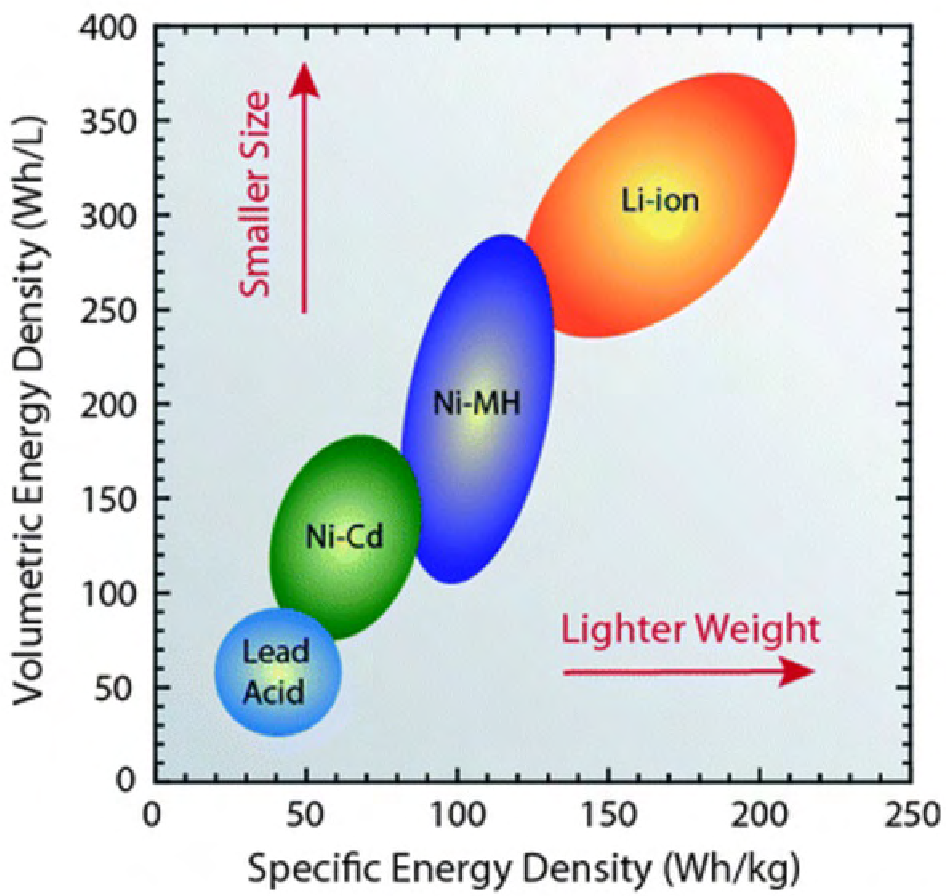

FIGURE 2.14: Energy Density Comparison of Size & Weight.

Source: http://www.epectec.com/batteries/cell-comparison.html

2. **Lead-acid Batteries**: Lead acid batteries are the most economical batteries with a short life if the battery is discharged below 30%. This leads to the reduction of energy density. They are commonly installed in interruptible power supply systems as well as in distributed power systems. Some of the basic characteristics of Lead-acid batteries are the following:

- Capacity (1 Ah up to 16,000 Ah)
- Energy Efficiency higher than 85%
- Cycle life up to 2,000 cycles
- Calendar life up to 20 years
- Operating temperature from -30 Celsius to +50 Celsius

3. **Nickel-based batteries**: Nickel-based batteries (Nickel-Cadmium, Nickel-Metal hydride, Nickel-Hydrogen, and Nickel-Zinc) are the second most used energy storage battery after Lead-based batteries. However, the Nickel-cadmium and Nickel-metal hydride are the most applied batteries for energy storage purposes. They are commonly used when the energy has to be stored under specific conditions such as fast charging, large scale energy storage, high life cycles. They serve special markets where energy must be stored in extreme climate or cycling or fast charging conditions. Some of the basic characteristics of the Nickel-based batteries are the following:

    - Capacity (0.5 Ah – 2,000 Ah)
    - Normal energy density
    - Cycle life up to 3,000 cycles
    - Calendar life up to 25 years
    - Operating temperature from -40 Celsius to +60 Celsius

Characteristics

For selecting the best technology for a specific application, the analysis of specific fundamental characteristics is required. Some of them are described in detail below (Leou, 2012):

- **Storage Capacity**: Storage capacity is the available energy into the battery after the charging mode. Actually, it is defined on the basis of total stored energy. The usable energy is limited by the Depth of Discharge (DoD) and it corresponds to the limit of discharge depth.

- **Depth of Discharge**: In many types of batteries, the full energy stored in the battery cannot be fully discharged without causing issues to the operation of the battery. The DoD of a battery is the part of the power that can be withdrawn from the battery. The batteries, used in renewable energy applications, are classified based on their capacity. However, it is worth mentioning at that point that the actual energy that can be used is less than the actual capacity. This is a way to increase the lifetime of the battery because if the full capacity of the battery is used its lifetime is reduced. For example, a lead-acid battery operated with a DoD of 20% has a lifetime of 2800 charging/discharging cycles, while with a DOD of 80% the lifetime is about 500 cycles (Figure 2.15).

- **Efficiency**: The efficiency of a battery is the ratio between the released energy and the stored energy. This definition is very simple; the efficiency can be discussed with the reference to rectifier efficiency, inverter efficiency, storage efficiency and storage self-discharge. For example, the efficiency of a lead-acid battery is 85%, 95% for Li-ion and 75% for Ni-Cd. With respect to other efficiencies, rectifier, inverter and self-discharge efficiency a lead acid battery has 95%, the Li-ion has 98% and the Ni-Cd has 0%.

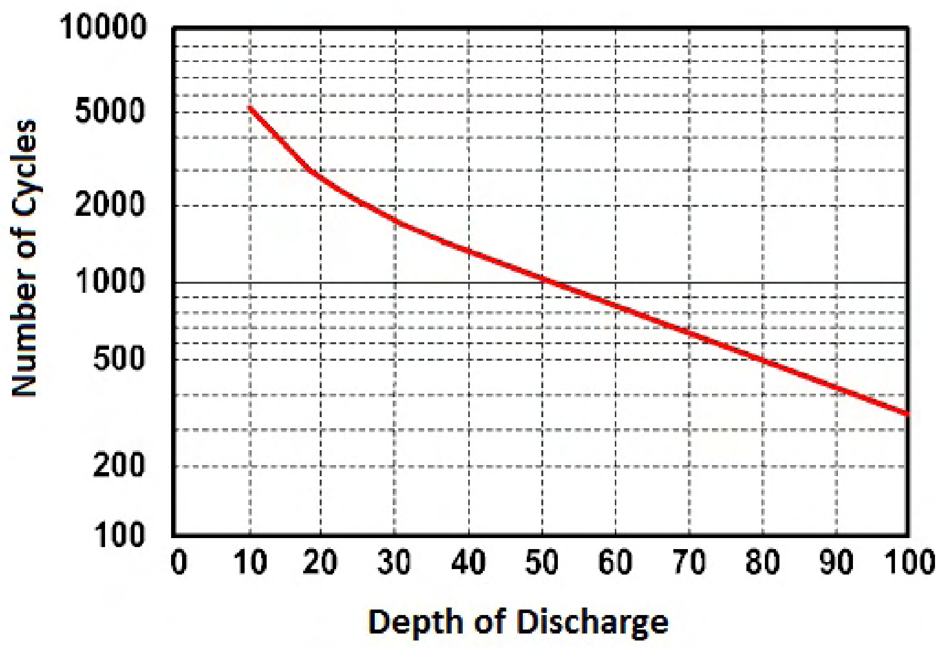

FIGURE 2.15: Depth of discharge for lead-acid battery.

Source: NORTHERN ARIZONA WIND & SUN

- **Discharge Time**: As discharge time is defined the max value of the power discharge of a battery. The discharge time depends on the DoD, the constant power, and the operation conditions and it is commonly referred as C-rate or E-rate. A C-rate or E-rate is a measure of the rate at which a battery is discharged according to its maximum capacity. A 1C rate means that the battery will be discharged within one hour. For example, a battery with a capacity of 100 Ah, 1C rate is equal to a discharge current of 100 A. A 4C rate for this battery would be 400 A, and a C/2 rate would be 40 A.

2.8.4 Distribution System

The distribution system is the part of the power grid that provides the electricity from the generation to the end-users. It starts from a distribution substation that is fed by a high-voltage transmission line or one or more sub-transmission lines. Each distribution sub-station can serve one or more primary distribution feeders.

Distribution Substation

A distribution substation transfers power from the transmission system to the distribution system. It is more expensive to connect the end-users directly with the transmission network to directly connect electricity consumers to the main transmission network. The distribution substation reduces the high voltage to medium or low voltage which is suitable with the distribution grid. The main components of a distribution substation are the following (Kersting, 2012):

- high and low side switch
- Transformer for reducing voltage from high to medium-low

- Voltage Regulator
- Protection
- Meter devices

Distribution Feeders

A common distribution system consists of one or more distribution feeders. The distribution feeder is actually a small distribution grid. The distribution grid in most of the cases except some rare cases is radial. The radial topology provides a grid where there is only one path power to flow from the distribution substation to the end-user consumers (Kersting, 2012). In Figure 2.16, the IEEE-13 test feeder (Kersting and Shirek, 2012) is depicted. In all the applications in this dissertation, besides the IEEE-13 test feeder, the IEEE-37 and the IEEE-123 test feeders are utilized. The main components of a distribution feeder are the following:

- Three-phase primary feeder
- Three-phase, two-phase, and single-phase laterals
- voltage regulators
- Transformers and distribution transformers
- Capacitors
- Three-phase, two-phase, and single-phase loads

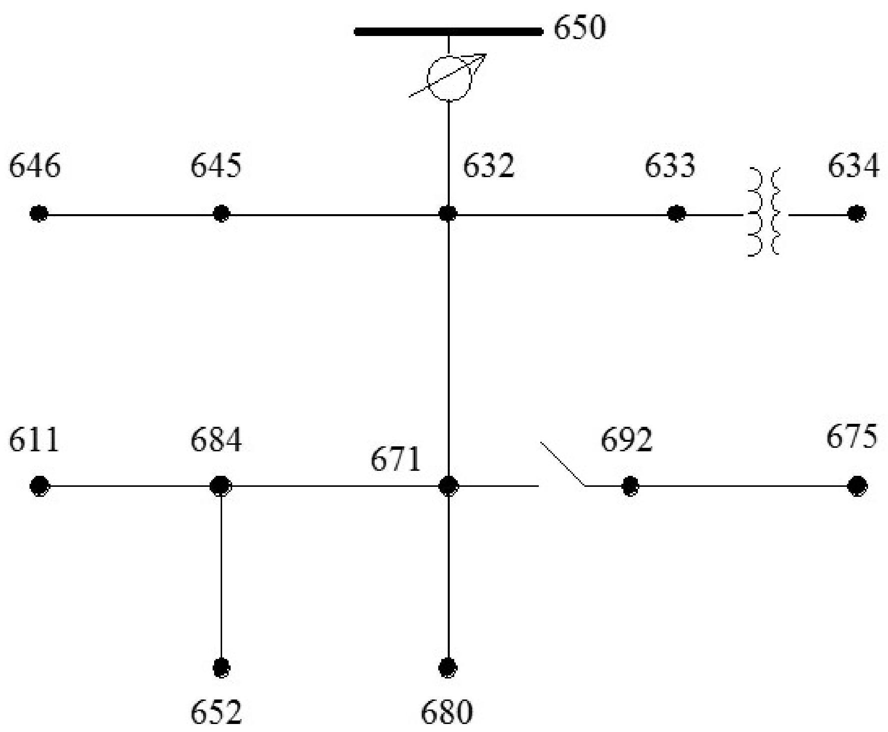

FIGURE 2.16: Architecture of IEEE-13 test feeder (Kersting and Shirek, 2012).

2.9 Power Theory

2.9.1 Real Power

The power that is consumed or utilized in an Alternating Current (AC) Circuit is called Real power or Active Power or Real power. It is worth mentioning at that point that power and energy are two totally different concepts. Power is measured in Watt (W) and indicates the electricity production instantly. A unit has a limited power production called unit capacity and it is measured in Watt as well. On the other hand, if a 1 MW unit works at its full capacity for one hour, then it will produce 1 MWh energy during this hour. Real power is an easy concept for someone to understand it but on the other hand, reactive power is a little bit more confusing concept.

2.9.2 Reactive Power

The power quality depends on the frequency and the voltage disturbances as well as the wave in the wave in the electrical power grid. The control of voltage and reactive power is crucial for the safe and reliable operation of power systems. Reactive power describes the background energy movement in an AC system arising from the production of electric and magnetic fields (R. Fetea, 2012). Reactive power is required to maintain the voltage within the appropriate limits in order to ensure the reliability of the power system and deliver the demanded active power the through transmission lines to the end-users. Voltage control in an electrical power system is important for proper operation for electrical power equipment to prevent damage such as overheating of generators and motors, to reduce transmission losses and to maintain the ability of the system to withstand and prevent voltage collapse. Decreasing reactive power causing voltage to fall while increasing it causing voltage to rise. A voltage collapse may have occurred when the system tries to serve much more load than the voltage can support. The concept of active power is more intimate than the reactive power, with the latter playing a vital role as ancillary service in the operation of power systems.

A factor that is strongly related with the reactive power is the **power factor** (Akagi, Watanabe, and Aredes, 2017). More specifically, power factor measures the efficiency of an AC power system. Power factor is the ratio of the real power to the apparent power (pf = W/VA). The apparent power is the power supplied to the circuit. The power factor is 1 when the voltage and the current are in phase, and is 0 when the current "leads" or "lags" the voltage by 90 degrees. Power factors are indicated as "leading" or "lagging" to show the relation between the voltage and the current. The lagging shows a positive sign which indicates that the load is inductive as the load "consumes" reactive power. On the other hand, the leading shows a negative sign which indicated that the load is capacitive as the load "supplies" reactive power. It is worth mentioning at that point that if a system has lower power factor than another system but the two systems transmitting the same amount of power, the former system will have high circulating currents. These currents will provide high losses and it leads to reducing the overall transmission efficiency.

The operation of each generator is directly related with a capability curve. In Figure 2.17, the capability curve of a cylindrical rotor generator is depicted. The capability curve defines the limits of the machine can operate safely. It is also known as Operating Charts or Capability Charts. The region of operation in which the machine can operate safely is restricted to the following points:

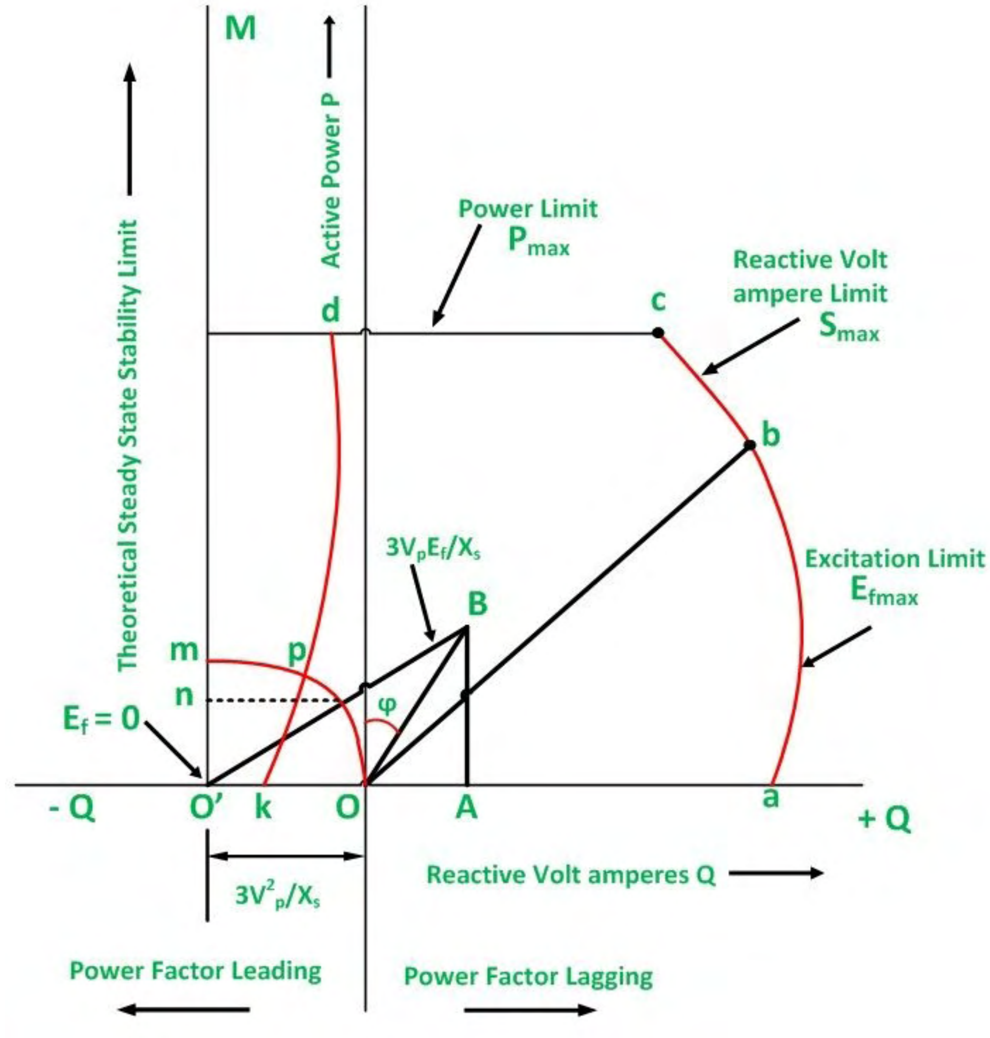

FIGURE 2.17: Cylindrical Rotor Generator Capability Curve.

Source: http://circuitglobe.com/capability-curve-of-a-synchronous-generator.html

- The MVA loading does not have to exceed the generator rating.
- The MW loading does not have to exceed the rating of the prime mover.
- The field current has to have specific values determined by the heating of the field.
- The load angle δ must be between -90 degrees and +90 degrees.

It has to be noted that if a generator is set to provide more reactive power instead of active power, it has to decrease its capabilities to produce the max capacity of active power by reducing the value of the power factor. This may have effects on the operation cost of the generator.

Sources of reactive power

The generated reactive power is equal to the demanded power. Some of the loads in the power grid need inductive reactive power to operate. For that reason, the demand for reactive power is much more than the reactive power that the generators can provide. So, there is a need for devices that have to be connected to the power grid in order to provide more reactive power. These devices can be the capacitor banks, synchronous motors, and power electronic sources of reactive power. The cooperation of these devices with linear circuits lead to the decrease of the supplying current.

Chapter 3

Software Platform

3.1 GridLAB-D Simulation Platform

In this dissertation, for experimentation and simulation purposes, the GridLAB-D[1] simulation platform is utilized. It is an agent based platform written in C++ for simulating power distribution grids using detailed models. GridLAB-D is an object oriented software and it provides the analysis capabilities needed in this dissertation to study the interactions of different components and models in the simulations. It also allows the users to add, modify or extract their own modules. Moreover, GridLAB-D is able to provide the simultaneous state of independent objects, each of which is associated by multiple differential equations. Besides, it can integrate new modules and communicate with third-party systems. It is developed by the Pacific Northwest National Laboratory (PNNL) in collaboration with the industry and academia support from the Department of Energy, Office of Electricity Delivery and Energy Reliability. Figure 3.1 shows the architecture of the GridLAB-D. It actual consists of four main, module based, components that describe the most representative distribution system operation, technically as well as financially. They are virtually connected with the **core** module which is the most important module of the GridLAB-D since it is responsible for controlling, managing and monitoring the parameters values of all objects of a simulation model at runtime. The GridLAB-D modules implement classes of objects, which are supported by the functions of this module and have similar characteristics such as the appliances and the generators.

GridLAB-D uses files with the suffix .glm (GridLab-d Model) for modeling, designing and simulating power grids. The .glm file is an internal simulation file and it contains an object-based pseudo-code which has information about the topology and the configuration of the under consideration distribution power grid. A simulation on GridLAB-D starts by giving initial values to the parameters of each object, a starting, and an end date. Moreover, it provides .xml files containing all the information associated with the output results of a simulation. It also provides a real time server with read-write actions on the parameters of particular objects.

Each of the four components (power systems, control systems, buildings, markets) includes the modules which concern various functionalities such as the power flow operation, distributed storage, and generation, load profiles in a residency and the operation of the market framework.

A brief description of the most important modules of GridLAB-D is given below:

- The **powerflow** module (Schneider et al., 2009) provides the model of the distribution system for power flow solutions. Power flow aims at calculating the voltage vector in the entire set of nodes of a distribution grid, that is the basis of further computation of currents, power losses, power etc. Newton–Raphson

[1] http://www.gridlabd.org/

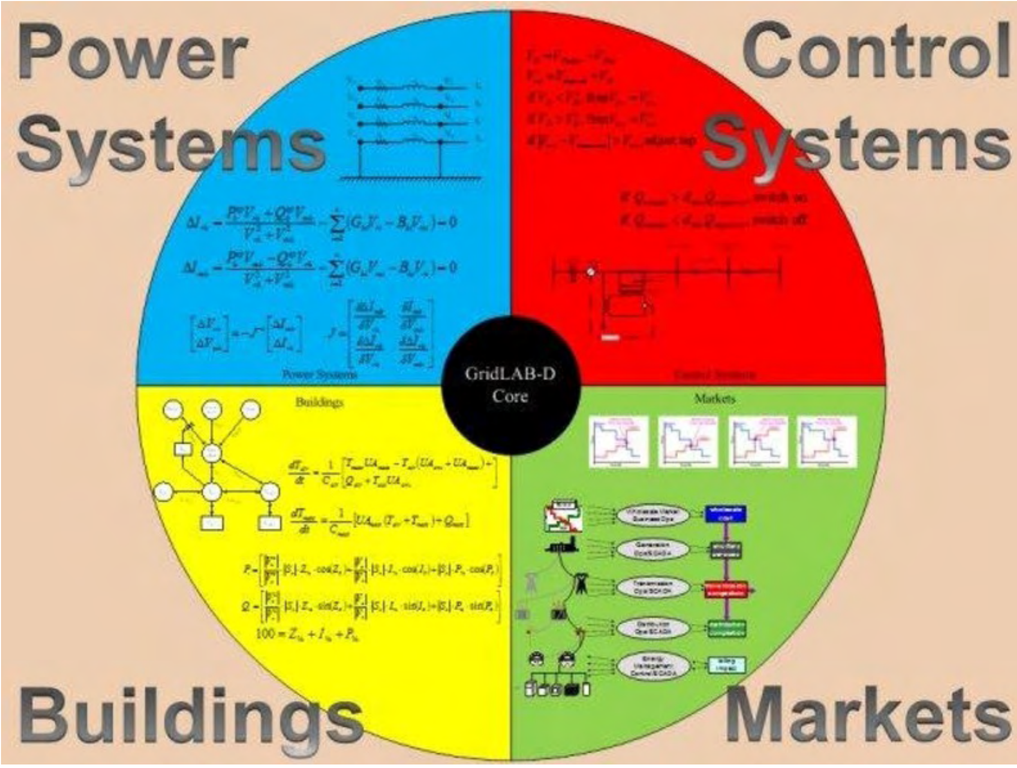

FIGURE 3.1: GridLAB-D Architecture.

Source: https://sourceforge.net/projects/gridlab-d

(NR) is the most widely-used, iterative technique for solving the power flow equations. GridLAB-D uses the NR, proved to be quite applicable either for balanced or unbalanced systems (3-phase distribution systems).

- The **residential** module (T. Taylor, 2008) implements the functionality of a residency's appliances and it provides the important characteristics to simulate the load demand of a residency.
- The **market** module (Chassin, Fuller, and Djilali, 2014) provides information about the functionality and the operation of an electricity market. The market in the implementation of the GridLAB-D is organized as a double auction. A double auction market is a two-way market, where both the producers and the consumers participate in it by submitting their bids (price and quantity).
- The **generators** module (Fuller, 2009) implements electricity generating and storage resources. Batteries, RES and other distributed generation sources are the main components of that module. The connection to the power grid is achieved by using specific objects of the **powerflow** module (*triplex_meters*, *meters*).
- The **climate** module (Tenney, 2008) provides weather files and the current weather data for a distribution model. The weather data is retrieved based on the .tmy file which is chosen by a user.
- The **tape** module (Chassin, 2008) implements objects that can be used to modify the conditions of a model, records the values of the properties of individual objects or of a group of objects.

In the next lines, more details of the most widely used objects of the aforementioned modules are given. The **tape** module is an important module. It consists of objects that are utilized to change and settle limits of a model. *Player* and *shaper* objects can be used to update the parameter values of a model at specific times driven from a file. *Recorder* and *collector* objects are used for collecting information about the parameters values of the objects in a .glm file every specific time period the user selects. Regarding the **powerflow** module, the objects *node*, *triplex_meter* and *meter* are the most significant. More specifically,

- The object *node* represents buses or connections in a distribution model.
- The object *meter* provides a measurement point of power and energy at a specific point in a distribution grid. More specifically, it provides a way to measure how much power and energy have been used by downstream connections and how much current is flowing through the meter at a particular time of a simulation.
- The object *triplex_meter* provides the same functionality as the object *meter* with the main difference that it works for triplex systems instead of three-phase systems.

For the **residential** module, the object *house* is the main class of objects, which can accommodate two types of domestic loads, non-thermostatically controlled appliances (lights, cooking, etc.) and thermostatically controlled appliances such as the HVAC system and the water heater. The *house* object is connected to the distribution grid through the *triplex_meter*.

The most important objects of the **market** module are the objects *controller* and *auction*. The *controller* compares the current clearing price with the average market price and bids the appliance's current demand as a function of the price back into the auction. The *passive_controller* object has the same configuration; it receives pricing signals from auctions but it does not have the capability to bid back. Details of the above-mentioned controllers can be found in (Aliprantis et al., 2010). The price, in which the controllers bid, depends primarily on the average clearing prices of the previous time steps. Regarding the *auction* object, it provides the functionality of a market. Both producers and consumers submit their bids for a specific time period, called market period, into the market framework. Once the market period ends, the bids of both consumers and producers are sorted, from the highest to the lowest price for the consumers and from the lowest to the highest price for the producers. The intersection of the curves of both producers and consumers gives the clearing price and the clearing quantity. In particular, the clearing price is the price, in which the consumers can buy and the producers can sell their energy. The clearing quantity is the available quantity of the market in that price. The clearing price, in which the market offers the clearing quantity, is the highest price that a buyer prefers to pay for energy and the lowest price that a seller prefers to sell its produced energy. The buyers and the sellers that belong to the area that is formed from the supply and the demand curves are included in the market dispatch of the market. Market dispatch is the determination of the producers and the consumers that will take the privileges of the market; buy at the clearing price and sell the clearing quantity at the clearing price.

In Figure 3.2, the most widely used scenarios [2] of how the market calculates the clearing quantity and the clearing price are depicted. In this dissertation, the

[2] http://gridlab-d.shoutwiki.com/wiki/Spec:Market

sellers are the producers and the buyers are the consumers, and more precisely the appliances of a residency. In particular, the left part of the Figure 3.2 depicts the marginal seller case, where the seller offers part of its produced energy. In that case, the clearing price of the market is the bid price of the marginal seller. The clearing quantity is the aggregation of all the buyers' quantities which submit bids with the bid price be higher than the bid price of the marginal seller. The right part of the Figure 3.2 depicts the exact opposite case, known as marginal buyer.

FIGURE 3.2: Marginal Seller Case (left figure), Marginal Buyer Case (right figure).

Source: http://gridlab-d.shoutwiki.com/wiki/Spec:Market

Moreover, the *stub_bidder* object of the **market** module is designed to submit the offers in the market on behalf of the non-thermostatically controlled appliances and the conventional generators. It is characterized by a fixed price and a fixed quantity during a simulation. It is worth mentioning at that point that the thermostatically controlled appliances participate in the market using the *controller* object. Regarding the distributed generators (the associated objects are developed in **generators** module of GridLAB-D), they participate in the market through the *generator_controller*. The price, in that case, is also fixed, but the bidding quantity depends on the size of the distributed generator and the weather conditions.

From the **generators** module, the *diesel_dg* object simulates a synchronous generator while the object *solar* simulates the operation of a solar panel. Similarly, the objects *windturbine_dg* and *battery* simulate the operation of the respectively distributed generator with the capability of having different types of batteries and wind turbines.

3.2 GridLAB-D User Interfaces

GridLAB-D provides only a text-based interface for controlling, modeling and monitoring the operation of a power grid. To the best of our knowledge, it certainly lacks a GUI depicting information about the power grid such as the simulation results of a specific simulation. Moreover, there no exists any friendly GUI which may help a user to demonstrate simulations in GridLAB-D from scratch. In the next lines, a detailed description of recent efforts (Nasiakou, Alamaniotis, and Tsoukalas, 2016a), (Foti et al., 2016) on implementing a GUI of the GridLAB-D is presented.

Occasionally, many efforts have been conducted for providing a GridLAB-D GUI. Song, Gallagher, and Clarke (2012) presented a generation-based approach for developing a GUI based on runtime model technology. In fact, the runtime model has the role of the mean between the system data and the GUI. It also organizes the data in specific types to be easily and efficiently accessible to the GUI which displays the selected by the user objects. The current simulation state of the power grid is available to users by providing information about the key attribute values of the objects and the graphical representation of the objects' connections. Each object of GridLAB-D is depicted with different shapes using the Eclipse GMF, while an object-shape is used to represent the attributes' values. The HTTP API which is provided by GridLAB-D is utilized for retrieving information about the current simulation state of the power grid. This approach aims to the representation of the information about the state of the power grid and not to the simulation results.

Maass et al. (2013) proposed an Electrical Data Recorder (EDR) for power grid analysis due to the fact that the volume of the data from the power grid is large and there is need of detailed information about the system state of a power grid. The proposed recorder captures almost real time (every minute) high-rate, low-voltage time series data, which are stored in a large database, concerning the load and the current on different locations across the power grid. A java-based electrical Simulation Modeling and Visualization Tool (eASiMov) with a graphical interface for GridLAB-D is developed with simulation, modeling, large scale data acquisition, analysis and visualization features. The data obtained from the EDR is visualized through the eASiMov tool. The OpenStreetMaps is utilized to match a grid model to geographical maps. The GraphML which is responsible for transforming the data in order to be accessible by the GridLAB-D plays the role of the "connection" point between the GridLAB-D and the eASiMov.

A simulation tool, called GridMAT, of integrating and combining the capabilities of both the GridLAB-D and Matlab is presented by Faruque and Ahourai (2014). Specifically, the capabilities of the GridMAT are the following:

- it provides a user interface for the GridLAB-D users.
- it helps the creation and debugging of models of a microgrid.
- it simulates, models and designs a power grid using the GridLAB-D and it controls using the Matlab the operation of the power grid in the residential point of view.
- it helps the users to change the simulation models and study the simulation results.

GridMAT focuses on DR programs, on the integration of the RES and on the developing of energy management controllers for reducing the load during peak load periods.

Gridspice (Anderson et al., 2014) uses parallel techniques to manage large scale simulations using existing tools; the GridLAB-D for the simulation of the distribution network, the MatPOWER for the simulation of the transmission network and the Amazon Web Services for the virtual machines for the parallelization parts of the simulation. GridSpice also provides, through a Python and a third party library, a GUI for providing the users options to edit and design their simulations. However, it does not provide visualization of the simulation results, but it uses a GIS editor that uses Google Maps.

To that end, all the aforementioned efforts have different objectives, which is different from the objective of this study. The objective is to implement a friendly GUI to represent in a graphical way the simulation results and to help any user to execute GridLAB-D simulations of a power system from scratch. In the next two sub chapters, the efforts to meet this objective are presented in detail. The MatGridGUI which is a user interface for performing and executing GridLAB-D simulations and the GridLAB-DVisor as a visualization application of the simulation results obtained from a GridLAB-D simulation are presented.

3.2.1 MatGridGUI

An easy and user-friendly GUI for executing GridLAB-D simulations of a distribution power grid, called MatGridGUI, is presented in detail in this part of the dissertation. It supports the design of residential power grids based on the IEEE-13 (Kersting and Shirek, 2012) or the IEEE-37 test feeder (Yun-Su et al., 2016). This toolbox enables the user to analyze and monitor in an easier way the state of a power system by focusing on specific characteristics such as voltages, current of a power system and residencies' load for every particular time instance. The use of the functionality, that the GridLAB-D provides, needs the familiarization with features of the configuration properties of the power grid components. However, using the MatGridGUI, the users can design its own power grid without having any physics background and without writing any code. The users have to give some values of various parameters needed for the development and the design of the distribution power grid.

MatGridGUI structure

In Figure 3.3, the architecture of the MatGridGUI toolbox is presented. MatGridGUI consists of three main components as follows:

- the selection of the IEEE test feeder
- the core of the MatGridGUI toolbox associated with the input data and the model creator.
- the data analysis module

The information starts from the toolbox and it continues to the GridLAB-D. The GridLAB-D takes the derived .glm file and starts the simulation. Since the simulation is over, the results are available through .csv files which can be analyzed and used for different purposes such as load forecasting.

MatLab Interface

The toolbox is straightforward for a user unfamiliar with power grid concepts and features and it is accessible as depicted in Figure 3.4. Once the user decides which of the two IEEE test feeders prefers to use for his simulations, a new window is popped up. In the new window, the user has to give the input information required for each feeder. For instance, in the IEEE-37 test feeder (see Figure 3.5) the number of required inputs is comparatively less than the ones for the IEEE-13 (see Figure 3.6). Since no congestion issues are considered for the IEEE-37 feeder, there is no need of input data associated with the thermal limits of the lines.

More specifically, each interface of each feeder consists of two control areas, the input, and the output. The input control area concerns the configuration information about the power grid model, such as the start and the end time of the simulation, the market period, which is associated with the frequency, the market participants

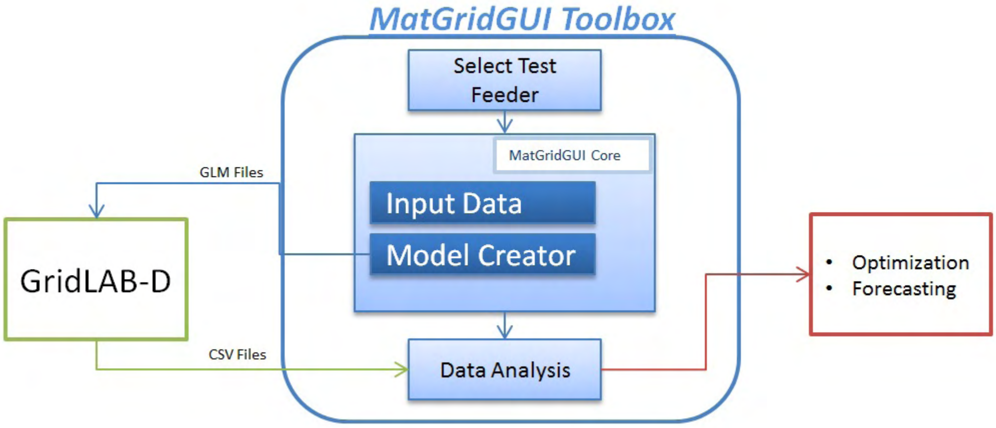

FIGURE 3.3: High-level architecture of the MatGridGUI.

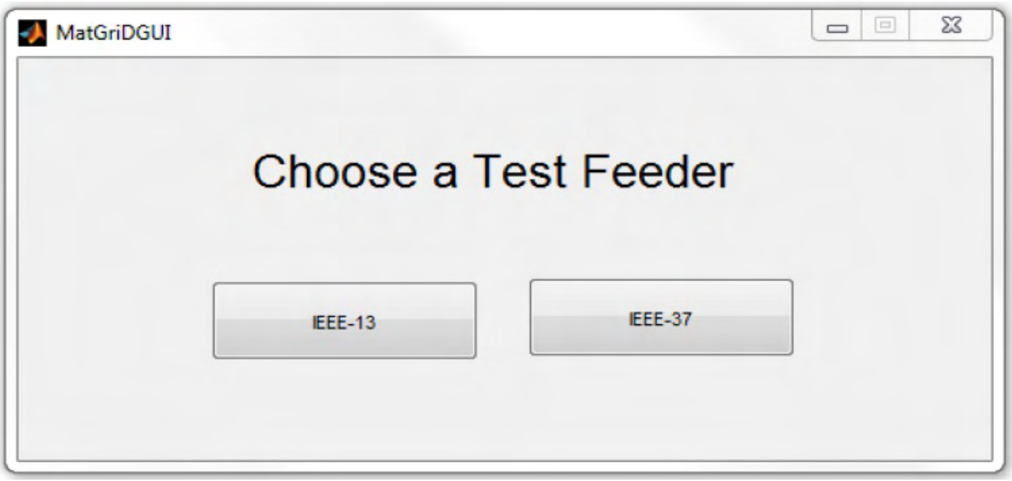

FIGURE 3.4: MatGridGUI feeder selection page.

which can make their offers, the number of the residencies that the user prefers to be accommodated in his power grid and the location of the RES; the nodes in which the RES can be connected. In the output control area, the user can select a) how often the measurements can be recorded, and b) the objects that he wants to have measurements, such as the total load of all the residencies or information about the solar panels. The user can request for creating the distribution grid by pressing the associated button, named **Create a simulation model**. After the submission, the .glm file is created and it could be executed from the GridLAB-D.

GridLAB-D backend

In the MatGridGUI, the user determines and submits his input, (i.e., the user creates its own power grid model). The user is able to initiate and perform a simulation by pressing the button associated with this action, denoted as **Run a Simulation.** After that, the simulation runs given that the executable of GridLAB-D is in the same repository with the .glm file. Then, the output data are generated and are available to the user after the end of the simulation. It should be noted at that point that when simulation time is longer than a day and the recording time interval is less than five minutes, then the output data is available before the end of the simulation.

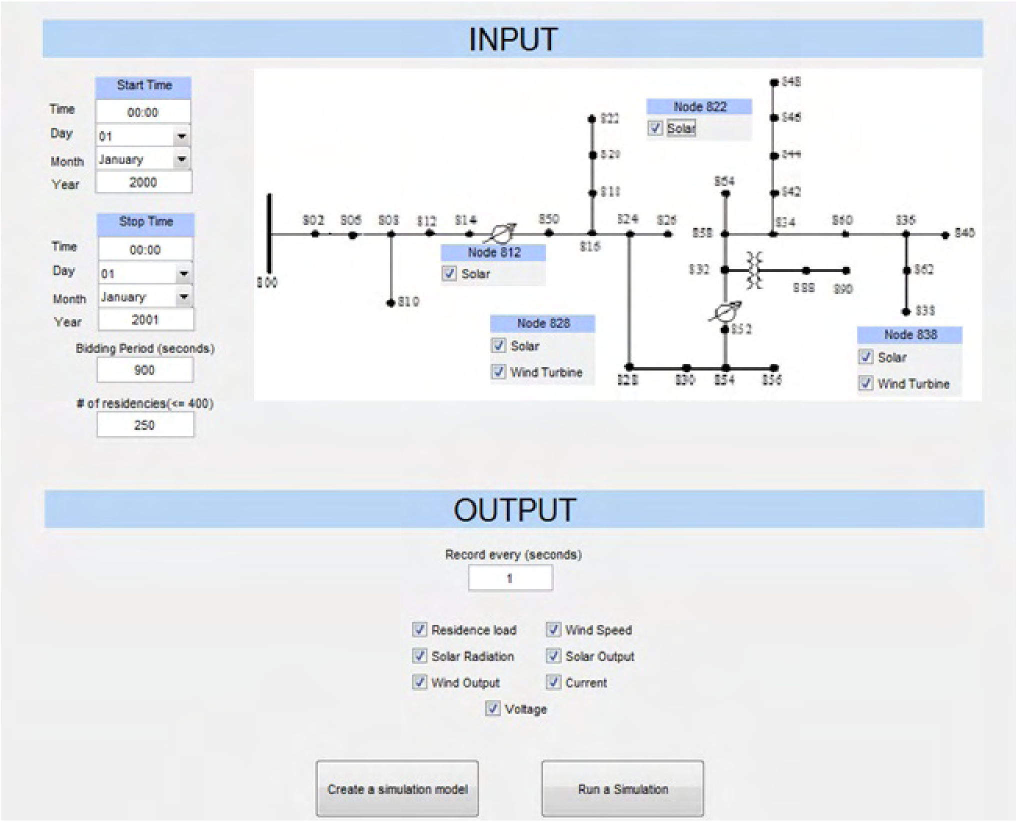

FIGURE 3.5: IEEE-37 test feeder interface.

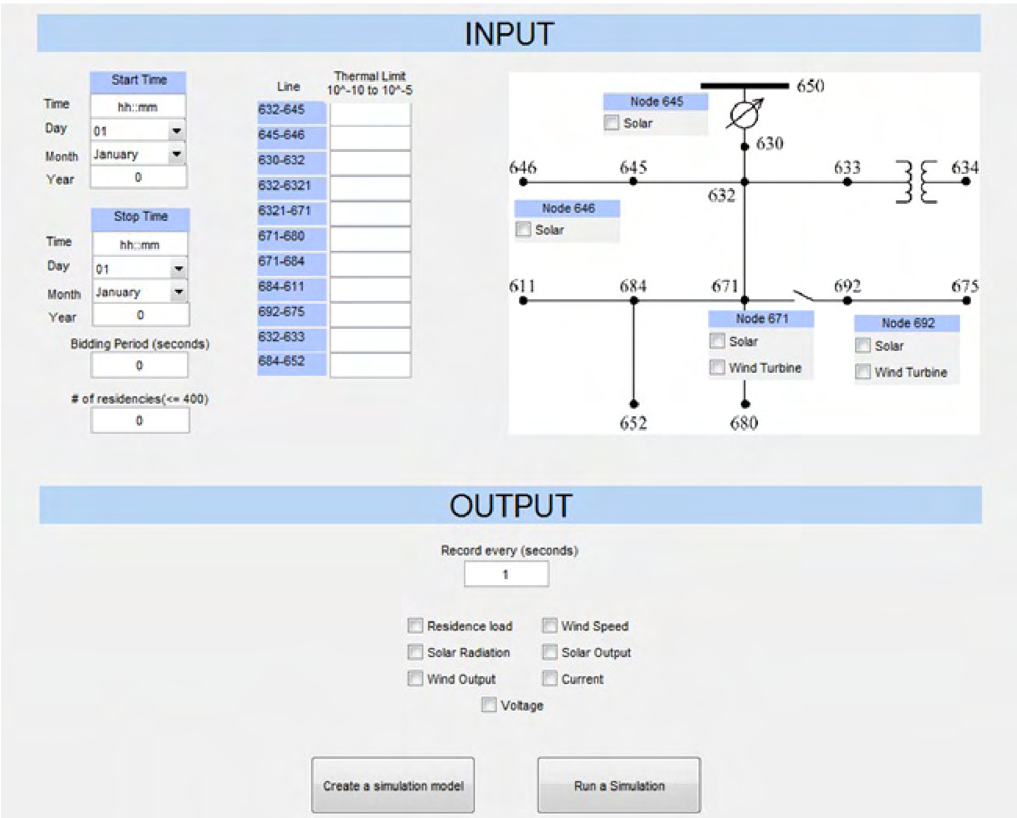

FIGURE 3.6: IEEE-13 test feeder interface.

Data analysis

Using the MatGridGUI the user can record: i) each residency's load, ii) the current of each line, iii) the voltage of each node, iv) the output of RES, v) the solar radiation and vi) the wind speed. After the end of a simulation, a set of .csv files are available to the users to be analyzed and processed.

3.2.2 GridLAB-DVisor

The proposed web application is based on the widely adopted Model View Controller (MVC) architecture. This application is part of the thesis **Visualizing New Generation Energy Systems** submitted by Lamprini Vasilaki in University of Thessaly. The MVC architecture consists of three layers with fully independent capabilities. More specifically, the Controller receives the application's requests and in cooperation with the Model layer prepares the data to be used by the View layer. In turn, the View layer utilizes these data to produce the response. You can find more details in the study submitted by Foti et al. (2016).

The GridLAB-D and the web application are both connected to the same database. When a simulation model is executed, the values of specific parameters are stored in real time in the database at specific tables for better and easy management. The Navicat, a database tool for MySql, responsible for the connection of a MySQL database with a single application is utilized. The distribution network consists of nodes and links. Nodes act as aggregation points of the links that are attached to it. In this spirit, the database is comprised of the following tables:

- **Nodes**: it contains information about node's features such as the nodeID, the phase, the voltage per phase, and the current per phase.
- **Houses**: it contains information about each residency such a unique id, the phase, the total load consumed and the reactive power.
- **Central triplex meters**: the triplex meter connects the house to the grid. The table contains information about the nodeID in which the triplex meter is connected to, the real and the reactive power at each particular time instance.
- **Devices**: it contains information about the appliances belonging to a house. Each device has a unique device id, related to the house that it belongs. The phase, the type of the device, the bid price, the bid quantity of each device, the total load consumed from the device are stored in that table.
- **Energy sources**: information about the node to which the energy source belongs, the type of energy source (e.g. Diesel, Wind or Solar), real power, reactive power and the bidding price and quantity are depicted.
- **Market pools**: the clearing price, clearing quantity, buyers' total quantity and sellers' total quantity are available to this table.
- **Transformers**: it contains information about the distribution transformers of the system. Only the power (real and reactive) per phase are stored in that table.

Details about the database tables and their characteristics are depicted in Figure 3.7. For each cell of the table, the type is indicated (to the right) while in the left each index has a different meaning. For example, the yellow icon of each cell indicates the existence of a primary key and the filled rhombus indicate a cell with NOT NULL value. The connection lines identify the dependence between the tables.

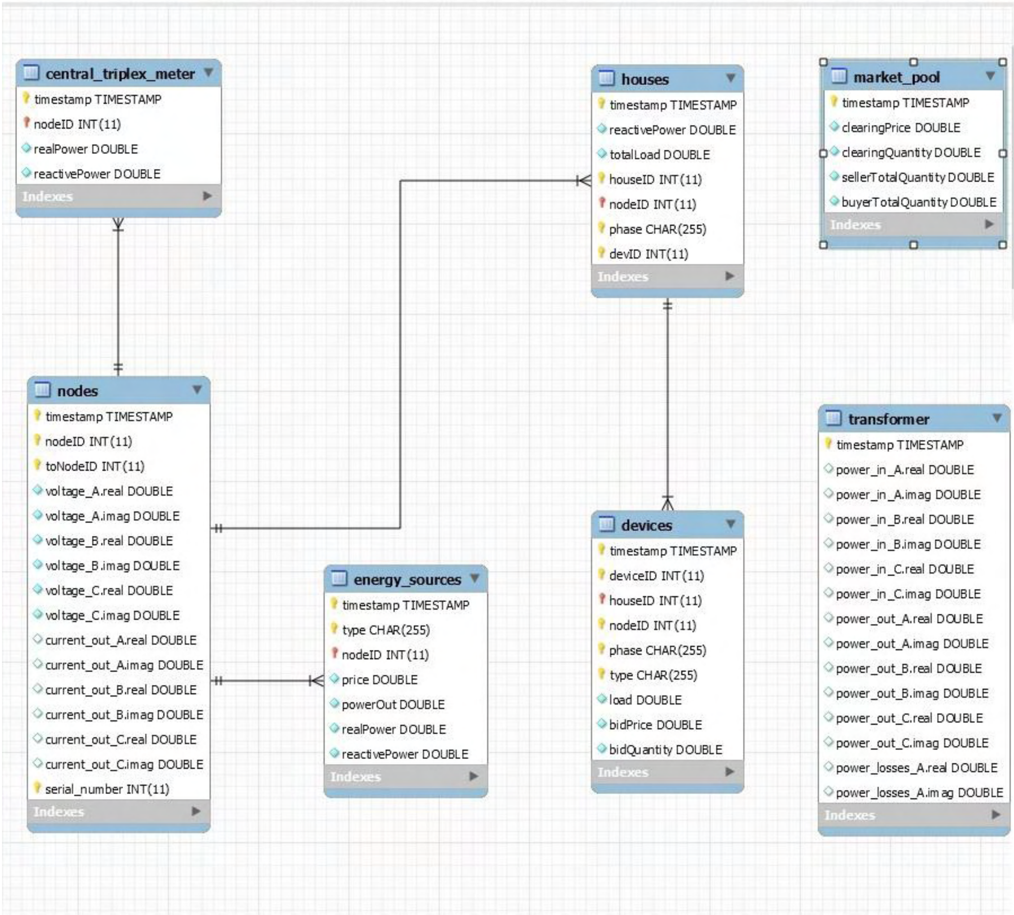

FIGURE 3.7: The Database Structure.

Prototype Implementation

The web application collects, manages and presents data information driven from the simulations conducted using the GridLAB-D. The GridLAB-D's functionality has been extended by first implementing its connection with the application. Various objects of the GridLAB-D modules are associated with a table in the database assigned to the application. For instance, the table **houses** which has information about a house in a simulation model, has entries that are associated with the parameters of the object *house* of **residential** module. Gridlab-D executes queries to the database to store information about power consumption characteristics, such as the total load, the timestamp of that load and values of the parameters that this table has as columns.

The contribution of GridLab-DVisor is that a user can have access to simulation results real time due to the fact that the results can be visualized while the simulation is still running. In the next lines, a description of the application's interface utilized by the user to navigate is presented.

The GridLab-DVisor presents the values, which are dynamically driven from the database, specific parameters of an object which is associated with a table in the database and a menu option in the proposed application (see Figure 3.8). The nodes of the grid are dynamically driven by executing a query which selects from the table **Devices** the records of the *nodeID* field. Each time a *nodeID* is selected, a query is executed. It reads from the database the phase which is associated with the corresponding node, the houseID for each house which belongs to that node. In a general view, the user has to select primarily the timestamp and the nodeID when the visualization of the results associated with a triplex meter, a house, a device and an energy source is requested. Additional information, like houseID and phase, are needed as input from the user to visualize results for a specific house or devices.

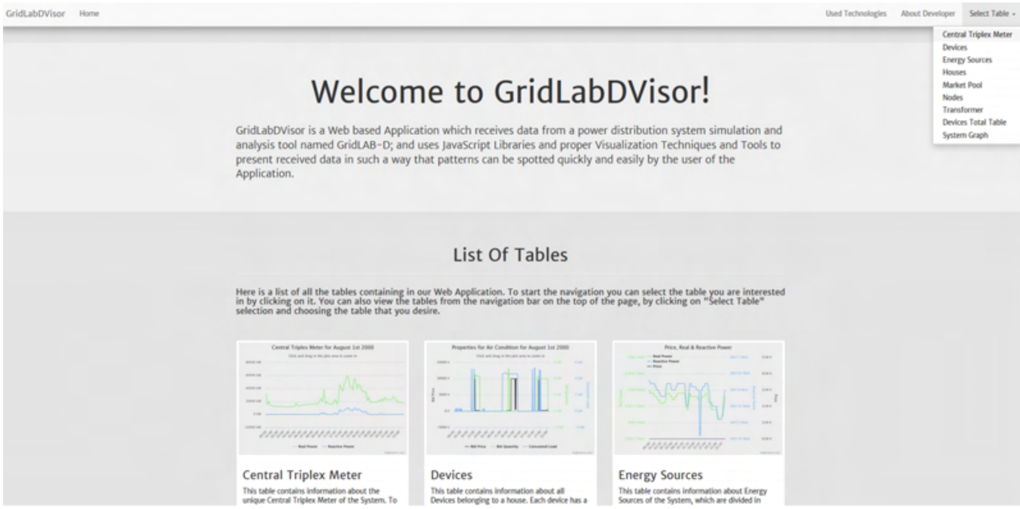

FIGURE 3.8: The homepage of GridLab-DVisor.

Some screen shots of the proposed web application focusing on how the data concerning the central triplex meter are presented. The procedure for visualizing the simulations results of other objects is the same. The user has to select the table from the index page (see Figure 3.9), in this case, the central triplex meter.

Before the appearance of the graph, a web page is popped up, the date and the identification numbers of the nodes (see Figure 3.10) of interest have to be identified by the user. When the user selects a date, a pop-up window containing a calendar has appeared. It is worth mentioning at that point that the date selected by the user has to be within the simulation time interval in order the information about the central triplex meter to be retrieved.

When the ”Node ID” input is selected, a list of all the nodes of the under consideration power system is dynamically displayed to the user. At this time, the user may select only one of the displayed Node IDs. When the user completes the form and he is ready to press the button ”Search”, a graph with the results (real and reactive power) for the selected day, which are retrieved from the database, of the Central Triplex Meter is displayed. In Figure 3.11, the Central Triplex Meter graph in a Multiple Axes diagram is depicted. A multiple axes diagram has a separate y-axis for each displayed variable. It is more accurate because of the number of its axes as well. In addition to Multiple Axes diagram, Line and Spider diagrams are also utilized to depict the information associated with that table. In the Line Diagram (Figure 3.12) the real and reactive power are depicted in lines. Since data is continuous, they can be easily presented as line diagrams producing a clear and appealing result. Regarding the Spider diagram (Figure 3.13), it is utilized to provide the user with a better view of system behavior within the time period.

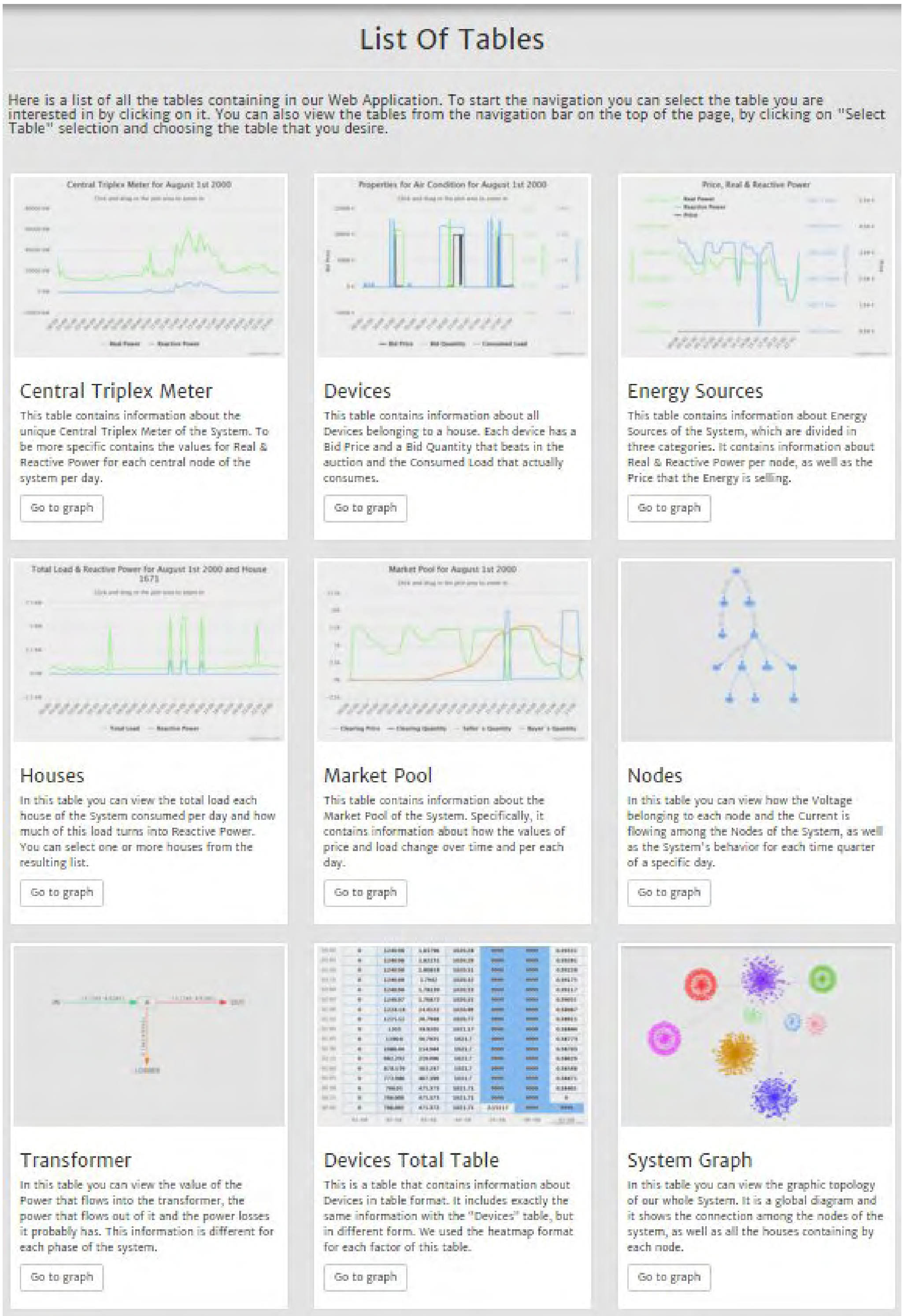

FIGURE 3.9: List of tables in homepage.

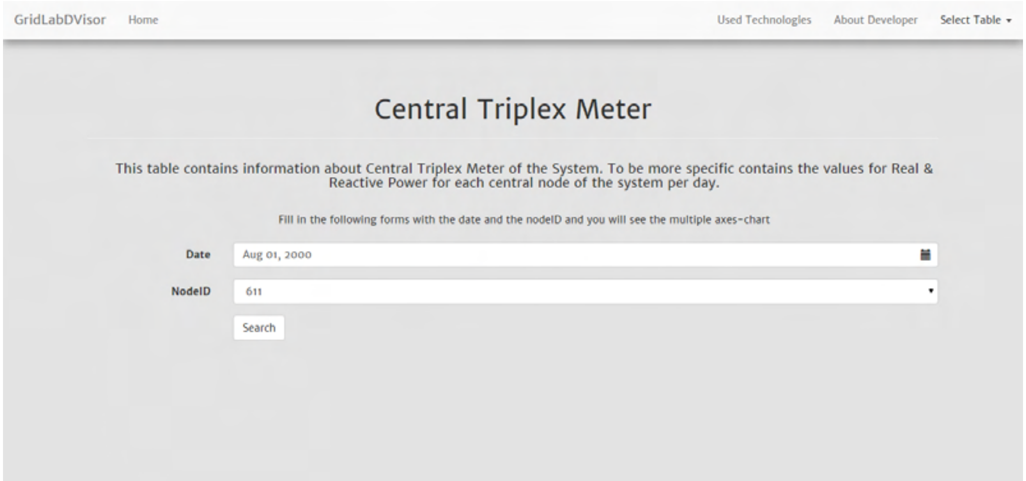

FIGURE 3.10: The selection form used for Central Triplex Meter table.

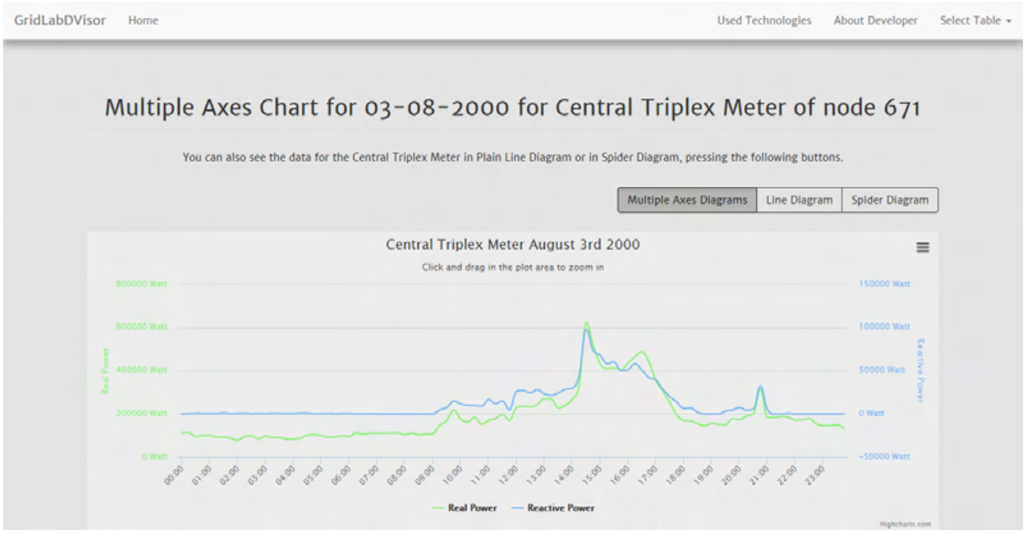

FIGURE 3.11: Multiple Axis Diagram for Central Triplex Meter.

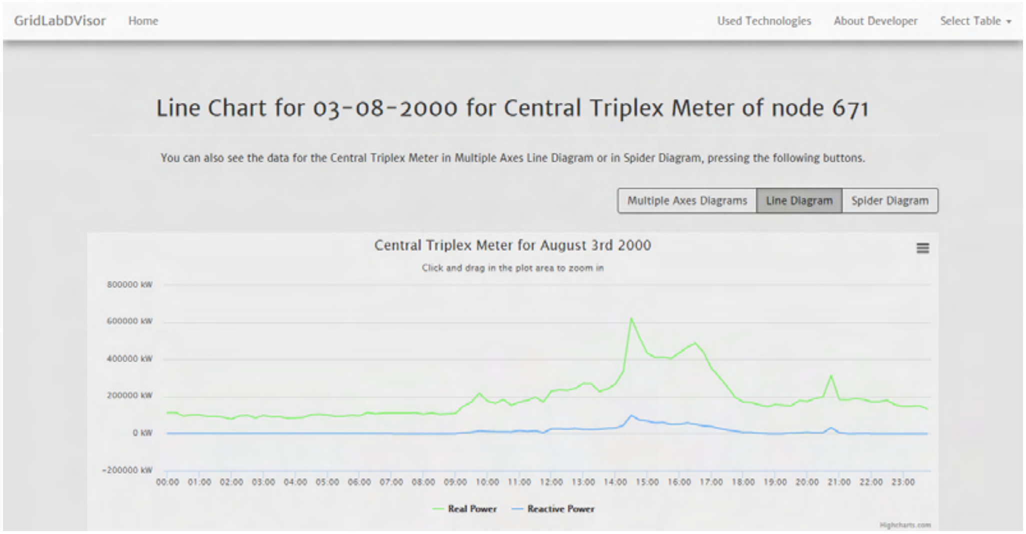

FIGURE 3.12: Line Diagram for Central Triplex Meter.

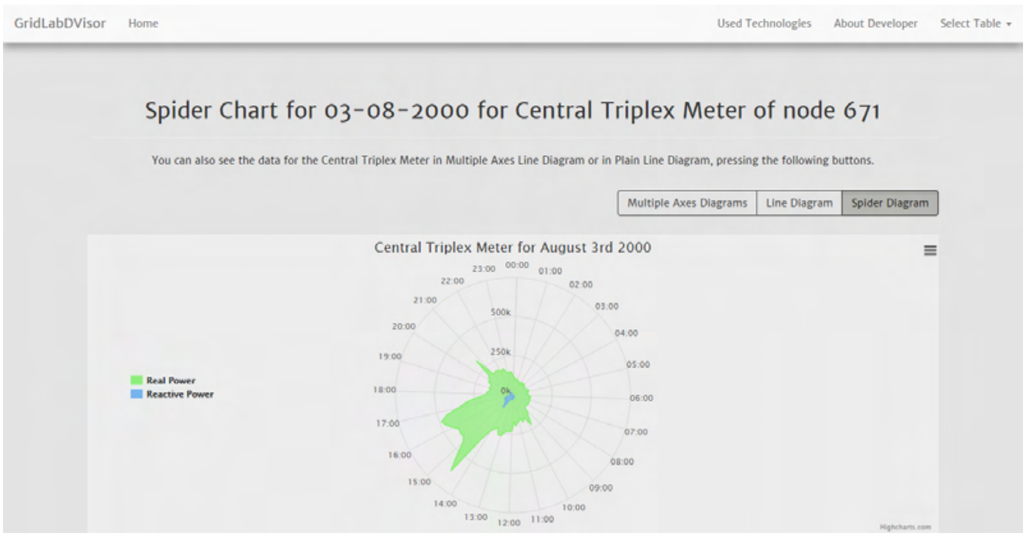

FIGURE 3.13: Spider Diagram for Central Triplex Meter.

Chapter 4

Active and Reactive Electricity Markets

In this chapter, an effort considering the analysis, the design and the implementation of next generation energy systems besides of the early proof-of-concept implementation and experimentation of a simulation engine aimed to active (Fainti et al., 2014b), (Tsoukalas et al., 2014), (Fainti et al., 2014a) and reactive (Nasiakou, Vavalis, and Bargiotas, 2015) power markets are described in detail. The proposed engine allows the authors to elucidate issues related to the open and smart participation of producers and consumers on large scale energy markets. The development of an engine capable of truly large scale simulations that covers real life scenarios and stress most components and modules of next generation smart energy markets is presented. Besides the active power market, the reactive power market, regardless its importance and its significance for the safe and efficient operation of the power grid has not been researched and developed in details so far. Hence, in this dissertation, a study of how the pricing of the reactive power affects the operation of a power grid in terms of pricing is conducted.

4.1 Related Work

The introduction to the competition in electricity markets has been paid significant attention the last years. Besides that, the consumers' active participation in the electricity market emerges a lot of applications. Moreover, both the participation of the producers and the consumers in a bid-based security constrained electricity market has been studied. First, William Hogan in Hogan (1996) introduces the concept of the pool market which as the electricity market designed in the US. Wholesale electricity markets in the USA follow the open access and the nondiscrimination characteristics. William Hogan proposes a market model which is bid-based, security-constrained and makes use of the market dispatch concept with locational prices. This model governs short-term operations, provides the foundation for financial transmission rights and supports forward trading. Now all the organized electricity markets follow the features of this design. Recent research efforts have been recently devoting to reactive power and in particular on issues related to its management and pricing and the coupling of both active and reactive power market. Specifically, Saebi et al. (2013), Feng, Zhong, and Gan (2008) and El-Samahy, Bhattacharya, and Canizares (2006) concern reactive power management mechanisms. In particular, Saebi et al. (2013) assume that the operation of reactive power market is decoupled from the active power market and in order to formulate the reactive power dispatch, reactive power reserve management issues associated with

"Voltage Control Area" and payments are considered. In the mechanism that is proposed in Feng, Zhong, and Gan (2008), a must-run capacity of a generator through an index called, **reactive must-run** index is used. This index is utilized for the reactive power calculation based on the network topology, the reactive power compensators, and the reactive power producers. Furthermore, El-Samahy, Bhattacharya, and Canizares (2006) concern about several issues associated with existing provision policies of reactive power services like procurement, energy price volatility, optimal provision for reactive power and payment mechanisms (auction, bilateral and long term agreements). The proposed mechanism is based on two levels, associated with long-term (seasonal) and short term management (real time). The necessity of incorporating, in addition to the conventional generic power generator and the reactive power providers that reduce the prices of the market are also discussed. Moreover, Li, Yu, and Chen (2012) considered and analyzed the production costs and the reactive power capacity.

The design of the reactive power market that is based on the clearing type of market (first price, uniform auction), the monopsony market and the long-term contracts are proposed (Zhong and Bhattacharya, 2002). The market is built on the providers' "Expected Payment Function" in which each component is associated with a region of the reactive capability curve of a generator. The proposed approach is also associated with (a) the minimization of total payment, (b) the minimization of transmission losses and (c) the minimization of deviations from contracts. The studies Rabiee and H. Shayanfar (2009) and Reddy, Abhyankar, and Bijwe (2011) are based on the above-mentioned study in order to develop a new market model. They study a coupled active and reactive market based on power system security and voltage dependent load models respectively. Methods concerning reactive power payments and reactive pricing schemes like "Total Payment Function", "Loss Minimization", "Load Served" and "Voltage Stability Enhancement Index" are discussed in detail by Reddy, Abhyankar, and Bijwe (2011). Furthermore, Rabiee and H. Shayanfar (2009) presented a coupled day-ahead active and reactive power market based on a settlement mechanism, which is cleared through an optimal power flow problem.

Regarding the pricing methods, a nonlinear programming method for reactive power pricing based on marginal costs in the competitive electricity market is presented (Ghazi and Asadi, 2010) . This method encourages the producers to procure reactive power by ensuring retrieval of their costs.

The above reactive market-related studies concentrate almost exclusively on the producers. To the best of our knowledge, the only study (Vaskovskaya, 2014) that investigated the role of the consumers in reactive power markets shows that the "Locational Marginal Prices" could be used as price signals to the consumers in order to compensate for reactive power by introducing compensation devices. The insertion of such devices leads to the decrease of the electricity cost.

4.2 Simulating Active Energy Markets

Going through the liberalization of the energy market, a simulation engine that

- focuses on decentralized energy markets
- provides efficiency and safety in the operation of the power grid and more especially of the distribution grid
- leads to the reduction of the wasted energy

is presented in detail.

A simulation engine which joins the features described above needs to be developed (a) on solid theoretical foundations and (b) through the integration of software modules.

4.2.1 Design and Implementation

As described in Chapter 3, GridLAB-D due its openness provides the chance to modify, add and extend its existing modules. For the design and the implementation of the simulation engine the **climate**, the **tape**, the **residential**, the **powerflow** and the **market** module are used.

The GridLAB-D's functionality in order to (a) allow the residency's appliances to change, when this is required, their state concerning the current state of the power grid and (b) implement the operations of the four energy market models described in the next lines has been extended. So, for demonstrating the capabilities of this engine, the design and the implementation of four different active power market models are presented. The first two models are based on a two level auction market framework while the third one is associated with the market framework described by Hogan (1996). Last, the fourth model is an extension of the second model in which a game theory model is developed and used for maximizing the profit of the market participants and more especially the profit of the producers.

The **first market model** is built through a hierarchy of two markets, a local and a global. In the first level, each city/town satisfies its energy demand by the local power plant(s) forming that way a local market. Each city/town is accommodated with one or more power plants and a specific number of residencies. The power plants have the role of the producers while the residencies and more specifically the appliances of a residency have the role of the consumers. Each residency accommodates four individuals and is equipped with four different appliances: a refrigerator, an HVAC system, a water heater and a light system. Both producers and consumers submit their bids - a certain amount of energy at a certain price at the local market. Both the local and the global market are organized as an auction (via the *auction* object of **market** module). In this model, the total required power of each of the appliances is scheduled via the GridLAB-D's *schedule* object that sets associated values of each appliance of the season, the time and the day. The process followed by the **first market model** is the following; when the demand of a city/town is not being met at the local market, a second level bid (via a properly configured of the global market *stub_bidder* object of the **market** module) to the global market for the remaining unsatisfied demand takes place. Using this approach, the demand of a town/city may be satisfied partially from the local market (paying the clearing local price) and from the global market at its clearing price.

Note that each town/city participated in the global market as a single buyer when its demand is not satisfied by the local market, or as a single seller in the case that the town/city has supply left from one or more local power plants that is not being dispatched. There are several ways that the bidding problem is executed into the global market. For instance, an assumption that the bid price of the towns/cities, which act as buyers, is the highest price in which the appliances bid at a particular time instance has been made. The towns/cities (that act as sellers) have as bid price the maximum price at the local market of the sellers that belong to the same local market.

The **second market model** has the same configuration with the first market model with the main difference that in this model when a residency is not satisfied by the local market then the state of the appliances with high priority is checked. If an appliance was ON during the previous simulation timestamp, the additional energy is requested from the global market. Furthermore, in such a case the consumers do not operate as stub_bidders but as real residencies' appliances which participate in the markets via the controller/passive controller, objects of the **market** module. The use of the controllers introduces the concept of DR in this model. The way the passive or the active controller works is described in detail in Chapter 3.

The **third market model** is simpler, just one global market (pool), where both consumers (residencies) and producers (power plants) submit their bids. This model uses also DR programs as the previous models, so if the price is above a certain threshold then the appliances of a residency are turned off.

The **fourth market model** is an extension of the second market model in which the game theory concept is utilized. When the demand is not satisfied by the local market, the towns/cities that have surplus of energy bid on the quantity that is the result of a game theory model as this is described below.

Game theory may support producers to estimate prices or quantities that optimize their profit. The state of the game where all the producers maximize their profit is called Nash Equilibrium. In the energy market, there are two commonly used types of Nash equilibrium; Cournot and Bernard (Al-Agtash, 2013). The former calculates the equilibrium based on the quantity while the latter based on the price. In this study, the Nash-Cournot Equilibrium is used where the quantity is calculated through the game theory procedure. The price is predefined either from an auction mechanism or from historical data. Following the methodology of Contreras, Klusch, and Krawczyk (2004) the function that characterizes the producer's cost is the following:

$$C_i(q_i) = c_i * q_i \tag{4.1}$$

where the q_i is the quantity and the c_i is the cost of the i-th producer.

The formula for the inverse demand is the following:

$$p(d) = \frac{d_0}{\lambda} - \frac{d}{\lambda} \tag{4.2}$$

where d_0 is the initial demand, λ the price elasticity and d the total demand.

The producer's profit is

$$F_i\left(q_i\right) = p\left(d\right) * q_i \ - \ C_i\left(q_i\right) \ \ with\ i = 1, ..., n \tag{4.3}$$

where n is the number of producers.

The Nikaido-Isoda function is described as follow:

$$K\left(\bar{q}, \bar{q}'\right) = \sum F_i(q_i^{'}|\bar{q}) - F_i\left(\bar{q}\right) \tag{4.4}$$

Equation 4.4 gives the total profit of all producers when the i-th player chooses the $q_i^{'}$ strategy instead of q_i. The Nikaido-Isoda function is used to convert Nash Equilibrium into an optimization problem for finding the strategies that optimize the profit of the producers. In fact, each factor of this function represents the improvement in the profit of each player when a player changes his strategy at a particular time, while the other players follow the same strategy as in the previous time.

$$M\left(\bar{q}\right) = argmax\ K\left(\bar{q}, \bar{q}'\right)\ with\ \ \bar{q}\ , M\left(\bar{q}\right) \in Q \tag{4.5}$$

Equation 4.5 determines the strategy that maximizes the profit of each player.

Each player's optimal strategy is computed using an iterative process. The iterative method can be implemented using two ways; through a form of a linear system of algebraic equations(Barroso et al., 2002) or through a fixed point iteration scheme with relaxation parameters (Contreras, Klusch, and Krawczyk, 2004). The producer's profit function for both methods is the same.

In this effort, the optimal strategy model is formulated based on the approach of Contreras, Klusch, and Krawczyk (2004) and the implementation of the algorithm that calculates the optimum solution using the methodology described by (Barroso et al., 2002) due to general compatibility issues regarding the GridLAB-D implementation details.

The configuration of the game theory model has been implemented as an additional C++ function of the **market** module of the GridLAB-D. Each *stub_bidder*, object of the **market** module, bids in the global market and it has the cost as an additional property. The function related to the configuration of the game theory model calculates the new bid quantity for each *stub_bidder* when the supply of all the producers in each local market cannot satisfy the local demand. This new quantity is the new bid quantity submitting in the global market.

4.2.2 Experimentation

For the experiments of this effort, the IEEE-13 test feeder is used. For demonstrating a large scale simulation, 12 of such feeders are connected to a single swing node. Hence, 12 local markets with different characteristics concerning the size, the number of power plants and the number of residencies have been designed. Two large local markets, namely from now on as cities, and ten smaller local markets, namely towns, have been considered. For all the presented experiments the market period is 15 minutes and the simulation time is 24 hours.

Getting into more details in the characteristics of each power plant, it should be noted that each city is equipped with four power plants, while each town only with one power plant. The price in which each power plant bids in a city is in average 45 cents and varies from the interval of 30 to 60 cents. For each town, the power plant bids at 50 cents. The bidding quantity and the bidding price of all the power plants remain constant during the simulation. Further configuration details are summarized in Table 4.1 and Table 4.2.

TABLE 4.1: The capacity of the power plants, the maximum load and the number of residencies for the two cities and the first four towns.

	C1	C2	T1	T2	T3	T4
Capacity (MW)	115	40	15	8	6	12
Max Load (MW)	100	43	14	1	7	7
Number of houses	22992	9744	3248	1624	1700	1890

TABLE 4.2: The capacity of the power plants, the maximum load and the number of residencies for the last 6 towns.

	T5	T6	T7	T8	T9	T10
Capacity (MW)	5	25	14	7	9	5
Max Load (MW)	5	9	10	8	6	12
Number of houses	1082	2080	2270	1728	1299	2706

For all the appliances in the first market model, the bid price is 50 cents while in the second, the third and the fourth market model the bid price of the thermostatically controlled appliances is calculated based on the business logic of the *controller* object of the **market** module. The configuration depicted in the Tables 4.1 and 4.2 concerns the first, the second and the fourth market model. Regarding the third model, in which the market is organized as a pool market, there are five power plants with 32MW capacity each of them and bidding price at 60, 50, 45, 40 and 30 cents.

4.2.3 Simulation Results

All experiments described here have been performed on a DELL Precision desktop PC with two Intel Xeon 6-core 2,27GHs processors. Since every run associated with the four market models took less than 3 wall-clock hours, with no any code optimization.

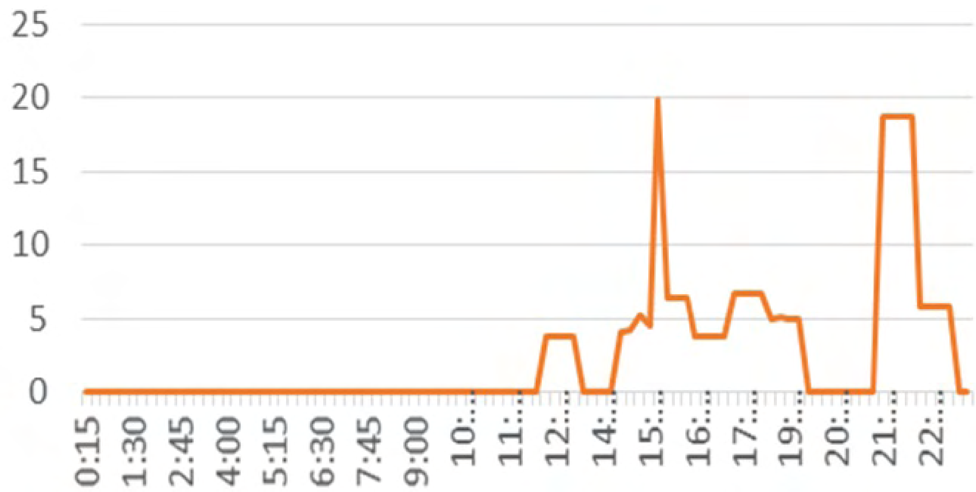

FIGURE 4.1: Clearing quantity of the global market of the first market model. The y-axis corresponds to the power in MW and the x-axis to the simulation time.

Selected experimental results are presented in this part of the dissertation. In the first market model (see figures 4.1 and 4.2) the demand that can be satisfied from the local market is clearly affected by the price. In Figure 4.1, the clearing quantity of the global market is depicted. In each city, where more than one power plants are located, the different prices of the power plants lead to clear unsatisfied demand. So, the city is directed to the global market, where it has the opportunity to buy energy either in higher or lower price depending on the bid price of the cities/towns of the other local markets. By comparing the figures 4.1 and 4.2, it is clearly observed that the unsatisfied demand (the yellow line is above the grey line (see Figure 4.2)) is satisfied from the global market at the respective times (see the peaks in Figure 4.1).

In the second market model (see Figure 4.3), it is observed that, when the appliances have their own priority and a DR program is applied, the demand during the peak hours is reduced. Furthermore, assuming that the clearing quantities are smooth, the demand curves (Figure 4.3) at the peak hours are smoother compared with those in the first market model (Figure 4.2) due to the DR. In towns T3 and T4, it can be noticed a big divergence between the demand and the clearing quantity. This is due to controller's specific business logic regarding the price. Moreover, the

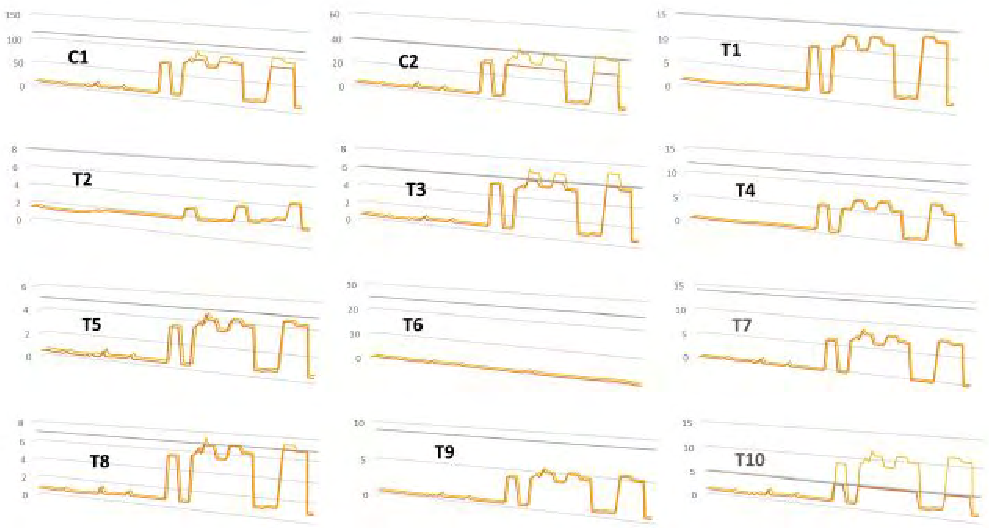

FIGURE 4.2: Simulation results of the local market of the first market model. In the x-axis, the simulation time and in the y-axis the power in MW are depicted. The grey, yellow and orange lines correspond to the producer's total quantity, consumers' total quantity, clearing quantity respectively.

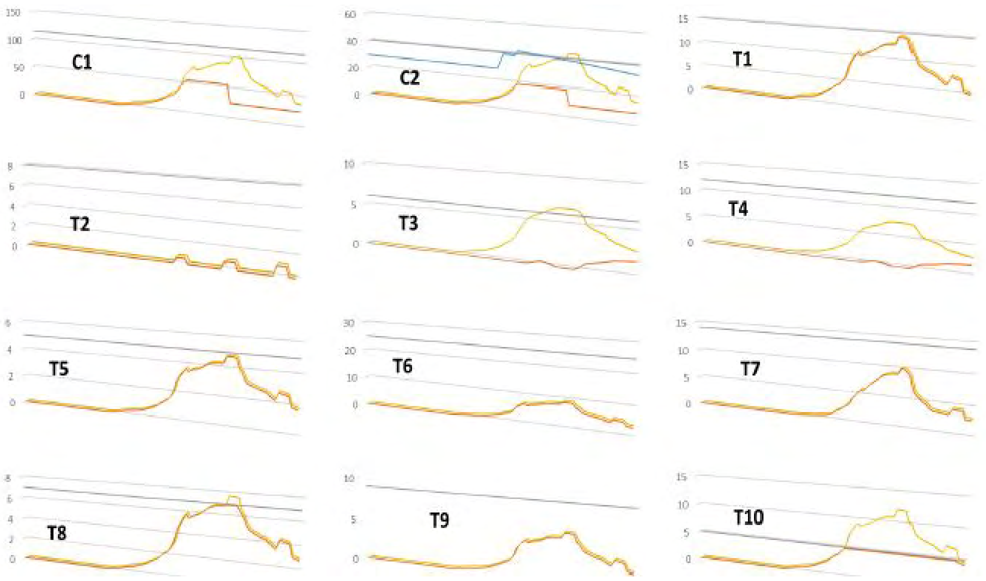

FIGURE 4.3: Simulation results of the local market of the second market model. In the x-axis and y-axis, the simulation time and the power in MW (or the price for C2) are depicted accordingly. The grey, yellow, orange and blue lines correspond to the producer's total quantity, consumers' total quantity, clearing quantity and clearing price respectively.

amount of energy that is asked from the global market is smaller and it depends on the state of each of the appliances.

In the third market model, which is simpler than the other three models, the demand is not satisfied when either the buyers or the sellers are interesting to bid in low/high prices in order to maximize their profit. In Figure 4.4, it is observable that during peak hours the demand is higher than the generation capacity. In that case, the appliances that bid for energy in a lower price than the clearing price (blue line) will not operate. As a consequence, the demand is decreased (orange line) ensuring that way the system's stability.

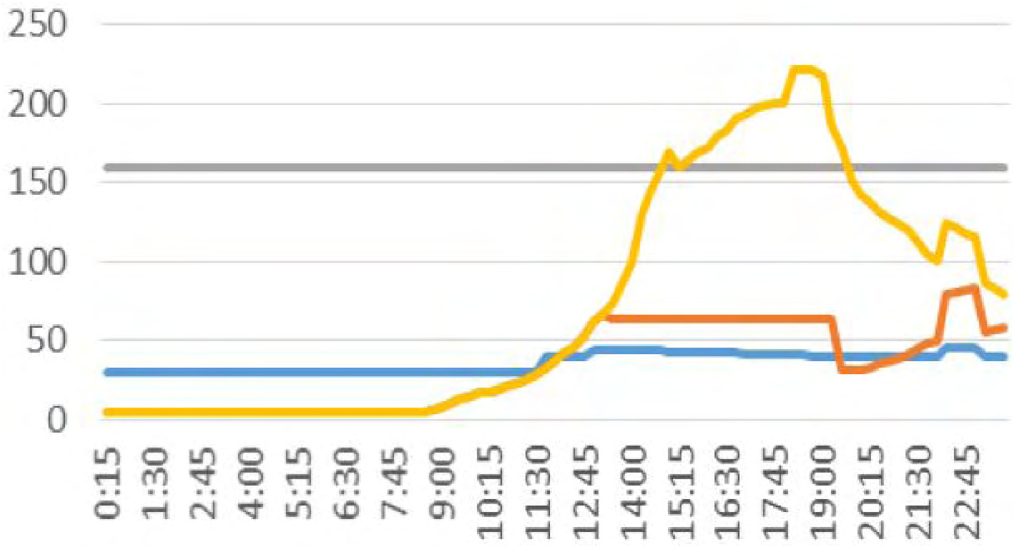

FIGURE 4.4: Simulation results of the third market model. The grey, yellow, orange and blue lines correspond to producers' total quantity, consumers' total quantity, clearing quantity and clearing price respectively.

In Figure 4.5, for comparison purposes, the results regarding the clearing price of the global market for the second and the fourth market model are presented. Specifically, the clearing price of the global market when the generator of the towns and the appliances bid for 50 cents for the second model and the clearing price of global market when the generators and the appliances bid for 70 cents in the fourth model is presented. It is worth to be mentioned here, that the price of the global market in the fourth model increases 70 cents from 60 cents due to the increased demand that has to be satisfied for the particular period.

4.3 Simulating Reactive Power Markets

In contrast to the attention that active power market has been paid, the reactive power market has not received significant research and development attention so far. Recent studies focus only on the producers' side giving less attention to the consumer's side and how the reactive power may affect them. Hence, a small research how the reactive power affects the consumers and the price that they have to pay is conducted. Any disturbances in the market of the reactive power may affect the state, the efficient and the safe operation of the power grid.

The necessity on separating the energy markets into two interrelated, one associated with active and the other with the reactive power helps us to achieve the integration of active and reactive power market. The separation of the two markets is motivated by (a) the emerging energy market mechanisms ,(b) the information available through smart grid and (c) the distributed local energy generation. In particular, a simple case in which each consumer bids into the market for real and reactive power is presented.

FIGURE 4.5: Simulation results for the second and fourth market model. In the x-axis and y-axis, the simulation time and the clearing price in cents are depicted accordingly. The blue and red lines correspond to the clearing price of C1, C2 respectively. The purple and green lines correspond to the global market clearing price for the second and fourth market case respectively.

4.3.1 Design and Implementation

The aim of this effort is the development, the implementation and the analysis of a coupled (active and reactive) power market. In Figure 4.6 the architecture of the proposed model is presented. For supporting the operation of such a market, the extension of the functionality of specific components of the GridLAB-D was required.

The local reactive power sources used in this effort are either diesel engines or wind turbines. Both diesel generators and wind turbines bid into the market through the *generator_controller* object of the **market** module. Similar with the market models presented in sub-chapter 4.2, the reactive power market is organized as an auction with the main difference that the producers bid for both active and reactive power at a single price. Actually, they bid only for active power; the reactive power is produced anyway. Therefore, the price that the consumers receive, after the clearing of the market, is both for reactive and active power. After the clearing of the market, the MO aggregates the reactive power from all the producers; the producers that are included in the market dispatch. It is worth mentioning at that point that the reactive power that each producer offers depends on its capability curve (see sub chapter 2.9.2). Moreover, the bid price of each producer is associated with the reactive power that it could give. More specifically, if the producer reduces the production of its active power to produce more reactive power, then the bid price is higher to balance

FIGURE 4.6: Architecture of the proposed market model.

the losses of producing more reactive power.

From the consumers' side, they bid only for active power. However, the MO aggregates the reactive power needed by all the consumers as well. After the market clearing, the MO collects the supplied reactive power and the demanded reactive power. After that, it compares those quantities to decide if more reactive power is needed. So, if the producers, that are included in the market dispatch, are not able to satisfy the demanded reactive power then the producers that are not included in the market dispatch because of their high bid price offer only their reactive power at the bid price. In that case, the consumers called to pay (a) the clearing price of the market and (b) an additional price per kVAr of each additional unit of the more reactive power that the producers offer.

The clearing price of the market is the price for both the reactive and active power and it is calculated via the auction procedure which is part of the simulation framework. On the other hand, the price of reactive power that the consumers have to pay is calculated through a simple procedure for which the MO is responsible for. That procedure is taken place after the clearing of the market and it calculates the price of reactive power as the average price of the producers' bid prices; these are the producers that are not included in the market dispatch. This procedure does not affect the simulation and it's not part of the simulation procedure. It is worth to be noted that, there is no actual bidding operation for reactive power. Consumers ask for active power but they have to pay for reactive power as well. Similarly, producers offer active and reactive power at a single price but they bid only for active power and they get paid for reactive as well.

4.3.2 Experimentation

For demonstrating the capabilities of the proposed reactive power market model several case studies have been designed that are as realistic as possible and of significant large scale. Based on IEEE distribution feeders, a distribution grid according to

TABLE 4.3: Configuration of Experiment II

C and T[1]	Feeders	Houses	Wind Turbines			Diesel	HVAC	Lights	RG[2]	WH[3]	MV[4]
			S	M	L						
C1	6	11500	1	1	7		1	0	0	1	0
C2	4	3600	6	2	1		1	0	1	1	1
T1	1	1600			3	1	1	1	1	1	0
T2	1	840		2	1	3	1	1	1	1	0
T3	1	850		2	1	1	1	0	0	1	1
T4	1	900	1			2	1	1	0	1	0
T5	1	540			2	2	1	0	1	1	0
T6	1	1040			4		1	0	0	1	1
T7	1	1135			4	2	1	1	1	1	1
T8	1	874			2	2	1	0	1	1	0
T9	1	650	3			1	1	0	0	1	0
T10	1	1350		2	3	5	1	0	1	1	0

1 Cities and Towns
2 Refrigerator
3 Water Heater
4 Microwave

various experimental design decisions is designed. Two properly configured experiments for the purposes of this study are considered. In the **Experiment I**, the IEEE-37 feeder is used providing a distribution grid of 790 residencies; each accommodated with an HVAC system and a water heater. Regarding the distributed generators, two small wind turbines and a diesel generator are utilized in that experiment. In the **Experiment II**, 20 IEEE-13 feeders are connected to a single node. A total of 30,000 residencies are supported and each residency may contain a water heater, an HVAC system, a refrigerator, a light system and a microwave. It has to be noted that from the aforementioned appliances only the HVAC system requires reactive power.

As regards the generator side, there are both distributed and transmission generators. The latter feeds, through the transmission lines, the distribution substation of the distribution grid, which offers active and reactive power as well. The generators have appropriate properties, which are associated with the active and reactive power at each particular simulation time instance. The *generator_controller* uses these properties to calculate its bid price and bid quantity submitting into the market every 15 minutes. Solar panels have not been used since due to the particular weather data selected, the WA-SEATTLE.tmy2 in the climate module (a .tmy2 file is a weather data file for a particular location of a given day and hour) resulted in the low contribution to total supply. Table 4.3 summarizes the configuration of Experiment II.

GridLAB-D provides various sizes of wind turbines regarding their power output. The characteristics and the bid prices for the distributed generators are given in Table 4.4. In the presented experiments, the power factor of the diesel generator is set to 0.85. The power factor of a small-sized wind turbine is set to 0.99, for a medium-size to 0.85 and for a large-size wind turbine to 0.97. It is worth mentioning at that point that there are transmission generators that offer only their reactive power for 0.75 $/kVAr. The price is high comparing to the other producers because these sources offer only their reactive power and because of the disadvantage

TABLE 4.4: Capacities-prices of the producers

Max capacity	Wind Turbines			Diesel
	Small	Medium	Large	
Active (MW)	.0058	.15	1.50	2.50
Reactive (MVAr)	.0025	.045	0.65	1.09
Price ($/KVA)	.40-.45	.40 -.45	.40-.45	.678
Price(($/KVA) (pf=0.85)	.60	.60	.60	.72

of reactive power to be transmitted over long distances due to its excessive losses. The price for bulk transmission generators is 0.40 $/kVA. The appliances act as consumers and they bid on average 0.50 $/kW for the refrigerator, 0.45 $/kW for the light system, and 0.45 $/kW for the microwave. The bids for these appliances are configured through the *stub_bidder* object based on properly configured scheduling. Furthermore, an assumption supporting that the reactive power that is needed for the lines is supplied by synchronous condensers which have not direct participation in the market but through forward contracts as it is usual in such cases has been made.

4.3.3 Simulation Results

Based on the above-described market model, selective experimental results are presented. Regarding the Experiment I, in Figure 4.7 (on the left) the red line represents the price of both the active and the reactive power (complex power) in $/kVA and the blue line corresponds to the additional price for reactive power in $/KVAr.

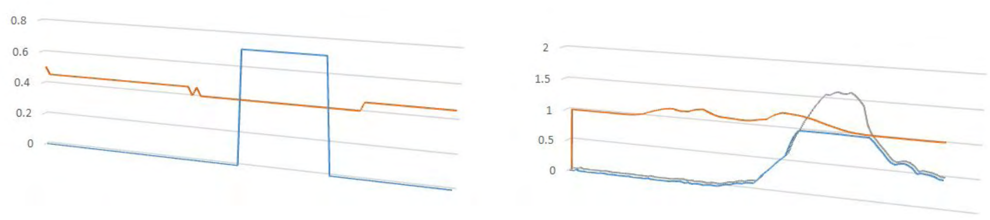

FIGURE 4.7: Active and reactive price (on the left) and demand, supply and clearing quantities (on the right) for the Experiment I.In the x-axis and the y-axis the simulation time (24 hours) and the price in $/kVAr (on the left) for the blue line and in $/kVA for the red line and the amount of active power in MW (on the right) are depicted accordingly.

The active power associated with the clearing quantity of the market (blue line), the sellers' (red) and buyers' (green) total quantity in MW are presented in Figure 4.7 (on the right). It is observed that there is unsatisfied demand for active power. This is due to the insufficient amount of active power that the producers provide. Additionally, the producers that participate in the market offer the insufficient amount of reactive power and therefore the consumers have to pay more for reactive power. Therefore, the consumers have to take part of the reactive power needed from the transmission network. The final price that they have to pay per each unit of reactive power, as it can be noticed in Figure 4.7 (on the left), is 0.71 $/kVAr; the average of 0.678 $/kVAr, and 0.75 $/kVAr. The former price belongs to the diesel generator which is the producer that is included in the market dispatch while the latter price corresponds to the price of the wind turbine or synchronous condenser of the transmission network.

In figures 4.8 and 4.9, the results that correspond to Experiment II are presented. Getting into more details, in Figure 4.8 at Town 5, it is observed that at 15:45 there is a big disturbance at the supplied power by the producers. This is mainly on account of the wind turbine operation and the fact that at that particular time instance the supplied power is dramatically decreased due to the low wind speed. Each town is equipped with different size of wind turbines as it can be observed by the curve between T5 and T4 in Figure 4.8. Moreover, there are cases that the demand is not fully satisfied despite the fact that the supplied power is enough to satisfy the demanded power. This is a side effect of the way that the controller of each appliance operates and how the residencies participate in the market. Actually, the bid price of each appliance, which is calculated through the *controller* object, is a function of two factors

- the average price and the standard deviation of the clearing price from the past markets
- the comfort zone of each appliance

Note that in figures 4.7 (on the left) and 4.9, there are time instances where the MO requests for more reactive power from generators that are not included in the market dispatch. In such a case the generators that offer only its reactive power have more profit because they can sell reactive power instead of nothing. In Figure 4.9, the grey lines denote the price for both active and reactive power, the blue lines the additional price per kVAr of reactive power, with power factor close to 1 and with orange or blue lines the case studies when the power factor for both diesel engines and wind turbines is changed.

Moreover, different cases where the power factor either from diesel generators or from wind turbines are changed to observe the effect of the power factor on the price of reactive power are considered. The power factor of diesel engine is set to 0.85 (blue line in Figure 4.9) and 0.97 (orange line in Figure 4.9) for the towns 1 to 5. For the City 1, the power factor for the five wind turbines is 0.97 and for the remaining four is 0.85 (blue line in Figure 4.9). No changes of the City 2 have been done regarding the power factor. For the rest of the towns, a combination of different values of the power factor of both wind turbines and diesel engines is considered. At the case where the power factor for both wind turbines and diesel engine is 0.85, it is observed (orange line for Town 8 and 10 in Figure 4.9) that the additional price is higher. When the power factor is less than 1 the price is higher because the cost of the losses and the opportunity cost of the producer have to be covered.

It is clear that for the Town 6 (Figure 4.9), which is associated with four large wind turbines when the power factor is 0.85 for two of the wind turbines and 0.97 for the last two, the price that consumers have to pay for reactive power is 0.576 $/kVAr. When the power factor is 0.97 for all the wind turbines, the price is 0.565 $/kVAr. It is worth to mention here that in Town 6 the transmission generators offer reactive power as well. In different cases, if the reactive power from the distributed generators is sufficient then the price is 0.452$/kVAr in the case where the power factor is 0.97 for the four wind turbines – the bid price of the local wind turbine.

In Figure 4.9 significant variations on the price in Town 9 can be observed. This is due to the fact that the transmission generators in that town bid at 0.38 $/kVA due to the operation logic of the controller. So, the variations in the price are due to the price that the consumers actually bid into the market which is affected by the high price of

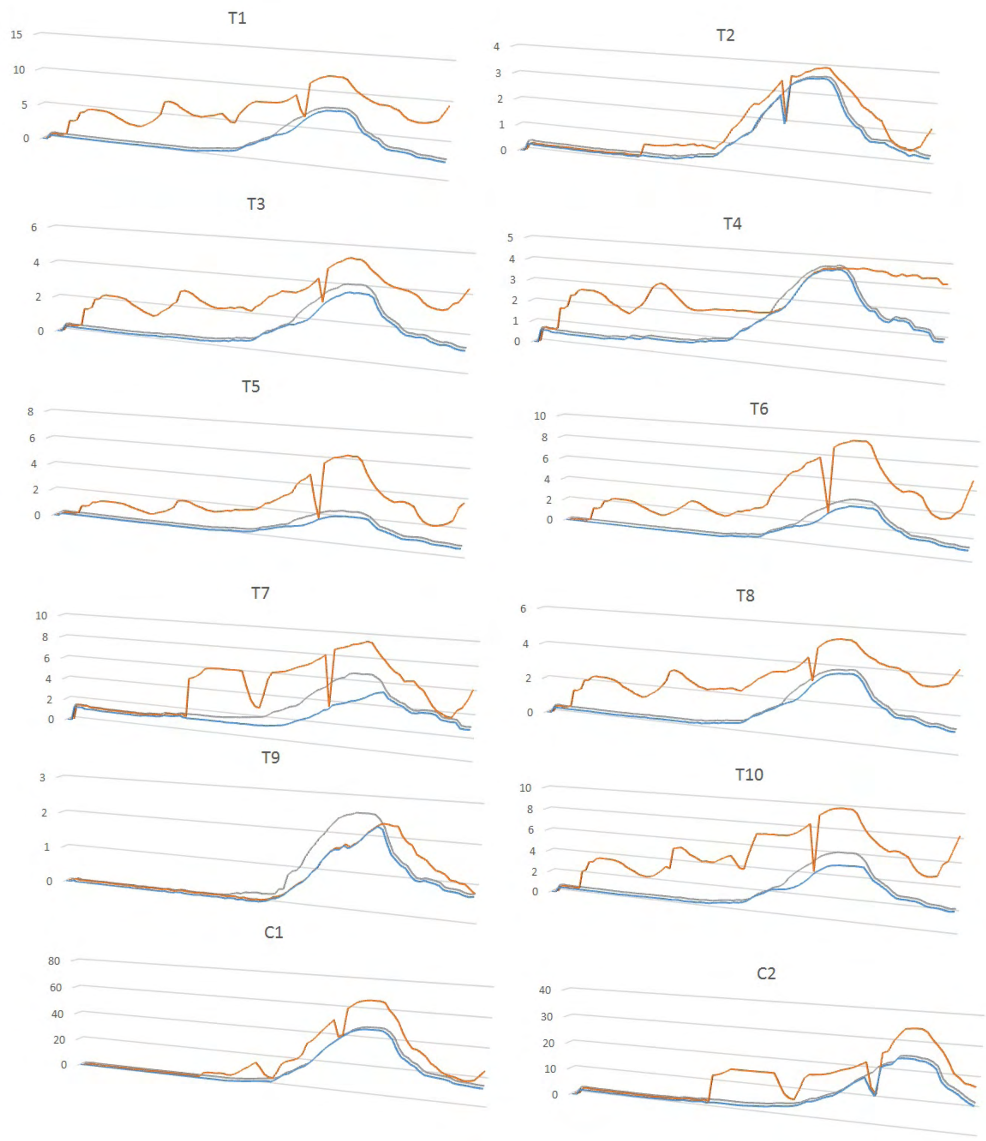

FIGURE 4.8: Demand, supply, and clearing quantities in the ten towns and the two cities for the Experiment II. In the x-axis and the y-axis, the simulation time (24 hours) and the amount of active power in MW are depicted accordingly.

the diesel engine. At Town 9, there are also variations on the price of reactive power. That town is equipped only with small wind turbines and one diesel engine and therefore the demanded reactive power is not satisfied from the generators that are included in the market. Similar results are observed at Town 4 where only one wind turbine and 2 diesel engines are installed. For the presented results, it is truth that the consumers have to pay less if there is local compensation of reactive power. If there is no distributed generators that can offer more reactive power the consumers have to pay more if there is need for more reactive power that has to "travel". So, the unsatisfied reactive power demand may lead the consumers to shift some of their consumption habits to the off peak hours. From the experiments, it is observed that there is need for more reactive power during the on peak hours. This is due to the influence of the wind turbine to the supply of the reactive power. Therefore, due to the stochasticity of wind turbines, the reactive power could not be exclusively provided from wind turbines. Instead distributed capacitors, diesel generators and synchronous condensers can be used.

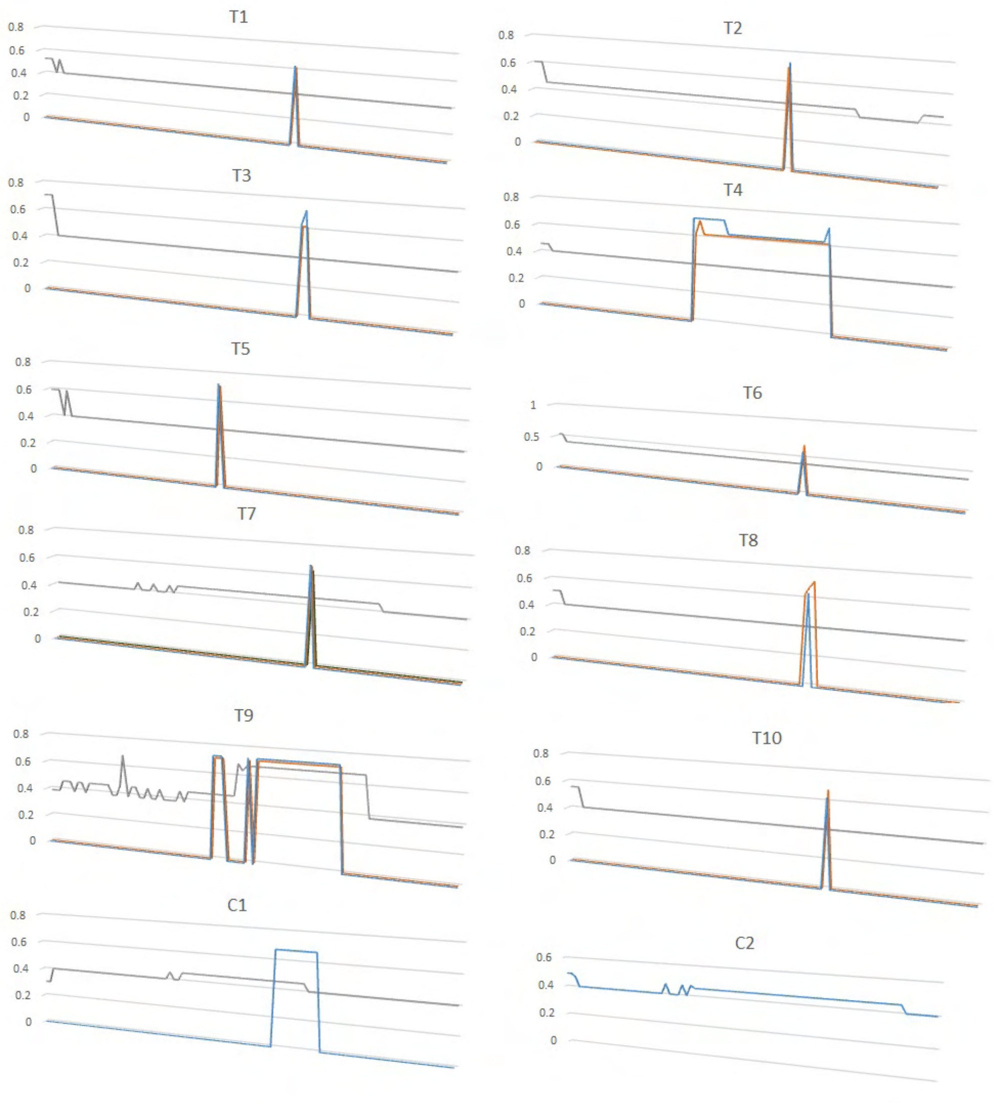

FIGURE 4.9: Clearing active and reactive prices per unit to the generators for Experiment II. In the x-axis and in y-axis, the simulation time (24 hours) and the price in \$/kVAr and the grey line corresponds to the price in \$/kVA are depicted accordingly.

This is due to the fact that the wind turbine has a significant influence on the supply of the reactive power. Therefore, due to the stochasticity of wind turbines, the reactive power could not be exclusively provided from wind turbines. Instead distributed capacitors, diesel generators and synchronous condensers from transmission network could be used. However, in the case of transmission generators, the price is a crucial factor.

4.4 Conclusions

In this chapter, the operation of the active and reactive power market is studied in details. Four different models for the active electricity market and a model which combines the characteristics of both the active and reactive power are developed. The primary objective is to (a) minimize the energy consumption during the peak hours and (b) show how the deficient supply of the reactive power affects the behavior of the consumers and more specifically the price that they have to pay. The simulating results concerning the four models of the active electricity market demonstrate that the load shifting to off-peak hours is achieved and it leads to a more stable and reliable system. The results of the model which makes use of game theory show that the producers can increase their profits. The coupling of active and reactive electricity market shows how the price is affected when the available reactive power is not enough to cover the demand of the reactive power. More specifically, the results show that in such a case the consumers have to pay more. For each model, a number of simulations to verify how each model works have been conducted.

Chapter 5

Partitioning of Distribution Grid

Continuous transmission and distribution of electrical power necessitates that the grid remains stable, reliable and efficient. Introduction of RES may induce disturbances and subsequent connectivity issues in the network due to their stochasticity. To that end, the presence of a huge disturbance may be handled by partitioning the grid at either the distribution or the transmission level, targeting at isolating the problem. However, partitioning should be done in an efficient and less costly way, i.e., isolating disturbance at a minimal portion of the network while avoiding pauses in electricity delivery. Occasionally, different clustering algorithms are used of partitioning the grid with the k-means to be the more promising in the clustering results. In this chapter, an extension of the standard k-means algorithm to improve its performance (Nasiakou, Alamaniotis, and Tsoukalas, 2016b) is presented. Additionally, the efforts on the grid partitioning using data either driven from the nodes (Nasiakou, Alamaniotis, and Tsoukalas, 2016c), (Nasiakou et al., 2017) or from the consumers of the distribution grid (Nasiakou, Alamaniotis, and Tsoukalas, 2017) are presented.

5.1 SkM, an improved version of the k-means algorithm

The most commonly used clustering algorithm in data mining applications is the k-means. Moreover, it is widely used in the smart grid energy management applications. However, the standard k-means has two main drawbacks that may affect the final clustering results. These drawbacks are (a) the random selection of the initial centroids or centers and (b) the a-priori definition of the number of clusters. It is of significance importance to the algorithm convergence that the k initial centroids are selected arbitrarily. To make it clearer, in standard k-means, the final centroids at the end of all the algorithm runs vary because of the random selection of the initial centroids. To that end, the clustering results differ after multiple runs of the k-means. The clustering results and the computation time can be improved using a heuristic based selection algorithm given that the initially selected centroids are likely to be closer to the real centroids compared to those taken by the random selection. The aim of the effort presented in this sub chapter is twofold:

1. the reduction of the computation time of the standard k-means algorithm
2. more compact and well-separated clusters compared to those provided by the standard k-means algorithm.

The proposed algorithm, called SkM, employs clustering validation measures (internal validation measures, the accuracy and the execution time) to prove its efficiency and its accuracy over the standard k-means algorithm. The SkM algorithm which actually extends the functionality, specifically the initialization step of the standard k-means algorithm, by adding a heuristic based selection of the initial cluster centroids is proposed. Extensive test cases for 2-D data points are considered to assess the validity and the efficiency of the proposed algorithm, whose features and characteristics are summarized as follows:

- The proposed algorithm uses the maximum value and the standard deviation of all the data points to calculate which of the data points could potentially be one of the initial cluster centroids.

- The proposed algorithm produces better clustering results as compared to the standard k-means, the k-means++ and the fuzzy c-means in the majority of the tested cases.

5.1.1 Related Work

Occasionally, several methods have been proposed in the literature for improving the performance of the k-means algorithm. Kathiresan and Sumathi (2012), for improving the selection of the initial cluster centers, proposed the Z-score factor to a set of sorted data points. Additionally, new methodologies focus on finding the proper initial cluster centers that ensure higher accuracy (Kathiresan and Sumathi, 2012) and less computation time (Manoharan and Ganesh, 2016). The authors in (Usman, U.Ahmad, and Ahmad, 2013) and (Vij and Kumar, 2012), using the information that can be extracted from the data points e.g. the average score, and the distribution of the data points, propose a naive algorithm for determining the initial cluster centroids.

Besides the aforementioned efforts, the algorithm proposed by the author M. Erisoglu (2011) uses a set of patterns formed by the data points as input of the clustering algorithm. More specifically, using the mean and the standard deviation of the data points, it tries to find better initial cluster centroids. It divides the data set into particular clusters and the similarity among the data patterns is assessed by using the density-based data condensation. The author Yedla (2010) uses two factors for the selection of the initial clusters centers and they are evaluated by creating a dual of axes in the data and by picking the max and min values from those axes. The author M. S. Mahmud (2012) proposes a simple technique close to one of the standard k-means for the selection of the initial cluster centers. The data points are assigned to those centers according to values obtained by a similarity measure. Once that ends, a process which computes the new centers takes place. The most significant improvement in the initialization step of the standard k-means algorithm is the k-means++ algorithm as described by David and Sergei (2007). The first cluster center is selected uniformly from the data points and every subsequent cluster center is selected from the remaining data points by computing the squared distance between each of it and the already selected cluster center. The next cluster center is selected using a weighted probability distribution function.

It has to be noted that in the literature there are a few efforts concerning the selection of the initial cluster centroids (called as initialization step) in order to increase the performance of the standard k-means algorithm using as measures, the accuracy, the execution time, or any other internal external validation measures. However, none of these efforts focuses on the compactness and the well-separation of the clusters, hence in this dissertation the validation based upon the accuracy of the proposed algorithm is considered.

5.1.2 Proposed Algorithm

The proposed heuristic based selection algorithm of the initial cluster centroids is implemented in C++. The implementation of the k-means algorithm used is the one provided by the dlib C++ library[1]. The data sets used for testing the proposed algorithm consist of 2-D data points, with both positive and negative values.

Before the execution of the proposed algorithm, some preliminary steps have to be conducted. The steps are the following:

- selection of the dimension which will be used for the clustering
- check if there are negative values
- find of the maximum value and computation of the standard deviation
- calculation of a representative value which will indicate the initial cluster center

It is worth to be noted at that point that, if there are data points with negative values, these values are replaced by positive ones. The procedure of replacing the negative values with positive ones is the following: the negative value is subtracted from the minimum value of the dimension which is chosen from the first step. Moreover, in the third step, if the data set has more than one data points with values equal to the maximum value, these data points are deleted from the data set. Hence, in that way the selection of the same cluster centroid or a close enough to the previous one between two different iterations is avoided. Using the trial-and-error technique, at each time a different number of data points close to the current center are removed; by computing the clustering performance with respect to the DB and S indices in order to avoid having cluster centers that are close enough to each other. Therefore, for all the data sets presented in this effort the lower bound of the number of the data points close to the selected cluster center that has to be deleted is two.

In the final step, a representative value, each for one of k initial cluster centroids is calculated based on the following formula:

$$c_i = max * (\frac{std}{N} - 1). \tag{5.1}$$

where i=1,...,k, with k to be the number of clusters, c_i is the initial cluster center, max is the maximum value of the under consideration data set, the std is the standard deviation of all the data points in the data set and the N is the number of data points.

Once the representative value c_i is computed, the data point, whose the value of the selected dimension is closest to this point, is selected to be the next cluster centroid. This procedure is executed for k times, as the number of clusters. Once the cluster centroids are selected, the k-means algorithm is applied to the data set. In particular, the algorithm works as described in Algorithm 1 where N is the number of data points, and k is the number of clusters.

Getting into more details, the new formula that computes the initial cluster centroids aims at including more information in the selection of the initial cluster centroids. First of all, the maximum value of the data points is used as a key factor to split the data set into groups. However, given that the factor max does not provide any essential information, it is multiplied by one more factor. This factor combines the standard deviation of all the data points and the number of the data points. The

[1]http://dlib.net/ml.html

Algorithm 1: SkM	
Input:	D={d_1, d_2,. . . .., d_N} – set of N 2-D data points
Output:	a set of k clusters – k is the number of clusters
Steps:	
1. Initialization step of the SkM algorithm, a. Transform the negative values to the positive ones by subtracting them with the minimum value of the under consideration dimension of the data set b. Find the maximum value of the under consideration dimension and delete all the data points with the same value in order to avoid having cluster centroids close to each other c. Find and take the absolute of a representative initial cluster centroid based on equation 5.1 d. Find the data point which is closest to the data point of the step **c**. This data point is one of the k initial cluster centroids e. Delete two more data points which are close to the initial cluster centroids f. Repeat step **c** and **d** for k times	
2. Assign each data point of the initial data set to the proper cluster using the standard k-means algorithm	

standard deviation includes information about the variation of the data points; if the standard deviation is small then the data points are close to the mean, while the high value of the standard deviation means that the data points are spread. So, the use of the standard deviation provides information about how spread or concise the distribution of the data points is. By dividing the standard deviation by the number of data points, the contribution of each data point to the overall standard deviation is computed. Moreover, the addition of the 1 to the $\frac{std}{N}$ factor is chosen for taking a value close to 1; the factor $\frac{std}{N}$ is close to 0 when the number of data points is too big. Therefore, by multiplying the standard deviation with the maximum value of the data points, the result of the equation 5.1 gets a high value if the standard deviation is low or gets a low value probably close to the maximum value if the standard deviation is high. To that end, the maximum valued is not used as the representative value which may give wrong clustering results. So, by considering the standard deviation in the proposed algorithm, information about the spread of the data points is taken into account. Moreover, when a cluster center is selected, the data point which represents this cluster center is deleted from the data set.

Some detailed examples of how the selection of the initial clusters is performed are presented in the next lines. The following simple data set is utilized:

(1,1), (1.2,1), (2,1), (3,1), (4,1), (5,1), (6,1), (7,1), (8,1), (9,1), (99,1), (98,1), (97,1), (100,1)

The number of data points in this data set is 14, the number of clusters is 3, and the selected dimension is the x. The representative values calculated based on the equation 5.1 for the three cluster centers are the **(203.40), (186.78), (0.50)**. The cluster centers are the **(100,1), (97,1), (1,1)** respectively. This case can be considered as an extreme case because the data points have both small and large values.

In the following case, the data set consists only of one data point with big value. The data points are the following:

(1,1), (1.2,1), (2,1), (3,1), (4,1), (5,1), (6,1), (7,1), (8,1), (9,1), (100,1)

The number of data points in this data set is equal to 11, the number of clusters is 3, and the selected dimension is the x. The representative values calculated based on the equation 5.1 for the three cluster centers are the **(150), (1.67), (0.35)**. The cluster centers are the **(100,1), (2,1), (3,1)** respectively. The reason for selecting as cluster center the data point (3,1) when the representative value is (0.35) is that two more data points close to the selected cluster center are deleted. So, it can be observed that the data point selected as a cluster center is not far from the representative value calculated by the equation 5.1.

5.1.3 Test Cases and Results

Datasets

For testing purposes, two-dimensional data points consisting of both negative and positive values are selected. The data sets retrieved from the UCI machine learning repository[2]. Table 5.1 depicts the labels of the clusters for the 20 data sets, the number of the data points and the number of clusters as well. The first dimension of the data points is associated with the **Quantity** of the energy and the second one with the unit **Price** based on the information given by the UCI library. The dimension which is used for the clustering is selected randomly. It should be noted that the labeling of the data points is known *a priori* (i.e., the ground truth).

Test Cases Results

The comparison results of the standard k-means and the proposed algorithm are presented in the next lines. Moreover, the clustering results of the proposed algorithm are also compared with the clustering results given from the fuzzy c-means and the k-means++ algorithms. The algorithms were run on a Dell Desktop Computer with Intel Core i7-2600 CPU @ 3.40 GHz processor and 8.0 GB RAM. It is observed from the results of the presented test cases that even for a different number of data points, in most of the cases the SkM algorithm provides i) higher accuracy and ii) lower execution time. Besides the compactness and the well separation of the data points, the computation/execution time is also considered due to the fact that the assignment of the data points to the nearest cluster center is conducted at every single iteration; this is a time-consuming procedure.

For assessing the validity of the proposed algorithm, the SkM algorithm to various data sets with different data points (as shown in Table 5.1) using validation indices which focus on the compactness and the well separation of the clusters is tested. These indices are the DB and S. The accuracy and the execution time of both the k-means and the SkM algorithm are computed as well. The standard k-means algorithm does not guarantee that provides the optimal solution. In each of different run of the standard k-means, the initial cluster centers were different. To that end, the standard k-means runs for 5 times for each data set. The value of both the accuracy and the execution time of the standard k-means algorithm is the average value of the accuracy and the average execution time of the 5 runs, respectively. The accuracy of the clustering results is determined by comparing the obtained from the SkM and the k-means algorithm results to the clustering information that the UCI library provides.

More specifically, the accuracy is calculated as follows: if a data point belongs to the same cluster indicated by both the SkM algorithm and the UCI library then an index, which is associated with the accuracy, is increased by 1. This procedure runs iteratively until this index is evaluated as many times as the number of the data

[2]https://archive.ics.uci.edu/ml/datasets.html

TABLE 5.1: Configuration of the data sets

# of data points	# of clusters	Labels of Clusters
50	3	Australia, Portugal, Switzerland
98	3	France, Germany, EIRE
105	3	Australia, Portugal, Switzerland
150	3	Channel Islands, Denmark, EIRE
200	3	United Kingdom, Germany, France
272	3	France, United Kingdom, Germany
303	3	United Kingdom, France, Australia
356	3	United Kingdom, EIRE, France
386	3	Italy, United Kingdom, Portugal
409	3	United Kingdom, Germany, Australia
489	3	United Kingdom, Germany, EIRE
602	3	United Kingdom, Spain, Poland
1000	3	United Kingdom, Spain, Poland
1034	3	RSA, Finland, United Kingdom
1152	3	EIRE, United Kingdom, Germany
1177	3	United Kingdom, Switzerland, EIRE
1628	3	Germany, Norway, United Kingdom
1738	3	Australia, France, United Kingdom
1902	3	United Kingdom, France, EIRE
3271	3	United Kingdom, Norway, Australia

points. For calculated the accuracy of the k-means algorithm the same procedure is conducted. Getting into more details, this index indicates how many times the SkM or the standard k-means algorithm assignments coincide with the ground truth.

The results of the SkM, the standard k-means, k-means++ and the fuzzy c-means based on the DB index are presented in the Table 5.2. Table 5.3 depicts the results of the SkM algorithm over the standard k-means algorithm based on the accuracy, the execution time and the S index.

It should be noted that in Table 5.2 the underlined cells indicate that the associated algorithm gives better performance with respect only to the DB index comparing with the SkM algorithm. In Table 5.3 the bold cells indicate the worse performance of the SkM algorithm as compared to the respective performance of the standard k-means, k-means++ and fuzzy c-means algorithm. The lower the value of the DB index, the better clustering results the algorithm provides. A high value of the S index indicates that the data points match to the assigned cluster and they do not match to the other clusters. If its value is close to 1 the best partitioning of the data points is truth. The S index indicates the consistency and the compactness within the cluster and it is commonly used to provide the optimal number of clusters essentially maximizing the separation measure.

The SkM provides better accuracy and execution time in the 18 tested cases out of the 20; an error of 10% for both the accuracy and the execution time. Besides the accuracy and the execution time, the SkM algorithm, compared to the other algorithms gives better clustering results according to the DB index. In Table 5.2, it is observed that the DB index gives better performance in 18 out of 20 cases; the SkM prevails of standard k-means in 90% of the test cases. On the other hand, the S index provides better performance in 16 out of 20 cases; an error equals to 20%.

TABLE 5.2: Comparing results among the SkM, the k-means, the k-means++ and the fuzzy c-means based on DB index

# of data points	DB			
	SkM	k-means	k-means++	Fuzzy c-means
50	0.49	0.55	0.50	0.50
98	0.35	0.58	0.73	0.65
105	0.52	0.57	0.62	0.51
150	0.96	0.46	0.57	0.41
200	0.37	0.53	0.46	0.52
272	0.42	0.52	0.48	0.48
303	0.80	0.68	1.28	0.82
356	0.32	0.48	0.45	0.42
386	0.32	0.59	0.5	0.52
409	0.24	0.50	0.66	0.44
489	0.15	0.68	0.62	5.2
602	0.12	0.54	0.48	0.49
1000	0.36	0.20	0.20	0.36
1034	0.27	0.43	0.4	0.28
1152	0.08	0.23	0.41	0.23
1177	0.28	0.33	0.55	0.007
1628	0.05	0.12	0.09	0.09
1738	0.03	0.28	0.35	0.35
1902	0.09	0.28	0.41	0.15
3271	0.12	0.41	0.36	0.37

TABLE 5.3: Comparing results between SkM and k-means based on accuracy, execution time and S index

# of data points	Accuracy		S		Execution Time	
	SkM	k-means	SkM	k-means	SkM	k-means
50	0.86	0.76	0.80	0.72	0.017	0.036
98	0.52	0.20	0.67	0.67	0.028	0.025
105	0.83	0.71	0.56	0.54	0.02	0.03
150	0.49	0.46	**0.54**	0.68	0.029	0.041
200	0.48	0.48	0.91	0.76	0.033	0.04
272	0.92	0.90	0.96	0.88	0.04	0.05
303	0.92	0.07	**0.82**	0.87	0.044	0.059
356	0.90	0.82	**0.73**	0.76	0.046	0.062
386	0.88	0.87	0.91	0.06	0.066	0.076
409	**0.22**	0.67	0.87	0.70	**0.094**	0.07
489	0.86	0.79	0.90	0.67	0.06	0.09
602	0.53	0.53	0.93	0.82	0.107	0.12
1000	**0.92**	0.99	0.99	0.99	0.113	0.13
1034	0.85	0.70	0.77	0.70	0.12	0.146
1152	0.98	0.97	0.99	0.99	0.13	0.16
1177	0.98	0.98	0.95	0.93	0.13	0.15
1628	0.83	0.74	**0.70**	0.97	0.17	0.18
1738	0.82	0.74	0.97	0.97	0.18	0.20
1902	0.99	0.99	0.99	0.97	0.205	0.23
3271	0.86	0.86	0.99	0.93	0.317	0.35

To achieve the goal of this effort, the SkM algorithm is compared with the k-means++ and the fuzzy c-means. The DB index is used to validate the SkM algorithm over the k-means++ and the fuzzy c-means because this index focuses only to the compactness and the well-separation of the clusters while the S index focuses to the compactness of the clusters and the optimal selection of the number of clusters. Comparing the SkM with the k-means++, the former has better performance in 18 out of 20 cases; an error of 10%. Additionally, the SkM has better performance in 17 out of 20 cases (this is an error equal to 15%) over the fuzzy c-means algorithm.

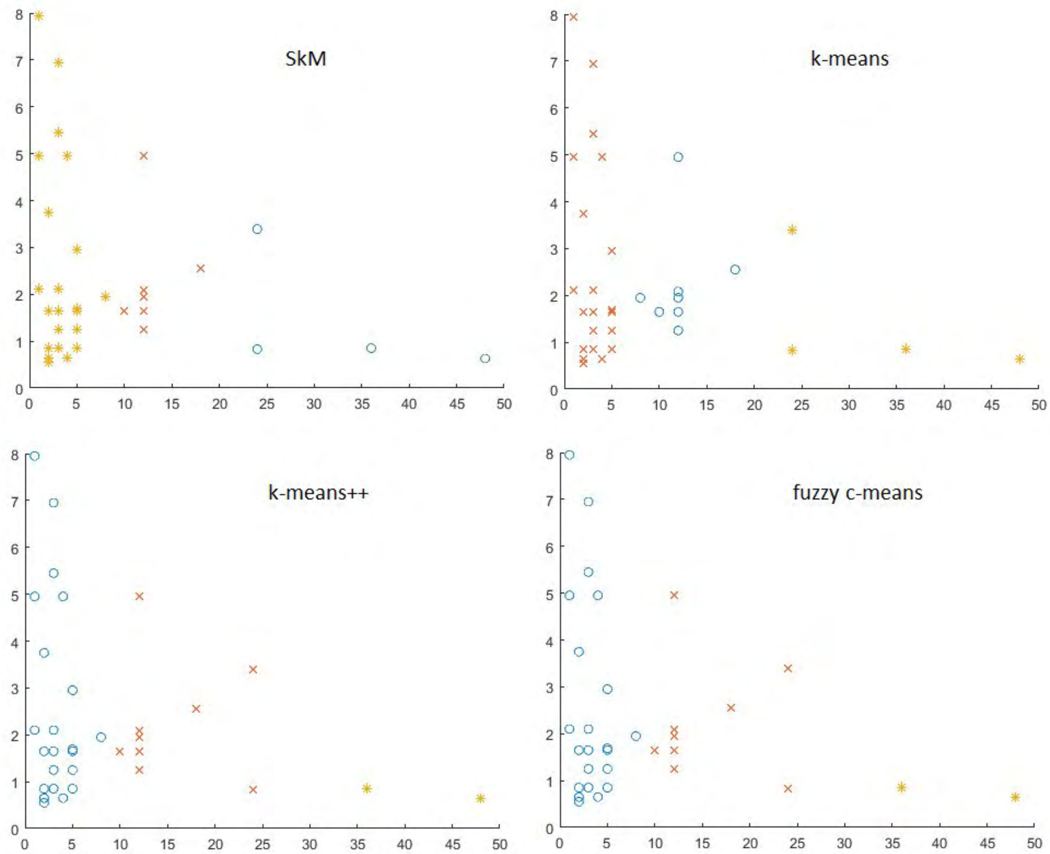

FIGURE 5.1: Clustering Results of the four algorithms for the case of 50 data points.

In figures 5.1 and 5.2, some of the tested cases, which give both better and worse performance over the standard k-means algorithm, the k-means++ and the fuzzy c-means with respect to the DB and S indices, are depicted. At both figures, 5.1 and 5.2, each sub figure depicts the clustering results for the SkM, k-means, k-means++ and fuzzy c-means in clockwise rotation. More specifically, in Figure 5.1 the clustering results of the case where the data set has 50 data points for the four algorithms are presented. It can be observed that the results of the SkM provide well-separated clusters comparing to the ones of the standard k-means algorithm. More specifically, the main difference is that the data point (9, 2) belongs to different clusters. This affects positively both the values of the DB and S indices for the SkM algorithm. This means that this data point matches better in the cluster that the SkM algorithm is assigned to rather than the other three algorithms do. SkM algorithm also has better performance over the k-means++ and the fuzzy c-means algorithms which can also be observed due to the fact that for both algorithms there exists a cluster with two data points which may affect the values of the DB and S indices.

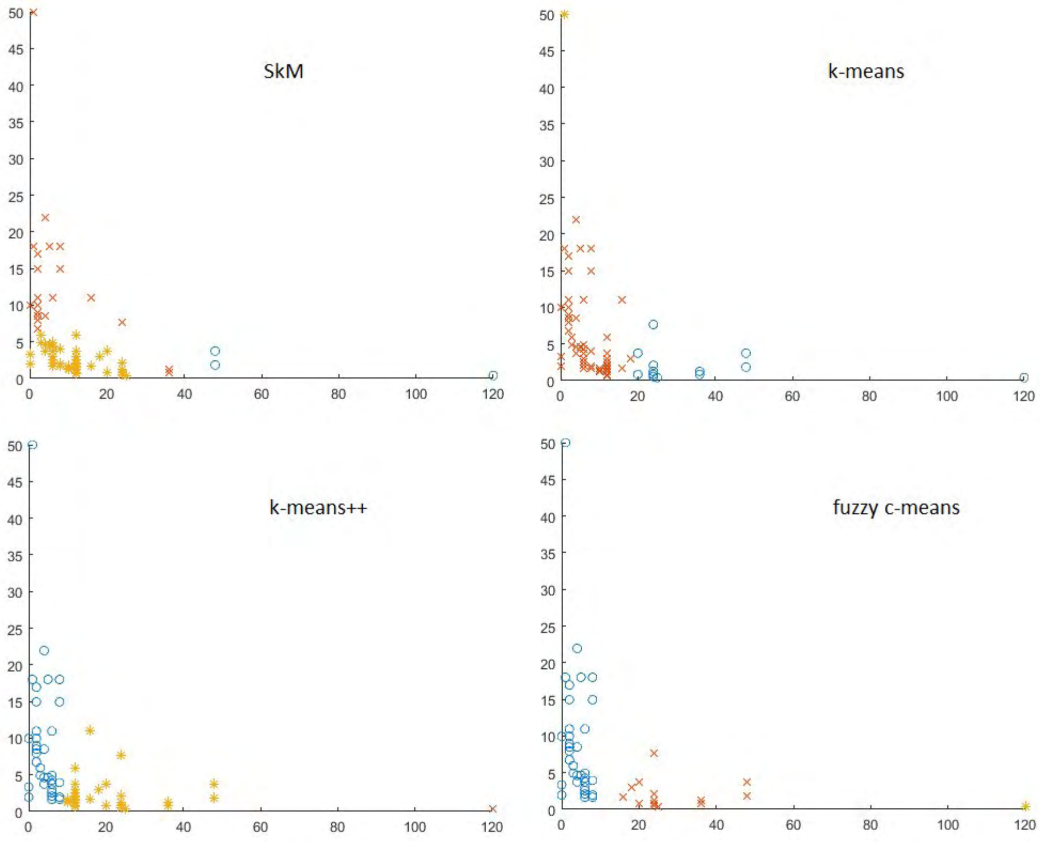

FIGURE 5.2: Clustering Results of the four algorithms for the case of 150 data points.

In Figure 5.2, the clustering results of the case where the data set has 150 data points with each of the four algorithms are presented. This test case corresponds to the case in which the SkM has worse performance than the other three algorithms. The k-means, the k-means++ and the fuzzy c-means provide clustering results with high values for both the DB and the S indices. This is due to the fact that the boundary data points (1,50) for the k-means and (120,1) for both the k-means++ and fuzzy c-means are indicated as separated clusters. On the other hand, the SkM algorithm assigns these data points to more populated clusters. Therefore, the distance between each of them and the other members of the clusters is long enough, leading to the reduction of the overall performance. This can be clearly observed from the low values of the DB and S indices of the SkM algorithm given that (a) the value of the S index is the distance between a data point and all the data points included in each cluster and (b) the value of DB index is the distance between each data point and the cluster center.

5.2 Nodal partitioning of distribution grid

A Dynamic Data Driven Application Systems (DDDAS) methodology for partitioning the smart distribution grid based on dynamically varying data is presented in this sub chapter. The proposed methodology uses the SkM algorithm presented in the previous sub chapter for performing partitioning and a fuzzy decision-making method for increasing the power efficiency and reliability in a distribution grid. The grid is divided in similar, regarding the energy characteristics of the nodes, sub grids who share the same characteristics pertaining to the energy needs but not necessarily the same geographic vicinity. A fuzzy logic decision-making method is utilized to make decisions on which of the resultant partitions by the partitioning procedure could be offered the energy available from RES at lower prices. Scenarios based on the GridLAB-D simulation platform exhibits how the operation of the smart grid is affected by the partitioning of the distribution grid. The connection of the IEEE-13, IEEE-37, and IEEE-123 test feeders into a single node composing a grid of 3004 residencies with both conventional and distributed generators is utilized for the presented experiments.

The proposed methodology consists of two main phases. The **first phase** is associated with the grid partitioning using the SkM algorithm. The SkM uses as input data representing the current state of the power grid. The members of the resultant partitions are the nodes of the distribution grid. The partitioning procedure is based on three factors which are dynamically driven: (a) the total load of all the residencies which are connected to each node of the distribution grid, (b) the average value of a factor called schedule skew (each appliance is characterized by this factor), that indicates the estimated time in seconds the householder "starts" his consumption in a given time interval, and (c) the total number of appliances in each node that are in operation the under consideration time instance.

The **second phase** is associated with the use of a fuzzy logic decision-making method which uses information (a) about the resultant partitions of the previous phase, (b) the number of appliances in operation and (c) the degree of satisfaction of the overall demand. The output of this method is the partitions which will be the recipient of offered benefits (usually in the form of price incentives); such that the stability of the power grid is also ensured.

5.2.1 Related Work

The DDDAS concept entails "the ability to dynamically incorporate data into an executing application simulation, and in reverse, the ability of applications to dynamically steer measurement processes", creating "application simulations that can dynamically accept and respond to 'online' field data and measurements and/or control such measurements" paradigm which was presented by Darema (2000).

The creation of a microgrid into the power grid is one of the most widely used and safest methods to integrate RES into the power grid. For example, for reducing the nine billion dollar cost of satisfying the demand for its installations, the US could adopt RES following a two stage decision-making method for reducing its cost as proposed by Moore and Celik (2014). A set of microgrids, equipped with RES and distributed communications, could provide energy resources to meet the level of demand satisfaction. The proposed method is based on a DDDAS methodology which allows a simulation model retrieving data from a specific database to dynamically predict the system's reaction under particular conditions. Moreover, Shi, Damgacioglu, and Celik (2015) proposed a DDDAS approach for operation planning of microgrids. It consists of three main components: (a) a database with information about the state of the power network (total demand, market price) and weather conditions, (b) a simulation platform for simulating the microgrid and (c) a multi-objective optimization problem. Based on the DDDAS concept, Thanos et al. (2015) propose a novel framework which consists of three main components associated with (a) the information about the grid and the weather, (b) an agent based simulation platform and (c) an optimization algorithm. The goal of this study was the development of a framework for microgrids focusing on the minimization of the computation time of the resource allocation and on the optimization of the operational costs. Additionally, Celik, and Saenz (2013) and (Thanos et al., 2014) developed a DDDAS framework for the power network to solve the economic load dispatch problem.

To increase the volume of the data for a set of sensors for the surveillance and crowd control through is proposed by (Khaleghi et al., 2013). This approach is conducted via a dynamic data driven adaptive multi-scale simulation framework. Blasch, Al-Nashif, and Hariri (2014) focused on retaining a balance between dynamic data driven solutions and static solutions. Blasch, Seetharaman, and Darema (2013) formulated the prediction of wide-area motion imagery data as a DDDAS problem. DDDAS applications are of paramount importance because they provide accurate analysis and predictions, more reliable and efficient outcomes, and more accurate controls.

5.2.2 Proposed Methodology

Both partitioning and selection procedures are dynamically driven and depend on the current conditions of the grid. The partitioning of the power grid forces the nodes with similar energy consumption patterns to be in the same partition. Moreover, in each run, the selected partition by the fuzzy logic decision-making method consists of different nodes. So, in each run of the proposed methodology different nodes take the privileges of buying energy at lower prices.

The two phases of the proposed methodology are the following:

1. The SkM is utilized to partition into three partitions the distribution grid using the dynamically driven data from each node. The nodes that are considered for the partitioning procedure are the ones in which residencies are attached to them. The other nodes are used for connection purposes.

2. The proposed fuzzy logic method is applied once the partitioning procedure ends and when the clearing quantity providing by the market is not enough to satisfy the overall demand the specific time period. In other words, when the demand is satisfied with the quantity provided by the market framework, the fuzzy logic decision-making method is not applied; only the partitioning procedure is applied but its results have not been used. On the other hand, when the demand is not fully satisfied, the energy produced by the RES that is not being dispatched will be offered at a lower price to the appliances of the partition which is selected by the fuzzy logic decision-making method.

The proposed FLS, utilized in this effort, uses as input:

- the total load of all the residencies attached to the nodes that are indicated as cluster centers after the partitioning procedure, named **Load**. Each node that accommodates residencies can potentially be one of the next cluster centers. The corresponding variable has three fuzzy sets; LOW, MEDIUM and HIGH.

- an index **Devices_ON** which indicates the aggregated number of appliances in operation of the nodes which are the cluster centers at the particular time that the proposed methodology is applied. The corresponding variable has two fuzzy sets; LOW and HIGH.

- an index, named **Satisfaction**, with values range from 0 to 2. A value in the range [0 1] corresponds to the satisfaction of the overall demand while a value in the range [1 2] that the demand is not full satisfied. The former fuzzy set is indicated as SATISFY while the latter is indicated as NO_SATISFY.

The output of the FLS is the partition that is selected. The output variable, called **Priority**, indicates which of the partitions has priority over the other partitions. Regarding the **Satisfaction** index, it is the ratio of the clearing quantity to the total appliances' quantity of each partition. If it is greater than 1, then the demand is satisfied and the value of the index is the inverse value of that ratio. On the other hand, if it is lower than 1, then the ratio's value is normalized between 1 to 2. If the new value is greater than 2 then the index's value is equal to 2 and if the normalized value is lower than 1 then the normalized value is accumulated by 1. The FLS implements the following fuzzy rules for associating the input variables **Load**, **Satisfaction** and **Devices_ON** to the output variable **Priority**:

- IF Load is LOW and Satisfaction is SATISFY and Devices_ON is LOW, THEN Priority is LOW
- IF Load is LOW and Satisfaction is NO_SATISFY and Devices_ON is LOW, THEN Priority is LOW
- IF Load is MEDIUM and Satisfaction is SATISFY and Devices_ON is LOW, THEN Priority is LOW
- IF Load is MEDIUM and Satisfaction is NO_SATISFY and Devices_ON is LOW, THEN Priority is MEDIUM
- IF Load is HIGH and Satisfaction is SATISFY and Devices_ON is LOW, THEN Priority is LOW
- IF Load is HIGH and Satisfaction is NO_SATISFY and Devices_ON is LOW, THEN Priority is HIGH
- IF Load is LOW and Satisfaction is SATISFY and Devices_ON is HIGH, THEN Priority is LOW
- IF Load is LOW and Satisfaction is NO_SATISFY and Devices_ON is HIGH, THEN Priority is MEDIUM
- IF Load is MEDIUM and Satisfaction is SATISFY and Devices_ON is HIGH, THEN Priority is MEDIUM
- IF Load is MEDIUM and Satisfaction is NO_SATISFY and Devices_ON is HIGH, THEN Priority is HIGH
- IF Load is HIGH and Satisfaction is SATISFY and Devices_ON is HIGH, THEN Priority is MEDIUM
- IF Load is HIGH and Satisfaction is NO_SATISFY and Devices_ON is HIGH, THEN Priority is HIGH

The membership functions for all the input and output variables are based on the triangular form as depicted in figures 5.3, 5.4, 5.5 and 5.6. The FLS runs independently for each partition (see Figure 5.7), and the partition with the highest value of the variable **Priority** is selected.

The proposed DDDAS methodology is executed iteratively until the simulation time ends. More specifically, in each run of the proposed methodology, the total

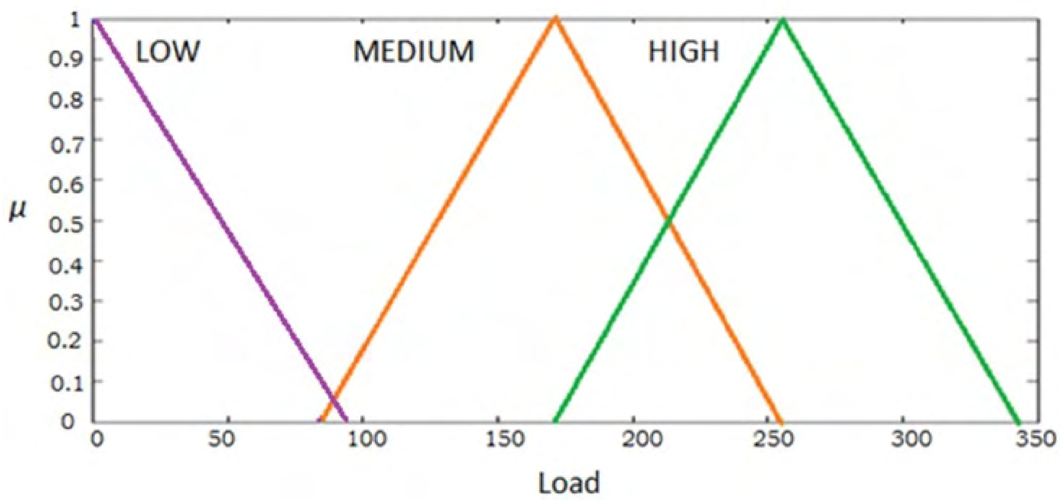

FIGURE 5.3: Membership function of the Load input variable.

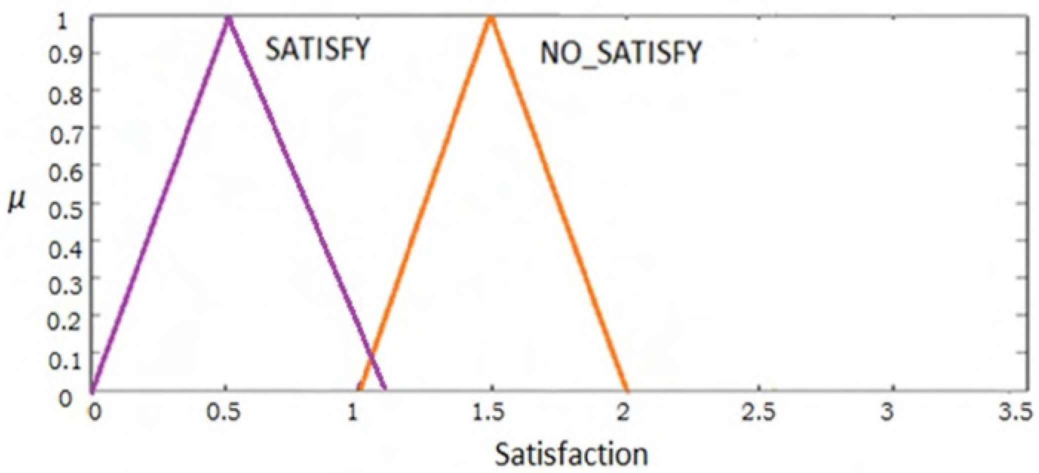

FIGURE 5.4: Membership function of the Satisfaction input variable.

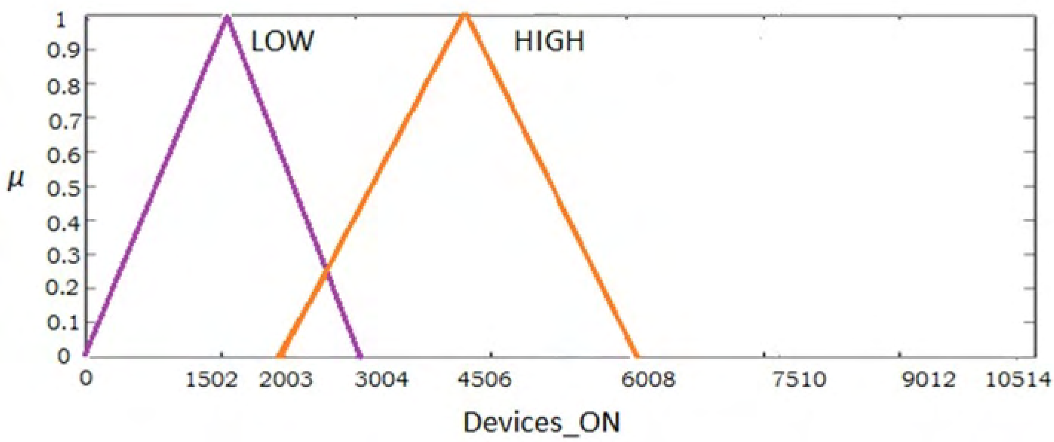

FIGURE 5.5: Membership function of the Devices_ON input variable.

load of all the appliances belong to each node, the total number of appliances in operation, the factor schedule skew, the Satisfaction index and the number of clusters

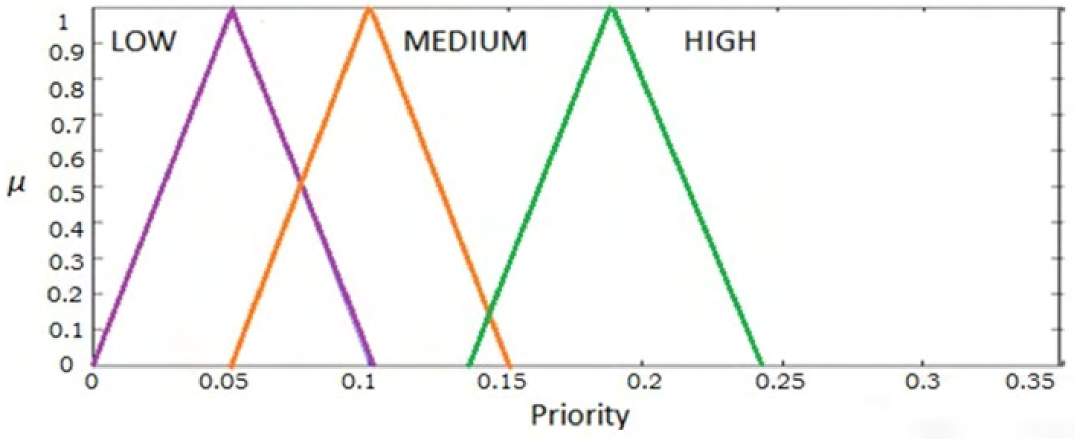

FIGURE 5.6: Membership function of the Priority output variable.

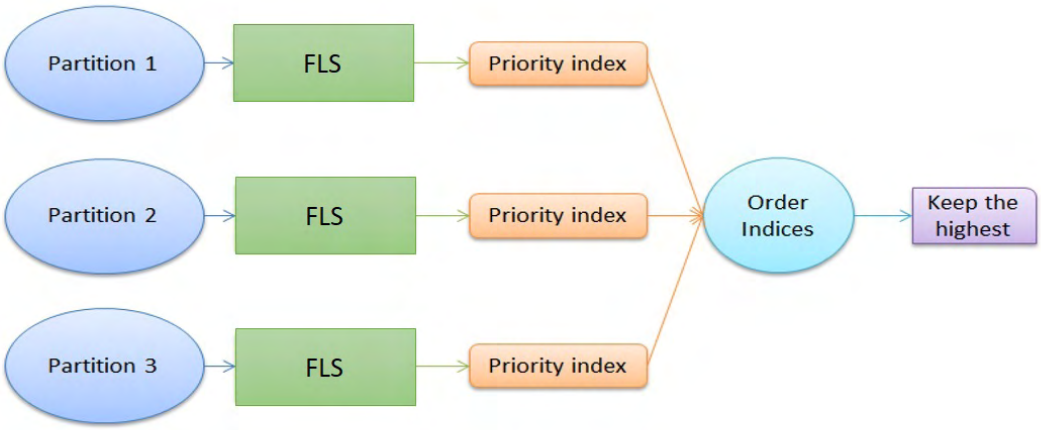

FIGURE 5.7: Architecture of a fuzzy logic inference system.

are needed for the execution of the proposed methodology (partitioning and fuzzy logic decision making method procedures). The data points used by the SkM algorithm are 2-D. One dimension corresponds to the sum of the aggregated load with the schedule skew factor while the second one corresponds to the number of appliances in operation. In other words, the first dimension is associated with the load at a specific node and the second one with the number of appliances of the same node that need the specific amount of load.

The fuzzy logic decision-making method provides which of the resultant partitions can take the privilege of buying energy at lower prices. The price of produced by the RES energy which is not being dispatched is defined to be 50% lower than the initial one. If the price of RES is reduced 50%, the new price is 0.05 $/kWh; 50% lower of the average price of RES. This price is lower than the clearing price and the price of the conventional natural-gas (see Table 5.4) and probably is closer to the preferred by the appliances' price. Therefore, the final price, that the appliances are called to pay for, is the average value of two prices: (a) the reduced price from RES when they belong to the selected partition, and (b) the market clearing price.

5.2.3 Simulation Results

The distribution grid is composed of 3004 residencies, 4 set of solar panels, 3 wind turbines, and 3 conventional natural gas based generators. Each residency

TABLE 5.4: Capacities-prices of the producers

Generators	Price	Max Capacity (kW)
Solar panel (116,000 sf)	0.11$/kWh	1917
Wind Turbine	0.10$/kWh	5500
Transmission G1	0.07$/kWh	3300
Transmission G2	0.07$/kWh	3000
Transmission G3	0.07$/kWh	3000

consists of non-thermostatically controlled appliances, like a refrigerator and thermostatically controlled appliances, like an HVAC system and a water heater. All the appliances participate in the market.

Regarding the producers, both transmission and distributed generators are simulated. Each RES or conventional generator acts as a different energy provider with his own offer price. The distributed generators are associated with RES. Four solar panels of total area 116,000 sf are simulated. The max power output of the solar panels is 1917 kW while the max power output of the three wind turbines is 5500 kW. Two types of wind turbines are used; the GE_25 and the GENERIC_SYNC_SMALL. Details about their configuration and their parameters can be found at the GridLAB-D's user guide[3]. In Table 5.4, the operating costs for each type of generator including the max power output as well are depicted.

For comparison purposes and for proving the dynamic nature of the proposed methodology, the simulation results of the three cases: when the proposed methodology is applied every one hour, called **Case I**, and when it is not applied, called **Case II** and when it is applied every four hours, called **Case III** are going to be presented, compared with other algorithms and analyzed.

For all of the presented test cases, the simulation time is 24 hours, and the simulation day is 08/01/2000. The market period is 15 minutes. The Table 5.5 presents the average total load of all the appliances belonging to each partition, for each particular time the fuzzy logic procedure is executed. The information depicted in this Table 5.5 corresponds to the **Case I**. It has to be declared at that point that the fuzzy logic decision-making method is executed when the demand is not fully satisfied with the quantity dispatched from the market framework. Because of that, in Table 5.5 the results are depicted every 2 hours and not every 1 hour as the **Case I** is characterized for. The partitions that are selected from the fuzzy logic decision-making method are presented in bold font.

In figures 5.8 and 5.9, the HVAC load of two residencies that buy energy at lower prices (**Case I**) and the HVAC load of the residency when our methodology is not applied (**Case II**) are depicted. The x-axis and the y-axis are associated with the simulation time and the HVAC load (in kW) respectively. It is clear that the HVAC system works more frequently in the case of applying the proposed methodology (**Case I**). Due to the fact that the HVAC system needs more energy to meet its energy requirements, it is the appliance that has been choosen to show the effects of the proposed DDDAS methodology. Moreover, it is observed that the operation of the HVAC system is shifted from the on-peak hours to the off-peak hours.

The clearing quantity for a whole day for both the **Cases I** and **II** is presented in Figure 5.10. The blue line is associated with the **Case I** while the red dots with the **Case II**. During the hours of applying the proposed methodology the clearing quantity of the Case I is much higher. The total load of each residency associated

[3] http://www.gridlabd.org/documents/doxygen/1.1/files.html

TABLE 5.5: The average load for the partitions for each run where the fuzzy logic decision making method takes place.

Time	Partition #1	Partition #2	Partition #3
2:00am	26.82	167.25	**61.52**
4:00am	165.62	26.44	**41.01**
6:00 am	115.67	31.98	**49.94**
8:00am	203.2	**26.16**	127.57
10:00am	7.74	16.94	**65.63**
12:00pm	8.68	**20.70**	5.27
2:00pm	4.44	**64.80**	15.23
4:00pm	17.14	**110.57**	28.28
6:00pm	**94.58**	138.26	25.39
8:00pm	**27.31**	160.25	69.16
10:00pm	78.22	161.60	**78.60**
12:00am	30.46	149.94	**90.28**

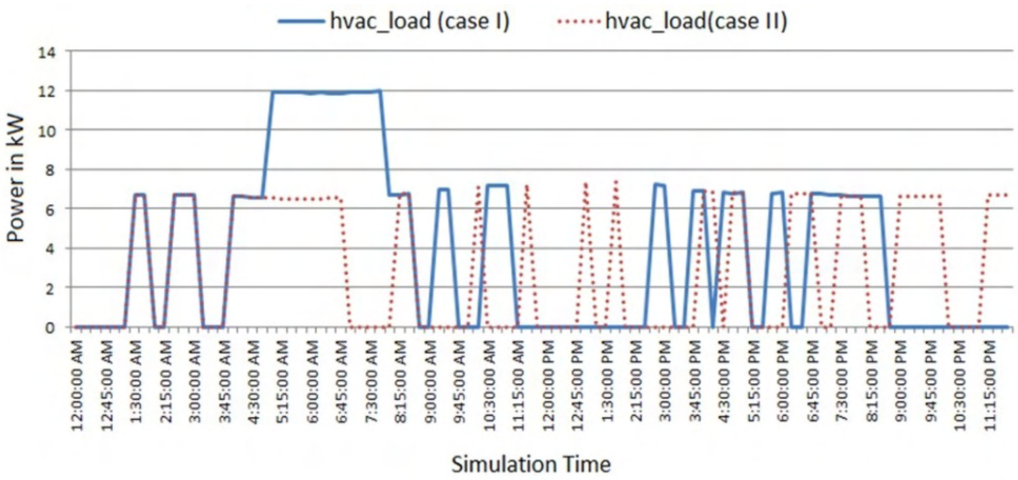

FIGURE 5.8: HVAC load for a residency that benefits from the lower prices in some time intervals when the partitioning process is taking place every one hour.

with the figures 5.8 and 5.9 of both the **Cases I and II** are depicted in Table 5.6. It has to be noted that the residential consumers have the opportunity to by energy at lower prices. More specifically, the consumer that belongs to the partition selected by the fuzzy logic decision-making method can buy one kWh at 0.061 $/kWh instead of 0.07 $/kWh.

TABLE 5.6: Comparison of the cost and the load of a residency for Case I and II.

		Case I	Case II
Total load(kWh)/Cost($/kWh)	Residency Fig. 5.8	7464 / 0.061	7176 / 0.07
	Residency Fig. 5.9	5184 / 0.058	4584 / 0.07

In Figure 5.11, it is clear that despite the fact that there is non-dispatched energy, this amount of energy is comparatively less. It is observable, that the clearing quantity, when the proposed methodology, is applied is closer to the supply curve

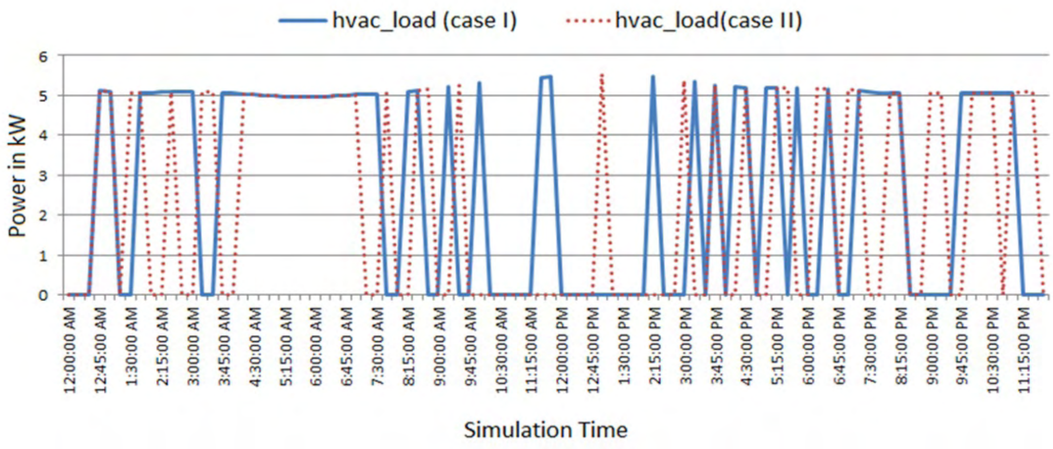

FIGURE 5.9: HVAC load for a residency that benefits from the lower prices in some time intervals when the DDDAS proposed methodology is taking place every one hour.

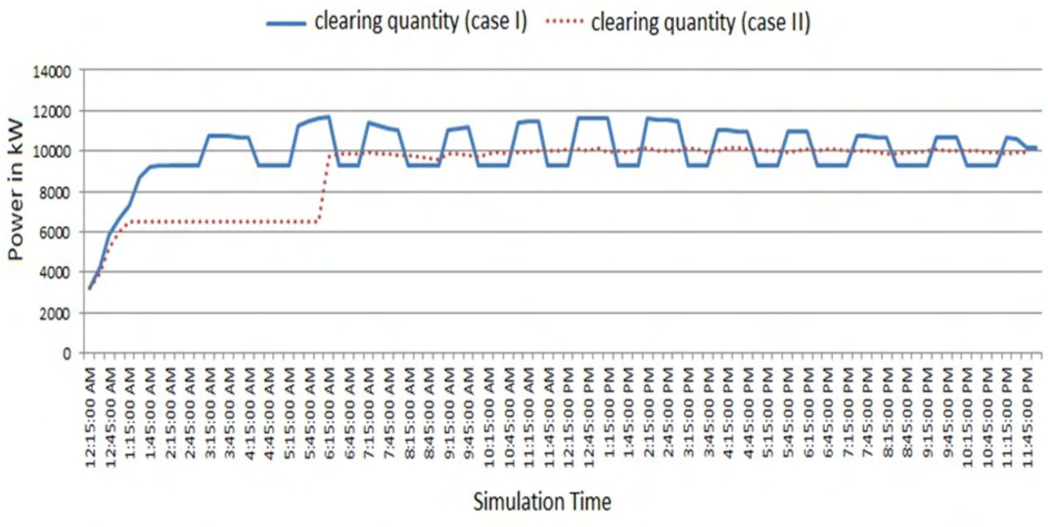

FIGURE 5.10: Clearing quantity offered by the market for the Case I and Case II.

(sellers' total quantity in Figure 5.11). Therefore, the amount of wasted energy is lower and this amount of energy can be stored in properly configured batteries.

In figures 5.12, 5.13, and 5.14, the results associated with the **Case I** are presented. In particular, each sub figure depicts a specific hour during the simulation and each point is associated with a specific node of the distribution grid. The points depicted with red circles are associated with the nodes, that are selected from the fuzzy logic procedure to buy energy at lower prices. The green square and blue rhombus points represent the nodes of the other two partitions. Figure 5.12 depicts the time interval 02:00 to 08:00 (early morning, off peak demand) with each sub figure (from left to right) to be associated with a specific hour within two hours' time interval. Note that as time increase, more nodes are able to buy energy at lower prices.

Regarding the figures 5.13 and 5.14, the results associated with the time interval 10:00 to 16:00 (daytime) and 18:00 to 24:00 (nighttime), respectively, are shown. It is clear that each time interval, different nodes have the privilege to buy energy at low prices. So, there are a few residencies that are excluded from the benefits that

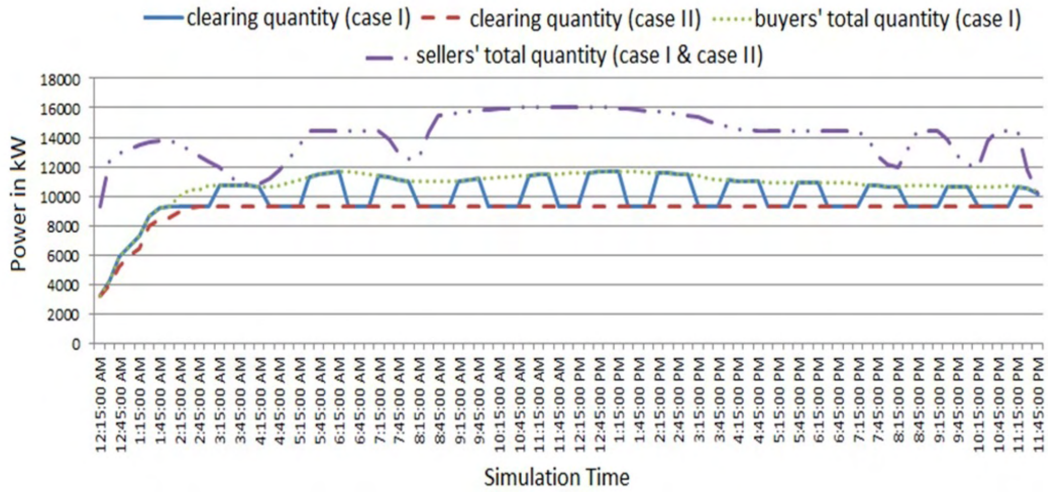

FIGURE 5.11: Clearing Quantity, buyers' (or appliances') total quantity and sellers' (or generators') total quantity regarding the Case I and Case II.

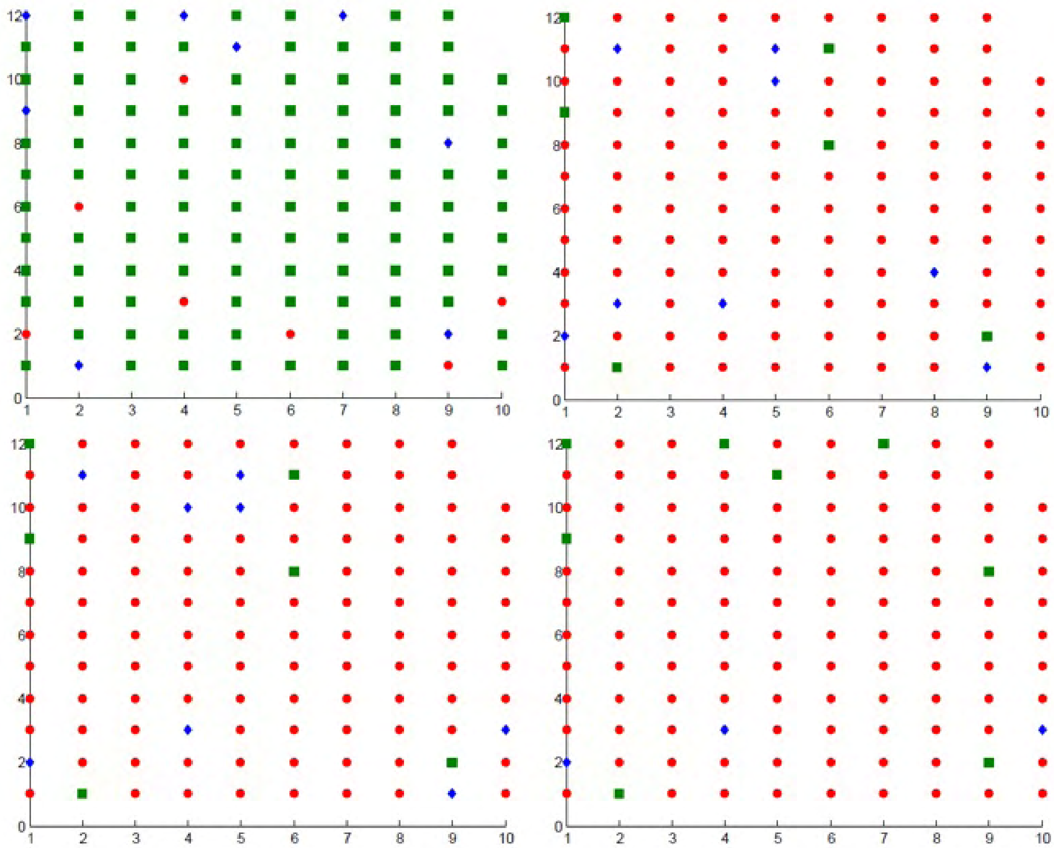

FIGURE 5.12: The every 2 hours results starting from 02:00 – 08:00 when the proposed methodology is applied every 1 hour. The red circle points indicate the nodes that are selected to buy energy at lower prices.

provided from the proposed methodology for a period of time.

For experimental reasons and for presenting the flexibility of the proposed dynamic data driven methodology, the results of **Case III** are presented in Table 5.7.

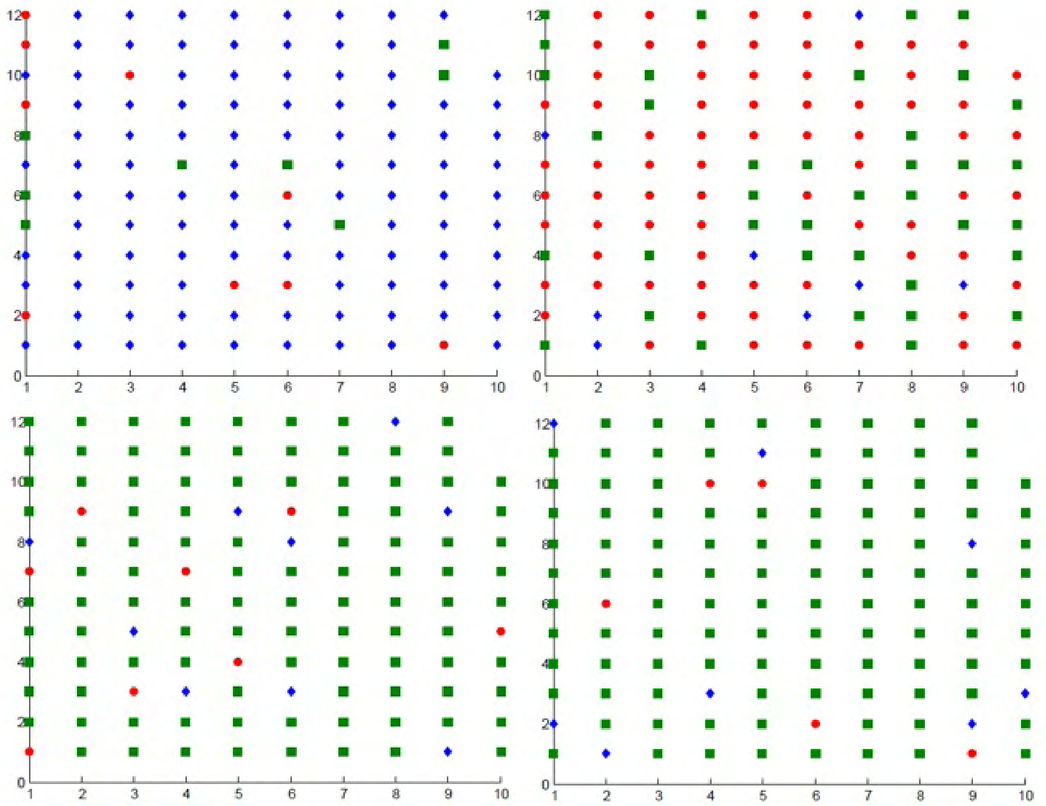

FIGURE 5.13: The every 2 hours results starting from 10:00 when the proposed methodology is applied every 1 hour. The red circle points indicate the nodes that are selected to buy energy at lower prices.

In fact, the average total load of each partition for each run and the partition that is selected (bold font) from the fuzzy logic decision-making method are presented. Moreover, it is clear, that in most of the cases, the partition with the biggest average total load is chosen from the fuzzy logic decision-making method.

TABLE 5.7: The average load for the runs where the proposed methodology is applied every 4 hours.

Time	Partition #1	Partition #2	Partition #3
4:00	75.21	**485.67**	19.22
12:00	53.46	**23.08**	7.66
18:00	**85.039**	12.65	34.41

TABLE 5.8: The aggregated clearing quantity in MW and the percentage of the unsatisfied demand demonstrated the 24 hours of the simulation in the case where the methodology is not applied (Case II), is applied every one hour (Case I) and every four hours (Case III).

	Case I	Case II	Case III
Clearing quantity (MW)	942	855	946
% of non-satisfied demand	7.28%	15.84%	6.88%

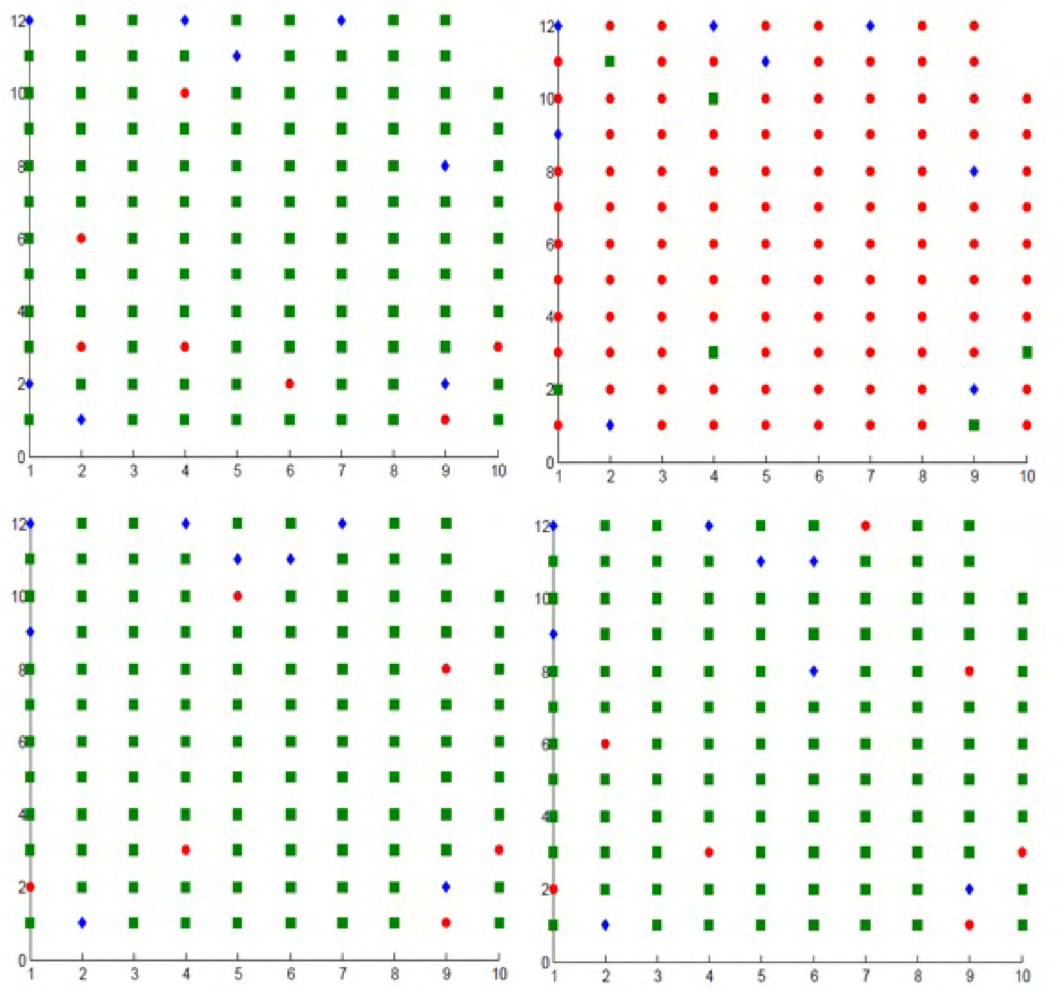

FIGURE 5.14: The every 2 hours results starting from 18:00-24:00 when the proposed methodology is applied every 1 hour. The red circle points indicate the nodes that are selected to buy energy at lower prices.

In Figure 5.15, the clearing quantity of both **Case I** (every 1 hour) and **Case III** (every 4 hours) is depicted. The amount of the demand satisfied when the methodology is applied every 4 hours is higher. The RES offer their energy in lower prices for longer time (4 hours in that case) thus more demand is satisfied. This is also can be noticed in Table 5.8, where the aggregated clearing quantity of all the cases (Case I, II and III) is depicted. It is also observed that the percentage of the non-satisfied power, when the proposed methodology is applied, is at lower levels in Case III. In Figure 5.16, the simulation results of **Case III** are presented.

5.3 Consumers' Partitioning: a three-stage scheme

In the sub chapter 5.2, a partitioning methodology of the nodes in a distribution grid is described in detail. A three-stage hierarchical scheme for residential consumers' partitioning using the Hierarchical clustering algorithm is presented. The objective of this effort is to cluster the consumers in well-separated and compact clusters using information from the near past (almost real time). The energy

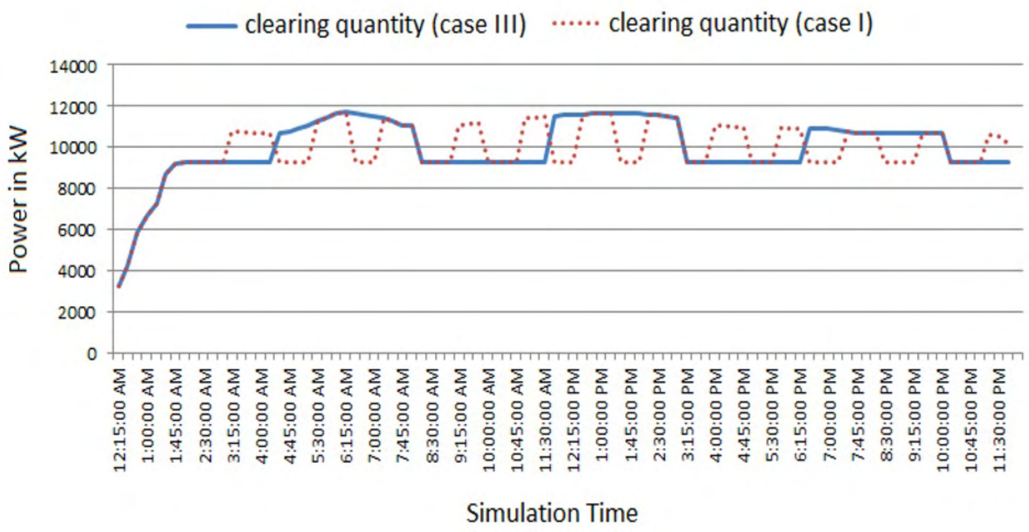

FIGURE 5.15: Clearing quantity offered by the market for the Case I and the Case III.

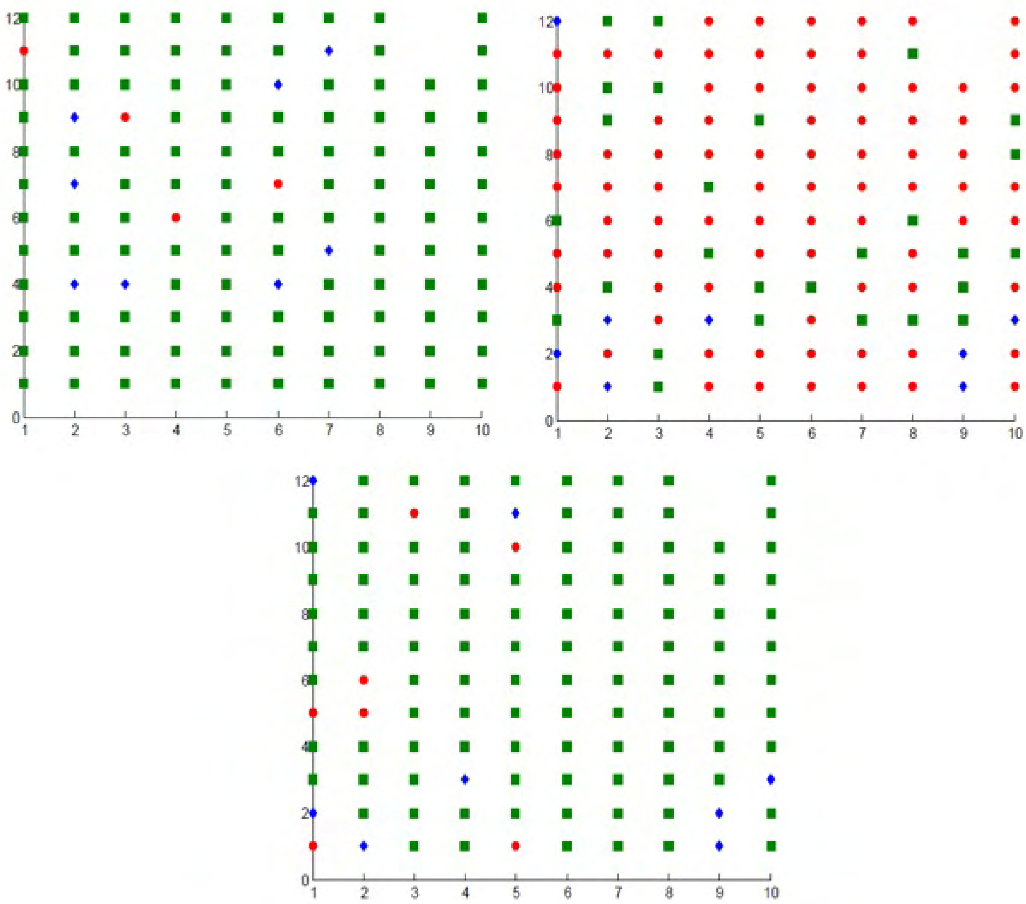

FIGURE 5.16: The results for the time interval 4:00, 12:00 and 18:00 when the proposed methodology is applied every 4 hours. The red circle points indicate the nodes that are selected to buy energy at lower prices.

consumption patterns of each consumer vary. This information can be used by the operator of the grid for improving the efficiency and the stability of the grid.

The first stage is associated with the consumers' clustering using data driven every three minutes (simulation time) from the meter of each residency. The procedure of the second stage takes part every h hours; it is defined by the user. The average value of each of the k*(60*h)/x clusters formed the last h hours is used as input for the hierarchical algorithm; k is the number of clusters, and x is the number of intervals that an hour is separated based on the time basis the information is driven. In the test cases presented in this dissertation, the x is equal to 3. In the third stage, the average value of the data of all the members of each cluster is calculated and each consumer is reassigned to the cluster using the minimum value of a distance metric. The results of the second stage provide deeper information about the load patterns existing each hour in the distribution grid. This information can be used from suppliers to design the energy tariffs for suiting better to the consumers' needs. This approach is tested using the IEEE-13 test feeder and the data are driven from 56 residencies.

5.3.1 Related Work

Occasionally, many methodologies and approaches have been studied to cluster the consumers of a distribution grid using various clustering algorithms, including k-means (Azad, Ali, and Wolfs, 2014), self-organized maps (Hernandez et al., 2012), hierarchical (Quilumba et al., 2015), and fuzzy c-means (Fernandes et al., 2016).

More precisely, the authors Mets, Depuydt, and Develder (2016) propose a two stage pattern clustering of customers in a distribution grid. In the first stage the authors try to cluster the daily load patterns of a set of consumers. In the second stage, the clustering of the consumers is performed using as input of the clustering algorithm a representative load pattern from each consumer. The Fast Wavelet Transformation for reducing the dimensionality of the data used for the clustering is used. Moreover, the g-means algorithm is utilized to adaptively select the number of clusters.

Panapakidis et al. (2013), tried to group the building in a Greece's university campus. Various validity measurements are used to compare several clustering algorithms. Moreover, Ben et al. (2014) use means of dynamic clustering to cluster and analyze the load patterns of a set of Spanish residential consumers. Three different types of consumers are used for the clustering; the first one is characterized by three peaks during a day, the second one is characterized by a quasi-flat load pattern during a day and the last contains consumers that use more electricity during the night.

5.3.2 Three stage Hierarchical Scheme

The proposed three-stage hierarchical scheme as its architecture is depicted in Figure 5.17 has the following three stages:

- In the first stage, load data driven every 3-minutes are used as input to the hierarchical clustering algorithm to provide uniform groups of consumers.
- In the second stage, the average value of each of the k*20*h clusters formed the last h hours is used as input for the hierarchical algorithm, where k is the number of clusters.

- In the last stage, a distance metric of the load of each consumer with the average values of all the members of each of the k clusters is calculated. The consumer is reassigned to the cluster with the minimum distance.

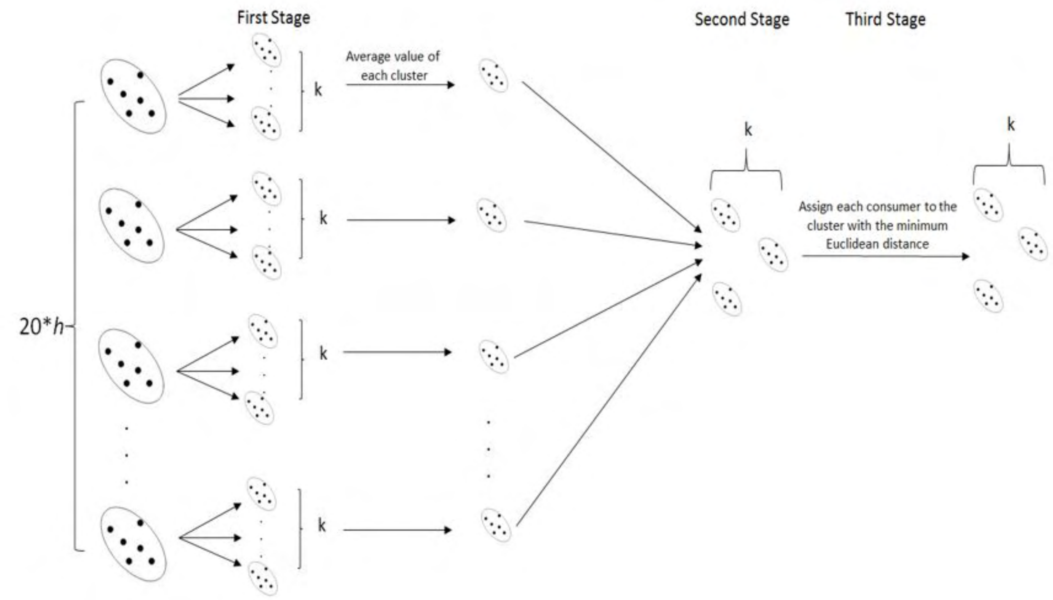

FIGURE 5.17: Architecture of the proposed hierarchical scheme, where h is the hours the proposed scheme is applied.

The main contribution of this effort is the proposal of at almost real time clustering of the consumers using:

- the hierarchical clustering algorithm
- a distance metric to reassign the consumers to the most fitted cluster

In other efforts, the consumers' clustering is performed using historical data. On the other hand, in this study data driven every 3-minutes from the meter of each residency is used. Moreover, the benefits of our proposed scheme are as follows:

- The proposed scheme groups the consumers in well-separated and compact clusters using information from the near past (almost real time).
- Consumers with similar energy consumption behavior belong to the resultant clusters. This achieved using the minimum distance for the final assignment of each consumer to the resultant clusters.
- The hierarchical algorithm is applied for the partitioning using as input the total demanded load of each consumer.

The **first** stage has two sub-stages. The first one is associated with the clustering of the residential consumers of the distribution grid and the second one with the calculation of the average value of all the members' values per cluster. More precisely, the data used as input of the clustering algorithm are obtained through the meters placed to the residencies. This data corresponds to time series data which represent the total load of the consumers at a particular time instance. The measurements of this stage are performed on a 3-minute basis. The time interval of the h hours, provides 20*h sets of k clusters. The average value of the load of the consumers belong to each cluster formed for the particular hours is calculated. So, in the end of this stage, the information that is going to be used for the next stage is 20*k*h item values.

In the **second** stage, before conveying through the next h hours of the simulation time, the hierarchical algorithm is applied. The input of the hierarchical algorithm is the output of the previous stage. The output of this stage is k individual clusters providing information about the average demanded load of the consumers the last h hours of the simulation.

In the **third** stage, the minimum distance is used for the assignment of the consumers in the final clusters. The data points in this effort are one dimensional. The equation 5.2 provides the formula of the minimum distance for 1-D data points.

$$d(x_i, C_j) = x_i - C_j \tag{5.2}$$

where x_i is the ith consumer of the distribution grid with the i ranges from 1 to N (the number of consumers) and C_j is the jth cluster with the j having values from 1 to k. In Algorithm 1, the steps of how the three stage scheme works are presented in detail.

Algorithm 1: three-stage scheme	
Input:	x_i's value and average value of all the members in C_j
Output:	reassignment of each x_i to the C_j with which x_i has the smallest distance
Steps:	
1. for each i=1, ..., N	
2. for each j=1, ..., k, a. Find $d_j = d(x_i, C_j)$ b. Take the minimum d_j with the j to be the id of the cluster c. Assign the x_i to the j cluster	

5.3.3 Simulation Results

For demonstrating and evaluating the proposed three-stage scheme, a dataset covering one day on 3-minute resolution load measurements at the IEEE-13 test feeder is utilized. The load that the feeder can handle is changed in order the IEEE-13 test feeder be able to be equipped with 56 low voltage consumers. For evaluating the proposed scheme the 23 residencies out of the 56 are equipped with an HVAC system, a water heater, and the refrigerator while the remaining 23 residencies are equipped only with refrigerators.

For comparison purposes and for assessing the validity of the proposed scheme two cases are considered. In the first case, called Case I, the proposed three-stage scheme is applied and the final clusters are the ones obtained every h hours in a 24-hour resolution. In the second case, called Case II, the proposed scheme is not applied but the partitioning of the distribution grid takes place every h hours using as input of the hierarchical algorithm the demanded load. So, a comparison of how the proposed scheme affects the clustering results is conducted. The DB index clustering validity measure is used for assessing the validity of the proposed scheme.

In the test cases presented, the number of clusters is two and the clustering of the second stage has taken place every 6 hours. Two clusters for easier presentation of the results are used. So, the number of inputs of the hierarchical algorithm at the second stage is 20*2*6. In figures 5.18 and 5.19 the resultant clusters of the Case I are presented.

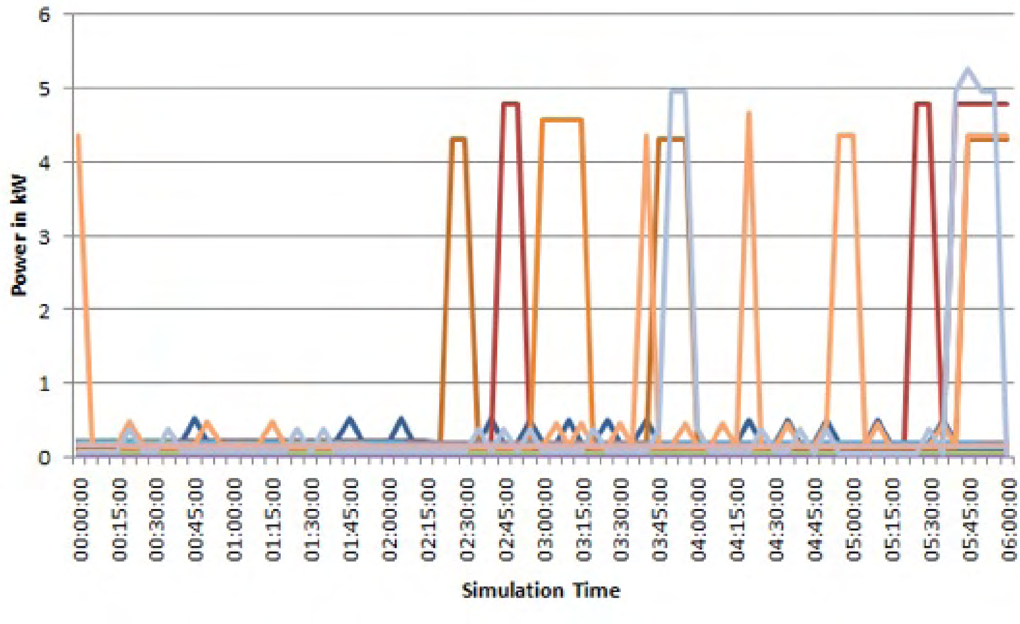

FIGURE 5.18: Items belonging to cluster A of the Case I.

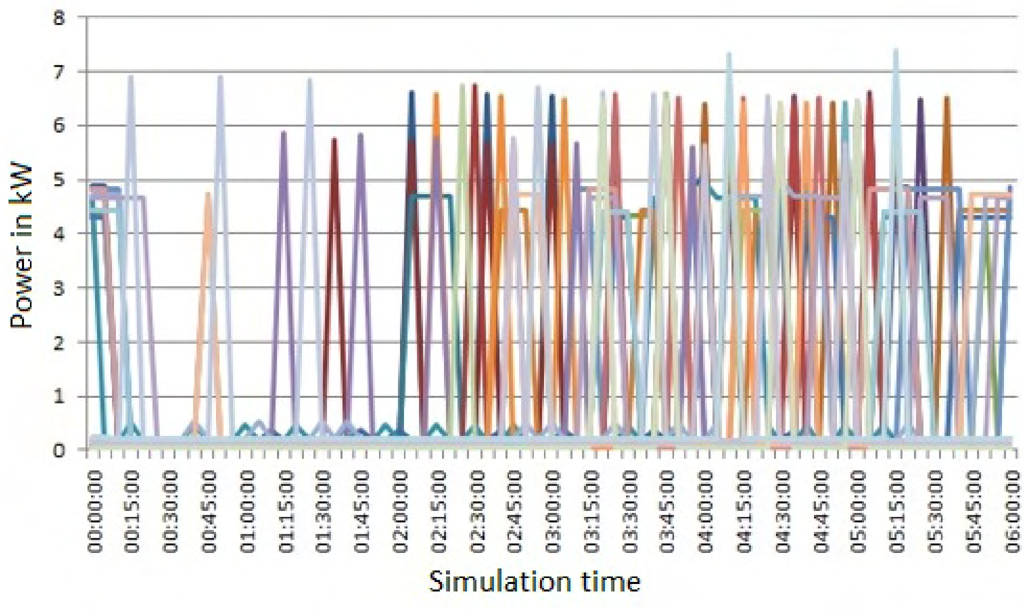

FIGURE 5.19: Items belonging to cluster B of the Case I.

The results of Case II for the same time interval are presented in Figures 5.20 and 5.21. It is obvious that when the proposed scheme is utilized, the consumers that belong to each cluster have almost the same energy consumption requirements. More specifically, the peaks of each consumer may indicate the use of the HVAC system, so consumers which use the HVAC at the same time belong to the same cluster.

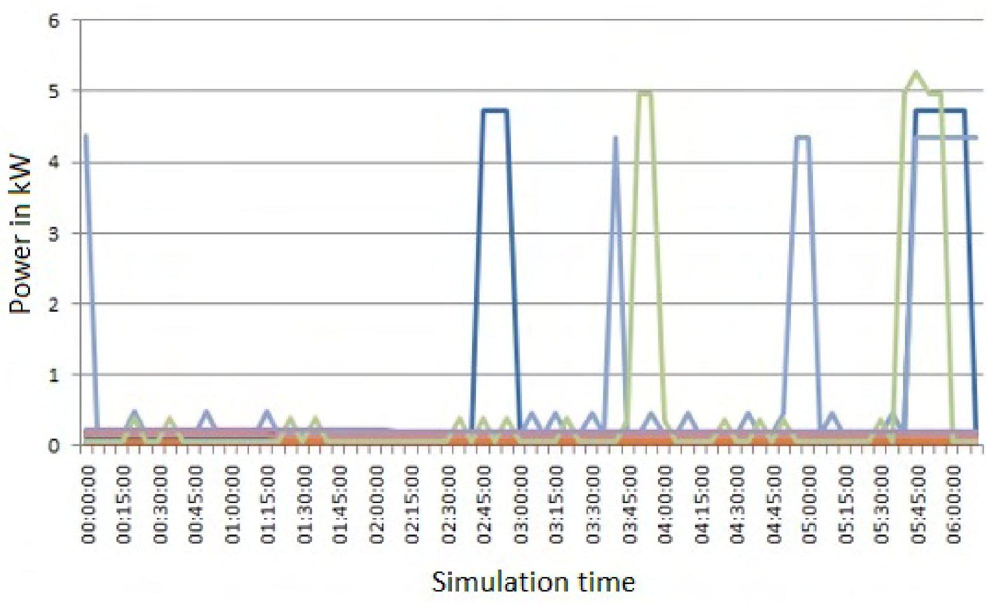

FIGURE 5.20: Items belonging to cluster A of the Case II.

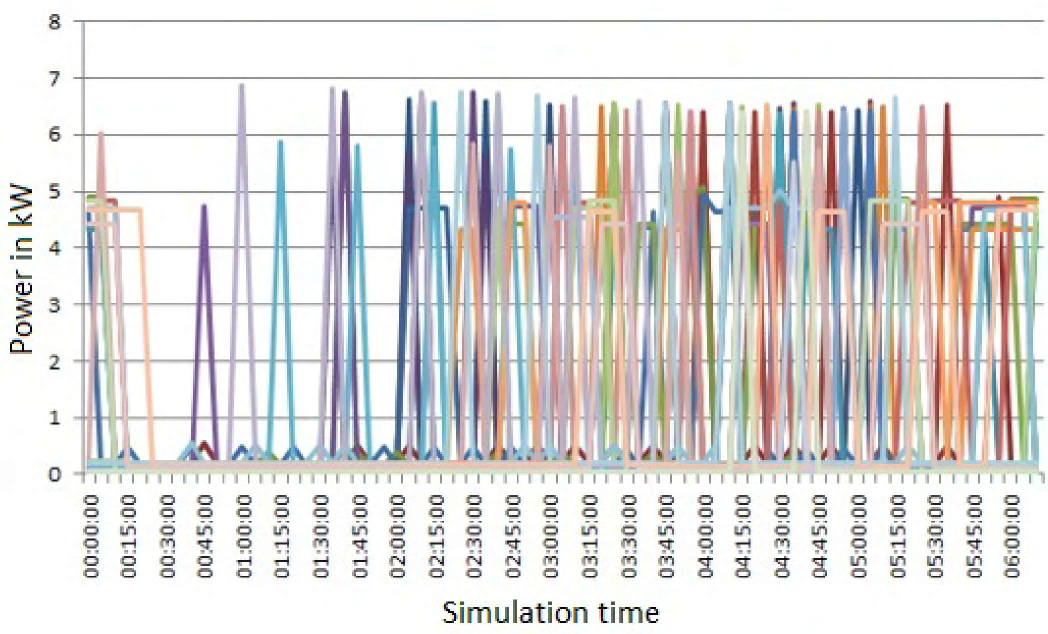

FIGURE 5.21: Items belonging to cluster B of the Case II.

The number of consumers of Case I that belongs to cluster A and cluster B is 20 and 36 respectively. On the other hand, the number of consumers of Case II is 11 and 45 for the cluster A and cluster B respectively. So, the consumers are evenly distributed in both the clusters regarding the Case I. Additionally, when the clustering procedure takes place at 12:00 pm, the number of consumers of Case I and II is 35 and 7 for the cluster A and 21 and 49 for the cluster B respectively. So, the proposed scheme assigns the consumers that have similar energy consumption preferences to the same cluster.

The distribution grid is modeled and designed in such a way to consider if the consumers with the same energy consumption preferences are assigned to the same cluster. Better clustering results obtained when the using information is driven every 3 minutes for the interval of 6 hours instead of using only the information which is available the specific hour. More specifically, in the Case I the clusters have approximately the same number of consumers. However, if each cluster has 23 consumers with exactly the same energy consumption characteristics, this will be the ideal case.

Moreover, it is clear from the Table 5.9 that the proposed three-stage scheme provides better clustering results using as validity measure the DB index. Therefore, using information for the last 6 hours instead of using the current information regarding the load of each consumer leads to clustering of the consumers in groups with almost the same energy preferences.

TABLE 5.9: Validation indices results

Time	DB (Case I)	DB (Case II)
6:00 am	1.41	1.47
12:00 pm	1.36	10.63
6:00 pm	1.59	2.58
12:00 am	1.24	1.46

In the next lines, some more experimental results, by changing the time interval that the final clustering takes place are presented. In these results, the final clustering took place every 1 hour instead of 6 hours. In figures 5.22, 5.23, 5.24 and 5.25 the results of the Cluster A and B of Case I and II respectively are illustrated.

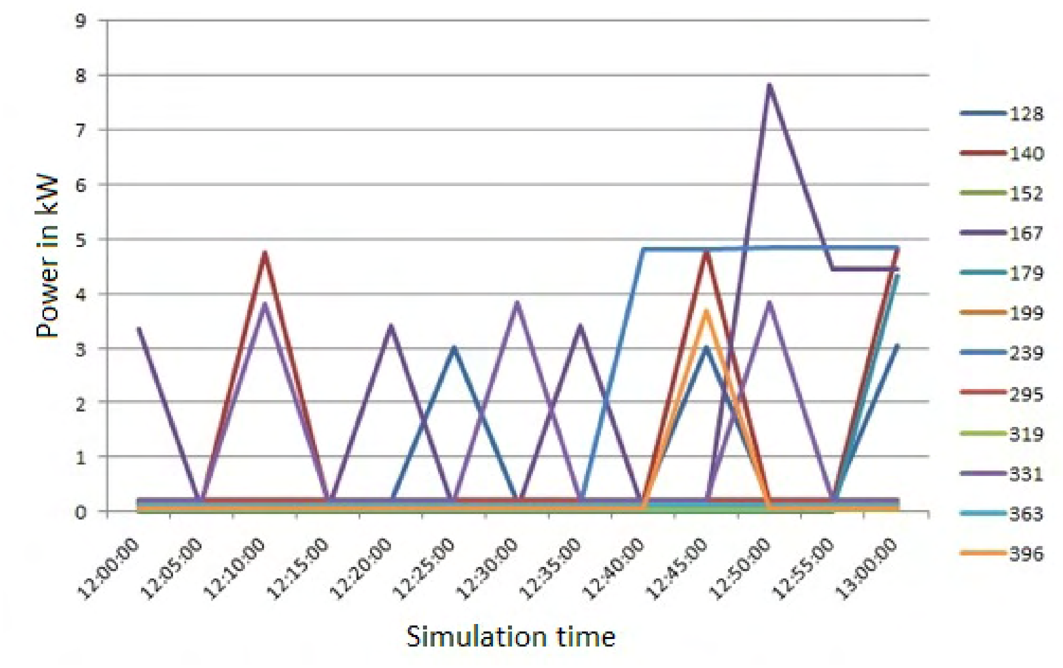

FIGURE 5.22: Items belonging to cluster A of the Case I.

The number of consumers is 44 and 52 for the cluster A of the Case I and II respectively. The difference is just 8 consumers that are assigned to the cluster B for the Case I. At 12:00 pm the cluster A has 35 and 7 consumers in the Case I and II, respectively. This is a side effect of the fact that the information is less than in the case where the time interval is 6 hours. However, it has to be noted, that the consumers are not distributed as well as they are distributed in the case where the time interval is 6 hours.

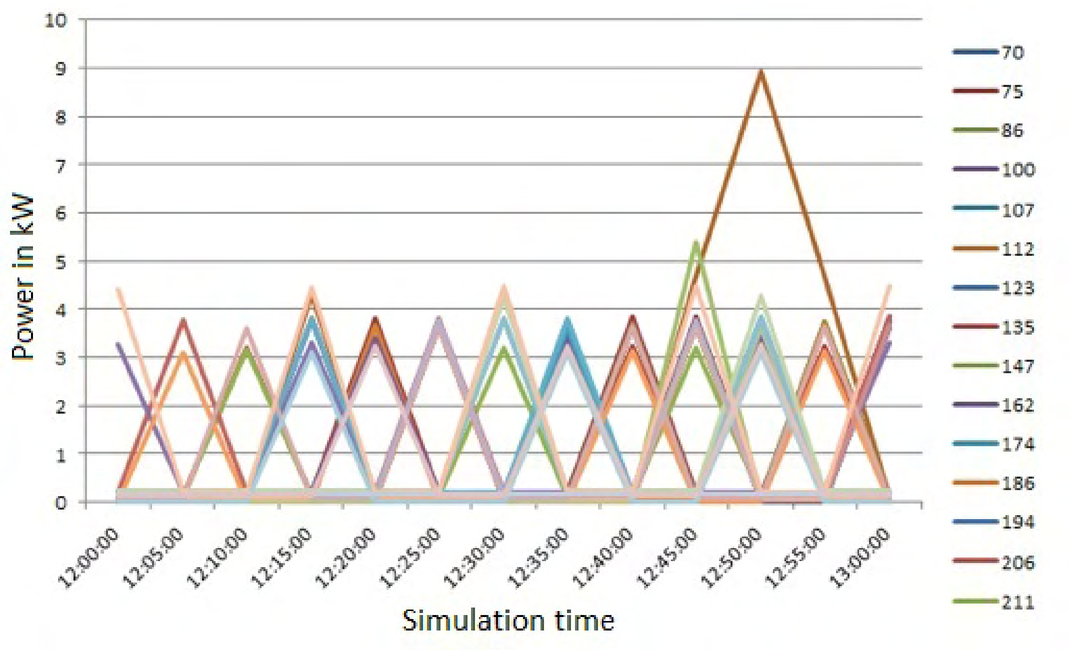

FIGURE 5.23: Items belonging to cluster B of the Case I.

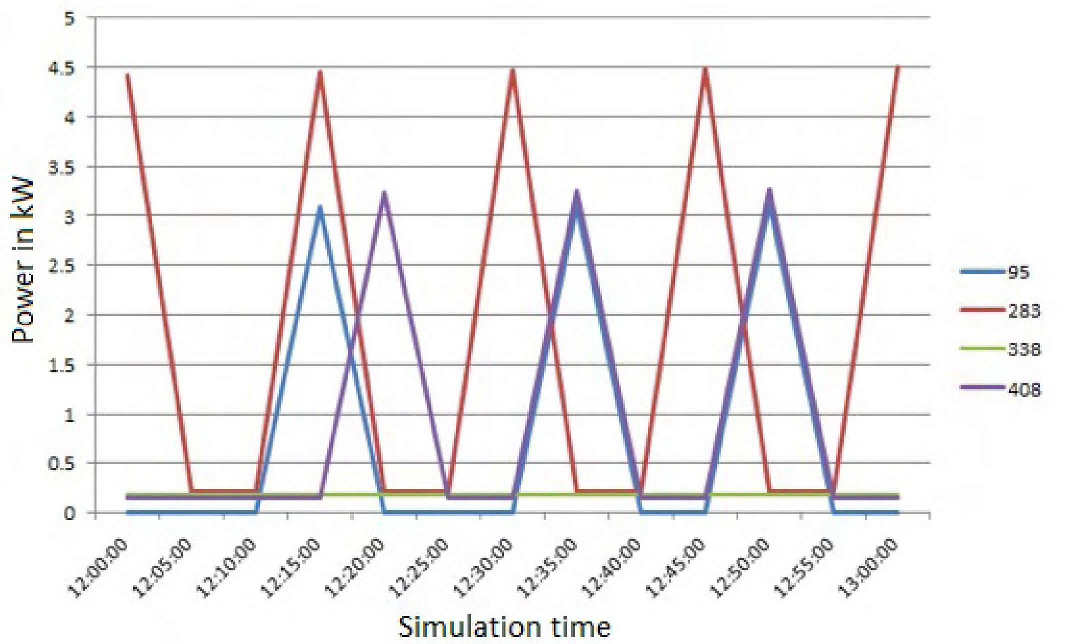

FIGURE 5.24: Items belonging to cluster A of the Case II.

5.4 Conclusions

In this chapter, the drawbacks of the k-means algorithm by proposing a new algorithm, called SkM, which deals with the optimal selection of the initial cluster centers are described. The accuracy, the execution time and DB and S validation measures to prove that the SkM algorithm prevails the k-means, k-means++ and the fuzzy c-means algorithm are utilized. The SkM prevails in 90% of the test cases comparing to standard k-means and the k-means++ and in 85% comparing to fuzzy

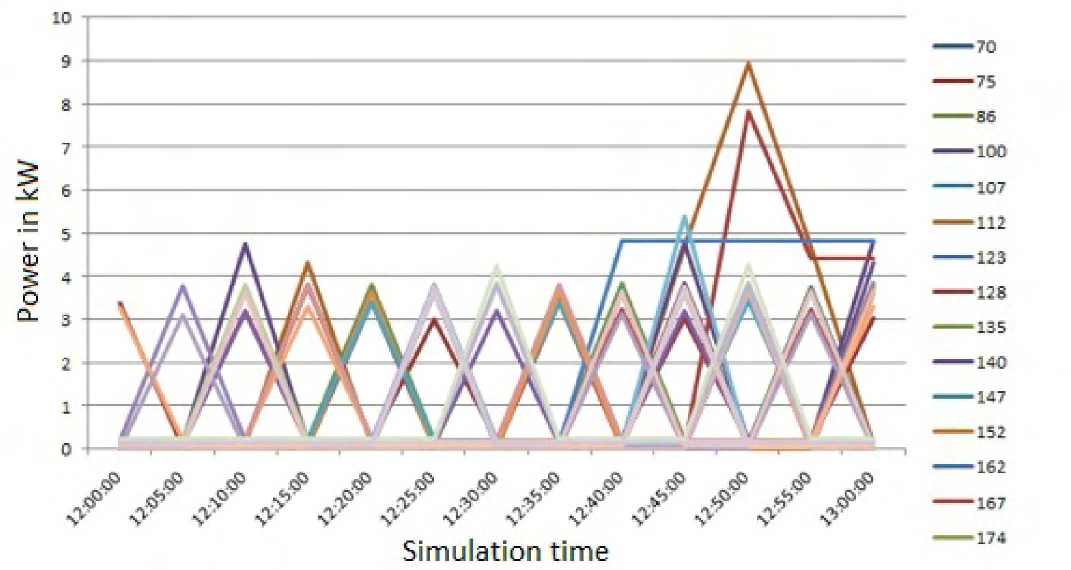

FIGURE 5.25: Items belonging to cluster B of the Case II.

c-means with respect to the DB index. Moreover, in this chapter, the partitioning of a distribution grid using data driven from the nodes of a distribution grid and data associated with the consumers is studied. Our objective is to provide methodologies that (a) integrates the RES and reduces the wasted energy and (b) leads to making use of data that are from the near past. The results for both methodologies show that their impact has positive side effects for the consumers and the producers of the distribution grid.

Chapter 6

Smart energy for smart irrigation

Irrigation systems and generally irrigation has been around since the humans have been used to cultivate plants. The irrigation system depends on the water availability, while water requirement is the main parameter for calculating equivalent energy demand. A simulation engine that couples both the smart energy systems and the smartness driven from the irrigation systems is implemented, designed and evaluated through different scenarios. The developed simulation engine allows large scale simulations with detailed configuration, that leads to the reduction of the irrigation cost through the use of RES. The primary objective of this effort (Nasiakou, Vavalis, and Zimeris, 2016), (Houstis et al., 2017) is the use of both RES and batteries to handle the agricultural load and the residential load. The agricultural load corresponds to the load of a set of irrigation systems for a specific time of period, while the second demanding the total load of both the evcharger station load and the residencies' load.

6.1 Related Work

Several agricultural projects around the globe emphasize on the need for further research and development towards the improvement of irrigation applications, in terms of better energy management and reduced cost.

A sizing procedure for calculating the variables involved in a solar panel installation for agricultural processes is presented by Morales and Busch (2010). Vick and Almas (2011) presented methods of the irrigation management via a central coordinator. Based on several parameters, this coordinator suggests various alternative implementations involving solar and/or wind energy, also taking advantage of storage ability of excess energy and water storage in surface reservoirs for future irrigation and other on-farm uses. It proposes a partitioning of the irrigation system between a winter crop and a summer crop that leads to an improved match between the wind turbine and the PV solar array power generation required for water pumping applications. The design and the operation of such an application depend on a set of crucial parameters that include monthly energy demand for different crops, the size of wind turbine or/and PV array, the average monthly evapotranspiration, the average monthly rainfall, average hourly wind speed for different heights, average monthly air density, power curve and many others.

Similar to the above, the authors T. Stambouli and Zapata (2014) present a modernized irrigation system, already in operation, coordinated by a central management operator. The central system involves elementary remote control that manages the automated frequency controlled pumping stations, while an on-demand program schedules irrigation processes for different farms. More specifically, it collects the desired schedule in terms of irrigation hours and days of the week and adjusts the pumping system and piping network operation by opening hydrants and valves accordingly. A district central managing office is responsible for paying all the electricity and maintenance expenses, which then distributes them according to the total water volume consumed by each farm.

Among other papers, Y. Shivrath and S. Thirumalasetty (2012), and Kelley et al. (2010) present sizing and optimization algorithms focusing on evaluating technical (dependent on soil, climate, geography, agriculture, and hydro-geology) as well as economical (dependent on the electricity price, life cycle cost of investment) feasibility of solar and/or wind energy installations for irrigation purposes. More specifically, a technical and an economic feasibility analysis for PV solar irrigation systems is performed by Kelley et al. (2010), in order to emphasize the use of solar energy for irrigating purposes compared to conventional diesel engine as well as grid powered irrigation systems. It is based on an analytical mathematical model, in a way that can be applicable to any geographic location and crop type. Technical feasibility is determined from the maximum power required for irrigation purposes which highly depends on the crop type and the geographic location. On the other hand, economic feasibility is determined in terms of life-cycle costs of the PV solar irrigation system compared to diesel and grid based irrigation systems. The results exhibit that the success of a PV solar application depends on the size of the overall farming size and the variations of the irrigation needs. The results of the economic feasibility analysis illustrate the superiority of PV solar applications compared to the diesel-engine and grid powered irrigation systems, despite the high capital costs of PV systems. As the price of solar panels decreases, the PV irrigation systems become more and more economically attractive. Several examples cases are examined, showing the benefits of the PVP applications compared to other system types.

The authors Y. Shivrath and S. Thirumalasetty (2012) propose a design and an optimization process of the cost of a hybrid wind/solar renewable system for drip irrigation purposes, based on various operating and design parameters. These design parameters are the pumping system specifications as well as the drip irrigation specifications. Meteorological data concerning the design of both the solar PV system and the wind turbine system are also utilized for sizing purposes. Finally, a battery bank calculation is presented. The cost optimization procedure is formulated by an objective function of minimizing a particular cost function consisting of the optimum number of solar panels and wind turbines, subjected to certain limitations and constraints obtained by the pumping and irrigation specification and requirements. The proposed design and optimization algorithm is executed on Microsoft Office Excel 2007. The authors conclude that when combined, solar and wind renewable sources enhance the reliability of the system, and lead to battery bank reduction increasing its lifetime.

The study Vishwa et al. (2015) is an up-to-day literature review of PV solar water pumping systems for irrigation purposes. The advantages of solar water pumps are illustrated compared to hand pumps or internal combustion engine pumps, that is, easy installation, zero maintenance, long useful life, no fuel, no contamination. It is shown that compared to diesel powered water pumps, solar powered pumping systems are less expensive over a life cycle of 10 years. This review focuses on specific issues like the evaluation, monitoring and performance improvement of the different components of a PV system. It considers

1. the dependency on the irrigation process of soil type, PV system sizing for the particular crop considered

2. meteorological data such as solar radiation air temperature, wind speed, relative humidity,

3. operational specifications like irrigation scheduling from which water requirements converted into electrical energy requirements.

The authors Martin and Levine (2007) propose a feasibility study concerning the renewable energy generation by PV solar panels and its storage for agricultural use in the San Luis Valley in California. Crop irrigation water requirements results are used for the estimation of the size of the PV solar array. A potato crop as the base case scenario, requiring approximately 12 inches of water per season, is considered. The growing season is assumed from April through August. Four major energy storage techniques are illustrated, namely, pumped water storage, compressed air storage, battery bank storage and grid storage. Two motor pumps are evaluated, a DC driven motor pump and an AC driven motor pump, with the later to be, as expected, the most practical solution.

Following the water storage concept, the author Martin (2007) proposes the so-called underground pump hydroelectric storage (UPHS) which is based on an agricultural well and an underground aquifer with a forward/backward water flow. The system includes an intermittent RES for pumping water, a surface water storage reservoir for irrigation and power generation water, a combined pump-turbine motor-generator unit and a system control interface. The alternative to the surface reservoir is the underground reservoir where of course such reservoir is geologically available. This new concept has the form of the aquifer recharge (AR) and aquifer storage and recharge (ASR) wells that already exist in California. These types of wells are designed to replace water in an aquifer by performing a backward water flow into a water well leading to an aquifer recharge. The basic principle of the latter idea is that water is traditionally pumped from the down aquifer by means of a renewable resource. The pumped water is stored in an upper underground well or in the surface reservoir. This water is used for the irrigation system when it is needed. When it is not needed, this water can be released from the upper underground or surface reservoir back to the down aquifer, reversing the operation of the motor and pump to generate electricity. To implement such a system, a pump-turbine coupled with an electrical motor-generator is installed as a single unit at the bottom of the well. To improve turbine efficiency, a reasonable height is assumed between the surface or underground upper reservoir and the down aquifer. UPHS systems can either be used with the standard utility power or for use with the engaged renewable energy sources. Even with only the central utility grid, the user can derive economic benefits by storing energy during the "off-peak" hours (when irrigation is not in operation) and releasing the generated energy on demand avoiding the cost of the "on-peak" hours where utility electricity tends to be more expensive.

Arruda et al. (2015) studied the effectiveness of a wind/solar power system with a battery bank for water pumping. The system involves a wind turbine, four PV solar panels, five electrolytic batteries, a converter/controller and a data-logger. The system operates mostly with the direct action of the renewable sources while the battery bank participation occurs only in the case of load demand and electricity produced unbalance. The wind/PV system operates in a complementary manner, that is, one resource powers the pumping system when the other is unavailable. Not very common but possible variations in indeterminacy are alleviated by the battery bank. Experimental results verify that the wind/PV + battery bank configuration needs less energy storage compared to the PV + battery bank one increasing the life cycle of the battery bank and of the whole renewable source system.

The authors Dursun and Ozden (2012) present an automatic irrigation system specifically designed for dwarf cherry trees. A solar PV array, equipped with a sun tracking system and a battery bank, generates the desired electricity demand for two brush-less DC motor pumps, one responsible for elevating the water from the aquifer to a surface pool and the other for transferring the water from the pool to the drip irrigation system. Solar panels are used for the purpose of providing the maximum solar energy according to the maximum daily solar radiation. The water demand is defined automatically with soil moisture sensors, that regulate the booster surface pump pressure providing effective management of the irrigation water requirements. The authors claim that the proposed system could provide a socioeconomic development in the region of application.

6.2 Prototype Implementation and Design

For developing the simulation engine, various modifications and insertions have been completed in the **market** module of the GridLAB-D. First of all, there was a need of a controller which can provide price responsive control of the agricultural load. This controller, called *irrigation_controller*, works as follows: the soil humidity measurements combined with the willingness of the customer to pay for the operation of the irrigation system are used in order the system's humidity set points to be adjusted. The primary task of the irrigation controller is to take information about the humidity from the sensors located on the under consideration field. For simulating the irrigation system and more precisely the agricultural load, the object of the **residential** module *ZIPload* is utilized. Moreover, the object *house* and its dependent objects (e.g. ZIPload) of the same module are used to describe new properties and definitions of terms engaged in agricultural engineering. In that sense, the parameter **humidity_setpoint** is added at the *ZIPload* object for use of its value from the *irrigation_controller*. The *soil_sensor* object, dependent of the *house* object, is implemented for providing information about the soil humidity needed by *irrigation_controller*. Moreover, a new parameter of the object *house*, namely **crop_type**, is added to define the different characteristics of each crop. The value of the soil humidity follows the normal distribution functions defined by equation 6.1.

$$P(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-(x-\mu)^2/2\sigma^2}. \tag{6.1}$$

Each consumer, through the implementation that the GridLAB-D provides, can define his comfort zone. Referring to the concept comfort zone, the limits, namely the parameter *comfort_low* and *comfort_high*, which declare the desired range of the consumer on soil humidity, dependent on the electricity price are defined. More specifically, on one hand, the parameter *comfort_low* corresponds to the tolerance of the consumer in a decrease of humidity. On the other hand, the parameter *comfort_high* indicates the upper limit of the humidity that the consumer prefers allowing the irrigation system to operate in saving money due to the low prices prevailing over the market. In Figure 6.1, it is clear that there are two areas; the right area indicates the willingness of the customer to save money by leaving the irrigation system to operate due to the low prices while the left area indicates the willingness of the customer to allow the humidity to decrease. For the determination of the irrigation system's bid price, two more parameters, namely *ramp_low* and *ramp_high* are determined. These parameters define the slope of the linear function that is used

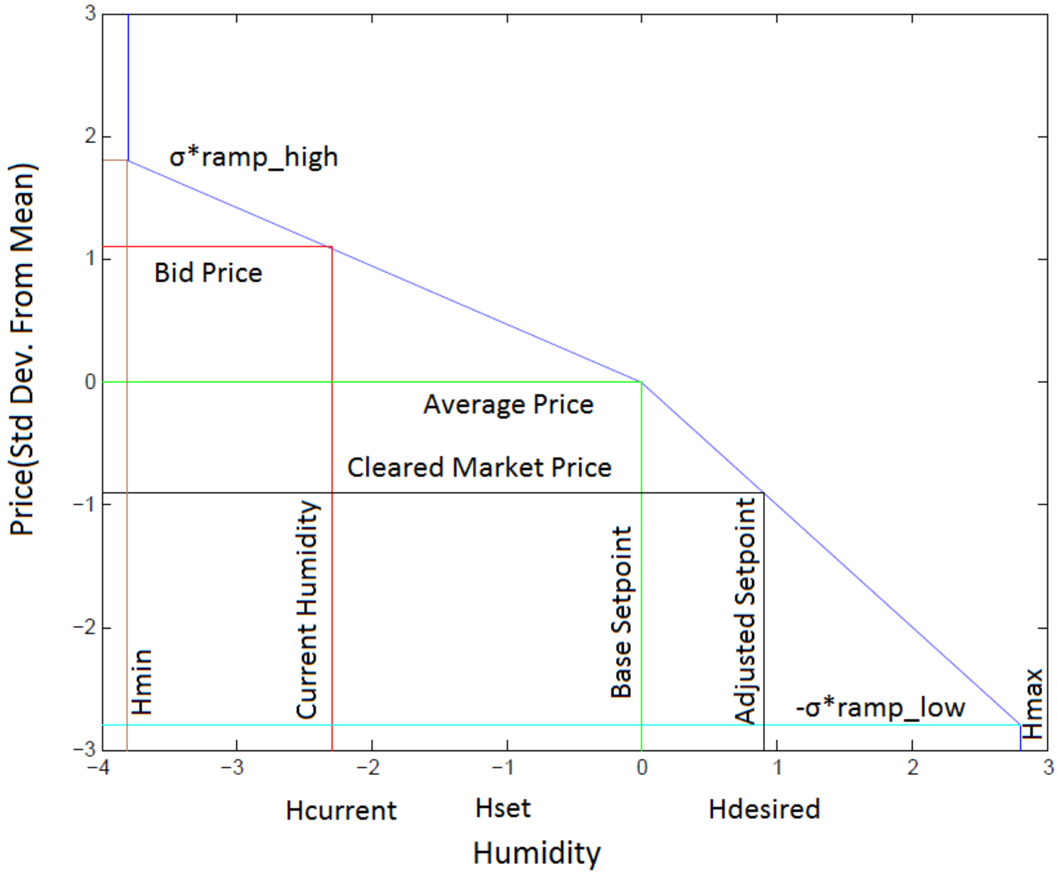

FIGURE 6.1: The model for irrigation loads in the case of using humidity sensors.

in order the irrigation system's bid price to be calculated. The price that the *irrigation_controller* uses as bid price in the market is calculated using the following equation:

$$bid\_price = avg\_price + \frac{(H_{desired} - H_{current}) * ramp\_high/ramp\_low * SD}{|comfort\_high/comport\_low|}, \quad (6.2)$$

where the avg_price is the average price of the market, $H_{current}$ is the current humidity of the soil as it is determined in equation (6.1), the SD is the standard deviation, $H_{desired}$ is the adjusted humidity set point of the irrigation system, which is calculated after the market has returned the clearing price. The equation of calculating th H_{set} and the adjusted humidity set point is presented below.

It is worth mentioning at that point, that the irrigation system has to start irrigating if the current soil humidity ($H_{current}$) is lower than H_{set} (see equation 6.3).

$$H_{set} = H_{desired} + term \quad (6.3)$$

where

$$term = \begin{cases} 0 & \text{if } SD = 0 \\ (cleared\_price - avg\_price) * \frac{|comfort\_high/low|}{(ramp\_high/low)} * SD & \text{otherwise} \end{cases}. \quad (6.4)$$

As previously mentioned, the irrigation system is modeled through the *house* object using the *ZIPload* object of the **residential** module. On the other hand, the residential load is modeled by using the *house* object as well which is controlled by active controllers and having *ZIPload* objects to model the non-thermostatically appliances.

6.3 Simulation Results

For demonstrating the capabilities of the developed simulation engine both the IEEE-13 and the R4_25.00_1[1] feeders have been utilized for supporting the residential network. The IEEE-4 test feeder is utilized for supporting the agricultural network. All the feeders are connected to a meter. The meter is the point in which the control and monitoring of both the amount of the needed energy needed and the amount of provided energy take place. The distributed generators are located in the agricultural network and the energy is transferred, through the properly configured equipment (fuses, capacitors and regulators) serving both residential and agricultural loads. The R4_25.00_1 is equipped with 68 residencies while the IEEE-13 bus test feeder is equipped with 70 residencies. Since, the standard version of these feeders does not support any connections with wind turbines, solar panels, batteries and evcharger stations, these new features for designing the simulation engine have been added. The characteristics of this engine are the following:

- The energy requirements of the irrigation-pumping system are met at most of the cases by the energy produced by the RES and by energy stored in the batteries in less cases.
- Hybrid or full electric agricultural or not vehicles are serviced by the evcharger stations that are located at specific points on the grid and close to the villages or/and the fields.

For experimentation purposes, two scenarios that are as realistic as possible are designed. The main difference between these two scenarios is the type of the irrigation system. For both scenarios, the load and the topology of the grid are the same. The energy requirements for both the scenarios are satisfied directly from the solar panels, the wind turbines and/or the batteries. Moreover, the satisfaction of the agricultural load has priority over the residential load. In a case, where the residential load is not satisfied by the RES and the batteries because of the low available energy, then it is satisfied by the main grid.

In the first scenario, the irrigation system corresponds to a system which uses the water from a surface water reservoir. The power need of such an irrigation system in order to operate is approximately 20 kW. In the second scenario, a coupling of a drilling and an irrigation system is simulated with the energy requirements to be approximately 60kW. On one hand, the drilling system is responsible for pumping the water to the surface; the amount of 40kW is needed for this procedure. On the other hand, the irrigation system is connected to a surface water reservoir (only surface irrigation is of concern) requires approximately 20 kW. The coupling of these two systems is referred to the rest of this dissertation as drilling system. It is connected to a surface water reservoir; surface irrigation is of concern requiring approximately 60 kW. Therefore, the drilling system's task concerns the whole irrigation procedure and not only the water elevation procedure. Since it consists of two different systems, it needs larger amount of power. The one system is associated with lifting up

[1] http://gridlab-d.sourceforge.net/wiki/index.php/Feeder_Taxonomy

underground water (approximately 100 m high), and the second one for the irrigation system itself.

A new concept, namely virtual battery, is introduced. The virtual battery is associated with the stored water in the surface reservoir. This association was considered since it allows to exploit GridLAB-D's market framework based on which the generators submit offers for generated power. The drilling system submits bids for water to be used for irrigation purposes. The stored in the reservoir water is translated into power in order to be available in the market since there is no any implementation of such a market; accepting water offers. This is achieved through the association of the drilling system to a virtual battery which is responsible to convert the stored in the reservoir cubic meters of water (m^3) into power (kW) and then to participate in the power market. The virtual battery is actually associated with the operation of the drilling system. So, the virtual battery is charging in slower pace when the drilling system operates. That happens due to the fact that the pumped underground water is used from both sources; the storage in the reservoir and from the irrigation process. For that reason, when the drilling system operates, the virtual battery bids high enough to stay at charging mode instead of selling energy to the market. So, the farmers use the water left in the reservoir and the underground water. In that way, the virtual battery is charging. However, there are cases where the underground water is not enough to satisfy the overall water demand, so the virtual battery participates in the market with lower bidding price by offering water in terms of power. Based on the assumption that the underground water is not enough to irrigate more than 7 fields, the virtual battery starts discharging by offering the water amount. So, it is obvious that at the same time the underground water and the water stored in the surface reservoir are used to satisfy the water demand for all the fields.

For the implementation of the virtual battery, the *stub_bidder* is utilized by adding new features to support its functionality. Moreover, the implementation of the virtual battery is based on the following assumptions:

- The water in the reservoir is enough to satisfy the overall water demand each particular time instance.
- Some of the farmers are willing to irrigate during the night.

Based on the aforementioned assumptions, the virtual battery charges during the day, and it discharges during the night by submitting bids to the market framework.

Continuing with the experiments' configuration, it is worth mentioning at that point, that one more assumption is made. In the beginning of each simulation, the surface water reservoir is full and there is enough water to cover the overall total demand. Moreover, at the beginning of each simulation, the value of soil humidity associated with each field is different for providing variability in the test cases. For all the test cases presented in this section, each residency is equipped with an HVAC system, a light system, and a water heater. An evcharger station consisting of three electric vehicle chargers with the same vehicle travel information but with the different sizes of evcharger batteries is also simulated. Both agricultural and residential loads submit their bids using the controllers (controller and irrigation_controller respectively) into the market.

The agricultural load is simulated through the *ZIPload*. The evcharger station is simulated using the *evcharger* object of the **residential** module. The energy requirements are satisfied through RES. Each solar or wind turbine is equipped with a basic and a reserve battery. The latter is used to store the excess energy when the basic battery is full. The batteries participate in the market framework. They discharge

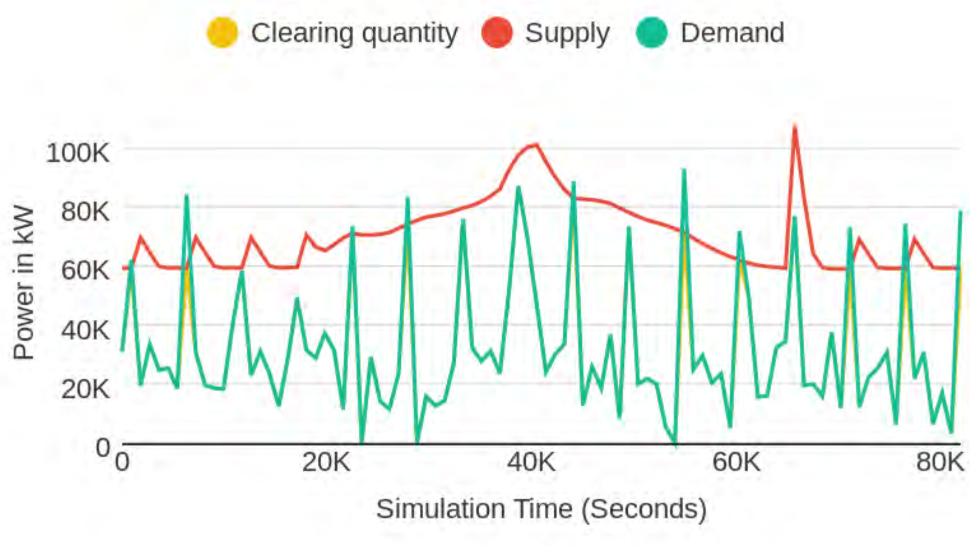

FIGURE 6.2: Market Clearing: The demand, supply and clearing quantity in kW correspond to y axis while the x axis represents the simulation time which is 24 hours for the first scenario with one field.

when they are included in the market dispatch and they charge when the energy produced by the RES is not totally used by the market for the energy dispatch procedure. Therefore, the amount of wasted energy from both the solar panels and the wind turbines is small.

In the next lines, the test cases are described in detail along with the simulation results. In the first test case, the R4_25.00_1 feeder is utilized for the residential network and it uses the configuration of the first scenario. The parameters of this test case described as follows:

- One wheat type field.
- The residential network accommodates 68 houses.
- The humidity set point of the wheat used by the irrigation_controller is 35.
- A solar panel with an area of 250 m^2.
- A wind turbine of 15 kW max power output.
- Two batteries of a maximum capacity of 20 kW h and initial available energy equal to 10 kW h.

In Figure 6.2, the results of the aforementioned test case are depicted. It is obvious that both the agricultural and residential load are satisfied from the DES (RES and batteries), with the exception of a few time periods where part of the demand is satisfied from the main grid.

A more realistic test case based on the first scenario which concerns important parameters such as the crop type is considered. Additionally, this test case involves also the calculation of the soil humidity taking into account the solar radiation. More specifically, the crop type affects the overall irrigation scheduling of a field; different crop types require different amounts of water. Therefore, the soil humidity is

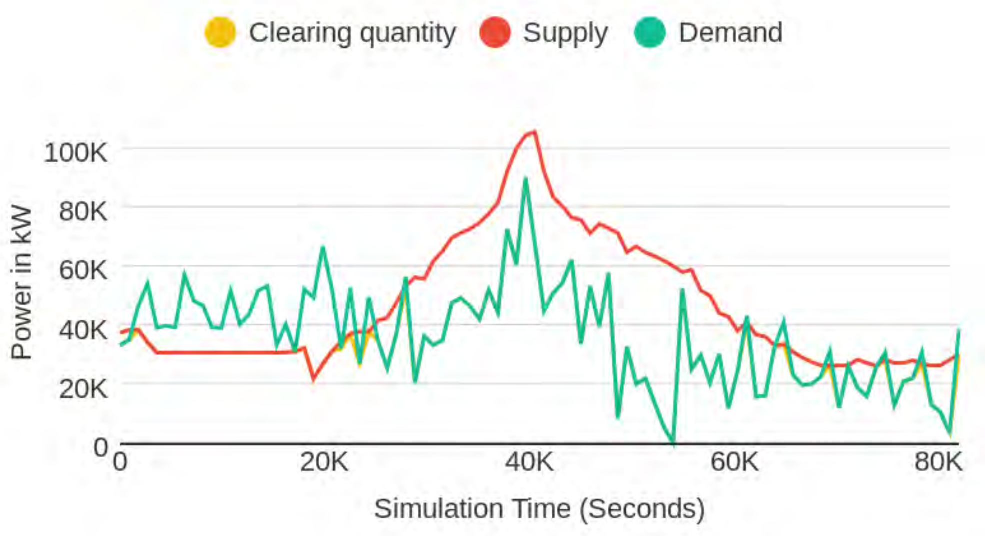

FIGURE 6.3: Market Clearing: The demand, supply and clearing quantity in kW correspond to y axis while the x axis represents to the simulation time which is 24 hours for the first scenario when the soil humidity calculation is based on the crop type and solar radiation.

affected in a different way; for example, cotton needs more water than wheat. Moreover, it is clear that the soil humidity is affected by the solar radiation and it is reduced faster during the day. The parameters of this test case are as follows:

- One cotton type field.
- The residential network accommodates 68 houses.
- The humidity set point used by the irrigation_controller is 45 for the cotton type.
- A solar panel area of 450 m^2.
- A wind turbine of 15 kW max power output.
- Two batteries of a maximum capacity of 20 kW h and initial available energy equal to 10 kW h are simulated.

The results of this test case are depicted in Figure 6.3. It is clear, that the energy demand for irrigating a field with cotton is smoother with fewer disturbances. More specifically, in that case, the irrigation system operates continuously for approximately four and a half hours instead of the seven 15 min discrete periods of time compared to wheat fields (see Figure 6.2). Moreover, the demanding energy to irrigate a cotton field is quite more by comparing the m^2 of the solar panel area that is needed; 450 m^2 instead of 250m^2 for the case with a wheat field.

A case more complex than the two aforementioned test cases is the case where the second scenario is used and more than one fields need irrigation. The following are the basic parameter characteristics of this case:

- Three fields are simulated. Two of them are wheat fields and the third one is a cotton field.

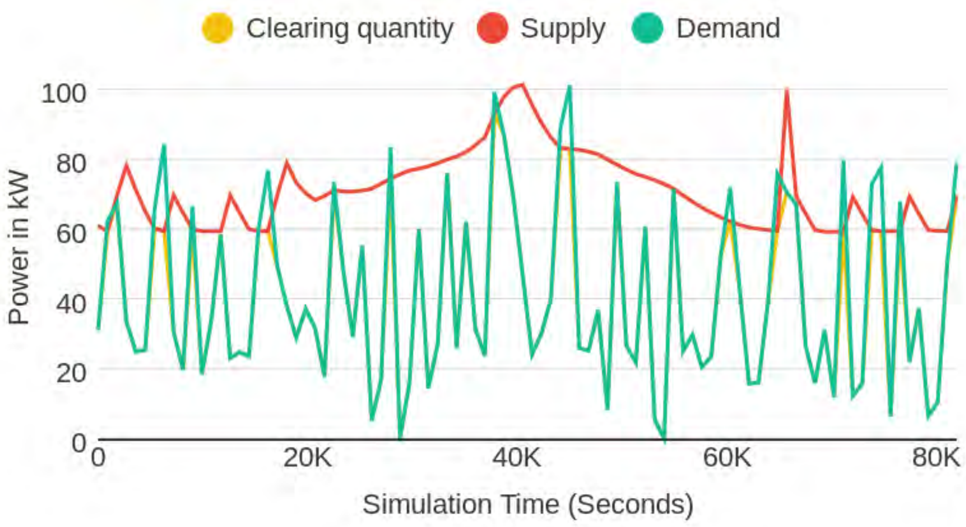

FIGURE 6.4: Market Clearing: The demand, supply and clearing quantity in kW correspond to y axis while the x axis represents the simulation time which is 24 hours for the second scenario with three fields.

- The humidity set point used by the irrigation_controller is set to 35 for the wheat field and to 45 for the cotton field.
- The residential network accommodates 68 houses.
- A solar panel area of 750 m^2.
- A wind turbine generator of 30 kW max power output.
- Two batteries of a maximum capacity of 70 kWh are utilized for storage purposes.

TABLE 6.1: Energy Resources Configuration

	Solar Panel		Wind Turbine		Battery	
Scenario	1	2	1	2	1	2
Max Capacity	51kW	85kW	15kW	30kW	20kWh	70kWh
Type	Single Crystal Silicon (SCS)		Small	Medium	Li-ion (Li) or Lead-acid (LA)	

The results illustrated in Figure 6.4, show that major part of both the agricultural and residential load is satisfied almost exclusively by the DES. Also, compared to Figure 6.2, the more the fields that are simulated, the less the wasted energy. The configuration of the DES for the last two presented test cases is depicted in Table 6.1, while the economic characteristics in terms of bid price are depicted in Table 6.2.

Since the fact that the consumer has to be price makers and the producers to be price takers, the price plays a dominant role in the operation of the market. This means that consumers participate in a type of a pool market playing the role of a decision maker. For instance, if the producers offer the energy at a price higher than the lowest price of the last consumer, this means that these producers are not included in the market dispatch. To that end, the demand is not fully satisfied in the

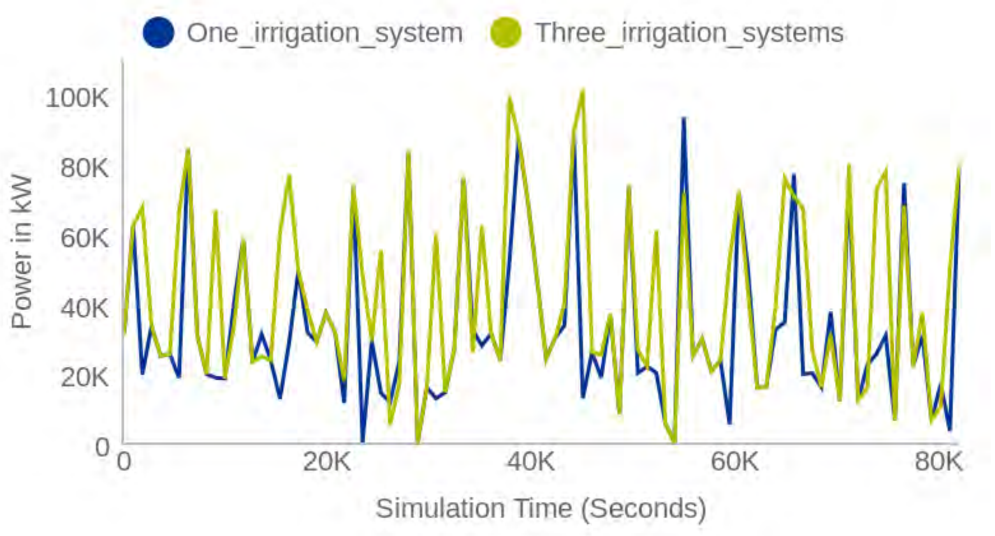

FIGURE 6.5: The residential and agricultural load for the cases with the one and the three fields.

TABLE 6.2: Energy Resources Price Configuration

	Solar Panel	Wind Turbine	Battery	Virtual Battery
Price($/kWh) (day/night)	0.25/0.30	0.25/0.25	0.30/0.25	0.30/0.25

market clearing price. Hence, more test cases by changing the price of the wind turbine during the night from 0.25 $\$/kWh$ to 0.30 $\$/kWh$ are tested. In that test case, it is observed that despite the fact that there is available energy to satisfy the overall demand, this doesn't happen. As aforementioned for experimental purposes, different values to the soil humidity of each field in order to change the operation of the irrigation system for each field have been set. Time shifting processes in water demand plays a significant role in an attempt to waste the least possible energy. Less amount of energy is wasted when the operation of the irrigation system varies from each field to another as it can be observed in Figure 6.5.

For the test cases presented in the next lines, the IEEE-13 test feeder instead of the R4_25.00_1 prototypical feeder is used. The test cases of both scenarios have the following characteristics:

- Ten fields with different crop types and humidity set points in the range of 25–45 in order to operate in different time intervals are simulated.
- The residential network accommodates 80 houses.
- Two areas of solar panels of 1500 m^2 for the first scenario and 3000 m^2 for the second one.
- A wind turbine of 150 kW max power output.
- Five battery sets of two batteries (basic and reserve) of a maximum capacity of 300 kWh for the first scenario and 400 kWh for the second.

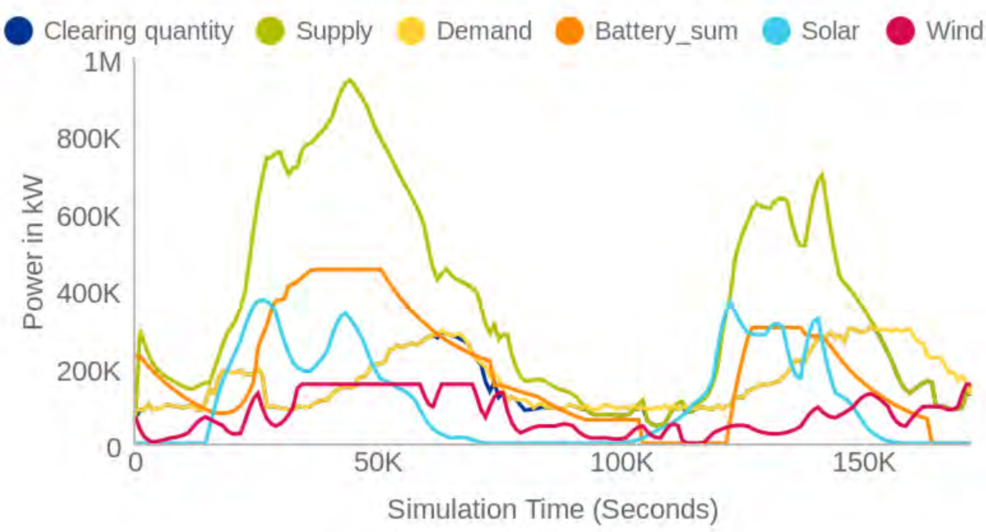

FIGURE 6.6: Market Clearing: Information about producing and consuming units (kW) corresponds to y axis while the x axis represents the simulation time which is two days for the test case associated with the first scenario and the IEEE-13 test feeder.

The results concerning the first and the second scenario are depicted in Figures 6.6 and 6.7 respectively. It is worth mentioning at that point that in these cases the 70% of overall demand corresponds to the agricultural load and the remaining 30% to the residential load (including the evcharger stations load). In the test case of the second scenario, three drilling systems associated with three virtual batteries of max capacity, 300 kWh each of them are simulated. The two out of three drilling systems serve 6 fields while the third one the remaining 4 fields. In Figure 6.8, the number of drilling systems that used each particular time instance during the simulation is presented. If the farmers prefer irrigating during the day, the pumped underground water is used by the irrigation system. On the other hand, when the farmers irrigate during the night, the virtual battery discharges by offering water in terms of power through its participation in the market framework as described earlier. The solar panels, the wind turbines, the batteries participating in the market for the whole simulation time and the price of each DES are illustrated in Figure 6.9. In this figure, the x-axis corresponds to the simulation time and the y-axis to each type of DES. More specifically, when the value of each DES is zero, this means that the DES does not participate in the market dispatch. On the other hand, when the value is not zero, it is included. In the former case, the price is 0.30 $\$/kWh$ while in the latter the price is 0.25 $\$/kWh$. As aforementioned, when a producer offers his energy at a price that is higher than the lowest price of the last consumer, this producer is not included in the market dispatch. In the experiments presented in this dissertation, this price is 0.30 $\$/kWh$.

For all the aforementioned test cases both the agricultural and residential load are satisfied at the same time on a yearly basis as depicted in Figure 6.10. More specifically, the agricultural network asks for energy for a time period of four months, called the agricultural season, while the residential network asks for the rest eight months, called the residential season. For this case, in the agricultural season, the agricultural load and the evcharger station load are satisfied by the energy from the DES, while in the residential season the load of the residential network and the load

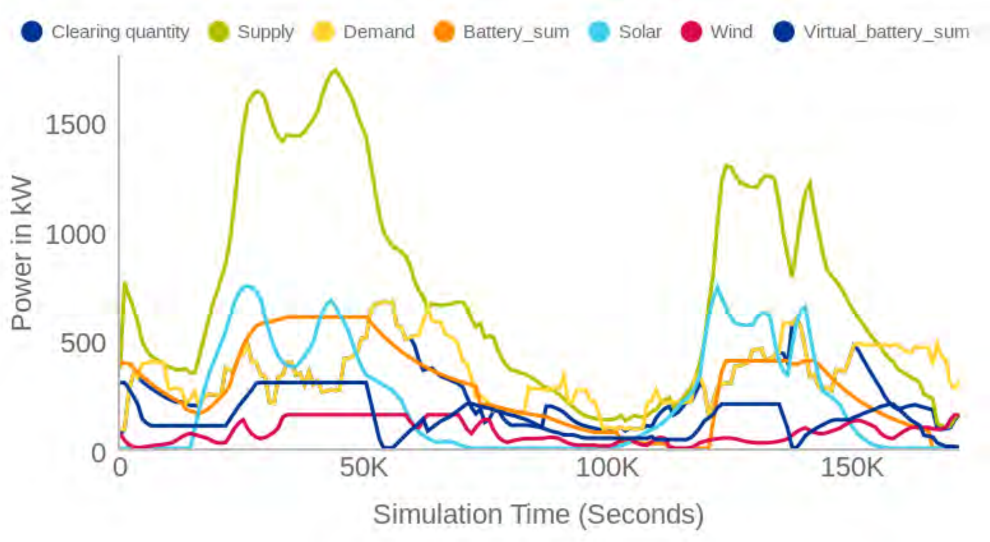

FIGURE 6.7: Market Clearing: Information about producing and consuming units (kW) corresponds to y axis while the x axis represents the simulation time which is two days for the test case associated with the second scenario and the IEEE-13 test feeder.

demanded by the evcharger stations are satisfied. Bids for the agricultural load, the evcharger station and the residential load are submitted through their controllers and stub_bidders in different seasons into the market. Bids for the agricultural load is submitted in the agricultural season while for the residential load in the residential season. The stub_bidder of evcharger station submits bids in both periods. The results of that case are illustrated in Figure 6.11. As a consequence of the above, there are a few cases where there is need of energy from the main grid while at most of the cases the energy provided by the DES is enough to satisfy both the agricultural and residential load. An isolated microgrid with these capabilities can be considered with the demand to be fully satisfied by RES and batteries. By locating larger batteries, there is less wasted energy because it can be stored in the batteries for future use.

6.4 Conclusions

In this chapter, the agricultural problem which needs further research and development in terms of better energy management and reduced cost is studied in detail. The under consideration problem is formulated based on two different scenarios regarding the way that the irrigation of the crops is handled; in the first scenario, the irrigation system corresponds to a system which uses the water from a surface water reservoir, while in the second scenario, a coupling of a drilling and an irrigation system is simulated. The objective is to integrate both the RES and the batteries to handle the agricultural load needed for irrigation purposes and the residential load necessary to cover the demand of households' appliances and evcharger stations. The results of the two different scenarios prove provided evidence that the RES can be used for completely covering the demand for both the agricultural and residential load. There are only a few cases where the energy offered by the RES and the batteries' capacity was not enough to cover the overall demand. In those cases, the energy is supplied by the main grid.

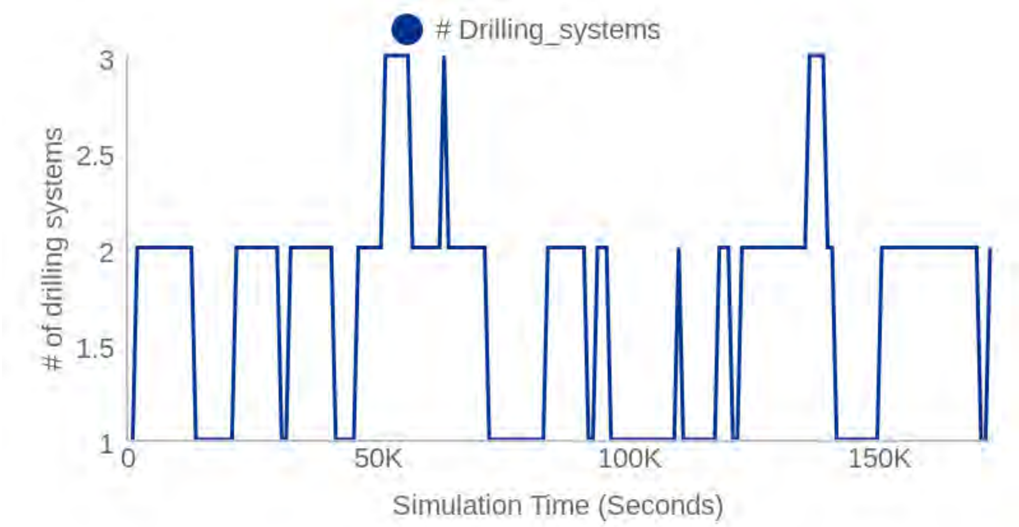

FIGURE 6.8: Number of drilling systems in operation for the whole simulation time.

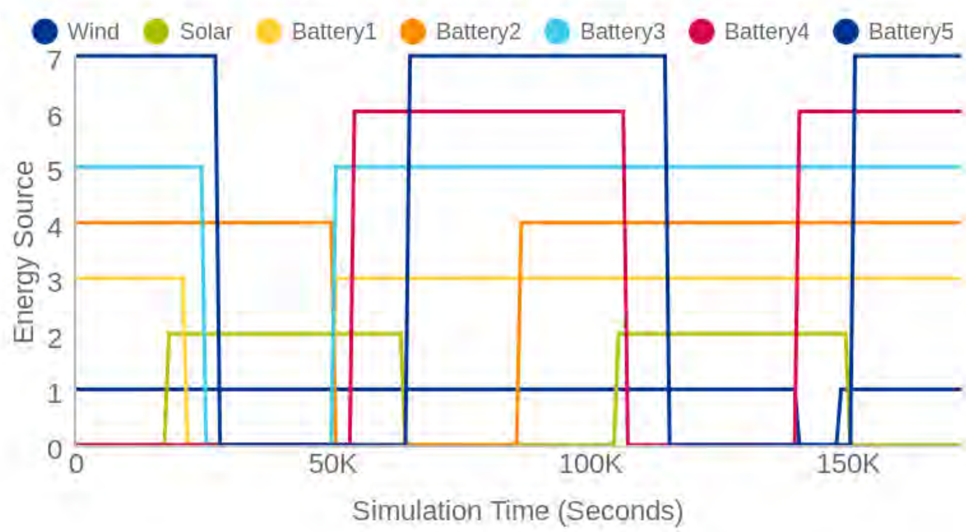

FIGURE 6.9: Energy sources that offer energy in the market each particular time for the second scenario.

FIGURE 6.10: Demand for the residential and agricultural load for both first and second scenario with the 10 fields for two days simulation.

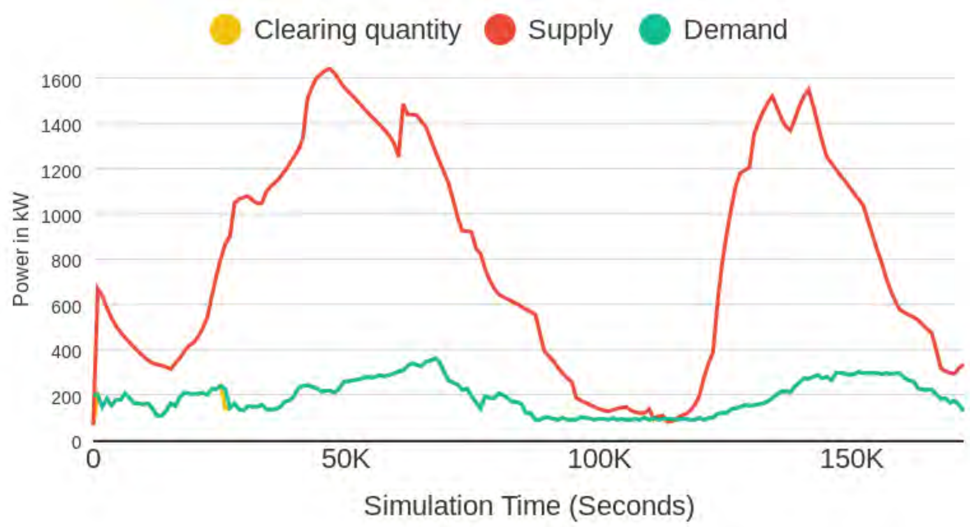

FIGURE 6.11: Demand for the residential and agricultural load with 10 fields for the two seasons. The first day (first half of the figure) corresponds to the agricultural season and the second one (second half of the figure) to the residential.

Chapter 7

Synopsis and Future Work

7.1 Synopsis of Contributions

The aim of this dissertation was the development, the implementation and the analysis of the simulation results in the models designed and developed based on the GridLAB-D simulation platform and machine learning techniques. In addition, the possible effects and benefits of the integration of the RES in the satisfaction of both the agricultural and residential load demand in the operation of the power grid are also examined. For this dissertation, we summarize in the following:

- GridLAB-D despite the fact that is one of the most prominent simulation tools for power grid simulations on the distribution level lacks a user-friendly graphical interface. Hence, we realize the necessity of developing a GUI for the GridLAB-D platform by introducing (a) two additional visualization layers and (b) a front end interface. The first layer corresponds to the visualization of the simulation results using techniques based on diagrams while the second based on Self- Organizing Maps approach. The front end interface helps every user regardless his background to create from scratch a distribution grid and use the GridLAB-D to demonstrate it.

- Four different energy market models to observe and analyze how the demand behaves with respect to the price are developed. Moreover, the ultimate goal was to check whether the interaction of the consumers in the energy markets contributes to the stability of the network, confirming indirectly the validation and the versatility of each model. The results confirm the assumption that the load shifted to off-peak periods (see Figure 1.2 in Chapter 1) through price elasticity, leads to a more stable system while accommodating the maximum possible demand and contribute to less unused energy. The results of the fourth model, which is associated with the use of a game theory model for increasing the profit of the producers, provide encouraging improvements regarding the amount that can be offered by the producers when the energy dispatched from the local markets is not enough. They verify that the cost affects the quantity that the producers could offer. Game theory is a very promising concept, which allows focusing on optimization of the existing implementation to maximize the profit of producers. Moreover, the second market model, in which DR techniques are used, also provides encouraging results by promising smoother demand curves when the appliances are controlled by properly with reasonable logic configured controllers.

- Besides the four energy market models, a model, which combines both the concepts of active and reactive power, is presented in this dissertation. More specifically, the design of the proposed model, the test cases, and selective experiments, which show how the price, that the consumers have to pay, is affected by the deficient of reactive power when the system needs more reactive power than it supplies are presented. The results show, as expected, that when the consumers need to buy additional reactive power from distributed and not distributed producers at a more “expensive” price.

- The k-means clustering algorithm is one of the most commonly used partitioning algorithms. Despite its wide use, it suffers from serious drawbacks. In this dissertation, an improved version of the k-means, called SkM, is proposed dealing with the selection of the initial cluster centers. The presented test cases results show that the SkM algorithm attains in most of the cases a better performance compared to standard k-means, the k-means++ and the fuzzy c-means. Furthermore, in most of the cases, the values obtained by the validity measures which measure the compactness and the well separation of the data in the same cluster are also better than those of k-means, k-means++ and fuzzy c-means. More specifically, the SkM prevails on 90% of the test cases comparing to standard k-means and the k-means++ and in 85% comparing to fuzzy c-means with respect to the DB index. Therefore, the proposed algorithm provides more compact clusters than the standard k-means algorithm, the k-means++ and the fuzzy c-means in an average of 88.3% of the considered test cases giving in that way the potentiality to use it for grouping the distribution grid in more compact clusters.

- A DDDAS methodology for implementing a smart grid management method is developed. It makes use of two intelligent tools, namely, the k-means and a fuzzy decision-making method trying to provide intelligent ways for high-level quality management. More specifically, an elitist approach, where the grid is partitioned into clusters and one of them is selected to benefit from buying energy generated by RES at a lower price, is presented. The data of both the partitioning and selection procedures are dynamically driven and depends on the current grid conditions. The results of the presented test cases show that the proposed methodology minimizes the unused energy and satisfies a higher number of residential consumers as compared to cases when the proposed DDDAS methodology is not applied.

- A three stage scheme for clustering a set of residential consumers is described in sub chapter 5.3. The main contributions to the proposed scheme are the fact the data used for the simulations are from the near past. The results of using the proposed scheme prove that the clustering of the consumers can be achieved even without using any historical data. To assess the validity of the proposed scheme, the DB index is used to measure the compactness and the degree of the well separation of the clusters. The DB index for all the considered test cases provides better values indicating in that way that the proposed methodology impacts in a way the better clustering of the residential consumers.

- The agricultural sector is very crucial to the safe and reliable operation of the power for many countries, in order to succeed further growth and development. In that sense and due to the high electric costs, the agricultural sector is a very promising area for smart grid applications. A very promising model which integrates the RES in the energy market is presented in this dissertation. The results provided evidence that the RES can be used for the complete satisfaction of the demand asked from the residential and agricultural sector.

7.2 Future Work

We conclude this dissertation by discussing some research challenges to the area of smart grid and simulations that lead to interesting research studies.

Besides the clustering in the power grid, there are more artificial intelligence techniques which are widely used in the energy sector. Some of them are the Particle Swarm Optimization, Ant Colony Optimization, Artificial Bee Colony which belong to the Swarm Intelligence group. Swarm Intelligence exploits the complex social behavior of birds, bees, ants and it formulates their actions to solve complex optimization problems. In the energy sector, it can solve energy market problems such as house appliances management, optimal allocation of reactive power and optimal allocation of spinning reserve. The motivation of swarm models is to mimic the interactions and the behaviors of swarm members either with their environment or with other members (Harison and Raja, 2012), (Hamedi, 2013), (Affijulla and Chauhan, 2011). These efforts can be extended for the more reliable and efficient operation of the power grid using DR techniques.

The concept of Web of Things (WoT) is recently used in the daily operation of an open energy market. A plethora of efforts have as a primary goal the smart grid management using this technology (Khan and Mouftah, 2011), (Ramakrishnan, 2013) and (Karnouskos and Izmaylova, 2009). The main concern is to focus on the operation of the energy market and the direct interaction of the consumers with the energy market using the concept of the WoT. The technologies that the Internet of Things (IoT) offers and especially the WoT, which is an extension of IoT, ensure ease development with existing third network software in accessing the functionality of the components of a smart grid. The heterogeneity among the components of the smart grid could be the motivation to use web services as a basis to build a framework of home networks. Moreover, in WoT, real world objects, such as appliances, are represented as web resources, which could be accessed using APIs based on REST and SOAP principles. So, the web services could be used to monitor, control and manage the operation of smart components or to send information to home networks about weather data, prices and power production. Also, web services may provide different functions such as the propagation of information about the network to compose energy usage patterns or providing information such as time of operation about the devices that are currently on the network.

A good idea is to focus on the development of frameworks that control, monitor and take information from components of the grid whose the functionality is published to the Web, something that has several benefits such as compatibility, flexibility, simplicity. Through a properly and friendly configured GUI, each user either a consumer or a utility provider can use the online web services which are associated with functions that characterize the operation of grid's components. So, the user can take information about the state of various appliances at a specific residency and various measurements from the smart meters. For instance, a consumer can exploit the information taken from an appliance or a smart meter to manage the energy usage at the household level or to control the operation of an appliance. A utility provider can control the operation of appliances if the consumer associated with that appliance permits the utility provider for this action.

Via web services, the utility provider also has the ability to state the change in price. In turn, the appropriate with this function web service is invoked by starting the procedure in order the appliances to adjust their operation. To the best of our knowledge, no actions for the consumers' side have been conducted in order to interact with their appliances when the energy price is changed. The need of running web services on embedded devices emerges the development of the Devices Profile for Web Services (DPWS) (Zhang et al., 2014) standard which is a set of the web services standards. As far as we know, the DPWS is solely used on smart meters. It is not currently used in home appliances in order to control them. DPWS devices could be coupled with a multi-agent system which enables on simulations with more non-deterministic decisions. DPWS also forwards the interoperability among heterogeneous devices that could be components of a wider network.

Finally, it is known that the stability of the power grid depends on the voltage. Few simulations have been conducted in order to figure out how the voltage, the angle and the reactive power in a single node are affected under various circumstances such as high energy demand at a particular time instance. We aim to extend this effort using monitoring techniques.

Bibliography

Affijulla, S. and S. Chauhan (2011). "Swarm intelligence solution to large scale thermal power plant Load Dispatch". In: *Emerging Trends in Electrical and Computer Technology (ICETECT), 2011 International Conference on*, pp. 196–199. DOI: 10.1109/ICETECT.2011.5760115.

Aggarwal, Charu C. and Chandan K. Reddy (2014). "Data Clustering: Algorithms and Applications". In: CRC Press.

Akagi, Hirofumi, E H Watanabe, and Mauricio Aredes (2017). "Instantaneous Power Theory and Applications to Power Conditioning". In: IEEE Press.

Al-Agtash, Salem (2013). "Electricity agents in smart grid markets". In: *Computers in Industry* 64.3, pp. 235 –241. ISSN: 0166-3615. DOI: http://dx.doi.org/10.1016/j.compind.2012.10.009.

Aliprantis, D. et al. (2010). "Integrated retail and wholesale power system operation with smart-grid functionality". In: *IEEE PES General Meeting*, pp. 1–8. DOI: 10.1109/PES.2010.5589594.

Anderson, K. et al. (2014). "GridSpice: A Distributed Simulation Platform for the Smart Grid". In: *IEEE Transactions on Industrial Informatics* 10.4, pp. 2354–2363. ISSN: 1551-3203. DOI: 10.1109/TII.2014.2332115.

Arruda, M. F. D. et al. (2015). "Solar and Wind powered Stand Alone Water Pumping system". In: *Renewable Energy and Power Quality Journal* 13.253, pp. 1–6.

Arthur, David and Sergei Vassilvitskii (2007). "K-means++: The Advantages of Careful Seeding". In: *in Proc 18th Annual ACM-SIAM Symposium on Discrete Algorithms(SODA)*, pp. 1027–1035.

Azad, S. A., A. B. M. S. Ali, and P. Wolfs (2014). "Identification of typical load profiles using K-means clustering algorithm". In: *Asia-Pacific World Congress on Computer Science and Engineering*, pp. 1–6. DOI: 10.1109/APWCCSE.2014.7053855.

Bahaj, A.S. and P.A.B. James (2016). "Design and Power Quality Improvement of Photovoltaic Power System". In: *Springer International Publishing AG, Switzerland*.

Bahramirad, Shaghayegh and Hossein Daneshi (2012). "Optimal Sizing of Smart Grid Storage Management System in a Microgrid". In: *Proceedings of the 2012 IEEE PES Innovative Smart Grid Technologies*. ISGT '12. Washington, DC, USA: IEEE Computer Society, pp. 1–7. ISBN: 978-1-4577-2158-8. DOI: 10.1109/ISGT.2012.6175774.

Balaji, A. Jayanth, D. S. Harish Ram, and Binoy B. Nair (2016). "Modeling of Consumption Data for Forecasting in Automated Metering Infrastructure (AMI) Systems". In: *Automation Control Theory Perspectives in Intelligent Systems: Proceedings of the 5th Computer Science On-line Conference 2016 (CSOC2016), Vol 3*. Ed. by Radek Silhavy et al. Cham: Springer International Publishing, pp. 165–173. ISBN: 978-3-319-33389-2. DOI: 10.1007/978-3-319-33389-2_16.

— (2017). "Machine Learning Approaches to Electricity Consumption Forecasting in Automated Metering Infrastructure (AMI) Systems: An Empirical Study". In: *Cybernetics and Mathematics Applications in Intelligent Systems: Proceedings of the 6th Computer Science On-line Conference 2017 (CSOC2017), Vol 2*. Cham: Springer

International Publishing, pp. 254–263. ISBN: 978-3-319-57264-2. DOI: 10.1007/978-3-319-57264-2_26.

Barroso, Luiz Augusto N. et al. (2002). "Market Power Issues in Bid-Based Hydrothermal Dispatch". In: *Annals of Operations Research* 117.1, pp. 247–270. ISSN: 1572-9338. DOI: 10.1023/A:1021537910823.

Ben, I. et al. (2014). "Dynamic clustering segmentation applied to load profiles of energy consumption from Spanish customers". In: *Electrical Power and Energy Systems* 55, pp. 437–448.

Blaabjerg, F. and K. Ma (2013). "Future on Power Electronics for Wind Turbine Systems". In: *IEEE Journal of Emerging and Selected Topics in Power Electronics* 1.3, pp. 139–152. ISSN: 2168-6777. DOI: 10.1109/JESTPE.2013.2275978.

Blasch, E., Y. Al-Nashif, and S. Hariri (2014). "Static versus dynamic data information fusion analysis using DDDAS for cyber security trust". In: *Procedia Computer Science* 29, pp. 1299–1313.

Blasch, E., G. Seetharaman, and F. Darema (2013). "Dynamic Data Driven Applications Systems (DDDAS) modeling for automatic target recognition". In: *in Proc. SPIE 8744, Automatic Target Recognition XXIII*.

Celebi, M. Emre, Hassan A. Kingravi, and Patricio A. Vela (2013). "A comparative study of efficient initialization methods for the k-means clustering algorithm". In: *Expert Systems with Applications* 40.1, pp. 200 –210. ISSN: 0957-4174. DOI: http://dx.doi.org/10.1016/j.eswa.2012.07.021.

Celik, N., and J. P. Saenz (2013). "DDDAMS-based dispatch control in power networks". In: *13th Annual International Conference on Computational Science*, pp. 1899–1908.

Chassin, D. P. (2008). *GridLAB-D Technical Support Document Tape Modules Version 1.0*. eng. Washington, D.C : Oak Ridge, Tenn.: United States. Dept. of Energy ; distributed by the Office of Scientific and Technical Information, U.S. Dept. of Energy.

Chassin, David P., Jason C. Fuller, and Ned Djilali (2014). "GridLAB-D: an agent-based simulation framework for smart grids.(Research Article)(Report)". In: *Journal of Applied Mathematics* 2014. ISSN: 1110-757X.

Conejo, Antonio J., Miguel Carrión, and Juan M. Morales (2010). "Decision Making Under Uncertainty in Electricity Markets". In: vol. 153. Springer US. ISBN: 978-0-387-77744-3.

Connell, Niamh O et al. (2014). "Benefits and challenges of electrical demand response: A critical review". In: *Renewable and Sustainable Energy Reviews* 39, pp. 686 –699. ISSN: 1364-0321. DOI: http://dx.doi.org/10.1016/j.rser.2014.07.098.

Contreras, J., M. Klusch, and J. B. Krawczyk (2004). "Numerical solutions to Nash-Cournot equilibria in coupled constraint electricity markets". In: *IEEE Transactions on Power Systems* 19.1, pp. 195–206. ISSN: 0885-8950. DOI: 10.1109/TPWRS.2003.820692.

Cupertino, A. F. et al. (2012). "A Contribution to Isolated and Grid-Connected Photovoltaic Systems under Shadow Conditions". In: *International Conference on Renewable Energies and Power Quality*, pp. 1–6.

Darema, F. (2000). "Dynamic data driven applications systems: A new paradigm for application simulations and a new paradigm for measurement systems". In: *NSF Workshop*.

David, A. and V. Sergei (2007). "K-means++: The Advantages of Careful Seeding". In: pp. 1027–1035.

Dursun, M. and S. Ozden (2012). "Application of Solar Powered Automatic Water Pumping in Turkey". In: *International Journal of Computer & Electrical Engineering* 4.2, pp. 161–164.

El-Samahy, I., K. Bhattacharya, and C.A. Canizares (2006). "A Unified Framework for Reactive Power Management in Deregulated Electricity Markets". In: *Power Systems Conference and Exposition, 2006. PSCE '06. 2006 IEEE PES*, pp. 901–907. DOI: 10.1109/PSCE.2006.296434.

Energy, U.S. Department of (2006). "Benefits of Demand Response and Recommendations". In:

— (2009). "The Smart Grid: An introduction". In:

Fainti, R. et al. (2014a). "Large scale simulations for electric energy markets". In: *IISA 2014, The 5th International Conference on Information, Intelligence, Systems and Applications*, pp. 160–165. DOI: 10.1109/IISA.2014.6878814.

Fainti, Rafik et al. (2014b). "Design and Early Simulations of Next Generation Intelligent Energy Systems". In: *Int. J. Monit. Surveill. Technol. Res.* 2.2, pp. 58–82. ISSN: 2166-7241. DOI: 10.4018/ijmstr.2014040104.

Fang, X. et al. (2012). "Smart Grid; The New and Improved Power Grid: A Survey". In: *IEEE Communications Surveys Tutorials* 14.4, pp. 944–980. ISSN: 1553-877X. DOI: 10.1109/SURV.2011.101911.00087.

Faria, P. and Z. Vale (2011). "Demand response in electrical energy supply: An optimal real time pricing approach". In: *Energy* 36.8. PRES 2010, pp. 5374 –5384. ISSN: 0360-5442. DOI: http://dx.doi.org/10.1016/j.energy.2011.06.049.

Faruque, M. A. Al and F. Ahourai (2014). "GridMat: Matlab toolbox for GridLAB-D to analyze grid impact and validate residential microgrid level energy management algorithms". In: *ISGT 2014*, pp. 1–5. DOI: 10.1109/ISGT.2014.6816479.

Feng, Donghan, Jin Zhong, and Deqiang Gan (2008). "Reactive Market Power Analysis Using Must-Run Indices". In: *Power Systems, IEEE Transactions on* 23.2, pp. 755–765. ISSN: 0885-8950. DOI: 10.1109/TPWRS.2008.919416.

Fernandes, M. P. et al. (2016). "Analysis of residential natural gas consumers using fuzzy c-means clustering". In: *2016 IEEE International Conference on Fuzzy Systems (FUZZ-IEEE)*, pp. 1484–1491. DOI: 10.1109/FUZZ-IEEE.2016.7737865.

Foti, Magda et al. (2016). "On Visualizing Distribution Systems for Next Generation Power Distribution Grids". In: *International Journal of Computational and Neural Engineering (IJCNE)* 3.1, pp. 16–27.

Fuller, J.C. (2009). *Modeling Wind Turbines in the GridLAB-D Software Environment*. eng. Washington, D.C. : Oak Ridge, Tenn.: United States. Dept. of Energy. Office of Science ; distributed by the Office of Scientific and Technical Information, U.S. Dept. of Energy.

Ganger, D. et al. (2013). "A review and simulation on real time reactive power spot markets". In: *2013 North American Power Symposium (NAPS)*, pp. 1–5. DOI: 10.1109/NAPS.2013.6666891.

Gantz, J. and D. Reinsel (2011). "Extracting Value from Chaos". In: *EMC Corporation*, pp. 1–12.

Ghazi, R. and M.B. Asadi (2010). "Reactive power pricing using marginal cost theory in competitive electricity markets". In: *2010 IEEE International Energy Conference and Exhibition, EnergyCon 2010*, pp. 369–372.

Grimm, Veronika et al. (2016). "Transmission and generation investment in electricity markets: The effects of market splitting and network fee regimes". In: *European Journal of Operational Research* 254.2, pp. 493 –509. ISSN: 0377-2217. DOI: http://dx.doi.org/10.1016/j.ejor.2016.03.044.

Gustavo, Molina Marcelo and Alvarez Juan Gimenez (2011). "Technical and Regulatory Exigencies for Grid Connection of Wind Generation". In: InTech. Chap. 01. ISBN: 10.5772/16474.

Hamedi, Hadi (2013). "Solving the combined economic load and emission dispatch problems using new heuristic algorithm". In: 46, pp. 10–16. DOI: 10.1109/ICCEET.2012.6203848.

Harison, D.S. and T.S.R. Raja (2012). "Application of particle swarm optimization for auction based dispatch in deregulated environment". In: *Computing, Electronics and Electrical Technologies (ICCEET), 2012 International Conference on*, pp. 116–121. DOI: 10.1109/ICCEET.2012.6203848.

Hashem, Ibrahim Abaker Targio et al. (2015). "The rise of big data on cloud computing: Review and open research issues". In: *Information Systems* 47, pp. 98 –115. ISSN: 0306-4379. DOI: http://dx.doi.org/10.1016/j.is.2014.07.006.

Hassani, Marwan and Thomas Seidl (2017). "Using internal evaluation measures to validate the quality of diverse stream clustering algorithms". In: *Vietnam Journal of Computer Science* 4.3, pp. 171–183. ISSN: 2196-8896. DOI: 10.1007/s40595-016-0086-9.

Hatziargyriou, N. (2014). "Microgrids:Architectures and Control". In: *Wiley-IEEE Press*, pp. 1–340.

Hejazi, H. A. et al. (2013). "Independent distributed generation planning to profit both utility and DG investors". In: *IEEE Transactions on Power Systems* 28.2, pp. 1170–1178. ISSN: 0885-8950. DOI: 10.1109/TPWRS.2012.2219325.

Hernandez, Luis et al. (2012). "Classification and Clustering of Electricity Demand Patterns in Industrial Parks". In: *Energies* 5.12, pp. 5215–5228. DOI: 10.3390/en5125215.

Hiko, E. A. T. M. (2011). "Electricity in new zealand". In: *Electricity Authority - Te Mana Hiko, Tech. Rep.*

Hogan, William W. (1996). "Markets in Real Electric Networks Require Reactive Prices". English. In: *Electricity Transmission Pricing and Technology*. Ed. by Michael Einhorn and Riaz Siddiqi. Springer Netherlands, pp. 143–182. ISBN: 978-94-010-3831-7. DOI: 10.1007/978-94-010-0710-8_7.

Houstis, Elias et al. (2017). "Linking Smart Energy and Smart Irrigation: Integration, System Architecture, Prototype Implementation and Experimentation". In: *3rd International Congress on Energy Efficiency and Energy Related Materials (ENEFM2015): Proceedings, Oludeniz, Turkey, 19–23 October 2015*. Ed. by Ahmet Yavuz Oral and Zehra Banu Bahsi Oral. Cham: Springer International Publishing, pp. 143–149. ISBN: 978-3-319-45677-5. DOI: 10.1007/978-3-319-45677-5_17.

Jain, A. K., M. N. Murty, and P. J. Flynn (1999). "Data Clustering: A Review". In: *ACM Comput. Surv.* 31.3, pp. 264–323. ISSN: 0360-0300. DOI: 10.1145/331499.331504.

Jarrah, Moath et al. (2015). "A hierarchical optimization model for energy data flow in smart grid power systems". In: *Information Systems* 53, pp. 190 –200. ISSN: 0306-4379. DOI: http://dx.doi.org/10.1016/j.is.2014.12.003.

Jayanth Balaji, A., D. S. Harish Ram, and Binoy B. Nair (2016). "Modeling of Consumption Data for Forecasting in Automated Metering Infrastructure (AMI) Systems". In: *Automation Control Theory Perspectives in Intelligent Systems: Proceedings of the 5th Computer Science On-line Conference 2016 (CSOC2016), Vol 3*. Ed. by Radek Silhavy et al. Cham: Springer International Publishing, pp. 165–173. ISBN: 978-3-319-33389-2. DOI: 10.1007/978-3-319-33389-2_16.

Kalogirou, Soteris A. (2013). "Solar Energy Engineering: Processes and Systems". In: Elsevier.

Kannan, Nadarajah and Divagar Vakeesan (2016). "Solar energy for future world: - A review". In: *Renewable and Sustainable Energy Reviews* 62, pp. 1092 –1105. ISSN: 1364-0321. DOI: http://dx.doi.org/10.1016/j.rser.2016.05.022.

Kaplan, Yusuf Alper (2015). "Overview of wind energy in the world and assessment of current wind energy policies in Turkey". In: *Renewable and Sustainable Energy Reviews* 43, pp. 562 –568. ISSN: 1364-0321. DOI: http://dx.doi.org/10.1016/j.rser.2014.11.027.

Karnouskos, S. and A. Izmaylova (2009). "Simulation of web service enabled smart meters in an event-based infrastructure". In: *IEEE International Conference on Industrial Informatics (INDIN)*, pp. 125–130. DOI: 10.1109/INDIN.2009.5195790.

Kathiresan, V. and P. Sumathi (2012). "An efficient clustering algorithm based on Z-Score ranking method". In: *2012 International Conference on Computer Communication and Informatics*, pp. 1–4. DOI: 10.1109/ICCCI.2012.6158779.

Kelley, L. C. et al. (2010). "On the feasibility of solar-powered irrigation". In: *Renewable and Sustainable Energy Reviews* 14.9, pp. 2669–2682.

Kersting, W. H. and G. Shirek (2012). "Short circuit analysis of IEEE test feeders". In: *PES T D 2012*, pp. 1–9. DOI: 10.1109/TDC.2012.6281539.

Kersting, William H. (2012). *Distribution System Modeling and Analysis, Third Edition*. CRC Press.

Khaleghi, A. M. et al. (2013). "A DDDAMS-based planning and control framework for surveillance and crowd control via UAVs and UGVs". In: *Expert Systems with Applications* 40, pp. 7168–7183.

Khan, A.A. and H.T. Mouftah (2011). "Web services for indoor energy management in a smart grid environment". In: *IEEE International Symposium on Personal, Indoor and Mobile Radio Communications, PIMRC*, pp. 1036–1040. DOI: 10.1109/PIMRC.2011.6139654.

Koliou, Elta et al. (2014). "Demand response in liberalized electricity markets: Analysis of aggregated load participation in the German balancing mechanism". In: *Energy* 71, pp. 245 –254. ISSN: 0360-5442. DOI: http://dx.doi.org/10.1016/j.energy.2014.04.067.

Lagari, P. L. et al. (2016). "Evaluation of Human Machine Interface (HMI) in Nuclear Power Plants with Fuzzy Logic method". In: *2016 7th International Conference on Information, Intelligence, Systems Applications (IISA)*, pp. 1–6. DOI: 10.1109/IISA.2016.7785345.

Leou, Rong-Ceng (2012). "An economic analysis model for the energy storage system applied to a distribution substation". In: *International Journal of Electrical Power and Energy Systems* 34.1, pp. 132 –137. ISSN: 0142-0615. DOI: http://dx.doi.org/10.1016/j.ijepes.2011.09.016.

Li, Gong, Jing Shi, and Xiuli Qu (2011a). "Modeling methods for GenCo bidding strategy optimization in the liberalized electricity spot market :A state-of-the-art review". In: *Energy* 36.8. PRES 2010, pp. 4686 –4700. ISSN: 0360-5442. DOI: http://dx.doi.org/10.1016/j.energy.2011.06.015.

— (2011b). "Modeling methods for GenCo bidding strategy optimization in the liberalized electricity spot market "A state-of-the-art review". In: *Energy* 36.8. PRES 2010, pp. 4686 –4700. ISSN: 0360-5442. DOI: http://dx.doi.org/10.1016/j.energy.2011.06.015.

Li, X.R., C.W. Yu, and W.H. Chen (2012). "A novel value based reactive power procurement scheme in electricity markets". In: *International Journal of Electrical Power and Energy Systems* 43.1, pp. 910 –914. ISSN: 0142-0615. DOI: http://dx.doi.org/10.1016/j.ijepes.2012.05.037.

M. Erisoglu N.Calis, S. Sakallioglu (2011). "A new algorithm for initial cluster centers in k-means algorithm". In: *Int. J. Pattern Recognition Letter* 32.14, pp. 1704–1705.

M. S. Mahmud M. M. Rahman, M. N. Akhtar (2012). "Improvement of K-means clustering algorithm with better initial centroids based on weighted average". In: *in Proc. 7th International Conference on Electrical & Computer Engineering(ICECE)*.

Maass, H. et al. (2013). "First Evaluation Results Using the New Electrical Data Recorder for Power Grid Analysis". In: *IEEE Transactions on Instrumentation and Measurement* 62.9, pp. 2384–2390. ISSN: 0018-9456. DOI: 10.1109/TIM.2013.2270923.

Mangiatordi, F., E. Pallotti, and P. Del Vecchio (2013). "A non cooperative game theoretic approach for energy management in MV grid". In: *2013 13th International Conference on Environment and Electrical Engineering (EEEIC)*, pp. 266–271. DOI: 10.1109/EEEIC-2.2013.6737920.

Manoharan, Jamesnd J. and Hari S. Ganesh (2016). "Initialization of optimized K-means centroids using divide-and-conquer method". In: *ARPN Journal of Engineering and Applied Sciences* 11.2, pp. 1086–1091.

Martin, G. (2007). "Aquifer Underground Pumped Hydroelectric Energy Storage". MA thesis. University of Colorado at Boulder.

Martin, G. and J. Levine (2007). *Renewable Energy Generation and Storage For Agricultural Use in the San Luis Valley: A Feasibility Study Report*. Tech. rep. University of Colorado at Boulder.

Mets, K., F. Depuydt, and C. Develder (2016). "Two-Stage Load Pattern Clustering Using Fast Wavelet Transformation". In: *IEEE Trans Smart Grid* 7.5, pp. 2250–2259.

Michalski, R.S., J.G. Carbonell, and T.M. Mitchell (2013). "Machine Learning: An Artificial Intelligence Approach". In: Springer- Verlag.

Moghaddam, M. Parsa, A. Abdollahi, and M. Rashidinejad (2011). "Flexible demand response programs modeling in competitive electricity markets". In: *Applied Energy* 88.9, pp. 3257 –3269. ISSN: 0306-2619. DOI: http://dx.doi.org/10.1016/j.apenergy.2011.02.039.

Moore, D. E. and N. Celik (2014). "DDDAS-based Communication in Distributed Smartgrid Networks". In: *In Proceedings of the Annual Industrial and Systems Engineering Research Conference*.

Morales, T. D. and J. Busch (2010). *Design of Small Photovoltaic Solar-Powered Water Pump Systems*. Tech. rep. 28. U.S. Department of Agriculture.

Nasiakou, A., M. Alamaniotis, and L. H. Tsoukalas (2016a). "MatGridGUI; A toolbox for GridLAB-D simulation platform". In: *2016 7th International Conference on Information, Intelligence, Systems Applications (IISA)*, pp. 1–5. DOI: 10.1109/IISA.2016.7785419.

— (2017). "A Three-stage Scheme for Consumers' Partitioning Using Hierarchical Clustering Algorithm". In: *8th International Conference on Information, Intelligence, Systems and Applications(IISA2017)*, pp. 1–6.

Nasiakou, A., M. Vavalis, and D. Bargiotas (2015). "Simulating active and reactive energy markets". In: *2015 6th International Conference on Information, Intelligence, Systems and Applications (IISA)*, pp. 1–6. DOI: 10.1109/IISA.2015.7388037.

Nasiakou, Antonia, Miltiadis Alamaniotis, and Lefteri H. Tsoukalas (2016b). "Extending the K-Means Clustering Algorithm to Improve the Compactness of the Clusters". In: *Journal of Pattern Recognition Research* 11.1, pp. 61–73.

— (2016c). "Power distribution network partitioning in big data environment using k-means and fuzzy logic". In: *IET Conference Proceedings*, pp. 89–97.

Nasiakou, Antonia, Manolis Vavalis, and Dimitris Zimeris (2016). "Smart energy for smart irrigation". In: *Computers and Electronics in Agriculture* 129, pp. 74 –83. ISSN: 0168-1699. DOI: http://dx.doi.org/10.1016/j.compag.2016.09.008.

Nasiakou, Antonia et al. (2017). "Dynamic Data Driven Partitioning of Smart Grid Using Learning Methods". In: Springer International Publishing.

Panagiotakopoulou, P. (2015). "Different types of electricity markets modelled using PLEXOS Integrated Energy Model – The UK Balancing Market example". In: *Press*.

Panapakidis, I.P. et al. (2013). "Analysis of the Electricity Demand Patterns of a Building in a University Campus". In: *IEEE 12th International Conference on Environment and Electrical Engineering*, pp. 1–6.

Prabavathi, M. and R. Gnanadass (2015). "Energy bidding strategies for restructured electricity market". In: *International Journal of Electrical Power and Energy Systems* 64, pp. 956 –966. ISSN: 0142-0615. DOI: http://dx.doi.org/10.1016/j.ijepes.2014.08.018.

Quilumba, F. L. et al. (2015). "Using Smart Meter Data to Improve the Accuracy of Intraday Load Forecasting Considering Customer Behavior Similarities". In: *IEEE Transactions on Smart Grid* 6.2, pp. 911–918. ISSN: 1949-3053. DOI: 10.1109/TSG.2014.2364233.

R. Fetea, A. Petroianu (2012). "Reactive Power: A strange concept". In: *In Press*.

Rabiee, A. and and N. Amjady H. Shayanfar (2009). "Coupled energy and reactive power market clearing considering power system security". In: *Energy Conversion and Management* 50.4, pp. 907–915.

Rajveer Mittal, K.S.Sandhu and D.K.Jain (2010). "An Overview of Some Important Issues Related to Wind Energy Conversion System (WECS)". In: *International Journal of Environmental Science and Development* 1.4, pp. 351–363.

Ramakrishnan, S. (2013). "WoT (Web of Things) for Energy Management in a Smart Grid-Connected Home". In: *Proceedings of the Informing Science and Information Technology Education Conference*, pp. 461–473.

Razykov, T.M. et al. (2011). "Solar photovoltaic electricity: Current status and future prospects". In: *Solar Energy* 85.8. Progress in Solar Energy 1, pp. 1580 –1608. ISSN: 0038-092X. DOI: http://dx.doi.org/10.1016/j.solener.2010.12.002.

Reddy, S.S., A.R. Abhyankar, and P.R. Bijwe (2011). "Market clearing of joint energy and reactive power using multi objective optimization considering voltage dependent load models". In: *Power and Energy Society General Meeting, 2011 IEEE*, pp. 1–8. DOI: 10.1109/PES.2011.6039652.

Rendon, Erendira et al. (2011). "A Comparison of Internal and External Cluster Validation Indexes". In: *Proceedings of the 2011 American Conference on Applied Mathematics and the 5th WSEAS International Conference on Computer Engineering and Applications*. AMERICAN-MATH'11/CEA'11. Puerto Morelos, Mexico: World Scientific, Engineering Academy, and Society (WSEAS), pp. 158–163. ISBN: 978-960-474-270-7.

Sadaaki, Miyamoto (2012). "An Overview of Hierarchical and Non-hierarchical Algorithms of Clustering for Semi-supervised Classification". In: *Modeling Decisions for Artificial Intelligence: 9th International Conference, MDAI 2012, Girona, Catalonia, Spain, November 21-23, 2012. Proceedings*. Ed. by Vicenç Torra et al. Berlin, Heidelberg: Springer Berlin Heidelberg, pp. 1–10. ISBN: 978-3-642-34620-0. DOI: 10.1007/978-3-642-34620-0_1.

Saebi, J. et al. (2013). “A new framework for reactive power dispatch in electricity markets”. In: *Research Journal of Applied Sciences, Engineering and Technology* 5.2, pp. 380–386.

Saini, Ashish and Amit Saraswat (2012). “Solving a Multi-objective Reactive power market clearing market using NSGA-II”. In: *International Journal of Advanced Information Technology (IJAIT)* 2.3, pp. 49–62. DOI: 10.5121/ijait.2012.2304.

Sanchez-Gomez, J. I. et al. (2016). “Fuzzy c-means clustering for steady state events classification of electrical signals”. In: *2016 12th Congreso Internacional de Ingenieria (CONIIN)*, pp. 1–5. DOI: 10.1109/CONIIN.2016.7498129.

Schneider, K. P. et al. (2009). “Distribution power flow for smart grid technologies”. In: *2009 IEEE/PES Power Systems Conference and Exposition*, pp. 1–7. DOI: 10.1109/PSCE.2009.4840078.

Sebastian, S. and V. Margaret (2016). “Application of demand response programs for residential loads to minimize energy cost”. In: *2016 International Conference on Circuit, Power and Computing Technologies (ICCPCT)*, pp. 1–4. DOI: 10.1109/ICCPCT.2016.7530345.

Shi, X., H. Damgacioglu, and N.A. Celik (2015). “Dynamic Data-driven Approach for Operation Planning of Microgrids”. In: *International Conference On Computational Science*, pp. 2543–2552.

Siano, Pierluigi (2014). “Demand response and smart grids: A survey”. In: *Renewable and Sustainable Energy Reviews* 30, pp. 461 –478. ISSN: 1364-0321. DOI: http://dx.doi.org/10.1016/j.rser.2013.10.022.

Song, Hui, Michael Gallagher, and Siobhan Clarke (2012). “Rapid GUI Development on Legacy Systems: A Runtime Model-based Solution”. In: *Proceedings of the 7th Workshop on Models@Run.Time*. MRT ’12. Innsbruck, Austria: ACM, pp. 25–30. ISBN: 978-1-4503-1802-0. DOI: 10.1145/2422518.2422523.

T. Stambouli, and J. M. Faci and N. Zapata (2014). “Water and energy management in an automated irrigation district”. In: *Agricultural Water Management* 142, pp. 66–76.

T. Taylor, z. (2008). *GridLAB-D Technical Support Document Residential End-Use Module Version 1.0*. eng. Washington, D.C : Oak Ridge, Tenn.: United States. Dept. of Energy ; distributed by the Office of Scientific and Technical Information, U.S. Dept. of Energy.

Tadelis, Steven (2013). “Game Theory: An Introduction”. In: Princeton University Press.

Tenney, Nathan D. (2008). *GridLAB-D Technical Support Document: Climate Module Version 1.0 - PNNL-17679*. English. Tech. rep.

Thanos, A. et al. (2015). “Dynamic Data Driven Adaptive Simulation Framework for Automated Control in Microgrids”. In: *IEEE Transactions on Smart Grid*.

Thanos, A. E. et al. (2014). “DDDAMS framework for real-time load dispatching in power networks”. In: *Proceedings of the 2013 Winter Simulation Conference - Simulation: Making Decisions in a Complex World*, pp. 1893–1904.

Thimmapuram, P. R. and J. Kim (2013). “Consumers’ Price Elasticity of Demand Modeling With Economic Effects on Electricity Markets Using an Agent-Based Model”. In: *IEEE Transactions on Smart Grid* 4.1, pp. 390–397. ISSN: 1949-3053. DOI: 10.1109/TSG.2012.2234487.

Tsoukalas, E. et al. (2014). “Towards next generation intelligent energy systems: Design and simulations engines”. In: *IISA 2014, The 5th International Conference on Information, Intelligence, Systems and Applications*, pp. 412–418. DOI: 10.1109/IISA.2014.6878829.

Tuballa, Maria Lorena and Michael Lochinvar Abundo (2016). "A review of the development of Smart Grid technologies". In: *Renewable and Sustainable Energy Reviews* 59, pp. 710 –725. ISSN: 1364-0321. DOI: http://dx.doi.org/10.1016/j.rser.2016.01.011.

Usman, G., U.Ahmad, and M. Ahmad (2013). "Improved K-Means Clustering Algorithm by Getting Initial Cenroids". In: *World Applied Sciences Journal* 27.4, pp. 543–551.

Vaskovskaya, T.A. (2014). "Market price signals for customers for compensation of reactive power". In: *European Energy Market (EEM), 2014 11th International Conference on the*, pp. 1–4. DOI: 10.1109/EEM.2014.6861235.

Vick, B. D. and L. K. Almas (2011). "Developing Wind and/or Solar Powered Crop Irrigation Systems for the Great Plains". In: *Applied engineering in agriculture* 27.2, pp. 235–245.

Vij, R. and A. S. Kumar (2012). "Improved k- means clustering algorithm for two dimensional data". In: *in Proc. 2nd International Conference on Computational Science, Engineering and Information Technology(CCSEIT)* 27.4, pp. 665–670.

Vishwa, N. M. et al. (2015). "Scope and Review of Photovoltaic Solar Water Pumping System as a Sustainable Solution Enhancing Water Use Efficiency in Irrigation". In: *American Journal of Biological and Environmental Statistics* 1.1, pp. 1–8.

Wang, Qi et al. (2015). "Review of real-time electricity markets for integrating Distributed Energy Resources and Demand Response". In: *Applied Energy* 138, pp. 695 –706. ISSN: 0306-2619. DOI: http://dx.doi.org/10.1016/j.apenergy.2014.10.048.

Wogrin, S., J. Barquín, and E. Centeno (2013). "Capacity Expansion Equilibria in Liberalized Electricity Markets: An EPEC Approach". In: *IEEE Transactions on Power Systems* 28.2, pp. 1531–1539. ISSN: 0885-8950. DOI: 10.1109/TPWRS.2012.2217510.

Wong, K. C. (2015). "A Short Survey on Data Clustering Algorithms". In: *2015 Second International Conference on Soft Computing and Machine Intelligence (ISCMI)*, pp. 64–68. DOI: 10.1109/ISCMI.2015.10.

Xiao, H. et al. (2014). "Sizing of battery energy storage for micro-grid considering optimal operation management". In: *2014 International Conference on Power System Technology*, pp. 3162–3169. DOI: 10.1109/POWERCON.2014.6993661.

Y. Shivrath, and P. B.Narayana and and E. L. Narsaiah S. Thirumalasetty (2012). "Design & Integration of Wind-Solar Hybrid Energy System for Drip Irrigation Pumping Application". In: *International Journal of Modern Engineering research* 2.4, pp. 2947–2950.

Yedla, M. (2010). "Enhancing K-means clustering algorithm with improved initial center". In: *Int. J of Computer Science & Information Technologies* 1.2, pp. 121–125.

Yoldas, Yeliz et al. (2017). "Enhancing smart grid with microgrids: Challenges and opportunities". In: *Renewable and Sustainable Energy Reviews* 72, pp. 205 –214. ISSN: 1364-0321. DOI: http://dx.doi.org/10.1016/j.rser.2017.01.064.

Yun-Su, Kim et al. (2016). "New Requirements of the Voltage/VAR Function for Smart Inverter in Distributed Generation Control". In: *Energies* 9.11.

Zadeh, L.A. (1965). "Fuzzy sets". In: *Information and Control* 8.3, pp. 338 –353. ISSN: 0019-9958. DOI: http://dx.doi.org/10.1016/S0019-9958(65)90241-X.

Zhang, L. et al. (2014). "Monitoring method of micro-grid based on DPWS network middleware and Petri nets". In: *Nongye Gongcheng Xuebao/Transactions of the Chinese Society of Agricultural Engineering* 30.20, pp. 163–171. DOI: 10.3969/j.issn.1002-6819.2014.20.020.

Zheng, Y. et al. (2014). "Electric Vehicle Battery Charging/Swap Stations in Distribution Systems: Comparison Study and Optimal Planning". In: *IEEE Transactions on Power Systems* 29.1, pp. 221–229. ISSN: 0885-8950. DOI: 10.1109/TPWRS.2013.2278852.

Zhong, H., L. Xie, and Q. Xia (2013). "Coupon Incentive-Based Demand Response: Theory and Case Study". In: *IEEE Transactions on Power Systems* 28.2, pp. 1266–1276. ISSN: 0885-8950. DOI: 10.1109/TPWRS.2012.2218665.

Zhong, Jin and K. Bhattacharya (2002). "Toward a competitive market for reactive power". In: *Power Systems, IEEE Transactions on* 17.4, pp. 1206–1215. ISSN: 0885-8950. DOI: 10.1109/TPWRS.2002.805025.

Zhou, Kaile, Chao Fu, and Shanlin Yang (2016). "Big data driven smart energy management: From big data to big insights". In: *Renewable and Sustainable Energy Reviews* 56, pp. 215 –225. ISSN: 1364-0321. DOI: http://dx.doi.org/10.1016/j.rser.2015.11.050.

Zhu, Q. et al. (2012). "A game-theoretic framework for control of distributed renewable-based energy resources in smart grids". In: *2012 American Control Conference (ACC)*, pp. 3623–3628. DOI: 10.1109/ACC.2012.6315275.