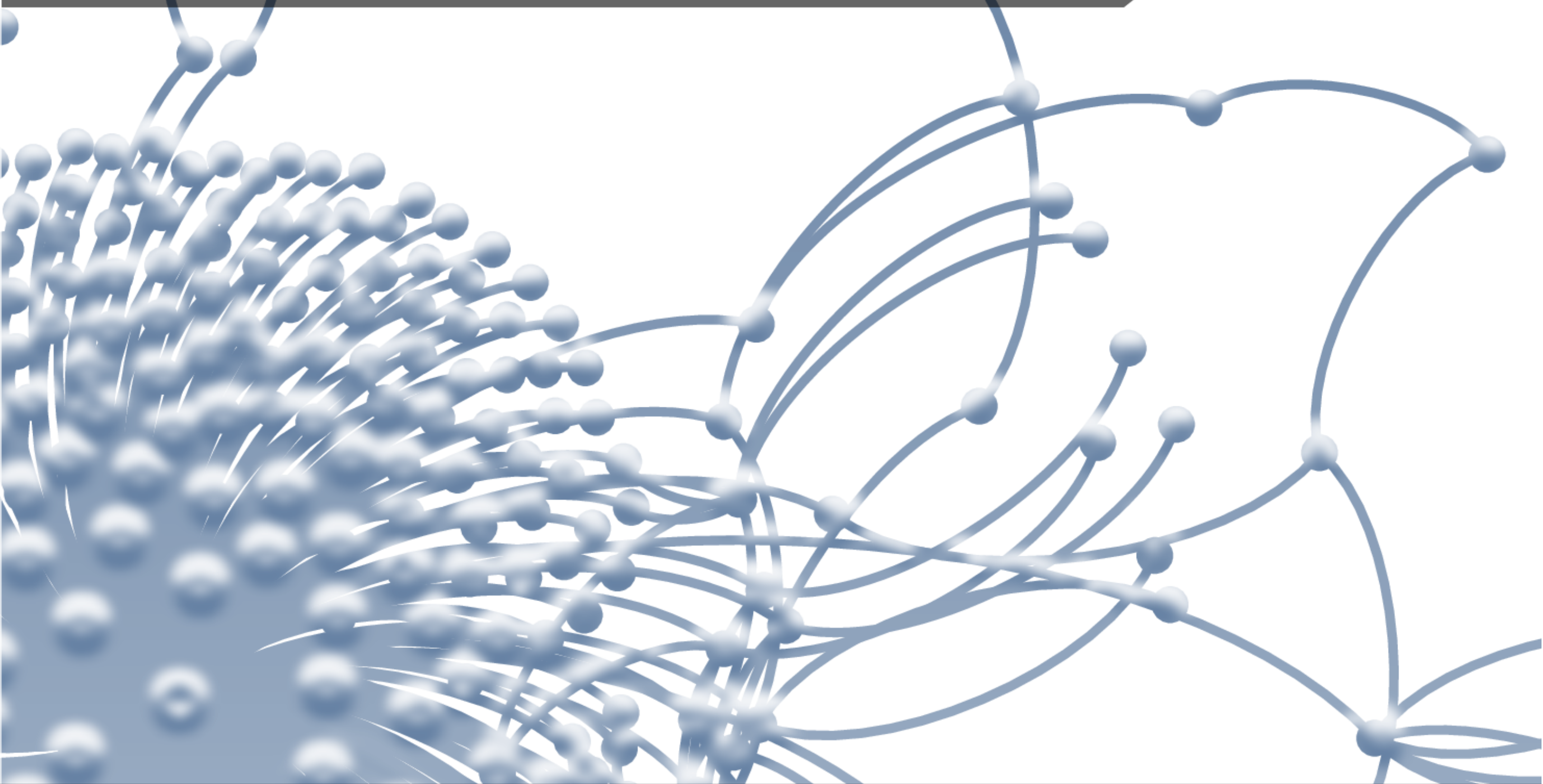
DOCTORAL DISSERTATION

A Methodological Research on Software Engineering applied to the design of Smart Grids using a Complex System Approach



José Évora Gómez

Las Palmas de Gran Canaria - October 2014

UNIVERSIDAD DE LAS PALMAS DE GRAN CANARIA

Instituto Universitario de Sistemas Inteligentes y Aplicaciones Numéricas en Ingeniería

UNIVERSIDAD DE LAS PALMAS DE GRAN CANARIA
Instituto Universitario de Sistemas Inteligentes
y Aplicaciones Numéricas en Ingeniería

DAVID J. GREINER SÁNCHEZ, SECRETARIO DEL INSTITUTO UNIVERSITARIO DE SISTEMAS INTELIGENTES Y APLICACIONES NUMÉRICAS EN INGENIERÍA (SIANI) DE LA UNIVERSIDAD DE LAS PALMAS DE GRAN CANARIA,

CERTIFICA

Que el Consejo de Doctores del Instituto Universitario de Sistemas Inteligentes y Aplicaciones Numéricas en Ingeniería (SIANI), en su sesión de fecha 30 de octubre de 2014, tomó el acuerdo de dar el consentimiento para la tramitación de la Tesis Doctoral titulada "A Methodological Research on Software Engineering applied to the Design of Smart Grids using a Complex System Approach", presentada por el doctorando **D. José Évora Gómez**, dirigida por el Dr. D. Francisco Mario Hernández Tejera y el Dr. D. José Juan Hernández Cabrera, propuesta para acceder a la mención de doctor europeo, a la vista de la idoneidad y calidad de su contenido, interés y relevancia del tema a nivel internacional, y la realización de una estancia de más de un trimestre en un centro de investigación europeo fuera de España.

Para que así conste, y a los efectos oportunos se expide el correspondiente certificado a treinta de octubre de dos mil catorce.

UNIVERSIDAD DE LAS PALMAS DE GRAN CANARIA
Instituto Universitario de Sistemas Inteligentes y Aplicaciones Numéricas en Ingeniería

David Greiner

UNIVERSIDAD DE LAS PALMAS DE GRAN CANARIA

A METHODOLOGICAL RESEARCH ON SOFTWARE ENGINEERING APPLIED TO THE DESIGN OF SMART GRIDS USING A COMPLEX SYSTEM APPROACH

PhD presented by José Évora-Gómez

Intelligent Systems and Numerical Applications in Engineering (SIANI)

Directed by Dr. Francisco Mario Hernández-Tejera
and Dr. José Juan Hernández-Cabrera

The PhD student | The director | The director

Las Palmas de Gran Canaria, October of 2014

A methodological research on software engineering applied to the design of Smart Grids using a Complex System approach

Author: José Évora-Gómez
Advisor: Francisco Mario Hernández-Tejera
Advisor: José Juan Hernández-Cabrera

Text printed in Las Palmas de Gran Canaria
First edition, October 2014

Dedico esta tesis a mis seres más queridos.
A mi madre, mi padre y mi hermano Chenko
por apoyarme siempre. A Paula por darme ánimos,
escucharme y muchísimo más. A mi abuela Mamen y
mi tía Margarita que siempre se han preocupado por mi.

Abstract

A transformation of power grids is taking place in order to address several concerns: reduction of dependency on fossil fuels; improvement of efficiency; liberalisation of markets and reduction of GreenHouse Gas emissions, amongst others. This transformation may involve the massive introduction of electrical vehicles, renewable energy sources, distributed energy generation and storage, etc.

The introduction of these technologies requires an increased flexibility of power grids to balance supply and demand. Nowadays, the balance between supply and demand is adjusted by modifying the production side. However, the idea for future power grids is to make the demand side flexible. To this end, a monitoring and controlling layer is being implemented in distribution power grids. This is a new paradigm known as Smart Grid.

Making demand more flexible requires the implementation of policies that manage the modification of this demand. These policies must be studied before their implementation since real testing is difficult on vital infrastructure such as the power grid. Policies can be studied through "in silico" experiments, also known as simulations. These simulations are normally carried out using a complex system based approach so that all consuming devices are represented.

The study of large Smart Grids as a complex system is challenging. The objective of these policies is to modify the demand in accordance to the grid state. Therefore, it is necessary that each decision maker acts by considering the actions of other decision makers to achieve the objective. This means, the behaviour each decision maker implements must consider the environment and possible interactions before acting.

The definition of a policy is refined using an iterative trial and error procedure. At each iteration, the policy must be defined, experimented and analysed. In these steps, the complexity has to be dealt with. On the one hand, there is the complexity of modelling and simulating the millions of heterogeneous elements that compose the power grid. On the other hand, the analysis of huge quantities of data coming from the simulations. In addition, there is the implementation of the policy in the decision makers which are heterogeneous and numerous.

This research work is devoted to explore the application of methodologies or formalisms to deal with these problems. Specifically, Model Driven Engineering, Business Intelligence and Swarm Intelligence are explored. Model Driven Engineering has been applied to the problems of complexity when dealing with modelling and simulation of large power grids. Business Intelligence methodologies have been explored for the analysis of huge quantities of heterogeneous data coming from large-scale simulations. Swarm Intelligence techniques have been applied to implement demand side control mechanisms.

The usefulness of these methodologies or formalisms to deal with the problems of Smart Grids cannot be fully verified. This means, it is not possible to validate that they are useful in all the Smart Grid case studies. However, case studies can be carried out to check whether they are useful or not. For this reason, in this research, several case studies based on real problems have been carried out.

The results of the application of the methodologies or formalisms in these real case studies have been positive since they reduced the time needed to carry out policy experiments. Model Driven Engineering has been useful in dealing with the engineering of simulators for large-scale models. Business Intelligence methodologies have been used to analyse data coming from simulations. They have demonstrated to be useful for providing a fast, multidimensional querying. The application of Swarm Intelligence techniques have also been shown to be useful for dealing with decentralised decision making.

Acknowledgements

The research of this thesis could not have been completed without the help of many people, to who I want to express my gratitude.

In first place, I want to thank to my two thesis advisors, Dr. Francisco Mario Hernández and Dr. José Juan Hernández. They have advised and helped me constantly. I thank them for their patience and dedication which have contributed to the quality of this research.

I also want to thank all the people that have made possible this research work in SIANI, EIFER and EDF. Their contributions have been crucial for developing this research. A special thanks to Enrique Kremers for his advising and review of this document, to Paul Andrews whose reviews have increased the quality of this document and to the researchers at the University of the Basque Country and Stuttgart with whom I have worked.

I want to thank my family their help, support, and encouragement as I was doing me to do this research.

Finally, I want to thank to Paula her unconditional support and patience throughout all these years.

This work has been supported by European Regional Development Fund (ERDF/FEDER) and Agencia Canaria de Investigación, Innovación y Sociedad de la Información (ACIISI) of Canary Islands Autonomous Government through the project whose reference is SolSub200801000137, and also through the ACIISI PhD grant funding to José Évora.

Cover image has been adapted from: "The rise and fall of a central contributor: Dynamics of social organization and performance in the Gentoo community". Authors: Marcelo Serrano Zanetti, Ingo Scholtes, Claudio Juan Tessone and Frank Schweitzerdoi.

Contents

List of Figures

List of Tables

Acronyms

ABM Agent Based Modelling

AI Artificial Intelligence

API Application Programming Interface

BI Business Intelligence

DLC Direct Load Control

DR Demand Response

DSM Demand Side Management

DSML Domain-Specific Modelling Languages

DSS Decision Support Systems

EDF Électricité De France

EIFER European Institute For Energy Research

EV Electrical Vehicle

GHG Greenhouse Gas

GIS Geographic Information Systems

HTML HyperText Markup Language

HVAC Heating, Ventilation and Air Conditioning systems

ACRONYMS

IEA International Energy Agency

IT Information Technology

KIT Karlsruhe Institute of Technology

MDA Model Driven Architecture

MDE Model Driven Engineering

MOF Meta-Object Facilities

MOPSO Multi-Objective Particle Swarm Optimisation

NIST National Institute of Standards and Technology on US

OECD Organisation for Economic Co-operation and Development

OLAP On-Line Analytical Processing

OMG Object Management Group

PSO Particle Swarm Optimisation

PV Photovoltaic

RES Renewable Energy Sources

SG Smart Grid

SI Swarm Intelligence

ULPGC Universidad de Las Palmas de Gran Canaria

UML Unified Modelling Language

XML Extensible Markup Language

XSD XML Schema Definition

Part I

Motivation and goals

CHAPTER

1

Contextualisation

1.1 The future of power grids

A sustainable energy future will require new ideas and new systems: essentially a transformation of the way we produce, deliver and consume energy [Org11]. There are many reasons for implementing this change: reduction of dependency on fossil fuels based energy production, market concerns such as the volatility of fuel prices and reduction of Greenhouse Gas (GHG) emissions.

Combustion-engine vehicles and power grid facilities produce respectively 28% and 33% of global GHG emissions [US]. It is foreseen that electrical vehicles, batteries and renewable energy sources will play a significant role in reducing GHG emissions [HHM11].

Therefore, the future power grid should be designed to support charging electrical vehicles, the massive introduction of renewable energy sources (RES) and the distributed energy storage and generation [Cla11]. In this way, both dependency on fossil fuel energy and GHG emissions can be reduced.

All of these factors are encouraging the transition from traditional power grids to Smart Grids [Org11]. The Smart Grid (SG) is a concept whose scope has not yet been agreed upon by the scientific and technological communities. What is agreed is that SGs must be used to look after an efficient usage of the available energy. Various definitions of SGs exist:

- According to the Organisation for Economic Co-operation and Development (OECD) [OEC], SGs include systems to balance supply and demand, automate grid monitoring and control, flatten peak demand and real-time communication with customers.
- According to the International Energy Agency (IEA) [IEA], SGs are networks that monitor and manage the transport of electricity from all generation sources to meet the varying electricity demands of end users.
- According to the Office of Electricity Delivery and Energy Reliability of the US [Off], SG generally refers to a class of technology people are using to bring utility electricity delivery systems into the 21st century, using computer-based remote control and automation.

The bottom line of all these definitions is that SGs will be more flexible since they will include a monitoring and controlling infrastructure. This flexibility is necessary since many emerging energy technologies like wind turbines and photovoltaic (PV) cells create problems for the power grid [Int10]. The main problem is due to the variability in their output and the difficulty of predicting it. It should be noted that these technologies produce energy as a result of unstable environmental conditions: solar radiation and wind speed and wind direction in these cases.

Let's imagine a power grid where half of the production is fossil fuel based and half is renewable based. In this scenario, half of the production may be fluctuating according to environmental conditions.

Therefore, the massive introduction of RES is a challenge. When these technologies are massively introduced into the power grid, an important part of the production capacity cannot be predicted. Since production cannot be predicted, an unbalance between production and consumption can arise. Power grids operate on the base of a balance between production and consumption. So, an unbalance is a serious problem for the power grids and must be avoided because it can provoke power failures [Ami01].

A classic solution to address this problem is the introduction of fossil fuel based generation units to back up the production in case renewable units do not produce sufficient electricity [Mas13]. This solution is effective but not efficient since these backup units must be operative even when renewable energy is being produced. On the one hand, infrastructures (buildings, power lines, transformers, generators) should replicate the renewable installed power. This infrastructure must be installed and maintained. On the other hand, backup generators keep burning fossil-fuels while renewable sources are producing energy. Therefore, renewable energy is not zero-emission, since it is indirectly producing GHG.

An alternative solution involves the use of batteries or other energy storage technology. This is the case of El Hierro island where an energy storage infrastructure has been built with a bidirectional hydroelectric power plant [BC05]. The main problem with this solution is the investment cost which is very high. However, the cost of energy storage technology could decrease in the coming years and this solution could become more competitive [And09].

However, SGs could provide alternative solutions that could be more efficient. That means, the same results could be achieved with less expense. For example, when the sun is hidden by clouds, PV production may be affected, making production lower than consumption. In this case, some loads can be selectively disconnected in order to rebalance consumption and production. In another case, when clouds disappear and production is recovered, loads can be reconnected. These actions are respectively known as load shedding and load shifting [Str08] and they are central topics in this thesis. Since these policies act directly on the consumption side, they are classified as Direct Load Control within Demand Side Management (DSM) [GMR+03].

From a general point of view, DSM includes all the policies that adapt the demand to grid requirements [GMR+03, NNdGW09]. Some of DSM policies are based on making the people aware of efficient energy use. Other policies are related to the modification of prices, encouraging consumption in off-peak hours and discouraging consumption in peak hours.

However, nowadays there is information technology (IT) that could automatically modify the consumption of specific loads in order to rebalance the grid [NNdGW09]. This technology is an opportunity that could benefit future grids. The main prerequisite for this solution is developing an IT network for monitoring and acting over the grid components. In this way, power grids could become smarter since they would be able to perceive the environment and act accordingly. This is a disruptive conception that introduces a new power grid paradigm, SGs, where the management of both sides, production and consumption, can be carried out.

Nowadays, many institutions related to power grids are researching and designing the transition towards SGs. This research and design process is not an easy task since there are many factors that have to be defined: infrastructure, market, device, communication, strategies and concerns, among others. All of these factors need to be well defined and studied since they will be implemented in real power grids which are critical systems that work around the clock. Considering all these concerns mentioned above, the focus of this thesis is made on the research of new DSM policies.

IT-based DSM policies need to be designed, evaluated and validated prior to their implementation in a real power grid infrastructure. This is necessary because they may involve risks for the stability and performance of power grids. The system could spin out of control, provoking malfunctioning and in the worst case, power failures. This could happen in the case of a production-consumption unbalance when DSM policies are modifying the consumption.

Since it is necessary to study the impact of DSM policies on SGs, this research contributes to the field of computer simulations that address this kind of studies. There are specific challenges that must be faced when simulating DSM policies.

1.2 Simulating future scenarios

To illustrate the need for simulations of DSM policies, some examples of DSM in SGs are presented. In all these examples, it is assumed that the grid under study has a communication network which allows the remote control of the appliances on the demand side.

Let us think of a policy where the criteria to make decisions is based on the state of the power grid. When there is an unbalance between generation and demand, this policy will react. When there is more generation than demand, the policy will try to increase consumption in appliances (e.g. switching water heaters on). In the opposite case, when there is more demand than generation, the policy will act on devices to decrease the demand (e.g. switching water heaters off).

In the case of a power grid unbalance in which demand is higher than generation, the application of the policy reduces the demand of all refrigerators of the grid. Since every household has a refrigerator, this action has a huge impact on the power demand causing the opposite unbalance: more generation than demand. In this situation, the application of the policy switches all the refrigerators on again, trying to balance the grid. Obviously, going back to the original position, i.e. when all refrigerators are on, unbalances the grid since demand is once again higher than generation. In this case, the application of the policy may lead to an unstable situation where refrigerators are continuously being switched on and off and the consequences could include a power failure or damage to the refrigerators [VKE+13]. An experiment of this situation has been done and more information can be found in chapter 16.

Another example of policy is reducing demand when there is an unbalance by controlling the intensity of lighting. Let's assume lighting intensity can be remotely modified by the policy applier. Then, whenever there is a grid unbalance, this policy reduces the intensity of all lighting by 30%. At first glance, this seems like a good idea and much energy demand can be reduced in the face of a power grid unbalance problem. Nevertheless, what happens if this occurs at night? Would all customers be in favour of having lights at 70% of their normal intensity? Lights may be considered as an essential power usage, especially at night. This policy does not take into account the quality of the service that is offered to customers.

These examples point out the need to study DSM policies prior to their implementation. Even though the policy designs are obviously wrong, there may be many other issues (e.g. technological, social, etc.) that are not as easy to infer as these. These issues are not easy to infer due to the fact that power grids are complex systems that contain many different actors whose own self-interested decisions affect the grid [PAB12]. Therefore, these issues must be identified and taken into account before implementing policies in real grids.

Simulation is a way to test these policies. Through simulation, policies can be tested on a large-scale making it possible to see effects at different levels of aggregation. Simulations for testing DSM policies may be developed by using disaggregated models since these policies, such as Direct Load Control, are devoted to acting at the lowest level of consumption (devices on the demand side). Examples of these simulations have been cited in many documentary items as in [CGLP94, PKZK08, EKM$^+$11, RVRJ11].

When many decisions are being made in a local and distributed manner, the aggregated effects of these decisions result in an emergent behaviour that cannot be easily predicted. In general, this is a common characteristic of complex systems and the best way to study them is through simulations. These simulations may be useful for analysing the policies in SGs and refining their design [PFR09].

Complexity in SGs arises when many components are acting and small variations in their behaviour may cause the system to evolve in an unpredictable manner. The challenges that are addressed in this document are related to deal with this complexity when developing simulations to test these policies. Power grids are huge engineering constructions which consist of many different elements linked to each other. In this sense, different approaches have been researched in this document to deal with these levels of complexity.

1.3 Problem definition

In the previous sections, the difficulties for simulating DSM policies have been described. Nevertheless, they are also challenges that must be faced. On the one hand, the need for disaggregating the demand in order to study DSM policies has been justified. In addition, the need for running the studies through simulation experiments has been justified as well.

The design of complex systems is usually approached through a trial and error method. In [Sim91], this is expressed: "the more difficult and novel the problem, the greater is likely to be the amount of trial and error required to find a solution". This is the case in the design of DSM policies which are both difficult and novel. This trial and error method is not completely random or blind; it is actually highly selective, since the knowledge acquired in previous experiments is used to design new ones.

Following this idea, the way in which DSM policies are designed can be based on an iterative process (Figure: 1.1). This process starts by defining the objectives

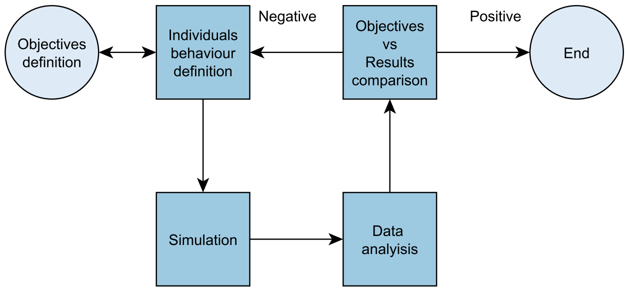


Figure 1.1: A trial and error method for designing DSM policies.

that the application of the policy must achieve. According to these, the behaviour of the individuals that make decisions is defined. This behaviour must be oriented to achieve the objective of the policy.

Thus, the power grid needs to be modelled and simulated in order to evaluate the policy. After the simulations are executed, the data they provide is analysed in order to compare it to the objectives previously defined. Depending on the results of this comparison, the iterative process may be completed or not, starting a new iteration.

The execution of the stages of this trial and error method may be effort-consuming. The main difficulties that these stages have are discussed in the following paragraphs through the description of a case study carried out in this research.

Some experiments developed within the context of the Millener project [EDFb, CSMJ13] were oriented to testing DSM policies on a French island called La Réunion [AER]. These experiments required the modelling and simulation of the island's entire grid involving the demand side.

Concretely, the experiment was oriented to simulating the effects of applying DSM policies over detached homes. To this end, every detached home has to be modelled including, at least, the appliances that can be remotely controlled. This involved a total number of elements, including their respective behavioural models, of approximately 2.5 million.

Facing this complexity is not a trivial problem. The management of such complex models is an arduous task. Modelling a scenario such as the one presented requires the integration of multiple data coming from multiple sources (Figure: 1.2).

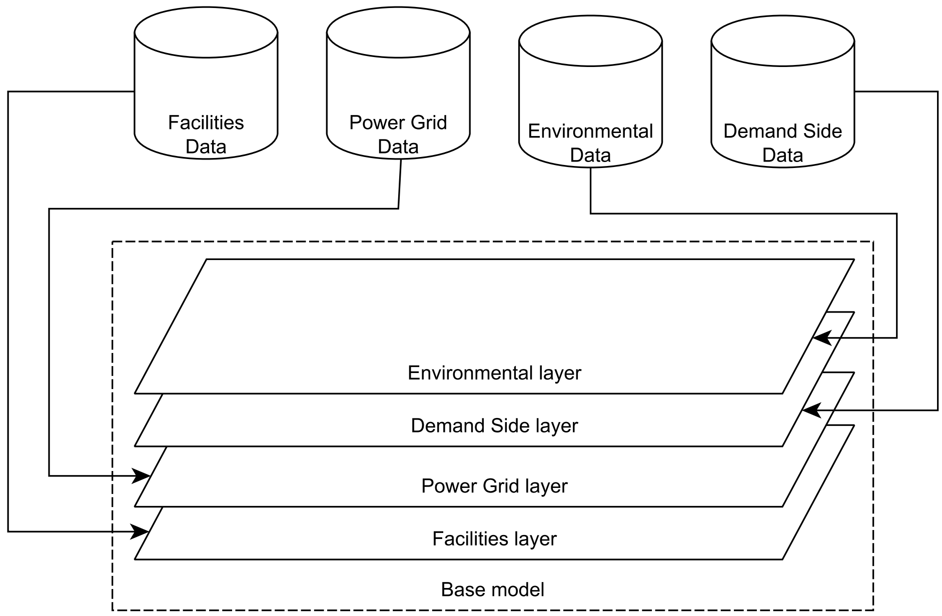


Figure 1.2: Base model enabled for the evaluation of Demand Side policies.

First of all, data for representing the power lines of the grid has to be integrated. Later on, a database of buildings throughout the island must be included in the model. At this point, appliances that are required for simulating the demand need to be modelled. Furthermore, the behaviour of the people living in homes must be designed and implemented in order to simulate the way people use the appliances at the residential level.

This modelling task described above is just one part of the whole process since it only provides the base scenario on which DSM policies can be tested. The inclusion of a policy like this may involve the implementation of a communication network which enables the transmission of messages among different actors in the grid to command the demand side (Figure: 1.3). In this network a set of devices that are responsible for making decisions, transmitting messages, applying commands, monitoring the power grid, etc. has to be placed. This is, the addition of the elements that are necessary depends on how the policy is designed and its requirements.

In this kind of simulation in which there are millions of elements involved the quantity of results provided by the simulator is huge. Every single element of

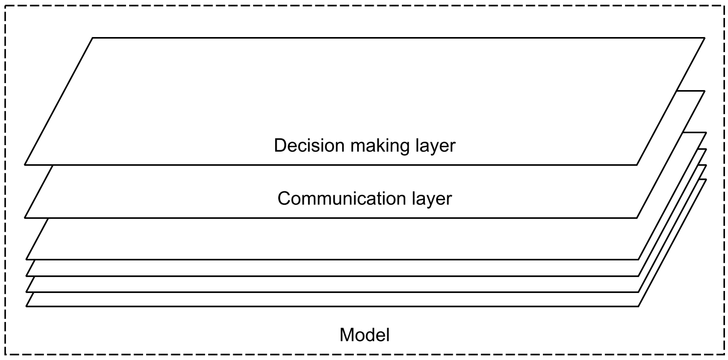


Figure 1.3: Addition of a communication and a decision making layer to the base model.

the simulation is providing its own variable parameters at each time step. Having 2.5 million elements that are outputting three different parameters each for one simulation day with a time step of minutes (so 1440 minutes / day), the total number of results reaches 10^{10}. With so many results, modellers usually focus only on certain features of the simulation since it is almost impossible to check all the output. Therefore, many important conclusions that may be drawn from the results may be missed since the whole data output is normally not being reviewed.

Another issue that has been observed is related to the conception of DSM policies. The root problem, and the reason why these policies are created, is that there are limited resources (energy) and many different actors that want to have access to those resources (customers). Furthermore, there is an environment which has limited resources and heterogeneous actors that have their own particular interests. Therefore, the challenge is how to efficiently distribute these resources according to certain objectives and constraints.

In conclusion, three major problems have been described in this section:

- SGs modelling and simulation: the complexity of dealing with a huge number of elements.
- Simulation results analysis: the complexity of dealing with a huge number of results from simulations.
- DSM policies design: the complexity of dealing with so many actors that want access to the limited energy resources.

1.4 Research questions

Complexity is a common issue in the problems described in the previous section. The questions presented in this section address the problems that have been presented.

In figure 1.1, the way in which DSM policies are designed is presented. First, after defining the policy's objectives, the way in which decision makers will act must be defined. Second, to test the policy, it is necessary to develop "in silico" experiments. Nevertheless, carrying out these experiments is an arduous task if the power grid represented is huge. Third, these simulations can output huge quantities of data which must be analysed.

Although the selective trial and error method is assumed to be the base for designing DSM policies, there is a general question that arises: how to improve the execution of this method? That is to say: how the effort to design a policy can be lessened. In other words, this is a question of efficiency and productivity. This can be addressed by making less iterations or reducing the effort to carry out each iteration.

This research is focused on the reduction of the effort to carry out the iterations as this issue can be dealt with methodological and technological approaches. The problem of reducing the number of iterations has to do with defining heuristics that suggest the paths that should be tried first [Sim91]. This has not been addressed in this research, although it could be further explored in future research. Thus, this research is oriented to exploring methods and technology for improving efficiency and productivity in the design of DSM policies.

To be more specific, this general question is elaborated through questions that address the problem of efficiency and productivity at each stage: definition of the individual's behaviour; modelling and simulation of complex systems; and analysis of simulation results.

Regarding the definition of the individual's behaviour: the approach researched is based on natural computing, regardless of other ideas that could be explored. "Natural computing" methods take inspiration from nature for the development of problem-solving techniques [RBK11]. These methods have been previously applied in other domains by our research group [1]. For example, Swarm Intelligence methods have been used in optimisation problems. For this reason, the question "can Swarm Intelligence help to deal with the complexity of designing DSM policies?" has been used to study these specific methods in the field of DSM.

In addition, this question is also related to the emergent behaviour of the power grid as the idea consists in using Swarm Intelligence to obtain the power grid objectives of consumption. In other words, Swarm Intelligence techniques could be used to obtain a desired emergent behaviour by acting on the individuals. Therefore, another question could be: "can Swarm Intelligence help obtain a desired emergent behaviour in a SGs?" There may be many different techniques to adress this issue. However, apart from the fact that we have used these techniques in the past, we have observed that other researchers have used them in the power grids field (Section: 4.4).

Concerning the modelling and simulation of complex systems, several questions can be defined. The most general one would be, how can the complexity be addressed? However, this question is too general and it must be more specific. A common strategy for solving a problem is using methods. So, a different question could be: "which method would be appropriate for defining complex models?" This question assumes as an axiom that it is better to apply a method rather than nothing. Nevertheless, this question focuses on searching for a method.

Since this research has been developed by a group [1] with research lines in Model Driven Engineering and Business Intelligence methodologies, it is logical to explore how these methodologies could be applied to the problem of complexity in SGs. Therefore, the above question can be precisely expressed: "could Model Driven Engineering help to deal with the complexity of modelling and simulating SGs?" and "could Business Intelligence methodology help to deal with the complexity of the data analysis from simulations?"

The nature of these questions defines an exploratory research, where these methodologies can be adapted to this problem. It is not possible to explore how these methodologies are applied to all the problems related to complexity in SGs.

[1] CES (Calidad, Eficiencia y Sostenibilidad - Quality, Efficiency and Sustainability) division of SIANI (Sistemas Inteligentes y Aplicaciones Numéricas en la Ingeniería - Intelligent Systems and Numerical Applications in the Engineering) institute, University of Las Palmas de Gran Canaria.

However, it is possible to explore them in certain cases. The goal that has been defined for this research is to provide evidence for or against these methodologies instead of verifying them. Furthermore, the problem cannot be expressed as the discovery of the best methodology for dealing with complexity. Such questions are not worthwhile because it is not possible to define suitable experimentation.

1.5 Structure of the document

The rest of this document is organised as follows: after this part that was intended to contextualise the research, the "state of the art" of topics related to this thesis are reviewed: SGs, DSM, modelling and simulation, model driven engineering and data analysis. This part aims to expand upon the information provided in the introduction. It also shows how other researchers are carrying out the DSM experimentation. Furthermore, it describes modelling approaches, applications of model driven engineering and technology for data analysis.

The third part presents the hypotheses. These hypotheses are related to exploring the questions previously defined. Therefore, this part is devoted to describing the hypotheses and their development in terms of how the research process has been addressed and the description of the tools that have been developed as a consequence of the research.

The fourth part is devoted to validating the research process described in the third part by applying it to case studies. In this part of the document, two synthetic case studies are presented to show features related to the modelling and simulation of SGs. The following case studies described here are based on real data. They are sorted from case studies in which only the demand is simulated to those in which there are policies that act over the demand.

The last part of the document, the conclusion, shows the results and discussions of this research. This is to say, in the "results" chapter, the outcomes of this research are described in terms of what has been done, frameworks, projects, papers, etc. The "discussion" chapter provides some interesting reflections and future considerations related to this research.

Part II

State of the Art

CHAPTER

2

Smart Grids

This review on SGs is structured in four main sections: concerns, definitions, scope and challenges. The first section, concerns, is devoted to explain the main forces that are making power grids evolve. Secondly, definitions of what SGs are supposed to be are provided. Later, the needs that SGs attemp to address are discussed by stating their scope. Finally, after discussing the scope of SGs, the main challenges they have to encounter are presented.

2.1 Concerns

There are many forces that affect the operation of power grids and these are the reason why power grids are being evaluated in order to improve their operation. According to this review [Eur07], these forces are classified into three different categories: internal market, security of supply and environment.

Figure 2.1 presents these three categories as sides of a triangle. These sides are being affected by different forces. These forces are making limitations for each of these three categories. The next sections are oriented to discuss the impact of these forces as they are factors that are leading to the introduction of new paradigms in power grids.

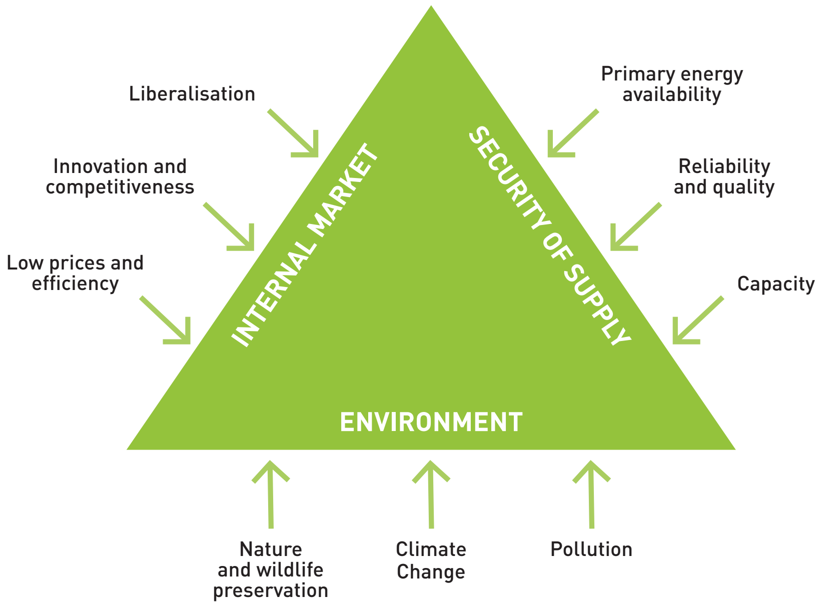


Figure 2.1: Forces that are affecting power grids. Source: [Eur07].

2.1.1 Internal market

The liberalisation of electricity markets in Europe was initiated by the UK (1989) and Norway (1991), followed by Sweden (1994), and Finland and Spain (1996) [Mey03]. This liberalisation has had an impact on what is expected from grids and the regulatory frameworks have been implemented accordingly [Gla02]. Nowadays, grids are still in process of being adapted to this new situation [Pol08, Gla12].

Innovative and competitive solutions must be developed in order to ensure the efficient use of the infrastructure assets [Eur07]. Examples of solutions that may improve the efficient use of these assets is the development of innovative energy management strategies.

Another challenge is the reduction of prices and the improvement of the efficiency as previously said. Power grids are critical systems for society and, for this reason, the reduction of energy production costs is economically beneficial for society. The improvement of the efficiency in power grids would not only reduce costs, but would also increase quality and reliability as well as reduce environmental impacts [Mom09].

2.1.2 Security

The availability of electric power depends partly on the availability of the power grid control systems. The introduction of new control devices in the power grid provides a better control and higher reliability [ME10]. However, these technologies may also have the potential to open up new vulnerabilities.

Power grids are facing increased frequency of unpredictable catastrophic events. This is due to a limited knowledge and management of complex systems and to the threat of attacks [Mom09]. For these reasons, among others, energy users are demanding increased quality and reliability. In USA widespread outages have raised public concern about grid reliability [Ami01].

Nowadays, the potential for large scale and more frequent power disruption is considered higher [Ami01]. This is due to the fact that now there are more ramifications which may lead to a network failure: transportation, telecommunications, oil, water and gas pipelines, banking and finance, among others. Furthermore, infrastructures are increasing their dependence on power grids.

New technologies are introduced in the grid so that energy may be exploited and used in different ways. An example of this is the introduction of electrical vehicles (EV). The introduction of EV may involve a problem of capacity in many current power grids as they are still not ready to support their introduction massively [HHS10].

2.1.3 Environment

Nature and wildlife is directly affected by the generation of energy. For instance, the operation of Nuclear power plants results in the production of long-lived radioactive waste [ZZGW05]. Damages that may be provoked by radioactive waste are extensive groundwater contamination, extensive soil contamination, buried soil or water containing harmful material or waste, and underground disposal facilities storing large volumes of hazardous, radioactive waste [PM98].

GHG emissions are one of the main contributors to climate change. These emissions are released by many different sources including power grid facilities which produce 33% of global GHG emissions [US]. Therefore, challenges to reduce GHG emissions are encountered by massively introducing EVs, RES and distributed energy storage and generation [Cla11].

Concerning pollution, the utilization of all fossil fuels is unsustainable by definition. Their combustion, especially coal, can produce heavy environmental impacts through smoke dust, sulphur dioxide, NO_x, and CO_2 emissions [ZZGW05].

2.2 Definition

There is currently no widely accepted definition of the term *SG*, with definitions varying across working groups and countries [XABO+14]. Some groups consider them in terms of technology alone. Other groups consider them as just an innovation on the demand side. Others perceive them in a broader view which includes the potential for smartness in the wider system.

In the introduction of this document, a review of definitions coming from different institutions were provided. In this section, apart from including these definitions, several complementary definitions are provided in order to give a more complete overview of the wide area of SG:

- According to the Organisation for Economic Co-operation and Development (OECD) [OEC], "SGs include systems to balance supply and demand, automate grid monitoring and control, flatten peak demand and real-time communication with customers".
- The International Energy Agency (IEA) [IEA] defines SGs in the following way: "SGs are networks that monitor and manage the transport of electricity from all generation sources to meet the varying electricity demands of end users".
- The Office of Electricity Delivery and Energy Reliability of United Stares of America [Off] establishes that "SG generally refers to a class of technology people are using to bring utility electricity delivery systems into the 21st century, using computer-based remote control and automation".

An interesting contribution to better understand what SGs are is made in [Cla11]. The author presents two different definitions, one coming from the SG European

Technology Platform and the other from the USA Department of Energy. They are cited below:

In [tpftenotf], the European Technology Platform for Smart Grids stated that SGs are “electricity networks that can intelligently integrate the behaviour and actions of all users connected to it – generators, consumers and those that do both – in order to efficiently deliver sustainable, economic and secure electricity supplies”.

In [Dep], the Department of Energy of US points out that SGs must include the following features: “self-healing from power disturbance events; enabling active participation by consumers in demand response; operating resiliently against physical and cyber attack; providing power quality for 21st century needs; accommodating all generation and storage options; enabling new products, services, and markets; optimizing assets and operating efficiently”.

As it can be seen, the Department of Energy is more precise when defining the aims assigned to a SG, highlighting the importance of addressing safety issues [Fre08].

2.3 Scope

SGs were initially conceived to address the improvement of the DSM, energy efficiency and safety through the construction of grids that are robust against sabotage and natural disasters [RPT07]. However, new requirements have expanded the initial scope of the SGs to something broader which includes the creation of frameworks to achieve the interoperability of all the actors within the SG [FMXY12].

An interesting way to analyse what SGs are supposed to address is through the proposal of frameworks. These frameworks are usually developed by institutions and one of the main points is the definition of the scope of SGs. The next paragraphs will summarise the scope that some relevant frameworks have defined for SGs.

a) NIST Framework proposal. According to a report from NIST (National Institute of Standards and Technology of US) [JG12] developed in 2010, SGs scope must be focused on:

- Improving power reliability and quality

- Optimizing facility utilization and averting construction of back-up plants
- Enhancing capacity and efficiency of existing electric power networks
- Improving resilience to disruption
- Enabling predictive maintenance and self-healing responses to disturbances
- Facilitating expanded deployment of RES
- Accommodating distributed power sources
- Automating maintenance and operation
- Reducing GHG emissions by enabling EVs and new power sources
- Oil usage by reducing the need for inefficient generation in demand peaks
- Presenting opportunities to improve grid security
- Enabling transition to plug-in EVs and new energy storage options
- Increasing consumer choice
- Enabling new products, services, and markets.

b) ETP framework proposal. The European Technology Platform (ETP) considers that SGs have been conceived to meet the challenges and opportunities of the 21st century [C$^+$06]. The use of revolutionary new technologies, products and services is considered as the main pillar for SGs. In particular, the SGs scope must be oriented to reduce peaks and waste, encourage manufacturers to develop more energy-efficient appliances and sense and prevent blackouts by isolating disturbances on the grid.

c) EEGI framework proposal. The European Electricity Grid Initiative (EEGI) considers that SGs scope must include increasing hosting capacity for renewable and distributed generation, integration of national networks into market-based networks, active participation of users in markets and energy efficiency, and opening business opportunities and markets [EE10] including the standardisation and interoperability.

d) IEA DSM Task XVII framework proposal. The International Energy Agency DSM task XVII considers that SGs must be oriented to face integration of DSM, Distributed Generation, RES, energy storages [Int09]. This integration is considered as important since Distributed Generation, distributed energy storages and Demand Response can be seen as distributed energy resources [HHM11] that may help to integrate intermittent RES.

Integrating all views and definitions from the literature that has been reviewed, SGs can be *in extenso* defined by listing their main observed features:

- SGs are devoted to improve energy efficiency, reliability and security.
- SGs will incorporate capacities for monitoring and controlling devices providing flexibility.
- SGs will redesign the way in which energy markets are working nowadays, including the participation of many new actors in the decision making process over the grid.

2.4 Challenges

In the literature review, different strategies that SGs may be addressing in the future have been shown. This last section of the review is intended to summarise SG strategies following the literature review made in [XABO$^+$14].

a) Smart meters and demand flexibility. Smart meters appear to be a genuinely cross cutting component of SGs, although they are not universally perceived as necessary for a SG [Eur10]. This technology is expected to help reduce demand based on better information and by shifting load consumption to off-peak times [LC11].

Nowadays, there is a massive introduction of thermal loads in the power grid such as heating systems, air conditioning systems (HVAC) and refrigerators. In a near future the massive introduction of Electrical Vehicles (EVs) is predicted. Both thermal loads and EVs may become a major driver for SGs [XABO$^+$14]. This is due to the fact that both may help satisfy demand peaks by either injecting stored energy or stopping consumption [CMG11]. However, their introduction in the grid involves a considerable increment in the overall energy consumption. In this situation, the challenge is divided into two sub-challenges: improving the capacity of power grids to support their introduction and adding mechanisms to control these kinds of devices to support demand peaks. The second one is related to a concept that is further developed in chapter 3 called DSM.

b) Security of supply. As mentioned before, this is an important topic since the introduction of RES may involve security of supply weaknesses due to their intermittent nature. As said in [Spe10], we are starting to become aware of the threat of supply disruption which may frame the development of SG around energy security.

c) Cyber security, privacy, and control. The introduction of an infrastructure to control and monitor makes it possible to capture data that may be sensitive. For this reason, data security becomes an important concern both from a data governance and a cyber-security point of view [Hor11]. The first one concerns the privileges to access this data, whereas the second one concerns how that data must be handled to prevent if from being accessed by intruders [TM11]. The introduction of Smart Meters makes this issue even more urgent and must be approached by SG strategies.

d) System defragmentation. The non-coordinated operation of companies that operate in the energy sector involves several different standard technologies and protocols. This leads to different business models that require to be merged in order to overcome such differences [Jac11]. At the same time, this makes a transition to a decentralised system more difficult, which is core for SGs [XABO$^+$14]. This fragmentation problem must be addressed by SGs.

e) Microgeneration and decentralization. Micro-generation is generally very important for low carbon electricity systems, e.g. to alleviate system congestion [BMW10]. Microgeneration offers benefits for reducing the demand, especially in a wider decentralised context. Individuals would be responsible for their energy production and consumption and, therefore, would be aware of how they have to use energy in order to optimise both variables [DW07].

However, apart from the high costs these technology have, there are two main factors which are preventing the development of these technologies [XABO$^+$14]. On the one hand, Distribution Network Operators discourage these investments on innovation. On the other hand, many energy markets are based on central generation. The inclusion of these technologies in the grid would open this market to small energy sellers which, on the aggregate, may become important players of the market.

An interesting idea [PRS08] to address this kind of regulatory issues is the use of Virtual Power Plants. Virtual Power Plants can emerge given the commercial and regulatory support. These plants will depend on the cooperation of those who are in control [Wol12].

f) Interoperability. Previously presented frameworks are not only intended to define SGs scope, but many more things. One of the most important is the definition of interoperability. For instance, NIST has defined a SG Interoperability Panel (SGIP). This panel is oriented to provide a forum that supports stakeholder participation and representation with the aim of developing and evolving interoperability standards [JG12]. SGIP has three primary functions:

- To oversee activities intended to expedite the development of interoperability and cyber-security specifications.
- To provide technical guidance to facilitate the development of standards for a secure and interoperable SG, and
- To specify testing and certification requirements necessary to assess the interoperability of SG-related equipment.

CHAPTER

3

Demand side management

In the literature, there are many sources that review DSM. From all of them, [Str08] is the one that has been followed, due to its clarity, to conduct the structure of this chapter. After reviewing DSM, different approaches to simulate DSM policies are presented as this is a core concept in this thesis.

Within the SG concept, monitoring and controlling the demand is one of the key aspects to improve power system efficiency. Then, the approach of new power grids may consist in, besides acting on the production, acting on the demand too. Power grids may act upon both the production and demand side, therefore they both become flexible grid elements. The objectives of DSM are, among others, the minimisation of peak demand and the improvement at the system operation and planning level [GMR+03].

3.1 Opportunities for Demand Side Management

There are different sectors where opportunities for the DSM can be found: generation, transmission, distribution and demand. These sectors are reviewed below.

3.1.1 Generation

Power grids are designed to support the maximum peak of demand. This peak of demand varies over time on a daily and seasonal basis. Apart from this, power grids are built with a 20% margin of generation capacity to deal with the uncertainty of generation capacity and unpredicted demand increases [Str08]. The average utilisation of the generation capacity in a year is below 55%.

There is another issue arising with the introduction of intermittent renewable sources as wind or photovoltaic. These sources of energy are not predictable and this causes the introduction of uncertainty in the power grid generation [BI04]. Nowadays, this is dealt by using backup generation units that are available to substitute intermittent renewable generation in case it is needed. This requires a high investment [Mas13].

This is an opportunity for DSM to apply load shifting (move some usages of energy) from peak to off-peak periods [GMR+03]. Load shifting can help, on the one hand, to reduce the required generation capacity since peaks are smaller and, on the other hand, improve the efficiency by increasing the average utilisation of generation capacity. Concerning the introduction of intermittent renewable generation, DSM can help to absorb energy in cases in which there is a high production of intermittent renewable sources and a low consumption.

3.1.2 Transmission and distribution

From the point of view of the transmission and distribution, power grids are designed to be robust in case a circuit is lost. When this happens, remaining circuits take over the load of the faulty one. However, these circuits cannot become overloaded. That is, under peak-load conditions, these circuits are usually loaded below 50% [Str08].

An opportunity emerges for the DSM in this field when power grids are designed to have a disaggregated generation. Nowadays, most of the generation capacity is centralised. However, the distribution of generation avoids the need of having low-used circuits just in case a circuit fails [Str08]. DSM may help in this task to make an active control of the distributed production according to the real-time needs of the grid.

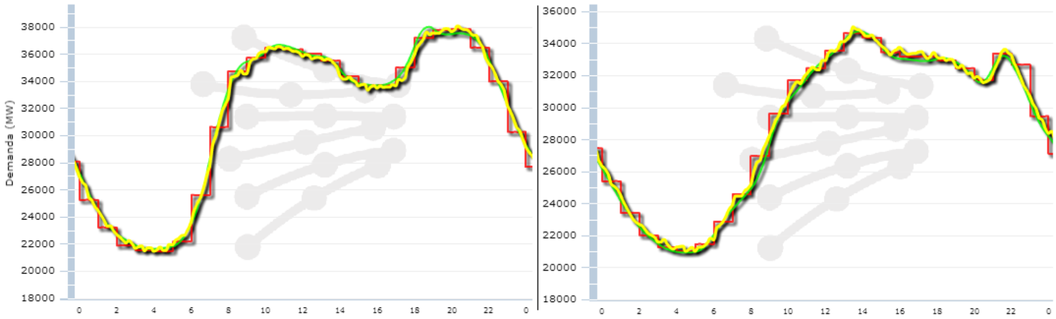


Figure 3.1: Left: the consumption of a winter weekday. Right: the consumption of a summer weekday. (Source: Red Eléctrica de España).

3.1.3 **Demand**

Demand side is largely uncontrollable in current power grids and varies with time of the day and season [Str08]. For instance, in Great Britain, the minimum consumption occurs in summer nights and is about a 30% of the winter peak. The variability of the consumption of a power grid depends on regional environmental conditions.

In Spain, as in Great Britain and many other power grids, there is a high variability in the amount of energy consumed during the day and between seasons. In figure 3.1, on the left side, the consumption of a winter weekday in the Spanish mainland is presented and, on the right side, a curve of a summer weekday (Source: Red Eléctrica de España [REE]). On the one hand, variability concerning the daytime can be observed since there are visible valleys and peaks. On the other hand, variability between seasons can be observed since not only the amount of energy changes but also the shape of the curve.

DSM can have an opportunity to balance this unbalanced condition that exists between peaks and off-peaks [PD11]. This can be made by shifting loads from peak periods to consume when the grid is in an off-peak period. A way to address a load shifting is through the use of electric energy storages (batteries) so that they can inject energy in peak-periods and consume energy in off-peak periods.

Since the massive introduction of batteries is nowadays very expensive, other energy storages that are found in the demand side can be used, such as thermal loads [PD11]. These loads can be shifted from peak-periods to off-peak periods.

The energy consumption of these loads is normally not reduced but postponed.

3.2 Review of Demand Side Management techniques

DSM is divided in several kinds of interventions at the customer level including Direct Load Control (DLC) and Demand Response (DR) [NNdGW09]. DLC is an intervention at the device level in order to shift customers' consumption regarding grid state objectives. DR consists in modifying the customers' energy usage from their normal consumption patterns in response to changes (e.g. price of electricity over time) [AES07]. The next subsections introduce different DSM techniques.

3.2.1 Direct load control

The next sections present several techniques in which a direct-load control approach is being used. In these techniques, appliances are directly addressed by remote devices in order to modify their consumption according to certain criteria.

a) Night-time heating with load switching. The idea of this technique is to set heaters to consume energy at night in order to balance the consumption of the grid. This technique requires the use of additional devices which allows to switch heating devices remotely by using radio tele-switching. This technique can fit into what is known as Direct-Load Control.

b) Commercial and industrial programmes. There are some programmes that are aimed at commercial and industrial purposes. Particularly popular are load-interruptible programmes which are oriented to provide reserve services enhancing the system reliability. Other programmes that are normally available for commercial customer consist in controlling loads by using building control systems for HVAC. These devices can be connected to power grid aggregators to command to either reduce or increase consumption according to the state of the grid.

3.2.2 Demand response

The next sections present several techniques in which a demand response approach is being used. These techniques are characterized for using external stimulus, such as price or grid frequency, to act locally and distributively over the loads [GWD$^+$11, PD11].

a) Load limiters. These devices limit the power that can be taken by individual consumers. This limit can be dynamically adjusted according to the grid state. Some flexibility is given to the customers in a way that they can choose which appliances they use and which of them can be postponed [Str08].

b) Frequency regulation. Frequency on grids is a measure that indicates how good the balance between generation and demand is, and normally it is working at 50hz. It is common knowledge that the smaller the power grid, the more problems of balance it will have. In the case of small power grids, frequency can be an interesting indicator that can be used by intelligent devices to modify the loads consumption [PD11]. A positive aspect of this approach is that it does not require to have a communication layer in a power grid as this information is available in all the points of the power grid so that local and distributed decisions can be made.

c) Time-of-use pricing. In the previous case, the frequency case, frequency was used to make local decisions in order to benefit the power grid. In this case, it is proposed to use the price of the electricity as a signal to make the regulation on the demand side [PD11]. Then, according to the prices that are different depending on the times of the day, devices are scheduled to consume energy in order to save as much money as possible.

d) Demand bidding. This approach is similar to the one seen above but in this case, price is used as a signal that dynamically changes over time. The previous case was oriented to establish different prices during the daytime, but they were the same prices everyday. In this case, it is proposed to have a fully dynamic pricing that is altered directly depending on the state of the grid [LG03]. This usually requires having a communication infrastructure to notify the prices, and intelligent devices that can make decisions on their own according to the price, which can be programmed by the customer priorly.

3.3 Demand Side Management technique experimentation

The usual approach to study DSM measures is through a modelling and simulation process where the tradeoffs can be analysed. This is an important task as those measures cannot be tested on a real grid because they can provoke blackouts. Several sources in the literature show experiments in which DSM measures were tested through simulation [CGLP94, PKZK08, EKM^{+}11, RVRJ11] and which are described in the next paragraphs.

As both DLC and DR measures lead to device switching (either through an automatic system or by influencing the customers) the modelling requires the demand disaggregation at the device level as well as the social behaviour which interacts with the devices by switching them. In [CGLP94] a disaggregation at the device level is achieved allowing the introduction of DSM measures. The social behaviour behind those devices is approached through statistical aggregated data mainly concerning income. An HVAC system model approach is made in [PKZK08]. The focus in the HVAC models is used in order to represent a high shiftable load of energy. An approach focused on both devices and social behaviour is shown in [EKM^{+}11], where almost every single appliance is modelled as well as the social behaviour starting from a small but geographically specific survey.

This conception of the electrical system shows analogies with other living complex systems, such as ant or bee colonies, in which there are several agents that are making decisions and acting locally. Actions that each agent are executing locally are aggregated, and these actions can lead to emergent phenomena.

In electrical systems, people living in the household are agents [Woo02]. They can be considered as intelligent [MPBL10] since they are self interested units and exhibit an adaptive behaviour to their environment. These agents are able to make decisions and eventually coordinate actions as a response to indirect grid signals (e.g. hourly prices) or through indirectly organising their actions with other agents through different coordination mechanisms. The main difference with respect to the study of living complex systems is that in the electrical system, we are not interested in the study of the emergent behaviour starting from the local actions, but the modification of local behaviours to get the desired emergent behaviour [SPT06]. Having said that, a complex system simulation model can be used to observe unwanted emergent phenomena and study these effects on the system. An approach to analyse complex systems from this second point of view, at the very bottom, can be seen in [PHS+08].

Focusing on DSM experiments, on the DR side, a DR implementation is shown in [GWD+11] aiming to minimise the customers' energy costs. Furthermore, they implemented an optimiser-based tool that determines the distribution of energy obtained from several sources. A decentralised DSM measure is presented in [RVRJ11], where a special effort is made to achieve a peak reduction in HVAC systems of up to 17%. A real-time DR implementation is made in [CMB10]. Based on hourly electricity prices, the objective of the model is to maximise the utility of the consumer subject to a minimum daily energy consumption level, maximum and minimum hourly load levels, and ramping limits on such load levels.

On the DLC side, an HVAC system controlling approach is reached in [CGLP94] where significant economic and technical benefits are achieved. Another interesting approach to a DLC application regarding HVAC systems is shown in [HPV+08]. They propose an intelligent energy management system to implement in commercial buildings where the schedules are usually well-known. In [KHTS09] a simple strategy is performed to control mainly heating and lighting.

DR is meant to be applied in places where the instability is not a big problem [PD11]. Then, DR can be applied in order to modify consumption patterns. However, in small grids, the instability problem arises especially when introducing RES. In the context of an island or small isolated grids, DLC policies must be considered since they allow to predict the impact of the generated measures facilitating the grid balance process [RCO09]. DR is not a feasible solution to facilitate the RES introduction since consumers' behaviour is not as flexible as needed to be adapted to the changes that may occur in the RES production. However, a very reactive DLC system could be an interesting solution worth evaluating.

In figure 3.2, an example of a DSM technique application is explained by stages. In this case, the technique is oriented to predict the different gaps that may exist between the production and demand. Based on these gaps, a scheduling is made

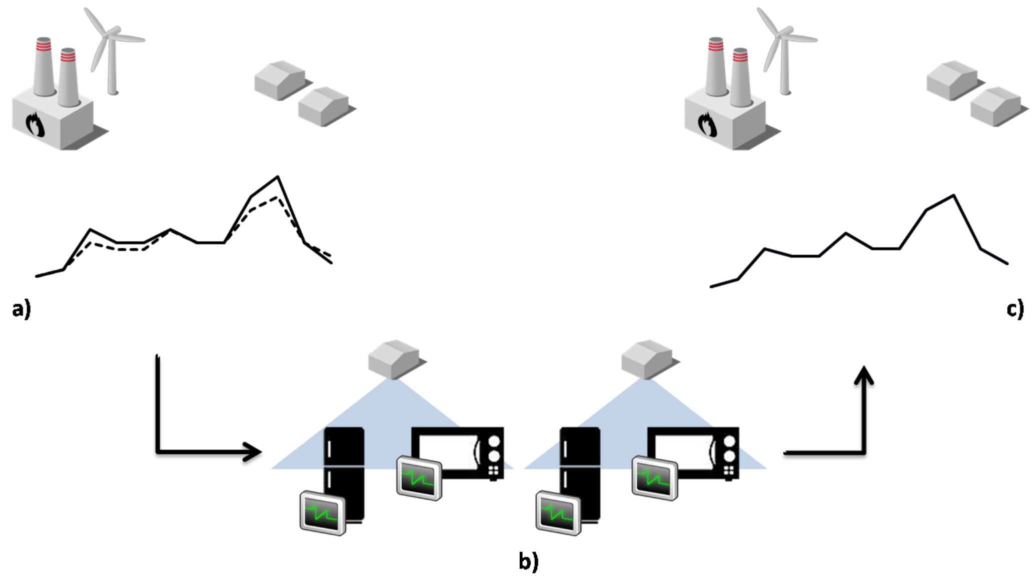


Figure 3.2: Cycle of a DSM technique example application.

in which the consumption of the devices of the demand are planned to be shifted accordingly to make the gaps closer.

In a first stage, the gaps are calculated considering RES production, conventional production and demand predictions (Figure: 3.2.a). Conventional production may be in this case scheduled according to, apart from the demand and RES production, economic constraints so that at certain moments it may be worth to reduce the conventional production and, therefore, act on the demand to balance the grid. Based on these gaps, this technique will schedule the devices of the demand side to adjust their consumption according to the production (Figure: 3.2.b). As a consequence, the power grid will be balanced (Figure: 3.2.c).

CHAPTER

4

Modelling and simulation for Smart Grids

The difficulty of testing possible implementations of SG policies in a real grid leads to the application of an approach based on modelling and simulation. Power grids are complex systems and their well-functioning is something essential for society. For this reason, the introduction of new technologies in the grid must be thoroughly tested as their inclusion may provoke problems in the power grid affecting many customers. In this sense, an important tool for testing power grid technology implementation has been simulators. Through simulation, the feasibility of the introduction of new technologies can be tested and polished.

Traditionally, simulators have been an essential tool for analysing and designing power grid systems. Many simulation tools have been developed for this purpose: UWPFLOW [CnA00], TEFTS [Uni00], MatPower [ZG97], VST [Nwa02], PSAT [Mil05], InterPSS [Zho], AMES [Tesa], DCOPFJ [Tesb], and OpenDSS [ADHM]. However, these tools are limited to simulating specific issues in SGs: some of them are limited to work in a specific level of voltage, others do not include a communication layer, etc. The modelling approach of these tools manage the production and demand in an aggregated manner.

In the context of the SG, the power grid can be considered a system in which different elements participate causing an effect on the grid. These participants are customers that consume energy, all the agents that are behind the generation unit scheduling, all the generation units (including non-renewable and renewable sources) and meteorological conditions, among others.

Based on these considerations, agent-based simulation models have experienced an increasing popularity amongst electricity market modellers [SGRM07]. The reason of the for this increase is the capability of ABM to represent different and various scenarios allowing for the analysis of economical systems, as compared to more traditional approaches like equilibrium models or optimisation models. Agent-based simulations have the potential to become a valuable approach for the analysis of electrical systems. However, this approach is very demanding in its data requirements and empirical validation is a crucial issue to be dealt with in order to reach a high level of realism.

The complexity of SGs leads to anticipating possible scenarios difficult [LVC09]. The uncertainty when applying new measures, policies or technologies on a power grid is high. In [BRS$^+$10], simulation models are believed to promote trust in SG solutions. The more accurate the model of the SG, including the whole power grid, the better to understand the advantages and risks. In this sense, in [BRS$^+$10], the design of procedures oriented to implement a SG with capacity of self-healing are successfully explored. These procedures are meant to be tested through a multi-agent system (MAS) simulation platform, which is composed of multiple interacting intelligent agents. Agents within this context are supposed to be distributed in a SG and capable of sensing, acting, communicating and collaborating with each other [BRS$^+$10].

4.1 Power grid simulators

This section provides a description of the previous mentioned simulators for power grids. As said before, most of them are devoted to analyse production side or distribution lines, but not demand side.

UWPFLOW [CnA00]: “is a research tool that has been designed to calculate local bifurcations related to system limits or singularities in the system Jacobian.”

TEFTS [Uni00]: “program has been designed to do transient stability and energy function analyses of reduced dynamic models of ac/dc power systems, with additional capabilities for voltage stability (bifurcation) studies based on continuation methods”.

MATPOWER [ZG97]: “is a package of MATLAB M-files for solving power flow and optimal power flow problems. It is intended as a simulation tool for researchers and educators that is easy to use and modify”.

Voltage Stability Toolbox (VST) [Nwa02]: “developed at the Center for Electric Power Engineering, Drexel University combines proven computational and analytical capabilities of bifurcation theory and symbolic implementation and graphical representation capabilities of MATLAB and its Toolboxes. It can be used to analyze voltage stability problem and provide intuitive information for power system planning, operation, and control”.

Power System Analysis Toolbox (PSAT) [Mil05]: “is a Matlab toolbox for electric power system analysis and simulation. The main features of PSAT are: Power Flow; Continuation Power Flow; Optimal Power Flow; Small Signal Stability Analysis; Time Domain Simulation; Complete Graphical User Interface; User Defined Models; FACTS Models; Wind Turbine Models; Conversion of Data”.

InterPSS (Internet technology based Power System Simulator) [Zho]: “is a free and open software development project. Simulation is key to enhancing power system design, analysis, diagnosis, and operation”.

AMES Market Package [Tesa]: “is our software implementation, in Java, of the AMES Wholesale Power Market Test Bed. Our objective is the facilitation of research, teaching, and training, not commercial-grade application”.

DCOPFJ [Tesb]: “is a free open-source Java solver for bid/offer-based DC optimal power flow (DC-OPF) problems suitable for research, teaching, and training applications”.

OpenDSS [ADHM]: “is a simulator specifically designed to represent electric power distribution circuits. OpenDSS is designed to support most types of power distribution planning analysis associated with the interconnection of distributed generation (DG) to utility systems”.

TSAT [Pow]: “is a leading-edge full time-domain simulation tool designed for comprehensive assessment of dynamic behavior of complex power systems”.

GridLAB-D [Pac]: "is a new power distribution system simulation and analysis tool that provides valuable information to users who design and operate distribution systems, and to utilities that wish to take advantage of the latest energy technologies".

Some of these simulators are able to simulate the demand side at the level of disaggregation that is necessary. However, they present some limitations that are discussed in the third part of this document.

4.2 Power grids as complex systems

Power grids are composed of many elements at different levels connected to each other, at a physical level, through a network infrastructure [Kre13]. The paradigm shift in the energy sector which is characterised by new market rules, together with the introduction of renewable energies and distributed systems have increased its degree of complexity. The implementation of SG technologies may involve the introduction of a network layer which enhances the capability of communication among the different actors of the grid.

SG concepts, such as DSM, distributed generation and energy efficiency usage, encourage new studies that require new approaches. Previous studies on power grids were mainly focused on how to be more efficient from the point of view of scheduling power generation units according to a demand which was considered as an aggregated load. New studies that are focused on demand side, distribution and local effects of applying demand side strategies require the use of approaches where the power grid must be represented as a complex system [PKZK08].

A complex system is comprised of a (usually large) number of (usually strongly) interacting entities, processes or agents, the understanding of which requires the development, or the use, of new scientific tools, nonlinear models, out-of equilibrium descriptions and computer simulations [Wor10] (Figure: 4.1).

An interesting concern is that the complex system refers to the way in which reality is represented. That is, every real system is complex on its own. For example, a system composed by a car going through a path is a complex system. However, its behaviour can be modelled as a simple equation which considers

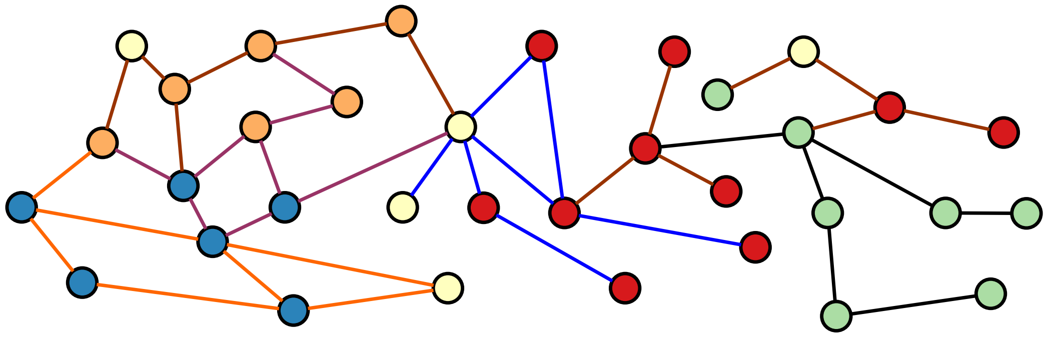


Figure 4.1: Complex system illustration. Network of different types of entities and links (colours) as example for a complex system.

some few factors. In this case, the real complex system is simplified into a equation. Nevertheless, this system can be represented as a complex system in which all main components are separately modelled and interconnected with each other. For example, the car can be modelled by its components (wheels, engine, structure, etc) and each of them will have an influence in the way the car behaves driving along the path. Models and simulators are normally developed to attend the requirements of the experiment hypotheses.

Similarly, this may be applied to power grids. Initially, they were simply modelled since the manageable section of the power grid was the production side. To this end, many models and simulations of these power grids were represented by a system of equations. Nowadays, the introduction of new technologies which enhance the capability of having a disaggregated control on the grid where many different actors can influence it boost the study of power grids to a higher level of complexity [KdH09]. This way, we can understand better the way in which SG concepts can impact on power grids concerning a wide variety of ranges: production, demand, markets, distributed generation, on-line microgrids, etc.

The next paragraphs are intended to provide some insights on the most important properties of complex systems in order to fully understand the interest of incorporating this approach to the simulation of power grids. These properties are heterogeneity, networks and emergence [BY97].

One of the main properties of complex systems is the heterogeneity of the elements that conform the system. This property is one of the factors that determine how complex a system is.

In this sense, referring back to the example of the car driving along a path, we can appreciate the participation of different and heterogeneous entities: four wheels, an engine, a car structure (which can be decomposed) and a path. Let's assume that the experiment is aimed at measuring how long it takes the car to go from one point of the path to another. In this sense, these elements which have concrete behaviours will influence on the total amount of time that it will take the car to do this trajectory.

Power grids are also composed of heterogeneous entities such as generators, consumers, distribution technologies, etc. Moreover, each category of them contains more heterogeneous elements inside. Each element within the power grid has an influence on the system behaviour (which is an emergent behaviour. Emergence in complex systems is further developed in this section).

Networks are another important property within complex systems. This property also concerns the heterogeneity property as connections within complex systems may be heterogeneous too. This heterogeneity at the network level is mainly due to two factors: type of connectivity and kind of connections. Connectivity types in networks may be any of the list below [Kre13]:

- Fully connected networks: every element of the system is connected to all other nodes of the system.
- Distance based networks: elements are connected among them according to distance criteria. For instance, every element is connected to its two closest elements.
- Random networks: in this kind of networks, elements are connected to each other following a random criteria which is based on a certain probability.
- Scale free networks [B+09]: all system networks that cannot be modelled following any of the approaches above. For example, social networks do not work following the kind of networks exposed before. In this sense, scale free networks have been identified in order to categorise all those system networks which do not fit in the other categories. In social networks, friendship among people does not follow the patterns described above. As they follow different shapes, they are considered to be scale free networks. Under

these networks, elements are connected freely to each other according to the system that is being modelled.

In the same way that there is heterogeneity among components within a complex system, connections can be heterogeneous from the point of view of the relation they express.

Referring back to the example of the social network, links among people within the social network may be of different nature: friendship, relationship, family, etc. In this sense, links are heterogeneous as they express different kinds of relation among entities and, therefore, they define the way in which components will interact with each other. Especially in the case of SGs, we can find many different kinds of links among components as, for instance, power lines, communication links or containment relations.

The cooperation among entities through these links leads us to the last property: emergence. Emergence is an important key issue within complex systems. Complex systems often behave in unexpected ways that cannot be inferred directly from the behaviour of their components; this is known as emergent behaviour [NN12].

Emergent phenomena are a consequence of complexity. The operating mode of a system that is being analysed following a traditional approach can be predicted from the system model as this operating mode has had to be programmed [SPT06]. In the case of complex systems, emergent behaviours are not predictable through tools devoted only to the data analysis, which is the reason why simulation becomes an essential tool for the analysis.

In the case of a complex system approach, every single component behaviour of the system is individually modelled according to the behaviour that each unit must have. However, the results at macro-scale are the consequence of all of these individual behaviours running and interacting through their links to each other.

In order to explain the main difference, a case based on the demand side of a power grid is used. Demand side on power grids can be modelled in many different ways that are discussed in the next paragraphs. A first approach can consist in taking into account profiles of consumption on aggregated levels of the grid to be simulated (e.g. profiles modelled through an equation). Even though this equation may provide different results with respect to the profiles of consumption, the behaviour of the system can be directly inferred. Another approach can be based on collecting data from appliances and consumers in order to fully represent the demand of a region. On the one hand, data from appliances can be used to, according to this data, model each individual appliance that is in the grid. On the other hand, data from customers can be used to model the way in which users interact with the appliance within a household.

In this sense, every single component of this complex system has been modelled from their own point of view. Nevertheless, the cooperation of all of them within the system provides different results in the overall consumption of the grid. This consumption cannot be inferred directly from the way in which appliances and customers can be modelled.

Then, the demand in power grids is the result of the aggregation of all the consumptions of the power grid. Should distributed generation be incorporated to the power grid, this generation can be considered as negative consumption. Taking into account that demand side policies are oriented to play a role at the highest level of detail (appliances), it is necessary to represent not only the model of each different kind of device, but also the social behaviour of the people living in the household [PKZK08]. Since people living in the household are self-interested, they are considered as agents whose acts cause an effect at the macro-scale. This way of conceiving the modelling of a power grid as a complex system in which there are intelligent agents is known as agent-based modelling (ABM).

4.3 Agent-based modelling and simulation

ABM is the computational study of social agents as evolving systems of autonomous interacting agents [Jan05]. ABM is a tool that allows for the study of social systems from the point of view of an adaptive complex system. Therefore, the researcher is interested in the way in which macro phenomena are emerging as a consequence of the heterogeneous individual behaviours that are taking place at the micro level [Hol92]. This approach makes it possible to systematically test different hypotheses that are related to the attributes of the agents, their behavioural rules, their types of interactions and the way in which they affect the system.

An interesting discussion which concerns the usefulness of this approach for running experiments is given in [Jan05]. There are researchers that wonder why an approach based on ABM may be needed for running experiments. Is it not possible to approach the experiments based on equations? The author's answer is that this decision may depend on the types of issue that are addressed. Many problems may be faced by using equation-based models. However, problems that concern coordination or strategy interaction where multiple agents are participating need to be addressed in a different way using ABM.

One of the main issues in ABM is the possibility to represent the complex structures of social interactions. In some systems (e.g. a power grid), the macroscale properties (e.g. overall demand of energy) are sensitive to the structure of interactions among agents and social networks (e.g. customers making decisions that affect energy consumption). However, using an equation-based approach, these agents are supposed to be implicit in these equation models making it impossible to research on the sensitivities of the structure of interactions [Jan05].

The research developed on multiagent systems in Artificial Intelligence (AI) has influenced much the architecture of agents present in ABM. Multi-agent research studies the adaptive behaviour of autonomous agents in a concrete environment [Jan05]. Intelligent agents are able to act flexibly and autonomously [Woo02]. This means that agents are goal-directed (satisfying or maximizing their utility), reactive (adapting themselves to environmental changes) and capable of interacting with other agents.

The use of ABM for research and management is growing rapidly in a number of fields [RLJ06]. This growth is due to the ability of these models to address problems by the theory of evolution [GRB+05] and strategies [GR05]. However, this kind of modelling is still an obstacle for researchers since it requires an intensive software development in order to be carried out. For this reason, simulation platforms have been developed to make experiments using an ABM approach. These libraries follow a framework and library paradigm, providing a framework [1] along with a library of software which implements the framework and providing simulation tools. These platforms have succeeded since they provide standardised software designs and tools enabling the simulation of different kinds of models.

[1] a set of standard concepts for designing and describing ABM

However, these platforms have well-known limitations. From [RLJ06], five different platforms are reviewed in the next paragraphs: NetLogo, Mason, Repast, Swarm for Objective-C, Swarm for Java. Furthermore, Anylogic and Flame are also reviewed as they are also popular solutions for agent-based simulation.

a) NetLogo [TW04]. It is suggested to develop models that are compatible with its paradigm of short-term, local interaction of agents and a grid environment which are not extremely complex. Highly recommended as a tool for prototyping models that can be, later on, implemented in lower-level platforms. However, its simplified programming environment may make experienced programmers feel uncomfortable since all code must be placed in just one file (which is contrary to good practices in object-orientation programming) and the lack of a stepwise debugger.

b) Mason [LCRPS04]. This is a good alternative for experienced programmers who work on models that are computationally intensive since they provide a good performance and the best execution time of the five platforms tested. Things that can be improved are its non-standardised terminology, its incompatible classes with the scheduler and its lack of a terminal window for debugging purposes.

c) Objective-C Swarm [MBLA96]. This version of Swarm is stable providing a fairly complete set of tools, a clear conceptual basis and clever design, where the model can be separated from the interfaces. This tool is oriented to help the model organisation by allowing the design and implementation of modules in separated swarms, each of one owns objects and schedules of their actions. This helps to manage the complexity of the models which is interesting when dealing with high complex systems. The drawbacks of this platform are the lack of friendly development tools, lack of garbage collection, weak error handling and low availability of documentation.

d) Java Swarm [MBLA96]. They provide the Swarm implementation for Java users but it contains significant drawbacks: clumsy workarounds to implement Swarm features in Java, difficulty on debugging errors that happen on Objective-C libraries and slow execution speed, among others.

e) Repast [Col03]. Apart from implementing most of Swarm's functions, they have added capabilities like reset and restart models from the graphical interface and the multi-run experiment manager. The execution speed is good when compared to other platforms. They also provide geographical and network support.

However, it presents weaknesses like the difficulty to get started on it, especially for amateur developers, and poor documentation.

f) Anylogic [BF04]. AnyLogic is a multi-paradigm simulator supporting ABM as well as Discrete Event modeling. It provides support to develop flowcharts, System Dynamics and stock-and-flow descriptions. AnyLogic can capture arbitrary complex logic, intelligent behaviour, spatial awareness and dynamically changing structures. AnyLogic is object oriented and based on the Java programming language. To a certain degree this ensures a compatibility and reusability of the resulting models [ZLB07].

g) Flame [HCS06]. FLAME (Flexible Large-scale Agent-based Modelling Environment) is an ABM framework which allows modellers from various disciplines like economics, biology and social sciences to easily write agent-based models and simulate them on parallel hardware architectures. The environment allows to create agent-based models that can be run in high performance computers and graphical processing units. The simulation code is generated by processing a model definition [KRH^{+}10].

4.4 Swarm Intelligence

The Swarm intelligence (SI) concept comes from the fields of AI, Distributed Intelligence and Robotics, where one of the main challenges is coordinating several robots. This concept of SI was firstly introduced in [BW89]. The author was interested in how robots programmed with simplistic behaviours may output intelligence as a result of the emergence coming from their collective behaviour. The next paragraph presents how the author defined these systems in which several robots were present:

"Systems of non-intelligent robots exhibiting collectively intelligent behaviour evident in the ability to produce unpredictably 'specific' ([e.g.] not in a statistical sense) ordered patterns of matter in the external environment." [BW93]

Based on these ideas, the author introduced a new concept: SI. Then, Swarm intelligent systems refer to those systems in which each individual, or agent, has been implemented with simple rules. The main characteristic of these systems is that their emergence is not predictable. Once more, the author's definition is presented for SI:

"Basically swarm intelligent systems are *unpredictable* for their definition of unpredictable (which is thorough) and they produce results that are *improbable* so are in some way surprising or unexpected." [BW93]

Another important author in the field of SI is Eric Bonabeau. In [BM01], he explains SI by exposing main characteristics of living systems as for instance an insect colony:

- Flexible: the colony can respond to internal perturbations and external challenges
- Robust: tasks are completed even if some individuals fail
- Decentralised: there is no central control(ler) in the colony
- Self-organised: paths to solutions are emergent rather than predefined

In [BM01], SI is considered a mindset rather than a technology that uses a bottom-up approach to control and optimise distributed systems. This bottom-up approach that provides an emergent behaviour is supported by the use of resilient, decentralised and self-organised techniques.

In this thesis, the evaluation of this kind of intelligence is made as a way to design and perform DSM policies. It can be considered that every single load of the demand side is an ant. Then, the system has a huge amount of ants which can be controlled at any time to balance the power grid. In the literature review there are different authors that have approached the DSM policy design and implementation through SI algorithms.

SI have been previously used to control different aspects in power grids. In [YKF$^+$00], the control of reactive power and voltage is made using Particle Swarm Optimisation. In [Cao04], the power grid is balanced through a collection of local interactions using Ant Colony Optimisation. An ecosystem of intelligent, autonomous and cooperative ants is presented in [SD06]. These ants make decisions when the system is unbalanced. An improved version of the Ant Colony Optimisation is used in [SS03]. In this version, multiple colonies are used for optimising the performance of a congested network by routing energy via several alternative paths.

Several SI algorithms have been developed to be able to represent this kind of systems. Most of them are normally inspired by natural living systems as for instance ant colonies or bee swarms. The next sections will present the two most known ones in order to illustrate how swarm intelligent systems work: Ant Colony Optimisation and Particle Swarm Optimisation.

4.4.1 Ant Colony Optimisation

The first example of a successful SI model is Ant Colony Optimization [AG12], which was introduced by M. Dorigo et al. [DMCM91, Dor92, CDMT94]. This algorithm was initially oriented to solve discrete optimisation problems. This algorithm is inspired in the social behaviour of ant colonies. From natural observation, it can be seen how ants are able to find out the shortest path between their food and their nest without visual information.

The idea consists in representing the problem as a weighted graph. These graphs are known as construction graphs. Then, artificial ants are used to search for quality paths (e.g. the shortest one) within the graph [AG12]. This graph allows ants to deposit their pheromone trails which will help to choose those nodes that are part of the solution.

This optimisation method has been used in a wide variety of experiments. Some of them are sequential ordering [GD00], scheduling [Blu05], assembly line balancing [Blu08], probabilistic Travelling Salesman Problem [BBS+09], DNA sequencing [BVB08], 2D-HP protein folding[SH05], and protein–ligand docking [KSE07].

4.4.2 Particle Swarm Optimisation

Another example of a SI model is Particle Swarm Optimisation (PSO) [AG12]. This was introduced by Russel Eberhart, an electrical engineer, and James Kennedy, a social psychologist, in 1995 [KE95, EK95]. Initially, PSO was used to solve non-linear continuous optimisation problems. Nowadays, its application has been extended to many practical, real-life application problems. PSO is inspired in the sociological behaviour of bird flocking. Birds are able to fly in large groups without colliding against each other.

PSO algorithm is based on a population-based strategy that finds optimal solutions. These solutions are found by a set of flying particles with velocities that are dynamically adjusted according to their historical performance and their neighbours in the search space [Shi04].

This optimisation method has also been used in a wide variety of experiments. Some of them are track dynamic systems [ES01], evolve weights and structure of neural networks [ZSL00], analyse human tremor [EH99], register 3D-to-3D biomedical image [WSZ$^+$04], control reactive power and voltage [YKF$^+$00], even learning to play games [ME04] and music composition [BB02].

CHAPTER

5

Model Driven Engineering

Approaching SG experiments through simulation can become an arduous task. In particular, considering DSM experiments it is hard to represent the whole demand in terms of society and appliances since demand requires to be disaggregated. To this end, not only ABM platforms for modelling demand side are needed but also mechanisms to make the modelling and simulation processes flexible. Third generation programming languages, as the ones that are used for most of the ABM platforms, do not provide enough flexibility to manage SG scenarios in which their composition evolves during the execution of the simulation and models cannot be reused in an easy way [Dou06].

In this sense, Model Driven Engineering (MDE) offers a promising approach to overcome the lack of flexibility of third generation programming languages. This methodology enhances the flexibility of the software creation as it is made based on models which represent the reality that is aimed to represent. This allows to express the simulation software in terms of high level languages (Domain Specific Languages) which is closer to the domain experts which are the responsible for determining the experiments. As software for simulation is expressed through models, changes on the models are propagated to the software so that modifying scenarios become an easy task. Additionally, this approach uses a metamodel which supports the generation of software based on models. In the context of SGs, the metamodel would define the components of power grids. In this sense, the metamodel may be used as a communication tool for different SG modellers enhancing the capabilities of reusing and sharing SG models.

5.1 Background and Definition

In the last two decades, many advances have been made in languages in order to increase the level of software abstractions that developers have available [Dou06]. At the very beginning, software had to be programmed by coding the processor instructions in bits. This evolved into what is known as assembly code which allowed the codification of software in a higher level of abstraction. This code makes transparent the bits that are used to refer operations and registers within the processor. Later on, a higher level of abstraction was made so that programs could be programmed in languages as, for instance, C. Another improvement came by introducing object oriented programming since this paradigm provides mechanisms to develop software making the focus more on the description of the reality (modelling) than on the procedures (known as third-language generation). Nowadays, class libraries and application framework platforms minimize the need to reinvent common and domain-specific middleware services (e.g. transactions, discovery, fault tolerance, etc.).

Despite these advances, there are still several problems. The main one is the complexity of the platforms (such us J2EE, .NET or CORBA) which makes it difficult for developers to master them [Dou06]. Moreover, these platforms are usually developed using third-generation languages which makes them difficult to maintain and thus turns them into a time consuming task. As a consequence of these problems, the software industry is reaching high levels of complexity, which makes developers spend years to master a technology.

MDE [Dou06] [BPSK02] [Poo01] is a methodological approach to develop software based on models. This is, the functionality of the software is defined in models and, by using a processor, software is generated according to these models. The main advantage of this approach is that the software development at the model level is conceptually described using a language that is close to the domain problem.

This methodological approach uses several abstraction levels between the software design and the implementation. These levels enable a design of the software in a high level language where the concepts are closer to the domain of the problem in which the software will be used. This way, not only developers, but also domain experts may have an important role in the development of software.

Therefore, a promising approach based on MDE technologies is proposed to address the problem of the complexity of platforms. These technologies must combine the following [Dou06]:

- Domain-specific modelling language (DSML). These languages must allow us to formalise the application structure, behaviour and requirements within particular domains (e.g. financial services, SG simulations, warehouse management, etc.). DSMLs are described using metamodels which define not only the concepts using in a concrete domain but also the relations among them. The idea behind this is to describe the elements declaratively rather than imperatively.
- Transformation engines and generators which analyse models written using a DSML in order to generate artifacts such as source code, simulation inputs, deployment descriptions or alternative model representations.

Existent MDE technologies apply lessons learned from developments made on high-level language platform and language abstractions. For instance, instead of common-purpose notation which was too far from domain concepts, the DSMLs can be built up based on metamodels which allow the expression of the syntax and semantic of the domain. By using graphic elements that are related to a domain, not only learning curves are less costly but also it helps a large set of domain experts to ensure that software systems reach the users' needs [Sei03].

Furthermore, MDE tools impose specific restrictions to the domain and carry out model checks in order to detect and prevent many errors in the first stages of the software life. Generators of MDE tools do not need to be so complicated since they can synthesise elements in order to make them coincide with the APIs (Application Programming Interfaces) and the frameworks of high-level middleware platforms. This results in an easier way to develop, debug and evolve MDE tools and applications based on these tools [Her09].

5.2 Model Driven Architecture

Model Driven Architecture (MDA) is a concrete specification which approaches a software development based on MDE. MDA has been defined by the Object Management Group (OMG). This approach is one of the most known, which is the reason why it must be developed within this document. This approach is mainly focused on architectures, artefacts and models [BPSK02]. Under this specification, models are understood as abstractions that are used to summarise the domain providing different perspectives. Models are mainly classified into two different types: platform independent model and platform specific model.

These models are proposed to be defined by using Unified Modelling Language (UML). UML is a modelling language for general purposes which is normally used in the field of software engineering. UML was developed by Grady Booch, Ivar Jacobson and James Rumbaugh. In 1997, it was adopted and managed by the OMG. These models must be described according to a language that must be priorly defined. To this end, a DSML is designed which is supported by Meta-Object Facilities (MOF).

5.3 Model Driven Engineering abstraction levels

The use of a metamodel is necessary in order to formally specify the modelling language of a concrete domain. While models describe concrete elements of an information system, the metamodel formalises the language to write these models.

To this end, there are different proposals and standards which are oriented to formalise the metamodel, such as Open Information Model [Coa99], [C^+94] or MOF [MOF02] which have four levels of abstraction in common (Figure: 5.1):

- M0. Object of the domain that is being represented.
- M1. Model of the domain. Description of all the objects that can be placed in M0.
- M2. Metamodel or language to define models of the domain. It provides a set of mechanisms to describe the model of the domain in M1.
- M3. Meta-metamodel or language to define metamodels. It provides a set of mechanisms to define concepts that can be expressed in M2.

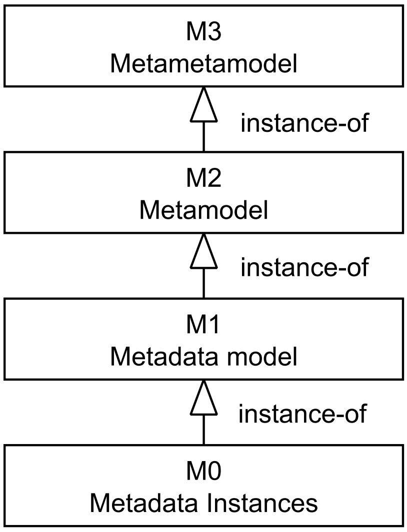


Figure 5.1: Proposal of MDE abstraction levels based on four layers.

In order to better understand these abstract concepts, an example related with films is explained (Figure: 5.2). In this example, the film *Casablanca* is in the lowest-level of abstraction (M0). In this level, there are objects of the reality, so this film, Casablanca, refers to a concrete instance of the film (e.g. a concrete DVD). This DVD is an instance of a film which is defined in M1. M1 may be devoted to represent the domain of the films. To do so, the film concept is defined there. However, this film is an instance of a concept which is coming from the level M2. In this level, a metamodel is defined in order to allow defining concepts of the film domain. Furthermore, in M2, a class called *Attribute* is defined. This class is oriented to provide the possibility of giving attributes in the level M1. This is the case of the attribute year in the concept film. Finally, attribute and concept are instances of classes, element that is provided from M3 allowing to describe and define the metamodel in M2.

Having these four levels is powerful from the point of view of the development. M0 and M1 levels do not need justifications as they are the common levels that are supported by all object-oriented languages in which the class definition would be similar to the concepts present in M1 and the object instances are in M0.

In this example, M2 contains classes that are necessary to define elements in M1. Thanks to these classes, the film concept can be defined. This level is neces-

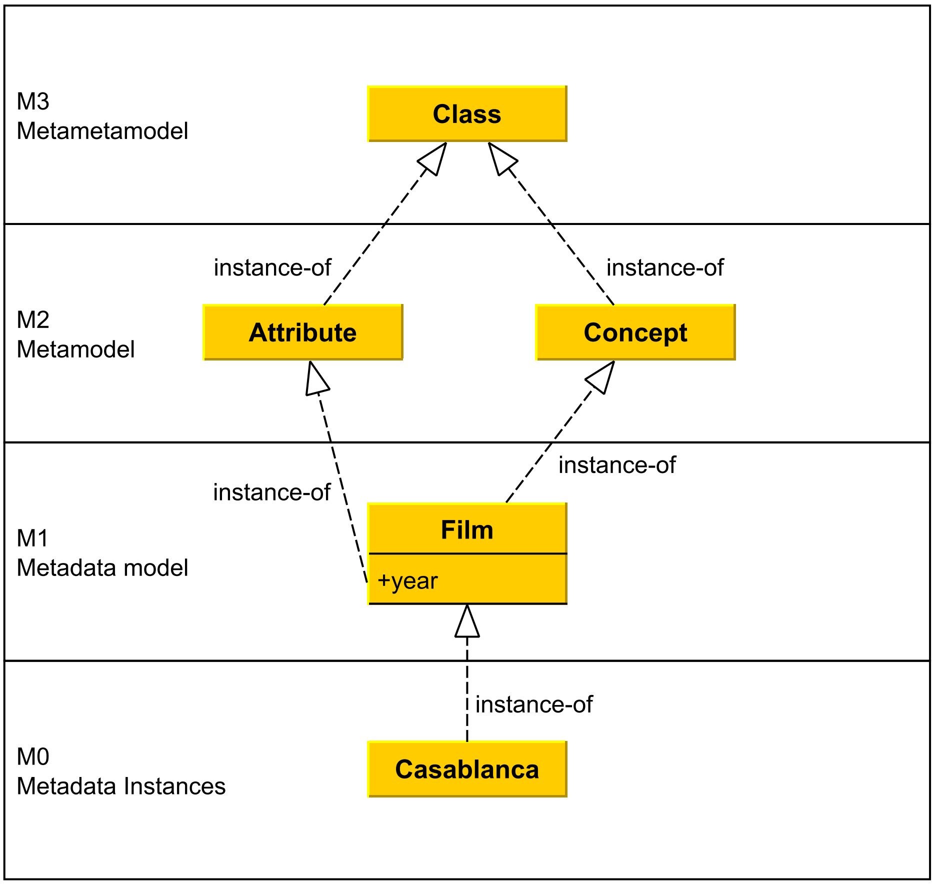


Figure 5.2: Example of the use of MDE abstraction levels (using UML notation).

sary not only for this reason, but also for having a formal description that will be used for generators and translators so that they can interpret M1 models. Therefore, based on this M2, generators and translators will be able to process all M1 models that are M2 compliance, generating as many software solutions as different M1 models are developed.

At the same time, M2 requires a set of mechanisms to define all its elements. To this end, the metametamodel, M3, provides one mechanism which allows defining classes in the M2 level. This M3 level is normally called metametamodel but it is usually a language that allows defining classes in M2.

Another example can be information systems. Under this context, M3 may provide a set of mechanisms to define classes in M2. These classes defined in M2 may describe all needed elements for building information systems. Then, M1 may describe a specific definition of an information system (e.g. information system of a specific city council) being supported by M2. M0 may represent the information system in exploitation. Therefore, many M1 models representing information systems of several organisations can be developed based on classes that are in M2. Generators and translators can build a software solution for each of these M1 models.

5.4 Model Driven Engineering application cases

There are many successful MDE application cases. For instance, Motorola has been working using MDE [BLW05]. In one way or another, they have been using MDE for nearly two decades. They have found that through the coordinated and controlled introduction of MDE techniques, significant quality and productivity gains can be consistently achieved, and the issues encountered can be handled in a systematic way.

Another example is the application of MDE in eBusinesses [Her09]. This research is oriented to help small organisations that have limited resources to define their organisational and technological projection. A conclusion of this research is that this approach is not only useful for this kind of organisations but also for modelling service structures of public administrations.

MDE has also been tried in the field of mobile applications development. In this field there are new challenges that must be faced, particularly, the prediction of performance of a given design [TWDS09]. The authors have seen in MDE a promising approach to address these challenges. Using MDE it is possible for developers to quickly understand the consequences of architectural decisions.

A research made in [RMMG08] evaluates the use of MDE in designing adaptive multi-agent systems. This is a topic that is really close to what this thesis is addressing. They conclude that it worked as expected for an ad-hoc case but that there is still some work to be done for generalising.

In [SCF+06], the authors tackle the main problems of auto-generated user interfaces. To this end, MDE techniques are evaluated. Their conclusion states that reusing MDE technologies may be promising.

CHAPTER

6

Data analysis and decision making support

Decision making processes take place in every company or organisation. In these institutions, many decisions that must be made are oriented to solve the problems of the domain in which they are working. Since these decisions have repercussions in each institution, they must be made with maximum accuracy so that benefits can be maximized.

One of the main keys for succeeding in decision making processes consists in having as much information as possible which concern the domain of the decisions to be made. However, getting this information may frequently be impossible or too costly. For this reason, the use of strategies to gather information or to reduce the uncertainty is important in order to improve decision making processes.

There are two key concept that must be defined: decision and decision making. A decision is defined as the choice of one among a number of alternatives. Decision Making refers to the whole process of making the choice [Boh03]

Over time, different concepts have emerged to support decision making processes [HW05]. There are many other classifications of all these concepts that are

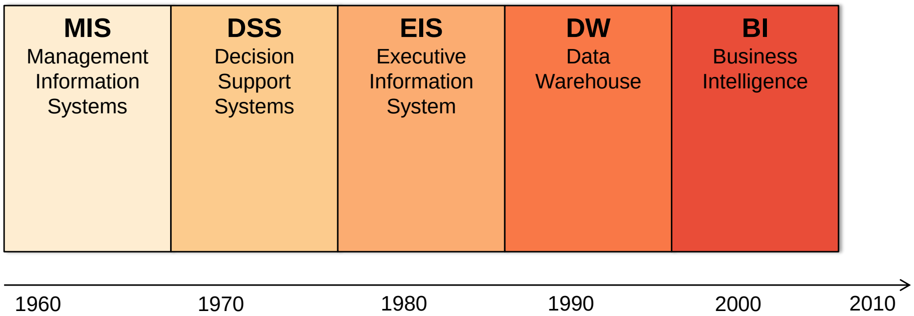


Figure 6.1: Evolution of concepts that have addressed the decision making support [HW05].

related to support decision making processes. However, checking the literature review, it can be observed that most relevant literature for each concept is mainly concentrated in the years that figure 6.1 presents.

In this chapter, two of the concepts are reviewed in a high level of detail: Decision Support Systems (DSS) and Business Intelligence (BI). The first has been included because it is one of the research topics that has had more relevance in the past for decision making processes. The second one has been included too because it is the one that is nowadays in vogue. Nevertheless, the other three are briefly described in the next paragraphs:

a) Management Information System [Swa74]. The main idea is to have a data bank containing all the relevant information concerning the company. To this end, each executive of the company will be equipped to remotely connect through terminal to a large scale computer that contains this data bank.

b) Executive Information System[Spr80]. EIS provides a set of capabilities including report preparation, inquiry capability, a modelling language, graphic display commands, and a set of financial and statistical analysis subroutines.

c) Data Warehouse[Gup97]. A data warehouse is a repository of integrated information available for querying and analysis [IK93, Wid95]. The information is stored in sets of views derived from the data retrieved from the sources.

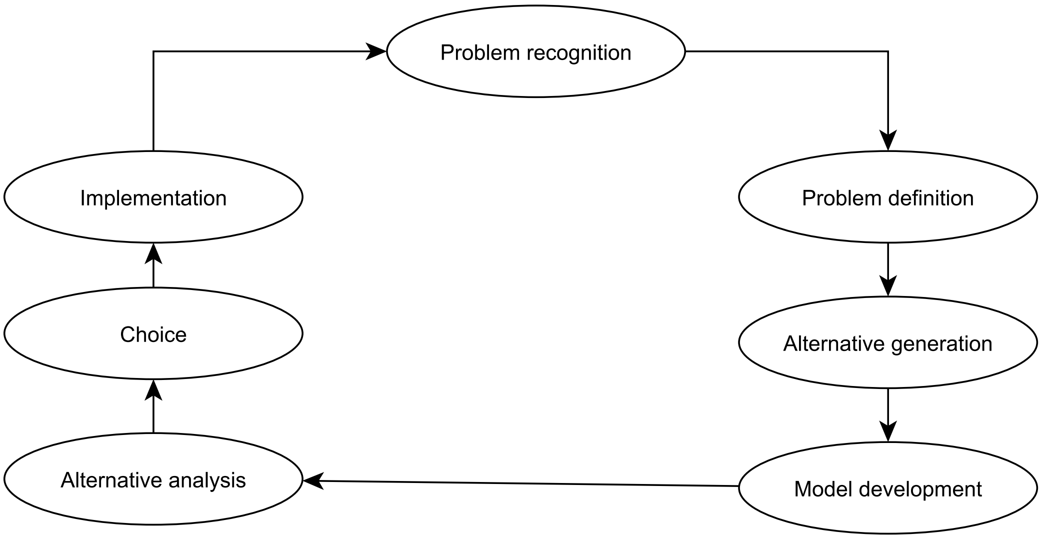


Figure 6.2: Decision making process under a DSS environment.

6.1 Decision Support Systems

DSS are computer technology solutions that are used to support complex decision making and problem solving [SWC+02]. Classic DSS tool design is comprised of components for (i) sophisticated database management capabilities with access to internal and external data, information, and knowledge, (ii) powerful modelling functions accessed by a model management system, and (iii) powerful, yet simple, user interface designs that enable interactive queries, reporting, and graphing functions. Much research and practical design effort has been conducted in each of these domains.

The process that is usually followed for making decisions under a DSS environment is shown in figure 6.2 [SWC+02]. The main focus is made on the model development and problem analysis. Once the problem is recognised, this is described in terms that make the creation of models easier. Models are created representing alternative solutions to the problem raised. Then, one of these solutions is selected and implemented. This process is iterative and has looping backs to previous stages according to the better recognition of the problem, solutions that fail, etc.

This research field has evolved for the last 50 years providing different approaches to DSS. The different approaches to DSS developed during this period can be classified in one of the following groups: model-driven, data-driven, communication-driven, document-driven and knowledge-driven [Pow07].

a) Model-driven DSS. The first approaches to DSS were model-driven. Examples of them can be seen in [MS84] and [FJ69]. Model-driven DSS is focused on the accessibility and manipulation of financial, optimisation and simulation models. This is, simple models that provide the most elementary levels of functionality. Model-driven DSS use limited data and parameters provided by decision makers in order to allow the analysis of a situation by the decision makers, not being necessary large data bases [Pow02]. IFPS (Interactive Financial Planning System) is the first commercial tool that allows to develop model-driven DSS based on financial and quantitative models. This tool, IFPS, was developed by Gerarld R. Wagner and his students at the University of Texas in the 1970s. Another DSS tool based on a model-driven approach was Expert Choice [exp]. VisiCalc [vis], a tool developed by Dan Brickling and Bob Frankston, provided the opportunity for the analysis and decision support at a reasonably low cost. This tool was the first *killer* application for personal computers and made possible the development of many model-oriented, personal DSS for managers to use.

b) Data-driven DSS. Data-driven DSS is focused on the access and manipulation of temporal data. This data can be either internal or external from the point of view of the company and sometimes it is required to deal with real-time data. In [CCS93], the authors consider that Data-Driven DSS with On-Line Analytical Processing (OLAP) provide the highest level of functionality and decision support to analyse large collections of data. In [Nyl99], the development of BI related to Procter & Gamble's effort was considered in 1985. They built a DSS that linked information of sales and retail scanner data. BI became popular as a term that was coined and promoted by Howard Dresner of the Gartner Group in 1989. BI was described as a set of concepts and methods that are oriented to improve the decision making of businesses by using fact-based support systems.

c) Communication-driven DSS. The use of the network and communication technologies in order to make easier the decision-relevant collaboration and communication is known as Communication-driven DSS. The dominant element of the architecture in these systems is the communication technologies such as groupware, video conferencing and based bulletin boards [Pow02]. Apart from these primary technologies, the massive expansion of the internet has made possible the increment of technologies for synchronous communication-driven DSS such as voice and video delivered through the internet.

d) Document-driven DSS. A document-driven DSS provides document retrieval and analysis using a computer storage in conjunction with processing technologies. These computer storages may contain scanned documents, hypertext documents, images, sounds and videos. This material can be accessed by document-driven DSS. An example of this kind of DSS is a search engine [Pow02].

e) Knowledge-driven DSS. Knowledge-driven DSS solutions provide recommended actions to managers. These systems are person-computer systems with specialized problem-solving expertise. Such expertise consists of knowledge about a particular domain, understanding of problems within that domain and skills for solving some of these problems [Pow07]. These systems are normally implemented based on AI techniques.

6.2 Business Intelligence and On-Line Analytical Processing

BI is a set of methodologies, processes, architectures and technologies that combine data gathering, data storage, and knowledge management with analytical tools to present complex internal and competitive information to planners and decision makers [EN08, Neg04]. The objective is, therefore, to improve the timeliness and quality of inputs to the decision process by making the access to the information easier. These systems are considered to be proactive and are composed by the following essential components [LV03]:

- real-time datawarehousing,
- data mining,
- automated anomaly and exception detection,
- proactive alerting with automatic recipient determination,
- seamless follow-through workflow,
- automatic learning and refinement,
- geographic information systems

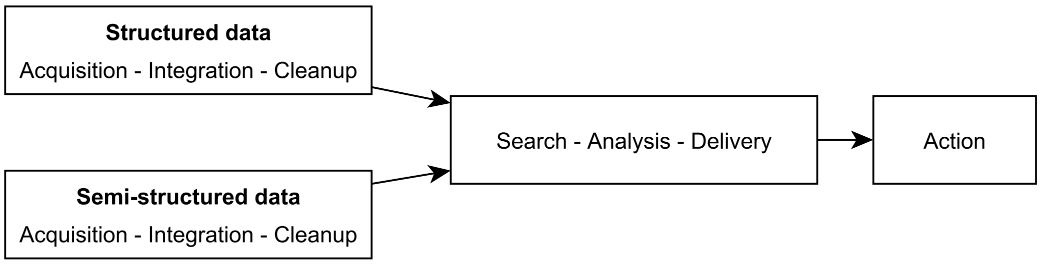


Figure 6.3: BI data framework.

- data visualisation

The way in which BI tools work is through converting data into useful information which, using human analysis, is then converted into knowledge [Neg04]. A main task consists in creating forecasts based on historical data, past and current performance, and estimates regarding future trends. Based on an analysis of the future, alternative scenarios can be designed in order to evaluate their impact (*What if* analysis). Another necessary task is accessing the data in order to answer specific questions. This is, that the data warehouse must be flexible enough for answering this kind of questions. Furthermore, BI tools must provide strategy insights as results of the queries that are launched on the tool.

BI is devoted to assist in strategical and operational decision making. An interesting summary of this kind of decision makings is provided in [Wil02]:

- Corporate performance management
- Optimizing customer relations, monitoring business activity, and traditional decision support
- Packaged standalone BI applications for specific operations or strategies
- Management reporting of BI

One of the main issues within BI is the management of semi-structured data. The tasks that are developed by the analysts in order to deal with both structured and semi-structured data is required [RC03].

Semi-structured data do not fit into relational or flat files, which is called structured data. A survey performed in [BA03] indicated that 60% of Chief Information Officer and Chief Technology Officer consider semi-structured data as critical for improving operations and creating new business models. Examples of semi-structured data that can be interesting to analyse are e-mails, images, letters, spreadsheet files, etc.

Concerning the solutions that are related to BI, some of the most known ones are Datawarehousing, OLAP, Data Mining and Information Visualisation. Inmon [Inm92], Devlin and Kimball [KR02] defined a datawarehouse as a solution which enables the integration of data from diverse operational databases to give support on decision making. When datawarehouses contain large amounts of data, the analysis of the data contained becomes a new issue.

This analysis may be carried out using an OLAP approach. In [OLA97], OLAP is defined as a category of software technology that enables analysts, managers, and executives to gain insight into data through a fast, consistent, interactive access to a wide variety of possible views of information that has been transformed from raw data to reflect the real dimensionality of the enterprise as understood by the user.

As a complement to these tools, data mining solutions must be considered when a more sophisticated analysis is necessary [Ede96]. Data mining is considered as a database exploration which intends to discover or extract information and knowledge. Data mining tools find patterns in data and infer rules from them [Pow99].

Information visualisation is the use of visual representations of data which brings human cognition to the next level [War12]. This is an important stage when analysing data since a proper visualisation may reveal information that could not be possible to extract using other visual representations.

Part III

Hypotheses and Approach

CHAPTER

7

Managing complexity

In the first part of this document, the problems that have been identified in studying SGs were presented. The first problem concerns the large amount of interacting elements that need to be considered when modelling SGs. The second problem is the difficulty of dealing with large datasets that may come from SG simulation. The third is was related to dealing with the design of DSM policies to manage requests coming from many different actors.

There are many other systems that present these challenges related to complexity. Decentralized market economies are complex systems, consisting of large numbers of agents involved in parallel local interactions [Tes03]. These interactions give rise to macroeconomic regularities such as behavioural laws. The result is a complicated dynamic system of recurrent causal chains connecting individual behaviour, interaction networks and social welfare outcome. Ecosystems are also prototypical examples of complex systems, in which patterns at higher levels emerge from local interactions and selection processes acting at lower levels [Lev98]. Understanding these interactions across scales is fundamental to resolving the problem of biodiversity and ecosystem function.

In [Joh07], complexity is defined as the study of phenomena which emerge from a collection of interacting objects. In [Hay64], the human capacity to predict the behaviour of simple systems is distinguished from the capacity to predict the behaviour of complex systems. The former, simple systems, can be easily modelled after the sciences that deal with simple phenomena. On the other hand, systems that involve complex phenomena need to be modelled through different approaches in which each element of the system and their interactions are individually modelled. The emergent phenomena of a system modelled in this manner is the result of individual behaviour and cannot be inferred from it. In [Mik01], complexity is discussed in depth by reviewing other authors' opinions expressed in [Hor95, Ros77]. In [Mik01], complexity is defined as a property of a real world system related to a collection of interacting objects.

The question that this research tries to answer is how to manage complexity in the domain of SG Engineering. Complex systems cannot be engineered using traditional formalisms and methodologies since they are not able to capture all of the properties of these systems. This means, traditional Power Grid Engineering cannot deal with complexity since the models are no longer simple. Individual models should remain simple and the system complexity must arise by combining many simple models and letting them interact. Nowadays, keeping in mind the three previously mentioned problems, it is necessary to manage large amount of heterogeneous elements which are interacting on different scales. After analysing the limitation of traditional approaches to study real systems in a complex manner [Mik01], a new approach to representing theses systems is suggested to build successful models.

Therefore, it is necessary to facilitate the construction of successful models that represent reality at the level of detail required by system analysis. Then, keeping in mind the three previously mentioned problems, it is necessary to research formalisms and methodologies that can manage large-scale scenarios, large amount of data and requests coming from all participants in a DSM policy.

This research focuses on how formalisms and methodologies from other fields can be applied to deal with these problems of complexity management in SGs.

7.1 Hypothesis #1

The demand in power grids consists of heterogeneous elements connected to the grid. Each of these elements has a specific behaviour which influences the overall system. However, the way in which the overall system will behave cannot be inferred from these individual behaviours. Therefore, as said before, this system needs to be studied as a complex system.

Furthermore, among all the elements, there are also agents which are characterised by being able to perceive the environment and act autonomously according to their particular goals. Thus, power grids may be represented using an ABM approach. This approach uses the same principles as complex systems, considering the inclusion of agents as an important part of the modelling and simulation process.

In the state of the art, several agent-based platforms were studied in order to perform SG simulations. As a result, we have found some limitations. Initially, the lack of an explicit semantic representation in models does not allow the reuse, sharing or combination. This is due to the fact that each model has its own semantic, so different developers can produce models that cannot be reused. Therefore, working in a modular way is not facilitated by these platforms due to their lack of semantic support.

Another limitation is the lack of support for developing large scale models. It is possible to build large scale models of these platforms, but the development process could resemble the complexity of the models. These platforms do not provide a methodology that supports the development process in large scale models. This is to say, we find there is a lack of architectural support for developing large-scale models.

Furthermore, these platforms provide languages for describing that are isolated from the problem domain. This means a semantic gap between the concepts of platform and modellers. Ideally, modellers would be more comfortable expressing their models in a language closer to the problem domain.

The approach of this thesis focuses on complexity. MDE has been explored as a methodology for developing an agent-based complex system for SGs. The application of this methodology in this field has been researched in order to represent large scenarios through models. This hypothesis is stated as follows:

"MDE helps to develop models for large power grids using an agent-based approach with a high level of disaggregation and can also generate simulators according to these models."

7.2 Hypothesis #2

Large-scale simulations provide a large amount of data that needs to be analysed. This analysis is necessary since problems in the modelling process can be detected and corrected, and consequences of DSM policies can be analysed. These consequences will refute or validate the ideas defined for a DSM policy and based on this, new ideas can emerge to improve its design.

New approaches to data analysis and management define the problem of dealing with data as the challenge of the 3 Vs: volume, variety and velocity [Lan01]. This "3 Vs" challenge was created by Doug Laney, a Gartner analyst. "Volume" refers to the large amount of data, "variety" to the data heterogeneity and "velocity" to the need to extract relevant information from the data as fast as possible.

The application of BI methodologies has been researched to address these problems of dealing with simulation data output. The use of these methodologies to analyse data coming from large-scale simulations may be helpful, from the point of view of the design of DSM policies, to identify problems in the design of the policy or "bugs" in its implementation or in the base scenario in which the policy is tested. This hypothesis is stated as follows:

"BI helps to analyse large amount of heterogeneous data coming from simulations."

7.3 Hypothesis #3

DSM policies aim to modify the way in which energy is consumed in order to improve the efficiency of power grids. Nevertheless, this modification of demand may affect the consumers by not meeting their needs. Therefore, these policies, especially those that directly act on demand appliances, must be designed to consider consumer needs in order to ensure a minimum quality of service. Thus, the environment in which policies are operating consists of many different self-interested agents.

This problem can be formulated as an optimisation problem in which resources must be distributed efficiently. In this sense, policies based on optimisation methods can be designed in order to deal with real problems on the demand side such as the massive introduction of EVs into the grid or the modification of the demand side according to grid states. SI offers a promising approach to dealing with this kind of optimisation problems. In this document, the application of this approach in the design of DSM policies is explored and evaluated. This hypothesis is stated as follows:

"SI helps to design DSM policies by dealing with the interests of many actors involved in the process."

7.4 Research concerns

The hypotheses presented involve the application of methodologies. Thus, the main research concern is how to validate that these methodologies are going to be helpful in solving the problems stated.

It is important to distinguish the concept of "methodology" from the concept of "method". In this section, both concepts are introduced and related to each other to clarify the semantic of the terms.

On the one hand, a "methodology" in engineering can be considered as a guideline for solving a certain kind of problem. Methodologies are generally comprised of the following four elements: description of the problem that needs to be solved, definition of which techniques are to be used and when they will be used, giving advice on product quality management as well as providing tools to facilitate the process [RH97].

On the other hand, a method is a procedure that defines a regular and systematic way of accomplishing something [How12].

Generally speaking, methodologies do not describe specific methods. A method thoroughly defines the steps that have to be performed in accordance to a methodology guideline. Therefore, many methods can be based on one methodology provided they follow the methodology directives [Vac12].

In general, methodologies cannot be validated through the scientific method. This means, the certainty of the hypotheses cannot be demonstrated with experiments since it is not possible to execute the same project with two or more different methodologies in order to compare them. Moreover, this concern becomes more determinant since the execution of the project depends on many other variables:

human knowledge, the nature of the system, budget, time, etc. Therefore, it is almost impossible to isolate the methodology variable in the success of a project.

Nevertheless, this work is an experimental research with the goal of providing evidence so as to the validity of these methodologies for solving the problems of engineering SGs through case studies. “Case study research design” is proposed as a valuable and important empirical approach for validating a methodology [LR04, ZW97]. This is a useful method for validating methodologies by using them in real world situations.

Generally, in experimental research, it is necessary to establish evidence for causality through *internal* validity [SAA+02]. However, this is not sufficient when conducting experiments using engineering methodologies. These experiments may also study the *external* validity of hypotheses, which ensures their application in a broader spectrum of cases rather than only in experimental situations.

An experiment has internal validity if it demonstrates a causal relation between two or more variables [Bre00]. A causal inference can be based on a relation when three criteria are satisfied [SCC02]: temporal precedence (cause precedes effect); covariation (cause and effect are related) and nonspuriousness (there are no more explanations for the observed covariation).

Internal validity is easy to demonstrate through the results. It is also important to demonstrate external validity. *External* validity is a generalisation of causal inferences in scientific studies [MJ12]. That is to say, results can be extrapolated from an experiment to other situations.

At this point, it is necessary to define what can be understood by “external validity” in the context of this research. In this sense, two different concepts of external validity can be considered. In the first, it can be said that the hypotheses of this research have external validity if they can be applied to more than one experiment related to SGs. The second, it could be that the hypotheses of this research have external validity if they can be applied to the engineering of other fields apart from SGs.

In this research, a set of experiments has been defined to validate the hypotheses. Some of these experiments have been jointly defined with our partners and colleagues from EIFER [eif] and EDF [edfa]. These experiments analyse real problems in real power grids. These contributions to the experimental part of this work are valuable because the validation has been made using experiments based on reality instead of only synthetic problems.

Furthermore, these experiments are important because they are real cases which helps to demonstrate both internal validity and external validity. The internal validity is demonstrated through the results of the execution of experiments. Concerning external validity, the first definition presented is also demonstrated since several experiments have been conducted using the same methodologies. The second definition is not demonstrated in this research but it could be proposed as future work.

CHAPTER

8

Smart Grid modelling

As said before, a simulator is required for each experimental study. For small experiments, the construction of a simulator is usually simple and fast. However, large experiments require the construction of a more complex simulator since different behaviours must be implemented. Furthermore, in large experiments, the initial experiment conditions are changing constantly, so it is necessary to modify the simulator in a fast and easy way [PHS+08].

As the simulator construction may be an arduous task during the experimental study, gaining productivity in this task is very important. From the point of view of Software Engineering, this simulator must be supported by a good architecture. This architecture should facilitate the simulator construction and allow the execution of continuous changing of requirements that take place during the experimental study.

Since the software development of these simulations is time-consuming, it is necessary to find a way to improve development performance. In the next sections, a framework to develop simulators based on MDE approach is presented. This framework, known as Tafat [EKM+11, EHHK12, EHH13c], is intended to speed up the creation of simulators for domains in which a complex system approach and a discrete timing can be applied. Among other advantages, the use of MDE enhances

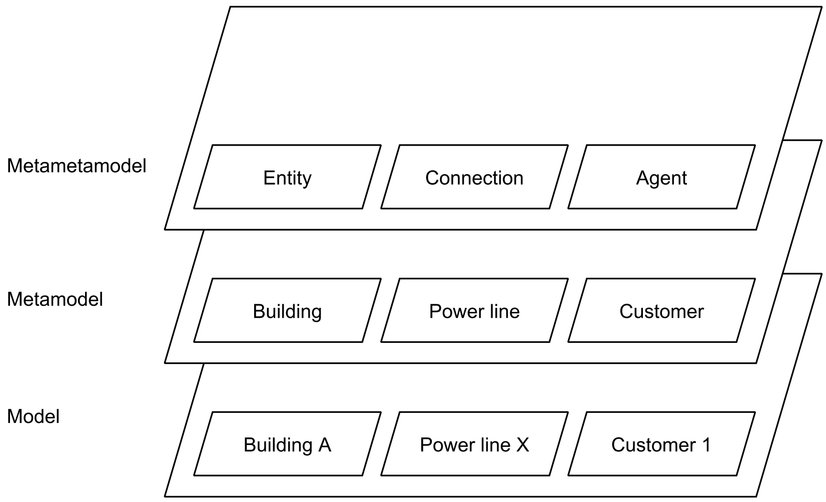


Figure 8.1: Abstraction levels for modelling power grids.

the capability of component reuse and hides implementation details providing a high level language to develop simulators.

The adaptation of MDE to the field of complex system simulations is the base that supports the Tafat framework. The design of the abstraction applied to this case consists of three levels of abstraction. The first level, considered the lowest abstraction level, consists of the concrete elements that can be found in a specific domain. The second and third levels propose a way to abstract the definition of this concrete element.

In figure 8.1 these three levels are represented for a concrete case which is related to the world of power grids. In a first level, which is related to the model description, concrete elements are located. This is, a concrete household, power line and customer. However, these concrete elements can be abstracted into a second level. The second level, known as Metamodel, the concept of a household, a power line and a customer is represented without considering specific details of a concrete instance. Based on this concept specification, the concrete elements can be defined by parametrising the elements of the second level. A third level of abstraction simplifies the complex system world into three different types of elements: entities, connections and agents. Therefore, in this case, a building is considered as a entity, a power line as a connection and a customer as an agent of the power grid system.

8.1 Model Driven Engineering for modelling complex systems

Tafat allows building models of complex systems where the overall scene is decomposed into different components. Each component of the model is statically represented by means of attributes and variables. In this way, it is modeled a structural view of the system. In order to model the behavioural view of the system, each component may include several behaviours that describe how this component changes over time and the way in which it interacts with other components. That is, the dynamic of the system.

This separation between structural and behavioural view is a major design concern in Tafat since it allows modeling a complex system in a modular approach. At the same time, this design technique is oriented to face the challenge of building large scale simulations. Basically, when a simulation is running, all the behaviours are concurrently executing and modifying the variables of their components.

8.1.1 Architecture

The core component of Tafat architecture is the Metamodel (Figure 8.2) which describes the types of instances that could exist in the complex system which shall be represented. The Metamodel is oriented to a specific domain and thus establishes a common semantic which allows modelers to share simulations and models. The Metamodel allows a representation in terms of entities, connections, agents, attributes, variables and context, among others.

The Metamodel can be translated into several formats: HTML, XSD and Java. The first one, HTML, allows the observation of the Metamodel elements, as well as their properties, in an easy and comfortable way by using a browser. The XSD translation returns a XML scheme model that helps to validate the model construction from a semantic point of view and provide valid next tokens to add when writing models. Finally, the most important one, the Java classes that implements the Metamodel elements and their features. These Java classes are used in combination with the Simulator Engine and behaviours (repository) in order to build simulators.

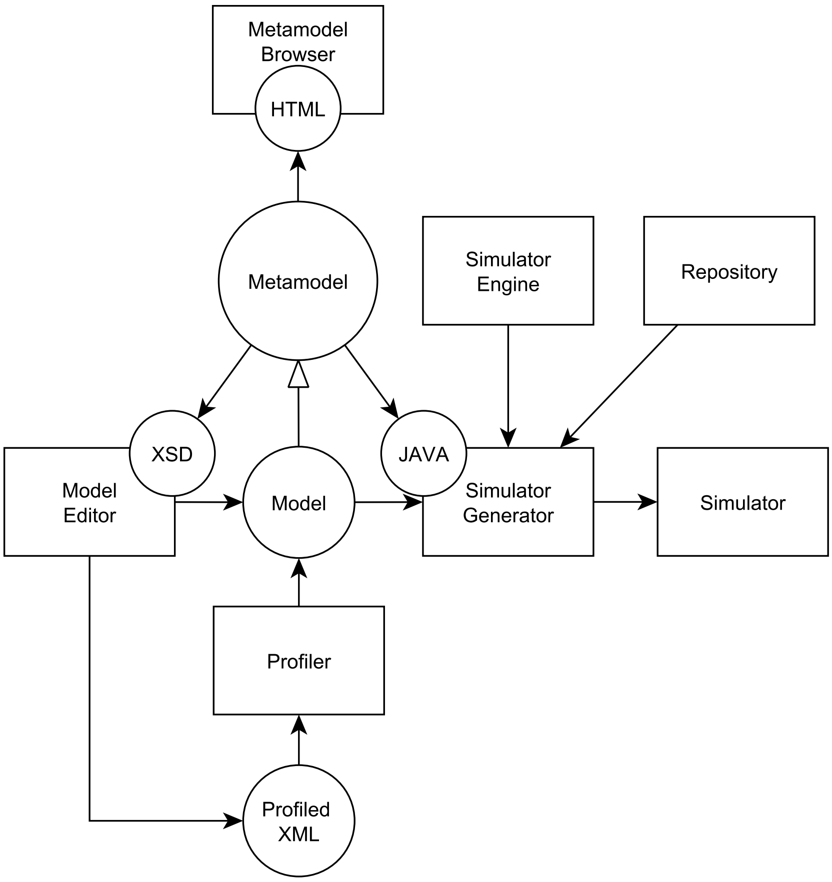


Figure 8.2: Architectural diagram of the framework.

The simulation model contains the description of the scenario and it is expressed in terms of components that have been defined as Metamodel concepts, i.e. metaclasses. Since components are always related to a metaclass, the components that can be instantiated in the model are those defined in the Metamodel.

Tafat includes different tools that support the simulator development process. Checker is a tool to validate syntactically and semantically the model; Profiler, a tool to assist the automatic construction of large scale simulation model; and Simulator Generator a tool that can automatically create a new simulator for this model.

The Simulator Generator parses the model in order to automatically generate the required Java classes and objects. The simulator is completed by including the Simulator Engine and all the behaviours that the simulation requires.

Behaviours are stored in the Behavior Repository and can be used in different simulations. A behaviour is a reusable element that describes the dynamic of a model component.

Therefore, model components are implemented by composing two different views: structural, contained in the Metamodel, and behavioural, contained in the repository. The structural view concerns the description of the components in terms of attributes, variables and context. The behavioural view describes how state evolves and their interaction with other components.

This separation of concerns enhances the possibilities to customize already developed behaviours or, even, create new ones and plug them into a model component.

8.1.2 **Metamodel**

The Metamodel constitutes the core of the framework architecture since it defines the metaclasses of components that can be simulated. A metaclass defines the structural composition of a component, which may include other components. In this way, a model is described by decomposing high level components into low level ones.

In other words, the Metamodel is a representation that describes how the model can be structured. However, there are different layers in the Metamodel that have been defined by means of an abstraction process over several complex system models. These layers constitute the basis for defining the metaclasses included in the Metamodel:

- Entities. Elements that are contained in the model scenario (e.g. understanding the power grid as a complex system, a washing machine would be an element)[Woo09].
- Agents. Interpreted as intelligent actors, they are able to interact with entities or other agents (e.g. people living in the household where the Washing Machine is)
- Connections. Elements that link entities (e.g. washing machine power connection to supply)

Metaclasses are classified into the Metamodel according to these categories (Figure: 8.3). This classification is useful since entities are used to model the scenario, agents to model the population and connections to model the topology. Besides, there are other categories within these major ones that have been defined according to the nature of the system.

Regarding the elements description, there are several aspects that can be used. The list below exposes the different points of view from where an object can be described:

- Features. Attributes of an element which do not change their values along the simulation (e.g. capacity of a washing machine)
- Variables. Attributes of an element which change their values along the simulation (e.g. power of a washing machine)
- Contains. Set of elements that can be placed within an element (e.g. a household can contain several appliances)
- Context. Elements that influence the element that is being defined (e.g. as a radiator influences the internal temperature calculation of a household, the radiator is considered to be in the context of the household)

The example below presents the static description of a household. This household description has one feature, one variable, two contains and one context. As

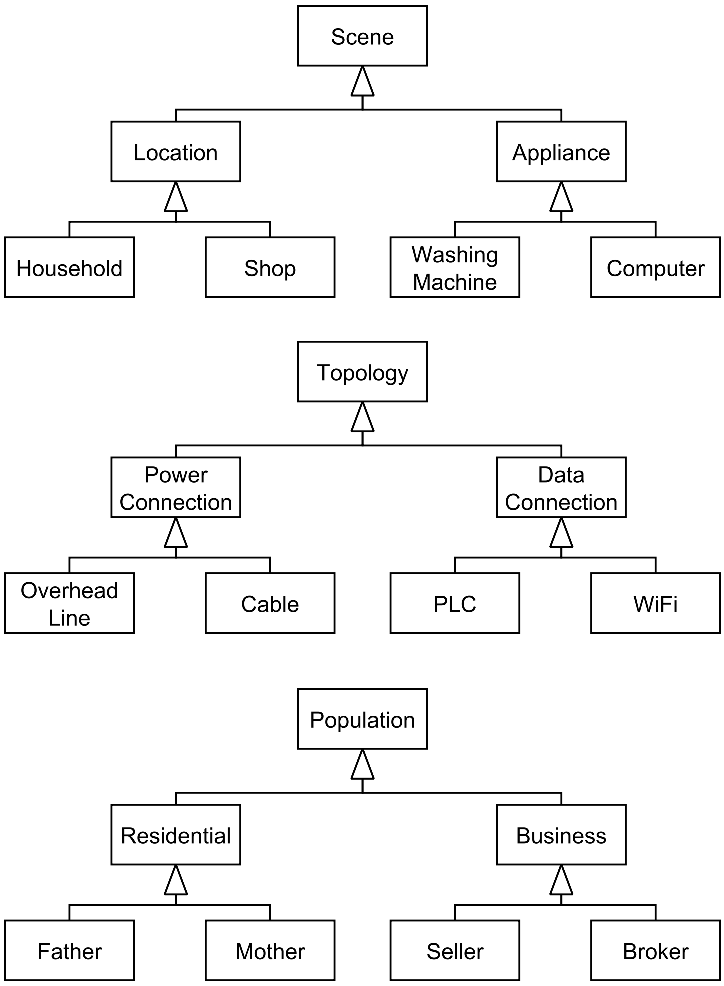


Figure 8.3: A Metamodel example for a power grid.

feature, the height of the household is expressed. The personCount variable indicates the number of people who are inside the household at a given time. This is variable since people leaving and joining the house would affect this value. The household can contain a powerMeters and powerEquipments. A list of those powerEquipments is part of its context since this list will be used by the powerMeters to calculate the overall consumption.

```
<class name="Household" parent ="Location">
 <feature name="height" type="double" unit="meter" initial-value=
     "3"/>
 <variable name="personCount" type="int"/>
 <contain name="powerMeter" type="PowerMeter"/>
 <contain name="powerEquipment" type="PowerEquipment"/>
 <context name="powerEquipmentList" type="PowerEquipment"
     replicated="true"/>
</class>
```

Listing 8.1: Metamodel entity example of a Household

8.1.3 **Model**

The simulation scenario is expressed through the simulation model. This model contains a set of instances of the elements that are in the scenario. The next example presents a simple scenario where a household containing a power equipment is represented. Furthermore, there is agent which is linked to this household.

```
<simulation>
 <scene>
  <outdoor>
   <building>
    <household id="h0" height="2.5">
     <washingMachine>
      <behavior name="WashingMachine" release="Operational"/>
     </washingMachine>
    </household>
   </building>
  </outdoor>
 </scene>
 <population>
```

```
  <familiarUnit linkedTo="h0">
   <behavior name="FamiliarUnit" release="CoupleWithChildren"/>
  </familiarUnit>
 </population>
</simulation>
```

Listing 8.2: Model example

The model construction can be addressed in several ways depending on the complexity of the scenario to represent. If the scenario to represent is composed by few elements, then, it is suitable to build the model by using a XSD scheme [1]. However, when simulating large scenarios, an automatic way to fill the simulation model would be really helpful. Based on this idea, the Profiler tool addresses this kind of problems where large scenarios are required. For example, a scenario which represents a whole city cannot be modelled by hand since too many buildings, facilities, roads, etc. have to be defined. In this case, the Profiler may parse databases for generating the whole city. Furthermore, this tool is helpful for parameters variation so that several models can be created where the only difference among them are that some parameters have different values. This allows the simulation of different configurations of a system in order to analyse the impact of the parametrisation. The Profiler also allows the parsing of models that are partially defined by completing the parts that are declared to be built.

8.1.4 Tools

Concerning the previous diagram where the whole architecture is presented, there are different tools which address different issues when simulating experiments (those tools are presented in rectangular boxes). The list below briefly presents the main issue that each tool addresses:

- Model editor. This tool is part of the architecture but it is not an own development of the framework. Its functionality is carried out by any model editor which supports XSD files, which will allow to write simulation models.

[1] The XSD is a schema that supports the model creation process by providing guidance.

- Profiler. Tool for creating models in an automated manner especially oriented to develop large scenarios based on statistical or incomplete data. The construction of large-scale models requires automation. For instance, the model of a large scale power grid in which the whole demand is represented in a disaggregated manner cannot be manually built. Furthermore, all required data to represent this scenario is usually not available at such level of disaggregation. The Profiler allows the integration of several sources of data which are used to build large-scale simulation scenarios based on statistical or incomplete data.
- Repository. Digital storage that contains the behavioural representation of the elements described in the Metamodel.
- Simulator Generator. Based on a simulation model, this tool compiles all required Java classes from the Metamodel translation and the instantiated behaviours from the repository. All this compilation is integrated with the engine generating a simulator which simulates the input model.
- Simulator. This tool is the one which parses the model, creates the instances as Java Objects and runs the simulation. This tool has the ability to export results into different formats so that posterior analysis can be performed.
- Metamodel browser. As well as the model editor, this tool is not a framework development. This functionality is provided by any browser since the Metamodel is translated into HTML format.

8.1.5 Simulation's life cycle

As previously said, in the design of DSM policies there are several tasks that must be preformed. One of the tasks concerned the execution of simulations from where the results are obtained. Based on many simulations we have developed, there is a common pattern about how to proceed to develop simulations. This pattern consists of four main tasks:

- Data preparation. When transferring the reality to a simulation model, part of such reality is neglected due to several reasons: no meaningful data, lack of data, computational issues, etc. However, the data used for each simulation can widely vary from one to another. Therefore, a main task for developing simulations consists in preparing data in order to use them in the model. Depending on the data format, coherence, completeness, complexity, etc. the effort to develop this task may vary. The final goal is to have the data with a format that allows it to be parsed by the Profiler. However, some small experiments may not have a complicated process for preparing the data.

- Model creation. This step is divided in two main parts: creation of the simulation model and creation/modification of model elements or behaviours. The first one is mandatory for developing a simulation since in this simulation model the scenario of the experiment is described. The second one depends on the experiment requirements (e.g. a new element that is not in the Metamodel may be needed or a new behaviour, etc.). Based on the data formatted in the step before, the Profiler could help to generate the scenario, especially, if it is a large one.
- Model simulation and calibration. The model simulation step finalises with the results gathering. However, this does not only consist of running the model in the simulator and waiting for the results, but also verifying and validating that this process is correct (known as calibration). This calibration process concerns the verification of the correct working of each simulation element by checking that the outputs they have are the expected ones. This can be approached in serveral ways: using testing frameworks, output checking by hand, etc. Furthermore, the experiment requirements must be confronted with the simulation features, so that, it is checked whether the simulation features match the expected requirements or not.
- Result analysis. According to the experiment goals, the results must be evaluated in order to obtain conclusions. Sometimes, this evaluation may involve the simulation of other simulation experiments in order to check how the system works under different conditions. Even if the experiment planning is thoroughly detailed, many conditions can be interesting to be modified after watching the simulation results so that new simulation experiments may be carried out.

8.2 Simulation performance

Most of the tools presented in this document execute the simulation with a synchronised approach. Synchronous simulations have the advantage of simple time management as all objects of the modelled system are running in the same time instant. It forces objects to always perform calculations, in every time step. Sometimes these calculations are unnecessary due to the fact they cannot provide new results. For example, a washing machine is usually waiting for an agent to be turned on, considering this as an event. Later on, it develops some washing cycles where the power may vary along the time. Whenever the washing machine state does not change, calculations could be avoided.

In this section, it is proposed that an asynchronous simulation approach is included in Tafat which would allow objects to develop their own time as desired. They could behave both event and time-based according to their nature. Furthermore, they could use variable steps from one calculation to another. For example, in an asynchronous simulation where the power consumption of a washing machine is analysed, calculations would be done only when the washing machine state changes. The advantage with respect to a synchronous simulation is clear since in the synchronised case, calculations are done every time step.

In the context of discrete event simulation the asynchronous concept has dual connotation. One of them consists in variable time-increment procedures as opposed to a "synchronous" or fixed time-increment procedures for simulation control. This connotation is related to the known concept Distributed Discrete Event Simulations [Kau87, Mis86]. For instance, Simula [Poo87], a simulation-oriented programming language, is based on this asynchrony concept where the time management is mainly event-based. This kind of asynchrony was already considered in Tafat through using different time steps for each mode of behaviour [EKM$^+$11]. On the other hand, the asynchrony can be understood as a non-sequential processing where simulation parts may not be executed in the proper temporal order. That is to say, later parts of the simulation may be executed before previous ones [Gho84]. The last connotation is the one to which we subscribe in this document. The objective is to apply the time-management to each model element allowing them to be in different time instants.

8.2.1 Tafat asynchronous simulation

This section examines a new approach to achieve asynchronous simulations with Tafat. This section introduces the concepts and constructions that Tafat architecture includes to model power grids. Theses constructions are focused on dependencies between objects that are massive and very relevant in a complex system simulation. In order to properly handle an asynchronous simulation, it is important to understand the dynamics of coupled objects. For the sake of clarity, a traced execution of objects interaction during an asynchronous simulation is demonstrated.

8.2.2 Tafat system modelling

Normally, a Tafat behaviour is coupled with other objects, both for querying their states or sending messages in order to change their states. In the Tafat model representation, defining behaviour which interacts with other objects is allowed.

This representation approach consists of interfaces that should be defined in the object which could be externally accessed. In Tafat, there are two types of interfaces:

1. event interfaces that handle messages and are responsible for modifying the object internal variables as requested, and

2. data interfaces that handle queries and provide the value of requested attributes

An example of these types of interfaces is shown in the figure 8.4. On the one hand, the thermal behaviour within a household has a data dependence with the temperature of the surrounding Outdoor. In this case, the Outdoor temperature data is requested by the associated object through the outdoor data interface. On the other hand, an sociological agent behaviour wants to turn on the washing machine. Then, this sociological agent must use the washing machine event interface to achieve this task. The washing machine event interface would change the washing machine mode to ˋˋON˝. The washing machine operational behaviour would calculate the proper power consumption based on this mode. Later on, when the cycle ends, the operational behaviour turns off the washing machine.

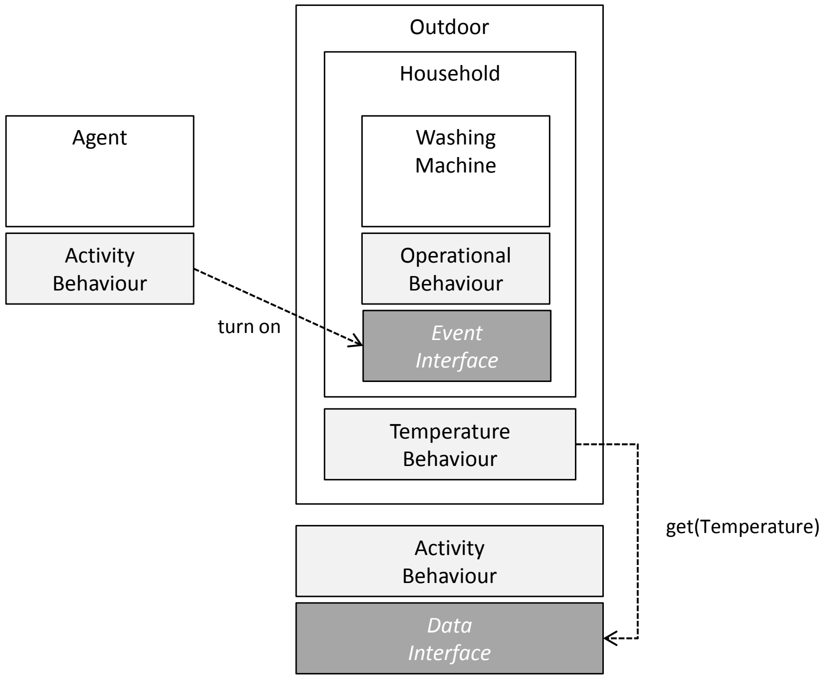


Figure 8.4: Dependencies examples between objects.

8.2.3 A power grid simulation case

In order to consider the main issues that involve asynchronous simulation a simulation case is proposed to show how objects interact when working in different times (Figure: 8.5).

The objects within this simulation case are an Outdoor, a Household, a Washing Machine and a Radiator.

- The Outdoor is the object that represents environmental conditions, in this case, the temperature. The Outdoor temperature behaviour is responsible for setting the temperature which can be loaded from an external database.
- The Household works as a container of the appliances of a household, a Washing Machine and Radiator in this case. The Household Behaviour is concerned with the thermal dynamics inside the household.
- The Electrical devices inside the Household are a Radiator and a Washing Machine. These devices are handled by an Agent.

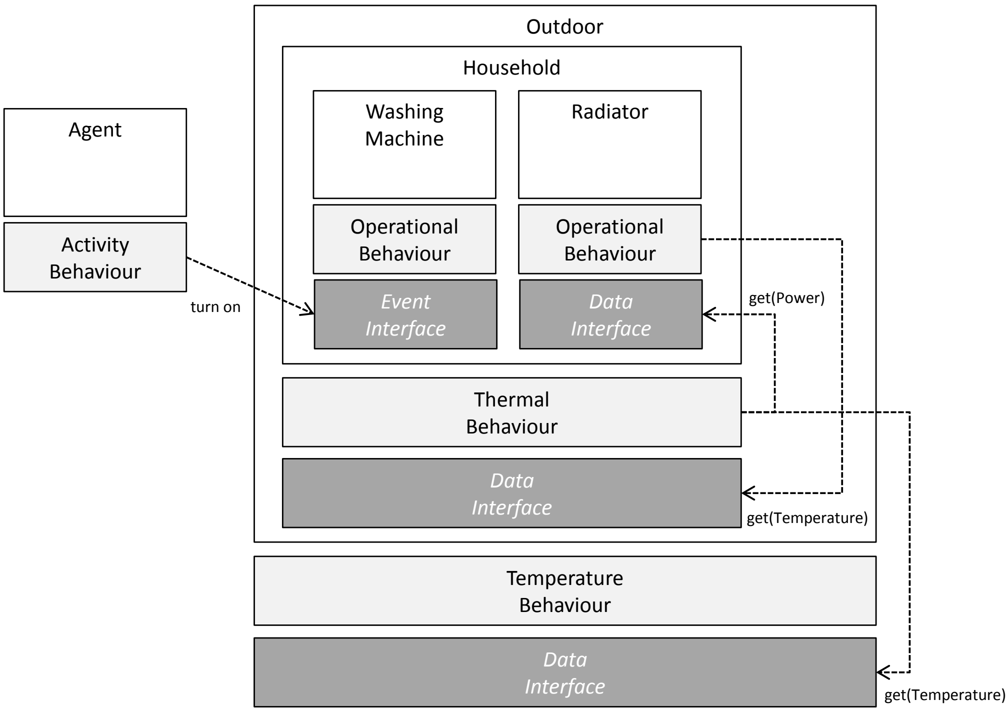


Figure 8.5: Model composition.

- Finally, the Agent represents the people living in the Household and the associated behaviour defines the actions that these people are performing. For example: a person turning on the Washing Machine.

The coupling in this model is represented by the dotted lines in the figure 8.5. This coupling is always defined from behaviours to interfaces. The Agent depends on the Washing Machine to change the operation mode of this device. The Radiator depends on the Household temperature, since the heat radiation is calculated based on the gap between the Radiator reference temperature and the Household temperature. The Household has two dependencies: with the Outdoor temperature and with the Radiator power, since the Household temperature is calculated by a numerical solution of a differential equation which includes these two variables. Note that, in this case, there is a cyclic dependence between the Household and the Radiator.

8.2.4 Asynchronous simulation dynamics

A system simulation requires time-management to ensure that temporal aspects are correctly represented and emulated. This temporal representation only exists during the simulation process and is referred to as "Simulation Time". Simulation Time is represented as a timestamp, a long integer where a unit corresponds to a millisecond of real time.

The time-management in a synchronous simulation is centralised while the time-management in an asynchronous simulation is distributed. That is, an asynchronous simulation involves that every object manages its time, so they could have different timestamps (Figure: 8.6).

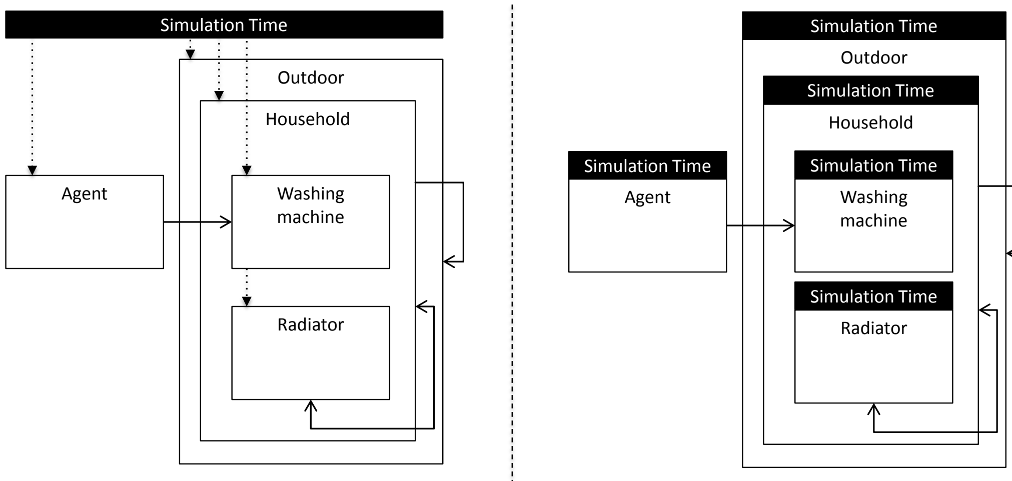


Figure 8.6: Synchronous vs Asynchronous simulation.

In this simulation paradigm, when an object is not coupled with other objects, its Simulation Time develops without considering other object Simulation Times. In this simulation case, the Outdoor is completely independent of other objects.

However, when objects are coupled, the challenge consists of correctly reproducing temporal relationships. The identified temporal relationships are as follows:

1. Coupling with a data interface
2. Cyclic coupling with data interfaces
3. Coupling with an event interface

In the following sections these relationships are discussed.

8.2.4.1 **Coupling with a data interface**

Since an object could access a variable of an external object which may be in a different time instant, every object must keep the different states that have been calculated during the simulation execution. So, when a variable is modified, a state snapshot is created in order to keep the object state in this time instant.

If an object is querying for a variable value in a time instant t_i, there are two cases: the object Simulation Time is delayed or ahead with respect to the external object Simulation Time. In the first case, the external object is able to provide the value by retrieving the last snapshot previous to this time instant (t_i). In the second case, the dependent object must wait until the external object reaches this time instant (t_i).

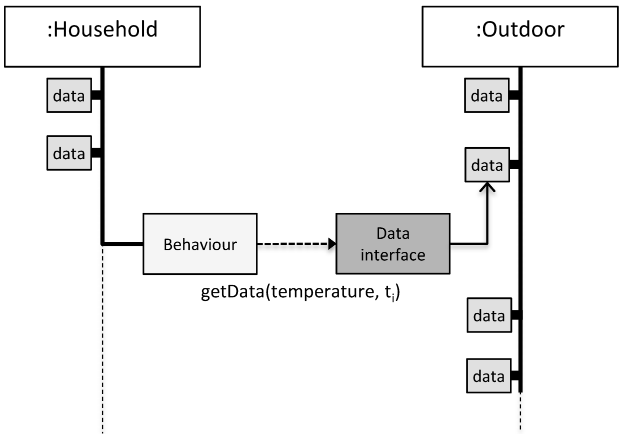


Figure 8.7: The Household asks the Outdoor. Time is vertically represented.

In the figure 8.7, the first case is shown. The Household Simulation Time is t_i and the Outdoor Simulation Time is t_j. Whenever t_i is lesser or equal than t_j, the requested data can be delivered since the data has already been calculated and stored.

However, when the Household Simulation Time (t_i) is greater than the Outdoor Simulation Time (t_j), the Household behaviour is blocked (Figure: 8.8) until t_j is greater or equal than t_i (Figure: 8.9) delivering the last Outdoor Temperature value stored in the last calculated snapshot.

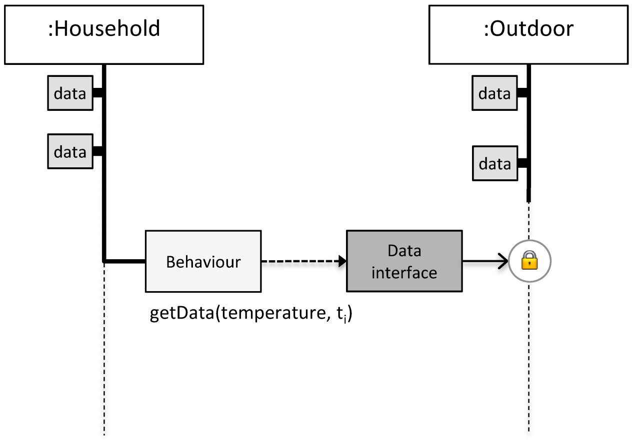


Figure 8.8: The Household behaviour request gets blocked.

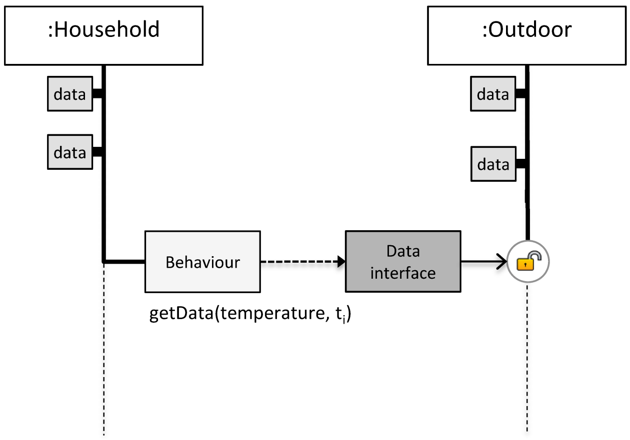


Figure 8.9: The data is delivered.

8.2.4.2 **Cyclic coupling with data interfaces**

The cyclic dependence is a concrete case of the data dependence. Two objects depending on each other whose Simulation Times are different, is handled with the following rules: the most delayed one will always retrieve the required data while the most advanced will be blocked until the delayed reaches its Simulation Time (Figure: 8.10). The mutual blocking is not possible since objects retrieve the value for the current Simulation Time to calculate the next Simulation Time value.

In the example shown in the figure 8.10, the Household requires the power consumption of the Radiator in order to calculate the new temperature value. On the other hand, the Radiator behaviour needs the Household temperature value to

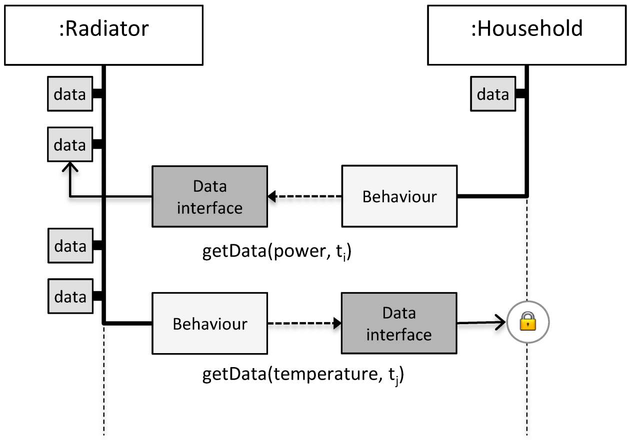


Figure 8.10: Radiator and Household cyclic dependence resolution.

modify the Radiator state, since the reference temperature at the Radiator thermostat serves as a control mechanism.

8.2.4.3 Coupling with an event interface

The event coupling means that an object receives external messages that contain orders for changing its internal variables. This is the case of objects which are managed by people that are represented as Agents in the model. The Agent interacts with these objects by sending a message using the object event interface. When the message is received by the object interface, the object Simulation Time is developed and then, a new snapshot state is created.

It could happen that the Agent develops its Simulation Time without the intention of sending an order to any object. In this case, the Agent behaviour must send a ``Notification Time Message´´to the object. In fact, when the Agent Simulation Time develops, the Agent behaviour must send a Notification Time Message to all objects the Agent is controlling. This notification determines how long an object can develop its Simulation Time. This type of relationship means that object's Simulation Time that is controlled by an Agent, will never exceed the Agent Simulation Time.

Figures 8.11-8.14 shows an event relationship between a social Agent that turns on the Washing Machine. In this example, the Washing Machine Simulation Time

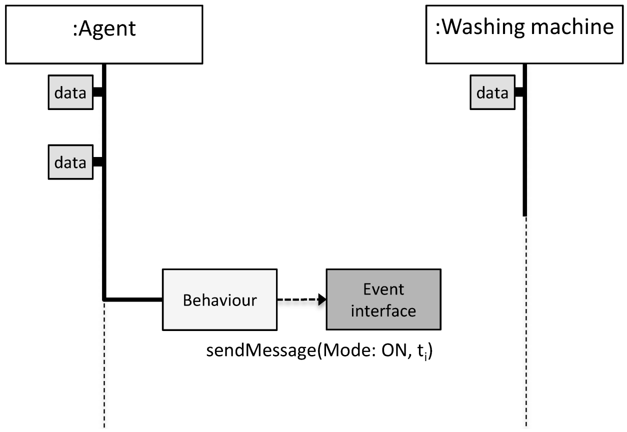


Figure 8.11: The Agent sends a message to turn on the Washing Machine.

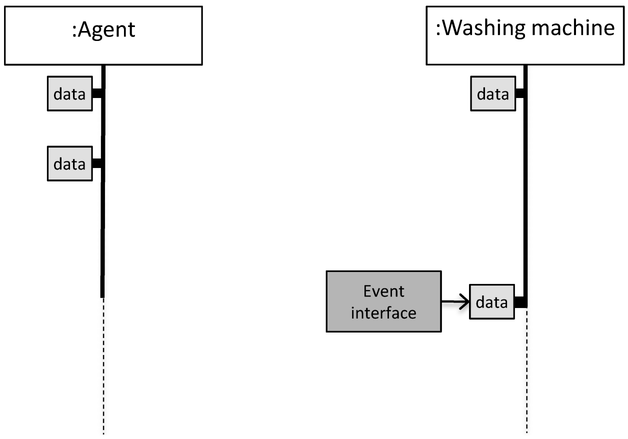


Figure 8.12: The Washing Machine event interface changes the object mode to on.

is always behind the Agent Simulation Time. In other words, the Agent Simulation Time sets a restriction for the Washing Machine Simulation Time.

In the case of the Washing Machine, its power consumption would be 0 at the beginning of the simulation as it's off. Therefore, a new snapshot is created when the Agent turns on the Washing Machine. From that moment, the Washing Machine behaviour will calculate the new power consumption with the restriction that the calculations development should not exceed the Agent Simulation Time, in case the Agent turns off the Washing Machine.

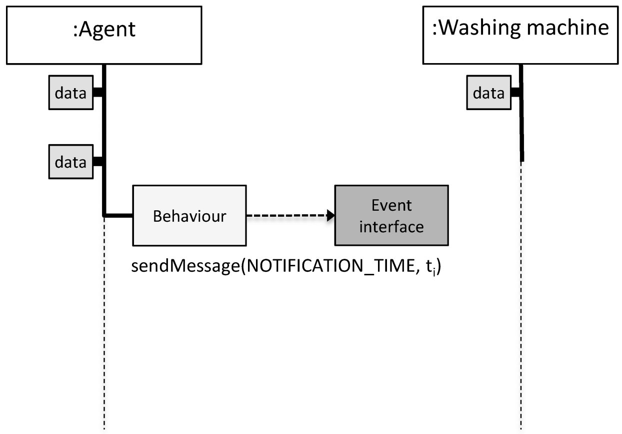


Figure 8.13: The Agent indicates the Simulation Time in which it is to its controlled objects.

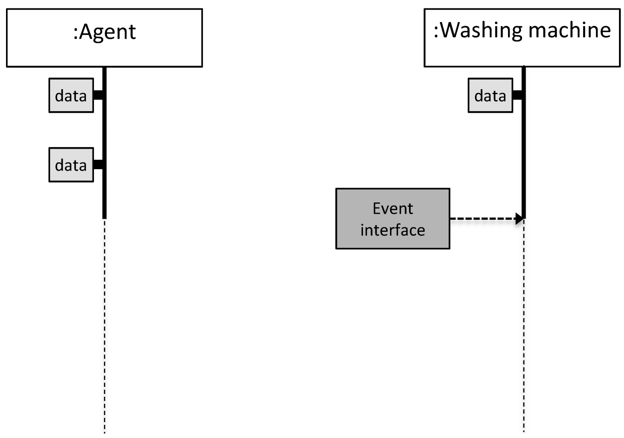


Figure 8.14: The washing machine receives the time notification.

8.2.5 Scales

The dependencies explanation has been focused on the low scale level. This is due to the fact more complex interactions take place at this level in the demand simulation of the power grids. Scaling up from the presented case to power grid levels demonstrates how the time would be developed following a bottom-up approach. In the figure 8.15, information flows are shown which indicate how the demand power is aggregated from the lowest levels to the highest ones at a concrete time slice. This aggregation is required to calculate the demand at every scale. Assuming that every element of a level makes the same calculations, it could be observed that each level may be delayed with respect to the lower one. This is the typical case since the upper elements are waiting for the information coming from the lower elements. However, it is possible for all of them to be in the same time instant. It is not possible for upper levels to be ahead of the lower ones.

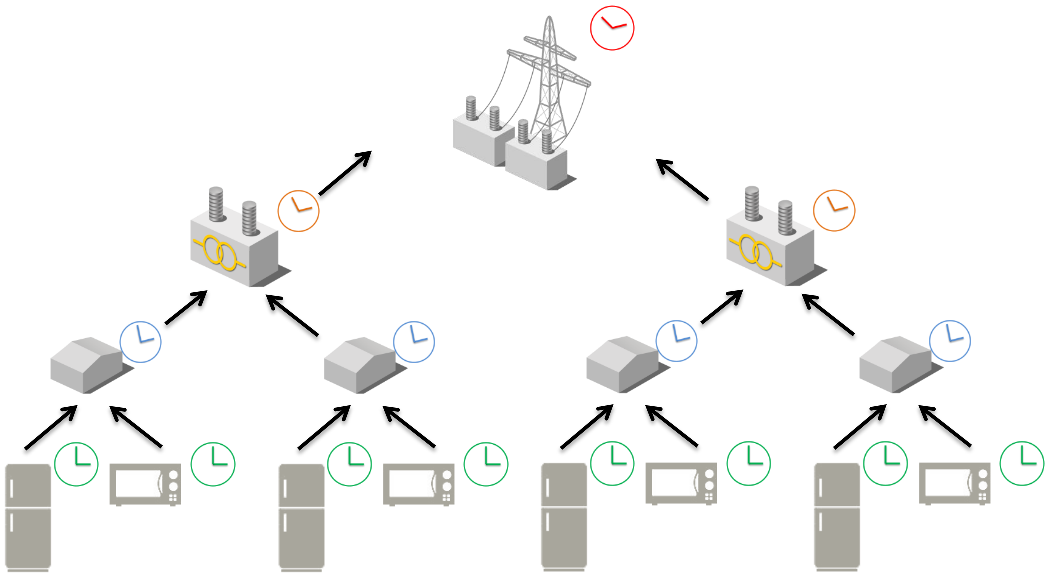


Figure 8.15: Demand simulation in a higher scale.

8.2.6 Object time management

In the previous cases, the discussion was focused on objects with a single type of behaviour. However the time-management of an object with several behaviours or/and several event interfaces must be dealt with. Every time a type of behaviour is executed, it registers the Next Time Execution, that represents when it should be executed. The object time-manager selects the behaviour with the nearest Next Time execution to the current Simulation time. The event interfaces are dealt with in the same way, so that the Interface Next Time Execution corresponds with the time defined in the last message received. Whenever a received message concerns a variable value modification, the object behaviours will be executed afterwards allowing a change in their Next Time Execution, according to the new state. Therefore, objects can dynamically develop their Simulation Times : that is, their Pace could vary from one Simulation Time to the next one. To illustrate the internal time-management, the photovoltaic cell behaviour is studied. This behaviour calculates the generated power, based on the environmental solar radiation. Therefore, the generation power variable will vary along the day until the sunset when the production will become 0. Then, this variable will not change until sunrise. According to this behaviour, three solutions can be proposed to avoid systematic calculations through the night:

1. When the sunset is reached the behaviour registers the Next Time Execution in the sunrise time, whenever this data is available.

2. When the sunset is reached the behaviour registers the Next Time Execution of the previous known sunrise time. This temporal jump may avoid the first solar radiation when the sunrise time is before the already known one. Therefore, the Next Time Execution could be the previous known sunrise time minus ten minutes.

3. The photovoltaic cell outdoor could send messages to the photovoltaic cell event interface whenever solar radiation changes. Following this, the photovoltaic behaviour could register its Next Time Execution to infinite (sleep mode). Therefore, solar radiation changes are received by the photovoltaic cell event interface-allowing mode of behaviour to access this information.

8.2.7 Implementation

In this section, architectural methods to implement this approach are presented. This architectural proposal takes into account the previously described requirements for simulating a power grid, using an asynchronous approach.

A Tafat Thread represents the execution of a single Model Object and from this point of view describes the execution state, awake or sleeping, and the Simulation Time in which it is (Figure: 8.16). During the execution of the whole simulation, Tafat Core request awake Tafat Threads to be executed. After this execution, a Model Object will have changed its Simulation Time and/or its state. In order to improve the performance, Tafat Core keeps a list of the awake threads and it is listening for state changes in threads to update this list.

A single Model Object has many controllers that can modify the Simulation Time. A controller factor could be either Behaviour or an Event Interface. These controllers, that implement the Develop Time interface, participate in the Model Object simulation, each of them proposing different Next Simulation times. When the Next Simulation Time of any of these controllers is undefined, the Tafat Thread that represents the Model Object turns into a sleeping mode. Once, the Next Simulation Time of all the Model Object controllers are defined, the thread will wake up. Next Simulation Time of Develop Time Controllers could be set to undefined or a value that should be greater than the current Model Object Simulation Time. A feasible value for a Next Simulation Time could be infinite, meaning that Behaviour is suspended, pending an external event.

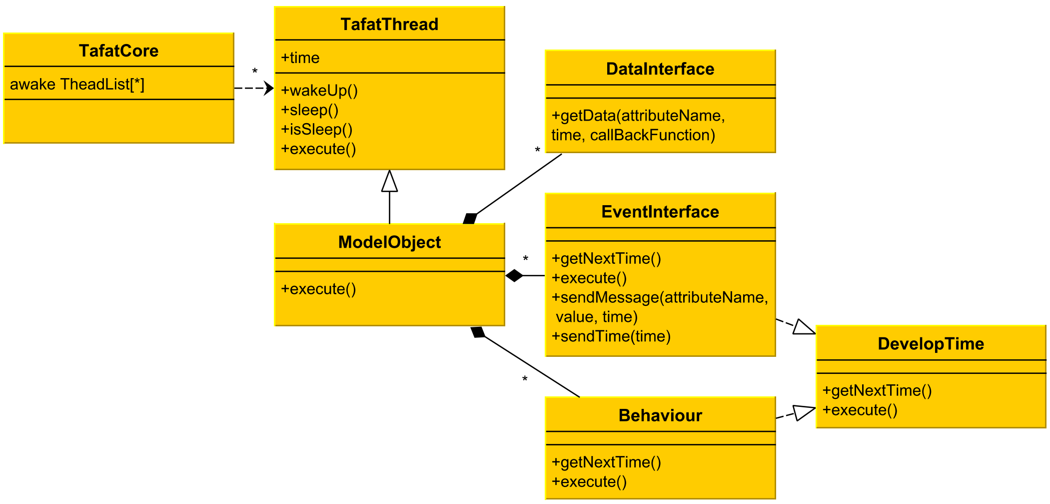


Figure 8.16: Tafat asynchronous simulation architecture.

For example, the Next Simulation Time of a Washing Machine Behaviour can be infinite, so that the washing machine is off and therefore, it is waiting to be turned on (Figure: 8.17). On the other hand, the Next Simulation Time of this Washing Machine Event Interface is undefined until other Model Object Behaviours that use it, set the Next Time Simulation. Since the Washing Machine depends on the Social Agent to be modified, the Social Agent must inform this device of this. This Current Time is transmitted through a message which arrives at the Washing Machine Event Interface. When the Social Agent Current Time arrives, the Washing Machine Event Interface will modify its Next Simulation Time from an undefined value to the one which has arrived in the message. Whenever an event for modifying the state of the Washing Machine arrives, the Washing Machine behaviour will be executed once, allowing it to calculate its Next Simulation Time based on this new state.

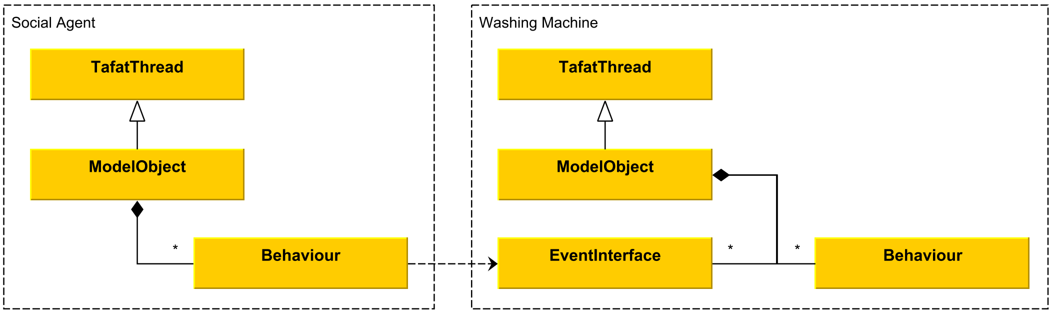


Figure 8.17: Interaction example between the agent and the washing machine.

Another improvement from the performance point of view is based on the Snapshots removing. A concrete Model Object may have dependences for requesting data or set values in external Model Objects. Similarly, other Model Objects could require this one to be accessed. For this reason, the Model Object must keep the snapshots for all the Model Objects which request data. As this Model Object knows the data requesters, it is able to find out the time in which the requesters are and, therefore, it could delete the Snapshots which are previous to the Current Time of the most underdeveloped Model Object requester.

CHAPTER

9

Simulation results analysis

As said before, simulations play a crucial role in the design of SG policies since they are a way to test them before their launch. However, the output provided by the simulations must be managed in a way that allows the policy designers to make decisions. This section explains the main concerns when analysing results obtained in a SG simulation. When facing a simulation of SGs based on a complex system approach, the results analysis becomes a difficult stage since the amount of entities is large.

All systems containing a large amount of entities and relations in simulation processes provide a large amount of results. The way in which these data are normally exported is through data files. These data files are usually designed according to the data that will be managed thus avoiding the possibility of querying this data beyond what was decided to export. Therefore, whenever we deem it convenient to extract data, which was not considered to export at the design phase, a new simulation must be configured and executed.

In order to exemplify this issue, a disaggregated model of a power grid system is used. This system only consists of the demand side, which is disaggregated at the device level. It is precisely at this level where we can find a layer consisting of heterogeneous elements, since the characteristics to extract from a radiator are not the same as the ones from a television (TV). If we want to preserve all variables that

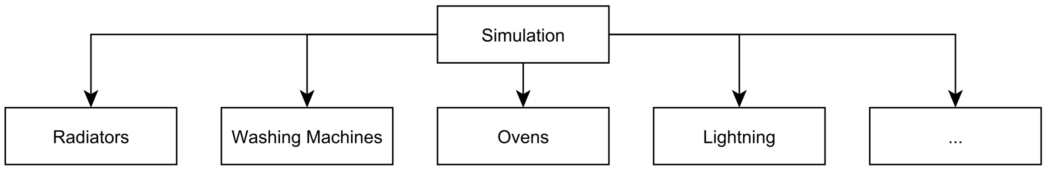


Figure 9.1: Structure example to export simulation results.

are not common to every device, it will be necessary to export each device type into a different data sheet (Figure: 9.1). At this point, once the data exportation process has been defined, we can start thinking about querying it. The list below states some query examples and how they should be dealt with according to this data exportation structure:

- **Querying the consumption of all devices.** This query is very likely to be required. According to our data structure, firstly we calculate the total consumption at each device type. This would involve opening as many files as device types and making the calculations to obtain the total consumption per device type. Secondly, those columns which have the aggregated value at each device type must be moved into a new sheet where the final calculation would be performed obtaining the query result. The more device types there are, the trickier this process becomes.
- **Querying the consumption of all devices in a specific household.** This process would consist in gathering the columns belonging to all the devices contained in the household from the data files. Once they are all together in a new sheet, the query result can be obtained by adding up.
- **Querying the consumption of all devices in a specific district.** The process to obtain this query is really tricky. Firstly, all the devices belonging to a specific district must be listed. Next, all the columns which refer to the devices consumption must be gathered from the device type sheets following this list. Finally, all gathered columns can be moved to a new sheet where the query can be obtained.

Taking these examples into account, it is possible to imagine how tricky the results management of more complicated queries can get. Probably, some of these queries are easier to obtain by redefining the simulation results format and running it again. However, it would also be really tedious, and depending on the simulation kind, the results may differ from the previous simulation and in the end it would be necessary to start the result analysis from the beginning.

All these difficulties in querying the output of a simulation could involve that many other queries are not made due to the fact that they involve a strong and time consuming effort to perform them. Unfortunately, this usually leads to focus on a small subset of variables of the simulation neglecting much information and wasting too much time in performing simple queries.

The root of the problem behind the result analysis is that such results have a multi-dimensional and a multi-scale (namely temporal and spatial) nature which cannot be managed by using conventional data sheets. The example of the demand disaggregation is multi-dimensional and multi-scale. Multi-dimensional, since each dataset (for example, a power measure) is related to a specific device, location (household, building, district, etc.) and time. Multi-scale, since the information can be aggregated at different time scales (per hour, per day, per month, etc.) and at different spatial levels (device, household, etc.).

9.1 Business Intelligence methodologies for analysing data

A framework, named Sumus, for applying BI methodologies has been developed. This framework for data analysis is based on OLAP. OLAP is a solution used in BI, the aim of which is to accelerate querying large amount of data. OLAP is based on cubes [CD97](Figure: 9.2), a multi-dimensional structure where data is stored. These cubes enable the insertion of data, namely facts, which are referred to several dimensions. For example, the measure of power taken from a washing machine can be referred to the device, the household where the device is and the time. Therefore, in this case, there would be three dimensions: devices, households and time.

The structure of a multi-dimensional cube which addresses our problem is

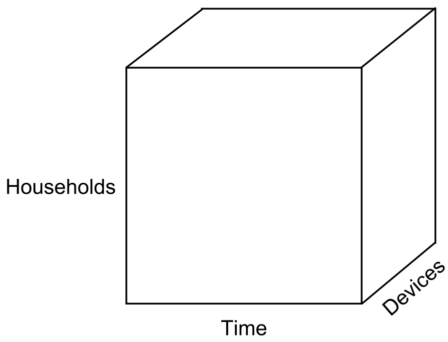


Figure 9.2: A multi-dimensional cube for residential consumption.

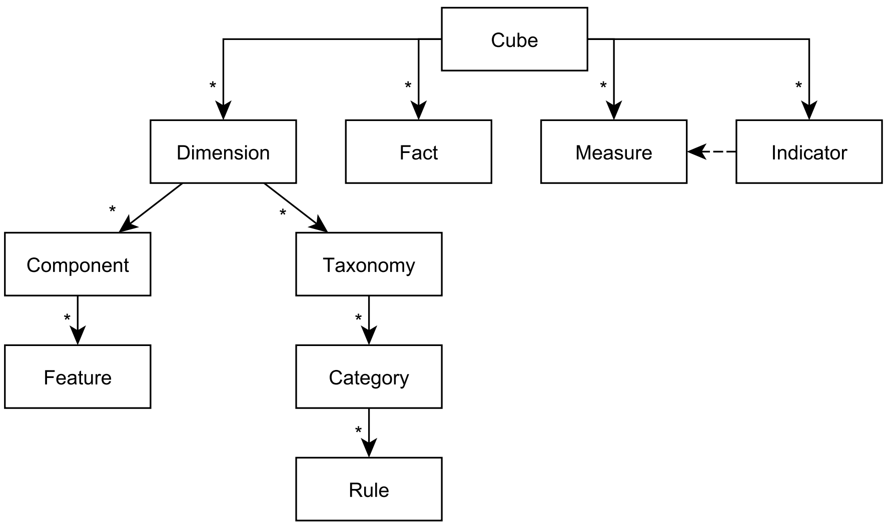


Figure 9.3: An OLAP Cube structure.

presented in the figure 9.3. Every cube consists of dimensions, measures and indicators. The list below describes every cube component.

- **Dimension:** it establishes a way to access the data inside the cube. Every single data is related to some elements such as when and where it happened. For example, a data of power consumption of a household would be related to the dimensions household and time.
  - **Component:** it is an element which is related to a dimension. For example, a dimension which concerns households would be filled by components which are households.

        * **Feature:** it is a property of the component. In case the components are households, a possible feature could be the number of square meters there is in each household.
    - **Taxonomy:** it is a way of categorizing a dimension. There are different ways to categorize the components inside a dimension. Each of these ways is known as taxonomy. In the example of the household dimension, a taxonomy could be the size or the orientation of the facade.
        * **Category:** it is a set of components that satisfy some specific conditions. For instance, possible categories for the size taxonomy could be small, medium or big. Therefore, each of these categories would contain a set of household components the relationship of which is having a similar size.
            · **Rule:** it establishes the condition that a component must meet in order to fall into the category that owns the rule. In the case of the small category, a possible rule could be: all the household components the feature of which *number of square meters* is below $80m^2$

- **Measure:** it provides a semantic to the data inserted in the cube, e.g. the power of the household mentioned above is just a number. However, the power measure is what provides the semantic to this number. A measure is usually related to a metric which enables the comparison among measures that are in different cubes. In this case, the metric of the power measure would be Watts.
- **Indicator:** it designates the way in which a measure or a set of measures are aggregated. For example, the power measure could be aggregated using an average function. This way of aggregating measures is known as indicator. It is possible to have several indicators for one measure, i.e. the integral operator over the power measure would provide a second indicator over this measure which could be designated as energy indicator.
- **Fact:** it relates the measures of a cube with the dimensions. A fact indicates that a certain combination of values (measures) took place for a specific combination of elements (components). In other words, a fact can be understood

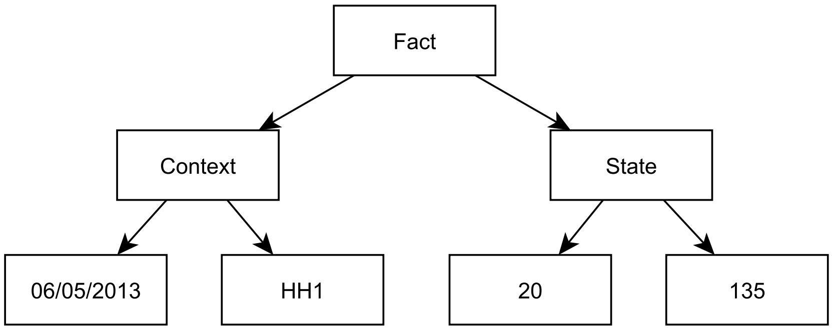


Figure 9.4: A fact consists of context and state.

as a relation of a state to a context. The state is a set of measures and the context consists of components including time. In figure 9.4, the state contains 20 (centigrades) and 135 (Watts) as measures. These measures are related to a context which indicates the time and household where those measures were taken.

9.2 On-Line Analytical Processing for Smart Grids

In this section, all concepts exposed previously will be used in a practical case. Assuming that a new SG policy is to be tuned, several simulations of power grid demand will be performed. To make decisions, these simulations must focus on the power demand and the temperature at the residential sector. Therefore, the scenario for those simulations consists of several districts with households (Figure: 9.5). Each household contains several devices and calculates the internal temperature.

To this end, several cubes have been designed so as to analyse the data coming from the simulation: first of all, the household cube which contains the facts regarding the temperature and, secondly, one cube per device type (TVs, Radiators and Washing Machines, among others) which contain facts about the devices. Since there are many kinds of devices in a household, in this example we are going to focus on two of them: TVs and radiators.

There are two dimensions in the household (HH) cube: one measure and one indicator (Figure: 9.6). The Time dimension is common to all cubes and configures a standard way of categorizing the timeline. Household dimension contains the

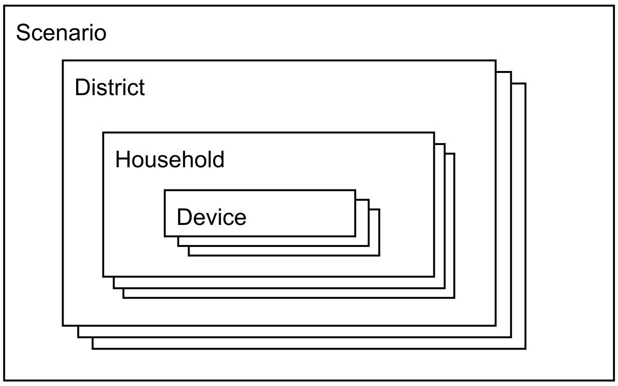


Figure 9.5: Scenario composition.

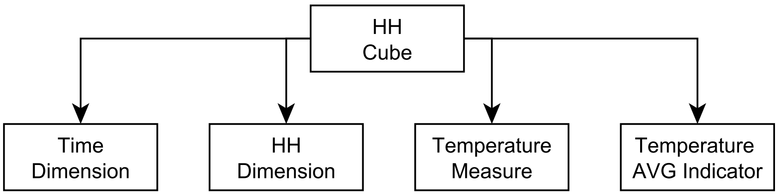


Figure 9.6: Household cube.

households transformed into components which are described by features. The temperature of the household is the only measure that this cube is going to store and it will be aggregated using an average criteria according to the designation of the indicator.

The household dimension contains a taxonomy which concerns the locations (Figure: 9.7). This taxonomy is categorized following several levels: country, city and district. For instance, two household components have been included, both of which contain a feature which is their location using UTM coordinates. Therefore, these location features allow the dimension to identify which district each household is located in.

The TV and Radiator cases are exposed in order to demonstrate why devices must be disaggregated into separated cubes. The main reason for this separation is due to the fact that both devices do not share the same features and, therefore, their classification methods are different. This separation enhances the capacity of making queries since it is possible to filter components by features that are only present in a specific kind of device.

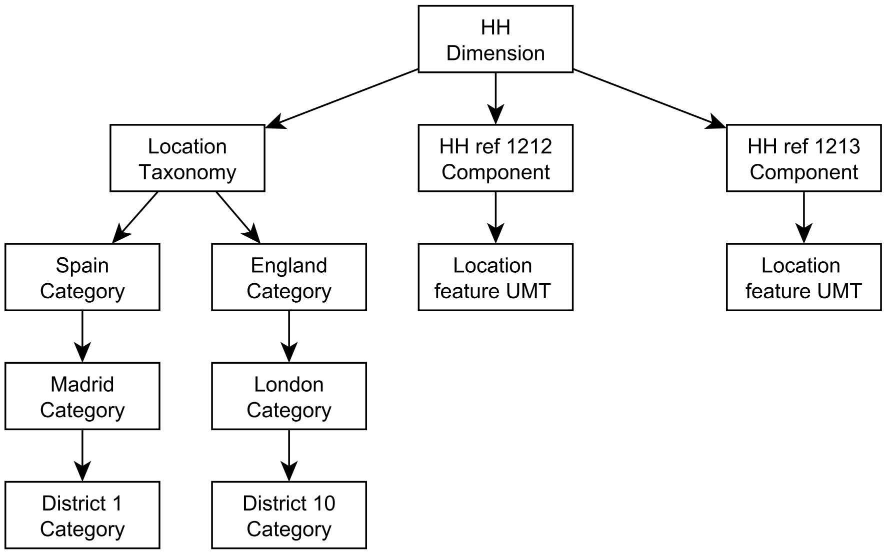


Figure 9.7: Household dimension.

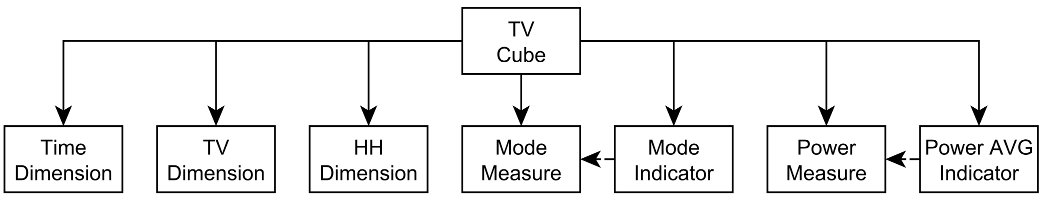


Figure 9.8: TV cube.

The TV cube registers data about power consumption as well as the TV mode (off, standby and on) (Figure: 9.8). Every set of measures (power and mode) is related to three dimensions: time, household and TV. Time and household dimensions are exactly the same dimensions as the ones detailed above. The TV dimension contains information about the TVs in a component format. Furthermore, there are two indicators which are responsible for aggregating measures: the mode indicator, which performs a calculation that provides the percentage of TVs that are turned on, and the power indicator, which aggregates the power measures registered using an average formula.

The TV dimension, like the household dimension, focuses on specific features related to TV components (Figure: 9.9). In this case, the possibility of filtering TVs using a technological criteria is considered relevant. Therefore, two categories

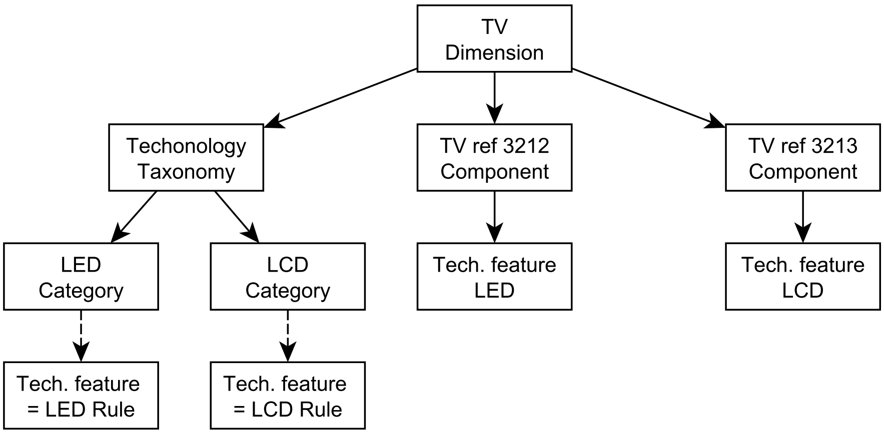


Figure 9.9: TV dimension.

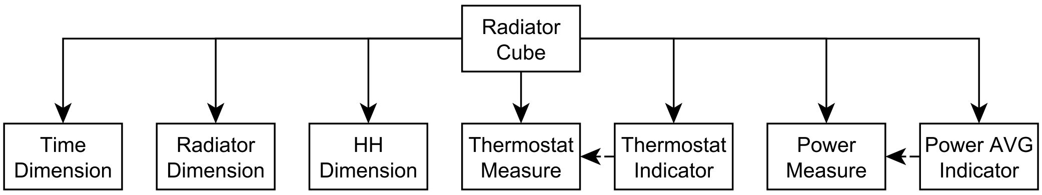


Figure 9.10: Radiator cube.

have been created so as to separate LED televisions from LCD televisions. This information will allow us to compare the consumption among the different TV technologies. Hence, TV components contain the technology feature which will be used to calculate whether a TV belongs to the LED or LCD category by using the rules that are related to these categories.

The radiator cube stores measures related to both the power consumption and the thermostat level (Figure: 9.10). These measures are related to three dimensions, as in the case of the TV cube. In this case, apart from time and household dimensions, a new dimension has been designed: radiator dimension. This dimension contains components that represent radiators and their features. In addition, there are two indicators which aggregate the measures. On the one hand, the thermostat indicator aggregates the measures stored using a gradient function which shows big changes in the thermostat level in short periods of time. On the other hand, the power indicator aggregates the power measures using an average formula like in the TV cube.

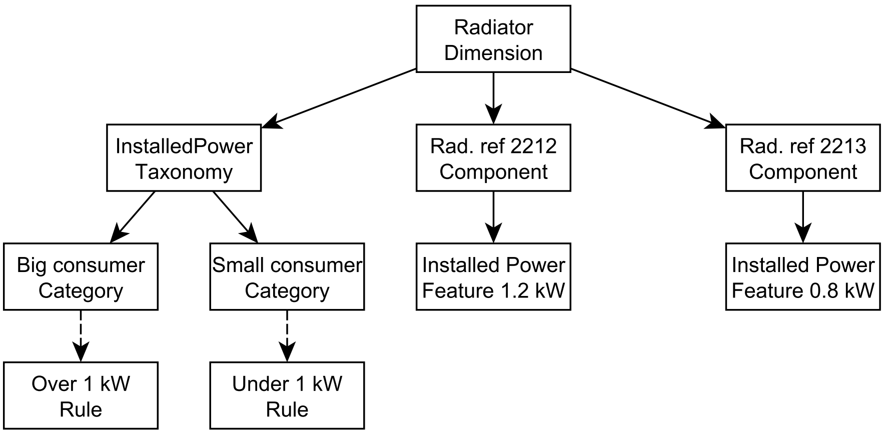


Figure 9.11: Radiator dimension.

The radiator dimension focuses on specific features which concern radiator components (Figure: 9.11). Since radiators are usually considered big consumers, a taxonomy to classify them into two groups has been designed. Indeed, this taxonomy will allow us to find out the amount of radiator components which are in what we consider a small consumer category (under 1kW installed power) or a big consumer category (over 1kW). Two components belong to this dimension and contain the feature installed power which is used to perform the classification in the installed power taxonomy.

9.2.1 N-Level indicators and Data mining

So far, some mechanisms which allow us to extract information based on the measures have been presented: indicators. These indicators are regarded as first level indicators since they are just based on measures. However, it is possible to define, a second level of indicators which are computations carried out based on previous level indicators. This idea can be extended to the concept of N-Level indicators. Through introducing this concept, data mining[RKU11] procedures can be used in order to find out patterns.

An example of this is presented in figure 9.12. In this case, a data miner has been designed in order to identify consumption habits which concern radiators. Using the thermostat indicator, which calculates the gradient based on the thermostat

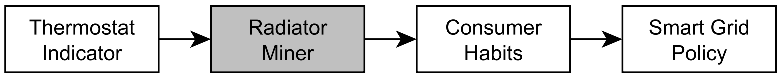


Figure 9.12: Radiator data miner.

level measures, this data miner is able to identify habits throughout time. Therefore, common patterns of radiator usage could be identified and used to feed the SG policy design.

Moving onto low level details, this miner queries, for each radiator, its thermostat indicator throughout time. Based on this indicator, it uses techniques to extract habit patterns. These habits can be used by the policy in order to exert a more personalised control over the demand which enhances customer quality of service.

The list below presents two other cases where miners can be used in order to improve the design of a SG policy:

- Based on the technology feature of TV components, a miner can calculate the average time to amortise a TV based on a low-consumption technology by comparing them to the consumption of other technological kinds. According to these results, a SG policy could subsidise the purchase of TVs with a lower consumption. This kind of policy applies to other device kinds such as fridges or washing machines, among others.
- In the household cube, a miner can correlate the temperature and energy consumption of a household with its isolation features, supposing the information is available. Based on this correlation, the improvement of household isolation could be proposed.

9.2.2 Information visualisation

Information visualisation is the use of visual representations of data which improve human cognition [War12]. This is an important stage when analysing data since an appropriate visualisation may reveal information that could not be possible to extract using other visual representations. Figure 9.13 presents different visual representations which are discussed in the paragraph below.

Time (h)	Temp. (ºC)	Power(kW)
0	14	0,7
4	13	0,5
8	16	0,9
12	19	1,2
16	18	0,9
20	17	1,5
24	15	0,9

a)

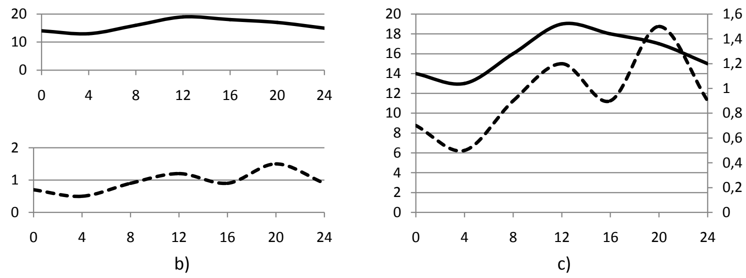


Figure 9.13: Visual representations of both temperature (temp) and power.

Part a in figure 9.13 presents the data in a table format. From there, it is cognitively difficult to extract temperature or power trends throughout the day, especially when there are many rows. In order to deal with this issue, both series can be represented in separated charts to enable the extraction of trends (part b). These trends can be extracted at each series but relational effects among them are neglected. However, part c represents both individual trends and relational effects among them. It is possible to extrapolate the relation among high temperatures and high consumption at noon.

The example presented in the section is a simple case which intends to provide insights of what is known as information visualisation. In this case, the representation required to show the information correctly is too evident. However, there are cases in which finding out the appropriate way to represent the information requires a deeper study.

9.2.3 Decision making

Using this approach to perform the simulation analysis enhances the ability of making decisions. The way in which the data is structured facilitates the interaction. From now on, queries can be as complex as needed in order to find out interesting conclusions which feed the decision making at the SG policy design. This structure is to be consumed using information visualisation patterns which could reveal interesting information that cannot be detected simply by analysing numbers.

In the previous example, important information can be extracted to be used in the SG policy design. The list below summarises some of the most relevant information:

- **Differences among districts:** Using the household dimension, it is possible to find differences among the districts located in the same city or, even, among cities. These differences can be noted in the way in which power is consumed, the devices are used or the temperature in the households. All of this could help in the design of a policy which provides enough flexibility in order to deal with these differences without losing efficiency when applied.
- **TV case:** it is possible to compare the differences in the consumption related to the TV technology. However, the TV dimension can be designed to take into account other aspects such as labelling and size. For instance, the labelling taxonomy could give information about whether it is worth promoting a SG policy which would subsidise the purchase of new high efficiency TVs.
- **Radiators:** using N-Level indicators allows us to identify consumption patterns that can be used to design more efficient SG policies which take into account customer usage. In other words, those patterns may be identified in order to build an intelligent control. On the one hand, this control could take note of the customer timetable in order to look after the quality of service. On the other hand, this control could take into account the grid state in order to reduce or increase consumption dynamically.

CHAPTER

10

Demand Side Management policy design

The design of DSM policies has to comply with the complexity of the decisions that are made by all the actors of the power grid. From the point of view of ABM, these actors are agents that perceive the environment and react according to their interest.

Generally, the problem that DSM policy has to address is the distribution of limited resources (energy) among all of the actors. This problem has been addressed previously in many other fields by applying AI techniques.

These techniques focus on dealing with the complexity of the many restrictions that are imposed by the actors of the system. In the case of power grids, core actors are those who are in charge of the production and consumption of electricity. Somehow, they need to agree in the way in which the available energy is going to be distributed.

In this work, the application of AI techniques in the field of energy distribution has been researched. In the next part of this document, several experiments are exposed where these techniques are studied and evaluated under large-scale power grids. In these experiments, the techniques presented in this part of the document are used to distribute the energy among all of the consumers of the power grid.

Since this distribution depends on the nature of the experiment, the way in which these techniques have been adapted is explained in each of the experiments.

10.1 Swarming Intelligence techniques for designing policies

A framework, named Bandada, has been developed for researching the application of SI techniques to the design of DSM policies. Bandada is a framework built on top of Tafat which intends to provide all necessary components to design and validate DSM policies. This framework contains a large set of power grid models for the residential sector which can be used to analyse DSM policies. Furthermore, Bandada provides an environment in which policies supported by SI techniques can be designed and evaluated.

The set of techniques that are currently available for designing demand side policies are Particle Swarm Optimisation and Multi Objective Particle Swarm Optimisation. These optimisers are fully described below. Apart from these techniques, modellers can also develop their own DSM policies. For instance, a DSM policy that is characterised for having a decentralised control (e.g. based on frequency) can be implemented by modelling the control of every individual component.

10.2 Particle Swarm Optimisation

PSO is an optimisation meta-heuristic based on the movement of bird flocks [KE95, EK95]. Each bird is a particle i in a d-dimensional space R^d and its movement is described by its position $\vec{x_i}$ and velocity $\vec{v_i}$ at each t. In every iteration t, these variables are updated following the equations below:

$$v_{id}(t+1) = w * v_{id}(t) + c_1 * u_1 * (pbest_{id}(t) - x_{id}(t)) + c_2 * u_2 * (pbest_{gd}(t) - x_{id}(t)) \quad (10.1)$$

$$x_{id}(t+1) = x_{id}(t) + v_{id}(t+1) \quad (10.2)$$

Where $v_{id}(t)$ is the velocity of the particle i in the dimension d at time t. $x_{id}(t)$ is the position of the particle i in the dimension d at time t. c_1 and c_2 are weight factors that adjust the movement. $pbest_i$ is the best position achieved by the particle i. $pbest_g$ is the best position of the neighbours of the particle i. u_1 and u_2 are random factors in an interval of [0,1] and w is the inertial weight.

This optimisation method is initialised with a randomly-generated population of particles in the decision space which try to converge into an optimal one by following the best particles at every iteration. The algorithm of this method is presented below:

```
Initialise swarm.
Initialise pbest. Update gbest. Initialise velocity.
iterationCounter = 0
while (iterationCounter < maxIteration){
  for (each particle){
    Pick random u1 and u2
    for (each dimension){
      Update particle velocity and position
    }
    Update pbest
  }
  Update gbest
  iterationCounter++
}
Report results in archive
```

Listing 10.1: PSO code. For further information please check the bibliography

10.3 Multi Objective Optimisation

A multi-objective optimisation problem can be expressed as follows:

$$\min_{\mathbf{x}} \vec{F}(\mathbf{x}) = [F_1(\mathbf{x}), F_2(\mathbf{x})...F_k(\mathbf{x})]^T \tag{10.3}$$

Subject to:

$$g_j(\mathbf{x}) \leq 0, j = 1, 2...m \tag{10.4}$$

$$h_l(\mathbf{x}) = 0, l = 1, 2...e \tag{10.5}$$

Where K is the number of objective functions and j and l are, respectively, the number of inequalities and equality constraints. $\mathbf{x} \in E^n$ is the vector of design variables and $\vec{F}(\mathbf{x}) \in E^k$ is the vector of objectives, criteria, fitness or cost functions to be optimised. Any comparison ($\leq$, $\geq$, etc.) among vectors applies to vector components [MA04]. Other primary concepts in multi-objective optimisation are non-dominated and dominated points [Ste89]. A vector of objective functions $F(\vec{x^*}) \in Z$ is non-dominated iff there is not another vector, $F(\vec{x}) \in Z$, such as $F(\vec{x}) \leq F(\vec{x^*})$ with at least one $F_i(x) < F_i(x^*)$. Otherwise, $F(\vec{x^*})$ is dominated.

[LTDZ02] proposed a relaxed form of dominance named ϵ-dominance. This acts as an archiving strategy to ensure both properties of convergence towards the Pareto-optimal set and properties of diversity among the solutions found. ϵ-dominance is proposed as an extension of the Pareto-dominance relation so that a point x^* not only dominates those points x lower or equal in all their objectives $F(\vec{x}) \leq F(\vec{x^*})$ and strictly lower in at least one objective $F_i(x) < F_i(x^*)$, but also all points close enough to x^* (i.e., those with a distance to x^* that is less than an ϵ). This value, ϵ, can be provided by the decision maker to control the size of the solution set [HDSQCCM07].

In multi-objective optimisation problems, the optimum solution is, in general, not as easy to formulate as in single objective cases. Since typically there is no single global solution, it is often necessary to determine a set of points that all match a predetermined definition of an optimum for every single problem [MA04]. In this sense, the predominant concept in defining an optimal point is that of Pareto optimality [Par06].

A point $\vec{x^*} \in \vec{X}$ is Pareto optimal iff there does not exist another point, $x \in X$ such as $F(\vec{x}) \leq F(\vec{x^*})$ and $Fi(x) < Fi(x^*)$ for at least one function. All Pareto optimal points lie in the boundary of the feasible criterion space Z [AP96]. A Pareto-optima hardly ever provides a single solution, but rather a set of solutions called non-inferior or non-dominated solutions. The minima in the Pareto sense are going to be in the boundary of the objective region, or in the locus of the tangent points of the objective functions, that is in the region in the space of values of the objective functions vector [Coe98].

Often, algorithms provide solutions that may not be Pareto-optimal but may fulfil other criteria, which can be useful for practical applications. It is the case of weakly Pareto optimal. A point $\vec{x^*} \in \vec{X}$ is weakly Pareto-optimal iff there is not another point, $\vec{x^*} \in \vec{X}$ such as $F(\vec{x}) \leq F(\vec{x^*})$. That is, a point is weakly Pareto optimal if there is no other point that improves all the objective functions simultaneously.

10.4 Multi Objective Particle Swarm Optimisation

The implementation of the MOPSO used in these studies is described in [SC05]. This optimisation method is based on PSO. In this method, the fitness of a particle is calculated considering several objectives which are defined through fitness functions. This method uses the Pareto-optimality [SC05] approach to address the multi-objective problem.

A set containing the best particles raised during the iterations is stored in what is named as *archive*, which retrieves the elements that reach the ϵ-dominance. This dominance relaxes the weak-dominance constraints. The next pseudo-code explains how the MOPSO works:

```
Initialise swarm.
Initialise leaders. Send leaders to archive. crowding(leaders).
iterationCounter = 0
while (iterationCounter < maxIteration){
  for (each particle){
    Select leader. Flight. Mutation. Evaluation. Update pbest.
  }
  Update leaders, Send leaders to archive. crowding(leaders).
   iterationCounter++
}
Report results in archive
```

Listing 10.2: MOPSO code. For further information please check the bibliography

Initially the swarm is populated through the creation of the particles or individuals. A particle contains a system configuration which is set up randomly. This configuration is the particle genetic code. Before starting the iterative process, the leaders are calculated using the fitness functions which evaluate the optimality of the particles. Those that reach the ϵ-dominance criteria will be stored in the archive. That is, the algorithm includes a crowding process that is used to establish a second discrimination criterion (additional to Pareto dominance) [CL02, SC05]. Along the execution this archive may be filled up and, in these cases, the archive deletes particles based on the crowding factor criteria. At every method iteration, each particle is moved according to its previous position and speed which are calculated as follows:

- The speed is calculated based on the previous particle speed, the distance to its best position and the distance to its leader. This method splits the search space in hypercubes, so that the particle leader will be the one that has the best fitness within the hypercube.
- The position is calculated adding the previous particle position to the speed calculated above.

After the particle movement, the particle could be mutated with a probability calculated as 1 / genetic code size (mutation rate). If a particle is going to be mutated, this mutation may be uniform or non-uniform. The uniform mutation allows the particle to explore the search space. However, a non-uniform mutation allows the particle to exploit the search space wherever it is. Later on, the particle fitness is evaluated and the best position of the particle is changed whenever the new position is better than the best achieved in the past.

Once all the particles have been processed, the whole population is evaluated in order to update the leaders of the archive. If this archive reaches the size limit, some of the particles inside it will be deleted following the same criteria as described above. After this step, the next iteration will be executed. The number of iterations must be set beforehand.

The reason why this method was selected instead of others such as Genetics Algorithms is that MOPSO has a good result quality and time response relation. The time response is important since the orders could arise regarding the current production-consumption balance in a real-time system where the speediness of applying the measurements is critical.

Part IV

Experimentation

CHAPTER

11

Experimentation considerations

This part of the document is devoted to show case studies that have been carried out in this research. As previously said in chapter 7, the hypotheses that have been stated in this document cannot be verified, but validated. That is, it is not possible to verify them as to do so it would be necessary to carry out all SG studies. Nevertheless, they can be validated through the experimentation with case studies. In this part, these case studies are presented and explained, providing evidence that validates the stated hypotheses.

In this chapter, some synthetic case studies are presented. The studies have been designed to show how the tools described in the hypotheses part work. That is, how a metamodel is defined, how a simulation model is defined, how to create large-scale scenarios, etc. Each of the chapters in this part refers to a specific case study. These case studies are oriented to solve real problems and, therefore, they can be considered as important evidences that validate the hypotheses.

11.1 Modelling in Tafat

The work, published in [EHH13c] [1], presented in this section concerns prospective experiments that were developed for studying the human flows on a shopping centre from the point of view of the amount of people. In order to run these experiments, the proposed framework will be customised. For this the development of the metamodel, the behaviors and models in which the experiments will be described is necessary.

However, as the aim of this case study is not to provide accurate results, but to test the framework for complex system simulation, a non-contrasted information has been used for simulating the human flows. The information we have designed behind the human flows going shopping is expressed in the list below:

- From 9.30 to 10.00, around 60-70 workers arrive to the Shopping Centre, leaving it between 16:00 and 16:50.
- From 15.30 to 16.00, around 80-90 workers arrive to the Shopping Centre, leaving it between 22:00 and 22:50.
- From 10.00 to 16.00, around 1000-2000 customers will arrive to the Shopping Centre, leaving it between 20 and 180 minutes later.
- From 16.00 to 22.00 around 2000-4000 customers will arrive to the Shopping Centre, leaving it between 20 and 180 minutes later. If such leaving time reaches the 22.00 limit, the leaving time will be 22.00

The next sub-sections focus on the steps for developing a simulator from scratch for running this kind of experiment. The first step is the metamodel development where the elements that are going to be used in the simulation models are described. Later on, the behaviours which address the way of acting of each model element are developed. Once the metamodel and behaviours are implemented, the Profiler tool is used for generating a huge scenario which contains many agents and two shopping centres. Once the model is ready, the simulation is executed and the results coming from it are analysed.

[1] People that participated in this experiment: José Évora, José Juan Hernández and Mario Hernández.

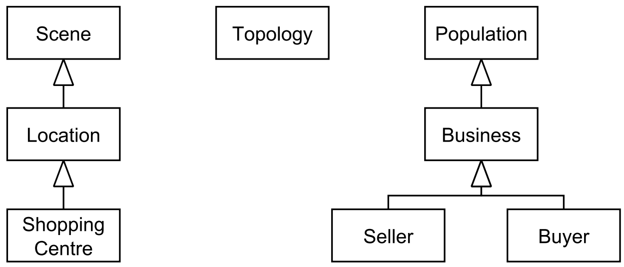


Figure 11.1: Case study's metamodel.

11.1.1 Metamodel development

For this experiment, the metamodel development is simple since it is only composed by three elements. Note that the metamodel design is one of the most critical processes since its good design will allow flexibility for adding or modifying elements without affecting the whole structure. In figure 11.1, the metamodel for this complex system is presented.

From the scene part, there is only one element: the Shopping centre; the topology is empty in this case and the population has two elements: seller and buyer agents. From the point of view of the experiment context, both do the same: get in and get out from the Shopping centre. However, they are separated in order to provide conceptual clarity. When running experiments, it is important to keep the concepts which exist in reality in the metamodel. The next XML codes represent the description of every metamodel element:

```
<class name="ShoppingCentre" parent ="Location">
  <feature name="address" type="string"/>
  <variable name="personCount" type="int" initial-value="0"/>
</class>
```

Listing 11.1: Description of the Shopping Centre metamodel class. This description contains the address where the Shopping Centre is located. Furthermore, it contains a variable which works as a counter of the number of people inside.

```
<class name="Seller" parent ="Business">
  <variable name="shoppingCentreId" type="string" required="true"
      > Id of the shopping centre where the seller works </
      variable>
  <context name="shoppingCentre" type="ShoppingCentre"> Shopping
      Centre where the seller works </context>
```

```
</class>
```

Listing 11.2: Description of the Seller agent metamodel class. When the Seller is instantiated in the simulation model, the shopping centre id is required in order to relate the agent to where it works. The init/configure code will look for this id in order to set in the context attribute the shopping centre provided.

```
<class name="Buyer" parent ="Business">
  <context name="shoppingCentreList" type="ShoppingCentre"
      replicated="true"> A list containing the Shopping Centres
      where the buyer usually go </context>
</class>
```

Listing 11.3: Description of the Buyer agent metamodel class. In this case, the buyer context is a list of shopping centres where the buyer usually goes. That's the reason why an id is not necessary to be provided since this context list will be filled by a discovery process where the shopping centres will be associated to this agent following a criteria (e.g. proximity, prices, etc)

11.1.2 **Simulation development**

The simulation development is guided by the presented Simulation Life Cycle which concerned four main steps. At this moment, all the ideas about how the experiment should be developed must be clear in order to decide how to proceed with the data, which elements are necessary to model, etc.

11.1.2.1 **Data processing**

Since the proposed experiments involve the simulation of many elements, the use of a tool for automating and generating simulation models is necessary. In this case, the Profiler is fed by a data store where the statistical information concerning the human flows on the shopping centre is stored. Based on this information, a model that represents a concrete scenario of the people flows is generated. At this point, and through the use of the Profiler tool, the scenario can be modified as needed according to the experiments design in order to test how the people flows would react concerning variables variation.

The table into which the hypotheses were translated contains the information about the population of the simulation model (Table: 11.1). Getting into the details

Table 11.1: Configuration of the agents that uses the shopping centres.

Agent	Start Time	Duration(m)	Amount	Shopping Centre
Seller	9:30 - 10:00	390-410	60-70	SC1
Seller	15:30 - 16:00	390-410	80-90	SC1
Seller	9:30 - 10:00	390-410	60-70	SC2
Seller	15:30 - 16:00	390-410	80-90	SC2
Buyer	10:00 - 16:00	20-180	2000-4000	-
Buyer	16:00 - 21:20	20-180	4000-8000	-

of the table presented, the columns can be observed: Agent type, Start Time, Duration time, Amount and Shopping Centre. The agent type column can have two values: seller or buyer. The Start Time expresses a temporal range in which the agent is going to arrive to the shopping centre. This range will be handled a posteriori by the Profiler by using a uniform distribution. The duration time is exposed in minutes and concerns the time that the agent is going to stay at the shopping centre. This time is also expressed in a range. The last column only applies for seller agents and specifies in which shopping centre the seller works. The buyers or customers have this column empty since they can go to any shopping centre.

11.1.2.2 Model development

Using the output of the previous step the simulation model development can be addressed.

```
<simulation from="15/01/1995 00:00:00" to="15/01/1995 23:59:00">
  <scene>
    <outdoorfull>
      <shoppingCentre id="SC1">
        <behavior name="ShoppingCentreBehavior" release="
            PersonCounter" step="60"/>
      </shoppingCentre>
      <shoppingCentre id="SC2">
        <behavior name="ShoppingCentreBehavior" release="
            PersonCounter" step="60"/>
      </shoppingCentre>
    </outdoorfull>
  </scene>
```

```
  <population>
    <seller shoppingCentreId="SC1">
      <behavior name="CommercialAgent" release="InOutTiming"
          startTime="9:32" durationTime="403" step="60"/>
    </seller>
    <seller shoppingCentreId="SC2">
      <behavior name="CommercialAgent" release="InOutTiming"
          startTime="9:45" durationTime="408" step="60"/>
    </seller>
    <!-- ... -->
    <buyer>
      <behavior name="CommercialAgent" release="InOutTiming"
          startTime="11:47" durationTime="37" step="60"/>
    </buyer>
    <buyer>
      <behavior name="CommercialAgent" release="InOutTiming"
          startTime="15:58" durationTime="178"step="60"/>
    </buyer>
    <!-- ... -->
  </population>
</simulation>
```

Listing 11.4: Simulation model example which contains two different shopping centres and several seller and buyer agents.

11.1.3 Simulations and results

This section is devoted to present the results of this experiment. As it was previously presented, the number of agents and their timing to get into or going out the shopping centre are stochastic. Therefore, we could expected having very different emergent behaviours across the simulation runs. Next paragraph discuss an averaged curve of the number of people obtained after executing 100 simulation runs in a shopping centre. Next subsections will present several curves in which the variability of these simulation runs will be discussed.

In the figure 11.2, an averaged curve based on 100 simulation runs is presented. Since this curve is an average, this curve is quite smooth without having oscillatory effects. According to how the experiment has been set up, there are more people in the evening than in the morning, reaching a maximum of 1,000 thousand people.

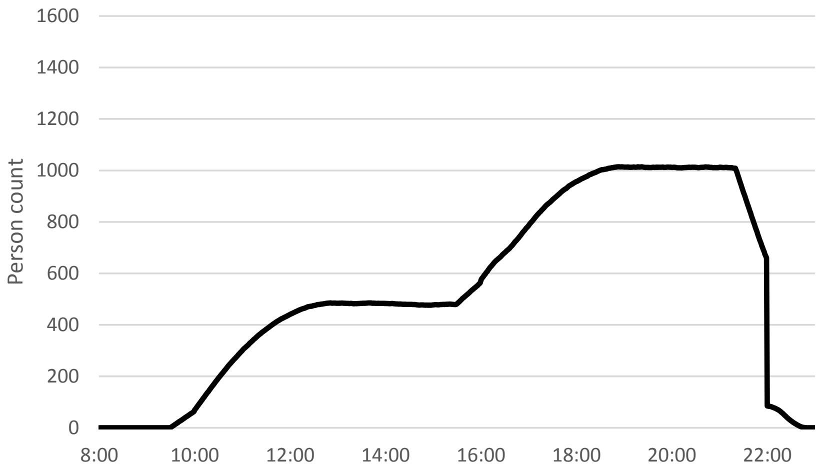


Figure 11.2: Averaged curve after 100 simulation runs.

According to this curve, it may look like there are many agents with coincident behaviours as there are two periods in which the number of people in the shopping centre gets stabilised: from 12 to 16 and from 18 to 21. However, we will see in coming charts that this is only a result of averaging the simulations, since the curve of a single simulation does not have this effect.

11.1.3.1 **Variability in simulation runs with stochastic population and timing**

In the figure 11.3, the curves of 100 simulation runs are presented. As can be seen, the variability across simulation runs is quite higher, providing different results. However, there is a common trend in which there are more people in evenings than in mornings.

In the figure 11.4, the curves of two simulation runs are presented. This chart shows two different simulation runs in which there are significant variations. As can be seen, their shapes are quite different. Therefore, it can be said that both parameters number of agents and timing are significant for the system's emergent behaviour.

11.1.3.2 **Variability in simulation runs with stochastic timing**

In order to evaluate which of the two parameters is the most important, charts are presented in this section, in which the number of agents has been fixed for each period (morning and evening) and the timing has remained stochastically. In the

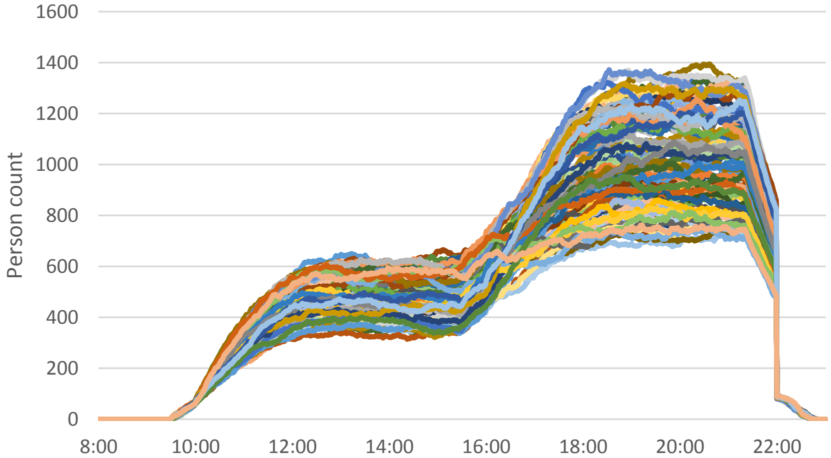


Figure 11.3: Curves of 100 simulation runs.

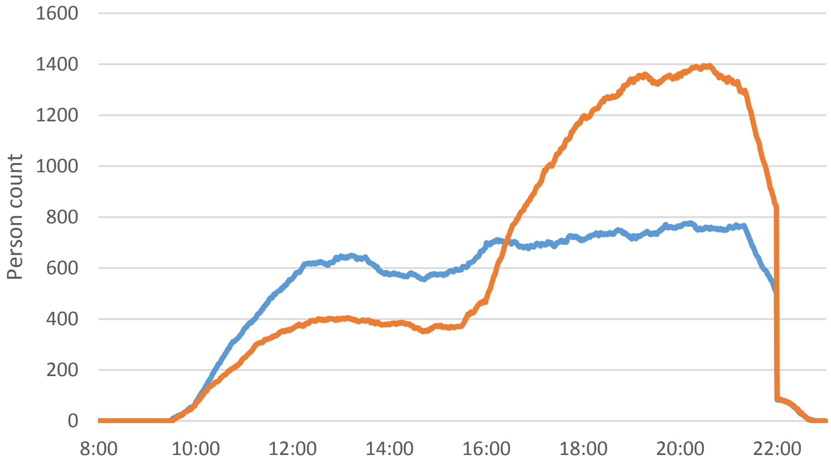


Figure 11.4: Curve of two simulation runs with significant differences.

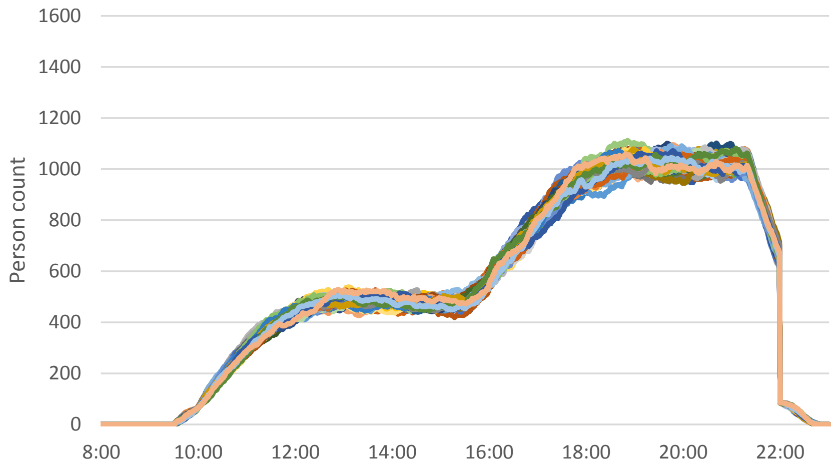


Figure 11.5: Curves of 100 simulations runs with fixed number of agents.

figure 11.5, it can be observed that the variability is quite lower compared to the case in which both parameters were stochastic. In this case, the trend of the curves is more defined than before and their magnitudes are in a smaller range.

As before, in the figure 11.6, curves of two different simulation runs are presented. In this case, the variability of these two curves is quite low, meaning that the number of agents is a more significant parameter to determine the systems' emergent behaviour.

11.1.4 Discussion

Summing up, the things that have been needed are the three metamodel classes, a script that is used by the Profiler to generate the scenarios based on the input statistical data and the behaviour of the agents. This behaviour only considers the information of the start and stop time to send a notification of entrance or exit to the shopping centre so that it can count the number of people inside. Concerning the results, the variability of this stochastic model has been evaluated determining that the number of agents is a more relevant parameter towards the emergent behaviour of the system.

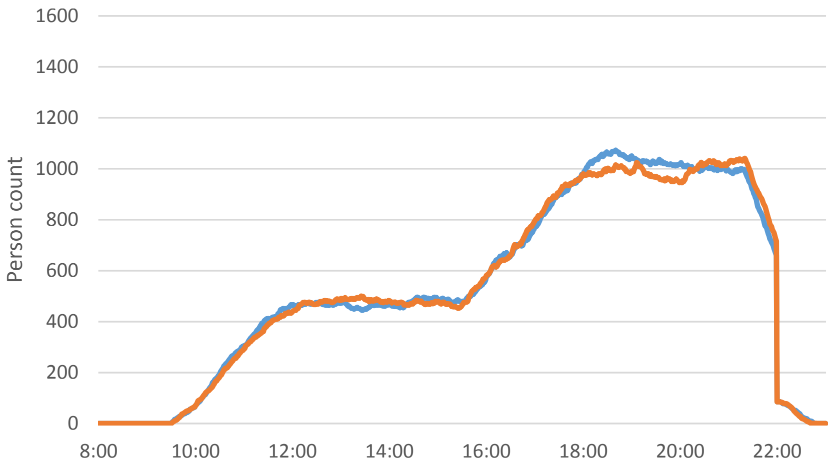


Figure 11.6: Curves of two simulations runs with fixed number of agents.

11.2 Experiments comparing different simulation timings

This case, published in [EHH13a] [1], is intended to show the advantages and disadvantages by comparing both synchronous and asynchronous implementations over a scenario that represents a power grid. This scenario describes the demand in a disaggregated manner which will enable the study of DSM policies and their effects can be analysed from the very bottom level of the grid. The composition of the experimental scene used to comparatively evaluate synchronous and asynchronous approaches is detailed in the list below:

- 1 environment which represents environmental conditions.
- 1,000 buildings.
- 20,000 households (20 per building).
- 20,000 agents (each one related to one household).
- 60,000 radiators (3 per household).
- 20,000 washing machines (1 per household).

These entities are coupled among them as follows (Figures 11.7 and 11.8):

[1] People that participated in this experiment: José Évora, José Juan Hernández and Mario Hernández.

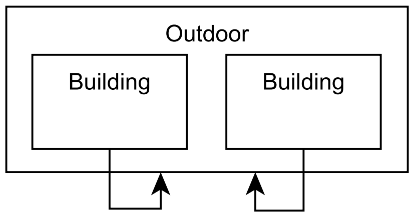


Figure 11.7: Coupling details at the outdoor-building level.

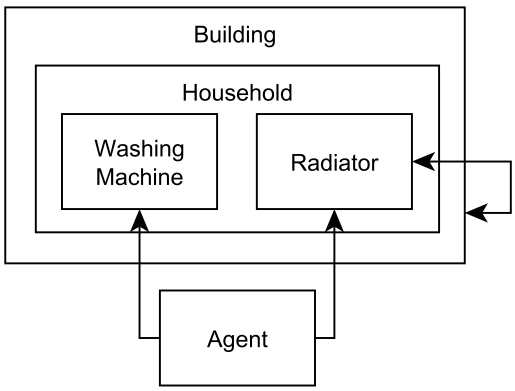


Figure 11.8: Coupling details at the building-household level.

- The environment has no coupling with other entities.
- Each building is informed of the external temperature by the Environment and the active power from the Radiators to calculate the thermal behaviour.
- Households play the role of a device container since they do not exchange information.
- Every Agent is related to a household so that they can switch on and off the devices inside.
- Radiators need to be informed of the internal temperature of the Building in which they are in order to calculate whether they have to heat it or not. Furthermore, they depend on the Agent to develop their time since they cannot be ahead.
- Washing Machines are bound by the same dependency as the Radiators are by the Agent.

Table 11.2: Timing of every device at each simulation

Behaviors/Step	**Synchronous**	**Asynchronous**
Outdoor	60 seconds	-
Building	60 seconds	60 seconds
Agent	60 seconds	Dynamic
Radiator	60 seconds	ON: 60 seconds / OFF: ∞
Washing machine	60 seconds	ON: cycles-based / OFF: ∞

11.2.1 **Timing**

Both synchronous and asynchronous simulations correspond to 24 hours. In the synchronous case, the simulation pace is fixed to the entity that requires the highest pace, which is the Building. Buildings need a higher pace due to the fact that the thermal calculation must be performed frequently. Since the asynchronous approach delegates the time management to each entity, the way in which they develop their time can vary. The table in 11.2 presents the timing configuration used at each simulation.

Moving on to the asynchronous case, the Environment does not have a pace since, in this case, it works as a database from which the external temperature can be extracted. Buildings have the same pace as the synchronous case. Agents have a dynamic pace which allows them to develop the time according to their decision making. Radiators have the same pace as the Buildings when they are on and the pace is not determined (considered as ∞) when off since they are inactive until the Agent interacts with them. Washing Machines behave the same way but when they are on the pace is based on the timing of their cycles.

11.2.2 **Results**

When considering the performance of an implementation, two important indicators are execution time and memory usage. Therefore, both synchronous and asynchronous implementations will be compared under both indicators. The table in 11.3 presents these indicators for each simulation averaged after ten runs.

There is a significant improvement when applying an asynchronous paradigm in the execution time to this experiment. However, there is a certain penalty in the

Table 11.3: Performance comparison between synchronous and asynchronous cases.

	Synchronous	**Asynchronous**	**Benchmark**
Execution time (seconds)	192	113	40.84% faster
Average memory usage (megabytes)	671	789	17.65% higher

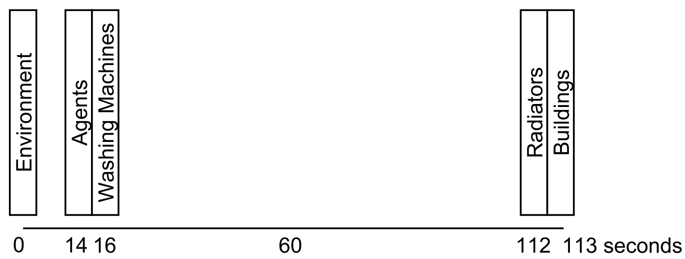


Figure 11.9: Times in which each entity kind finished.

memory as a result of the use of snapshots. It is also interesting to see the times at which every element finished the simulation. In the synchronous case, it is evident that all of them finished at the same time. Nevertheless, in the asynchronous one (Figure 11.9), this does not happen since every entity develops its time according to its pace.

Since the Environment does not have pace, it finishes at the very beginning. Agents finish when they stop sending the events. This makes it possible for Radiators and Washing Machines to develop their times. Seconds later, Washing Machines can finish. However, this is not the case for Radiators since they are coupled with the Buildings. Radiators and Buildings finish at the same time thus bringing the simulation to an end.

CHAPTER

12

Agent-Based Modelling of Electrical Load at Household Level

This experiment, published in [EKM+11] [1] is oriented to model the residential sector of Power Grids. To this end, electrical grids are analysed from the point of view of how to simulate them following an ABM. Therefore, this residential sector model is addressed using a bottom-up approach. This is, all appliances of the residential sector are simulated (bottom) and then their consumptions are aggregated (up) for validation purposes. It must be noted that this validation is made on the aggregated level since real power consumption data is not usually available at lower levels.

[1] People that participated in this experiment: José Évora, Enrique Kremers, Susana Morales, Mario Hernández, José Juan Hernández and Pablo Viejo.

12.1 Description

12.1.1 Simulation of electrical systems

Like other complex systems, electrical system simulations involve the behaviour execution of the many elements that exists in a real grid.

Since the electrical system is locally and massively affected by the human interventions, it is also required to model this behaviour along with other entities that belong to the electrical system architecture. Representing the human behaviour means that it is required to define what people do during a time period, which electrical devices they have and how they use them.

Until now, the human behaviour has not been analysed and electrical companies only have historic aggregated data of the consumption at global level. From now, it is quite important to have data at the household level to be able to model households and simulate the behaviour of people inside. So, the main challenge in modelling electrical systems is to describe all the electrical appliances in the households and the behaviour profiles whose actions generate electrical consumption.

12.1.2 Programming behaviours

To simulate the electrical system, all Tafat metamodel classes have to include behaviours that must be programmed. A separation between the different types of behaviours is done according to their nature which may be environmental, mechanistic or social.

a) Environmental behaviour. Environmental behaviours represent the change over time of some environmental variables. Environmental variables are normally common to a group of entities or devices and describe the surroundings or "outdoor". These can be for example solar radiation models, which represent the insolation and can be used for calculation of the thermal gains of a building, or further for energy production (photovoltaics, solarthermal use, etc.). These behaviours do not directly change the attributes of a device or agents, but rather allow some interactions in an indirect way (e.g. through heat exchanges, etc.).

b) Device behaviour. Most of the electrical devices used in a household are major appliances such as washing machines, refrigerators, etc. There are also some other, smaller appliances, such as CD players, TVs, HiFi Audio equipment, etc.

Usually, the major appliances cause a larger part of the electrical consumption. In order to recreate the individual load curves, EIFER [eif] (a research institute by KIT [kit] and EDF [edfa]) has developed individual models for the behaviours of electrical appliances, which were integrated into Tafat. Simplified technical models are used, which take into consideration different technical parameters of a specific appliance. So, for example, the load curve a TV will be characterised by the size and technology (CRT, LCD, Plasma, etc.). Major appliances also are modelled using the EU Energy Label as an input parameter, which is an indicator for the energy consumption of a device and is compulsory for appliance sold in the EU.

Different releases for the behaviours of the electrical devices were created. Using this modular approach, a behaviour of a single device can be *exchanged* in a simple way. The different releases include simplified technical models with varying degrees of accuracy, thus allowing for an optimisation of execution time vs. accuracy of the model. In Figure 12.1 an example of the load curve generated by the behaviour of a washing machine can be seen. This load curve is created by a simplified technical model of this appliance.

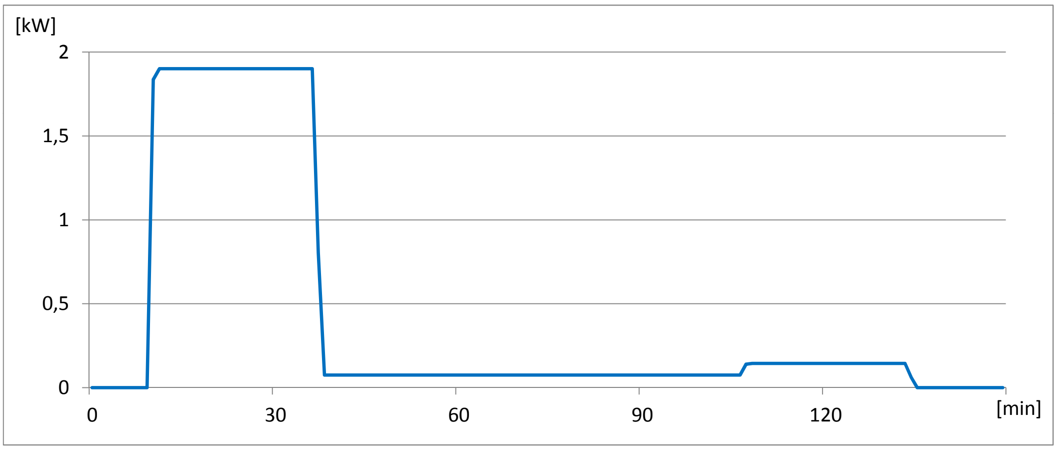


Figure 12.1: Simulated load curve of a washing machine.

c) Social behaviour. The architecture must be flexible to allow a range of behaviours from simple behaviour based on a list of tasks to a complex behaviour implemented as a neural network.

The mission maker is the intention launcher, the decision maker is in charge of choosing a recipe to accomplish the mission launched and the action maker is the executor of the recipe (Figure: 12.2). The recipe is a list of actions that executes to

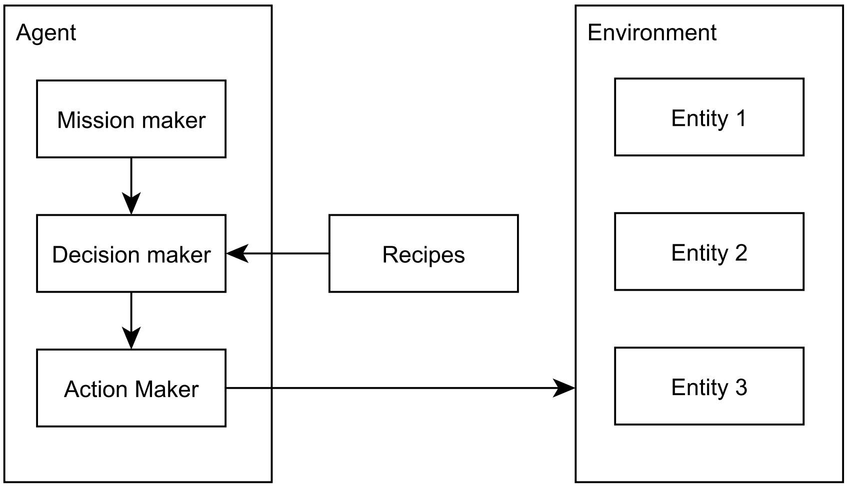


Figure 12.2: Agent architecture.

accomplish a mission. With this architecture, a simple behaviour can be developed by creating a big recipe in which all the tasks are described and having a mission and decision maker very simple. Otherwise, in a complex social behaviour, the task can be launched by the mission maker according to several parameters of its own agent or the environment, having a hard process to choose a recipe in the decision maker, but easier recipes that only describe how to arrange a task as, for example eat.

12.1.3 Simulation model

The main problem for developing good models that represents accurately a place (town) is the lack of data. Often, it is quite difficult to gather the needed data.

Ideally, models should be built using real data, since the simulation will help to understand what happens in the electrical system. However, since data is not available at all, a model approximation is done. The Profiler, the Tafat tool, helps to carry out this task. Using a high-level description of a place (for example, amount of buildings and population), the Profiler automatically generates an electrical system model that can be simulated directly.

Electrical models can be created to represent the load accurately but lightweight enough for use in large scale simulations, and handle DSM mechanisms through the use of an agent-based approach.

12.2 Case Study

The case study proposed is an analysis of the electrical load of households according to social group characteristics. The following five different social groups have been taken into account to develop the case study:

1. junior single,
2. senior single,
3. junior couple,
4. senior couple and
5. family with children.

These groups were taken as a part which represents about 70% of the population of Germany, being the most numerous groups identified by the German Socio-Economic Panel Study (SOEP). The constitution of this sample is shown in Figure 12.3.

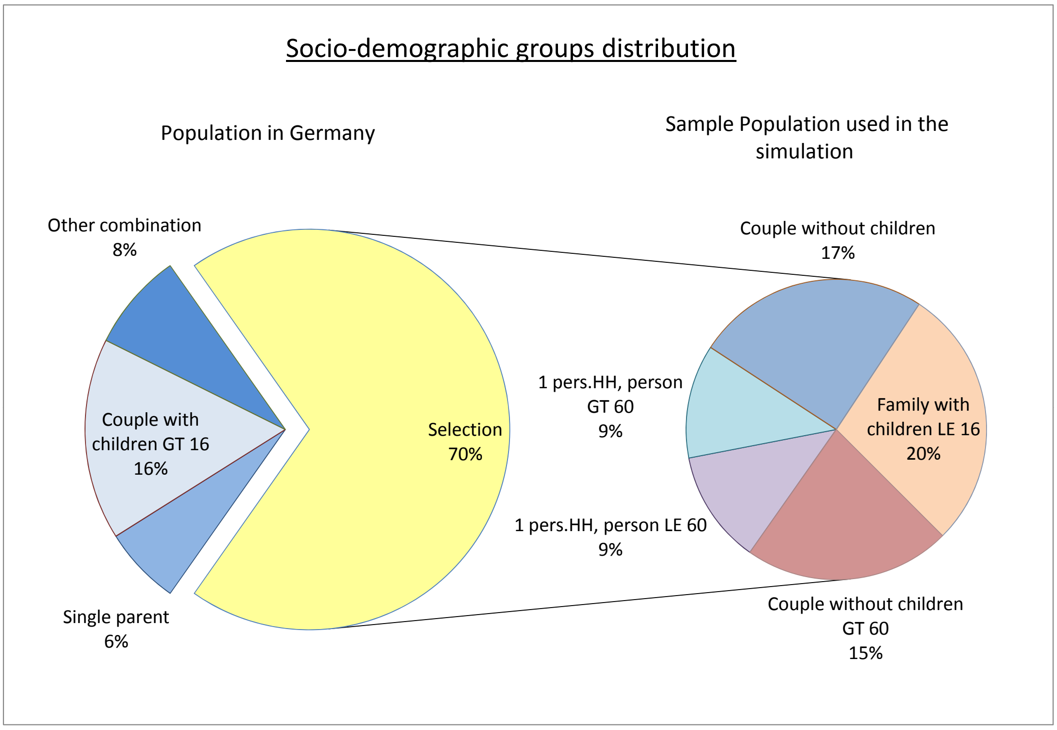


Figure 12.3: Socio-demographic groups used in the case study.

A survey [1] has provided the information about the timetables and appliance usage of a household according to the social group. Based on this information the agent behaviour which is related to the household is implemented. The electrical usage behaviour data employed for this study was obtained through a local survey. Thus the gathered data from a small sampling was used as input parameter in the Tafat model. Hence, 20 different social behaviours (i.e. 20 different model recipes) have been used to simulate the five socio-demographic groups. Each of these recipes includes some randomness through the definition of intervals at which devices are switched on or off. The exact switching time is obtained as a uniformly distributed random variable over this interval.

To arrange this study, the Profiler has been used. This utility uses statistical data from the SOEP and the survey to create a model scene in which the social groups are distributed in the households. Those households contain the electrical devices, and their amount, according to the social group.

12.3 Results

In the simulation, the device load curves within a household are generated. The curves were aggregated using an individual behaviour for each household taking into account some randomness (variation of the duration and use time in a defined range of the different electrical devices).

The simulation results show the load curve of a day of 1000 households with around 12 appliances each, composing thus an amount of approximately 12000 simulation elements. This type of simulation can provide the relevance of a determined power consumer in the household consumption as well as the influence of a specific type of consumer on the global load. The number of 1000 entities was chosen, as it was determined that larger amounts did not change substantially the results and only increased the execution time of the simulation. This is probably due to the use of a limited number of recipes (taken from the number of surveyed persons). In this case, the execution time was around 20 minutes on a standard

[1] The survey consisted in local interviews with around 20 persons in Karlsruhe, Germany in order to obtain data like usage times and durations of specific socio-demographic groups. It has to be noted that the survey is not representative but rather a sample of the user behaviours of those groups.

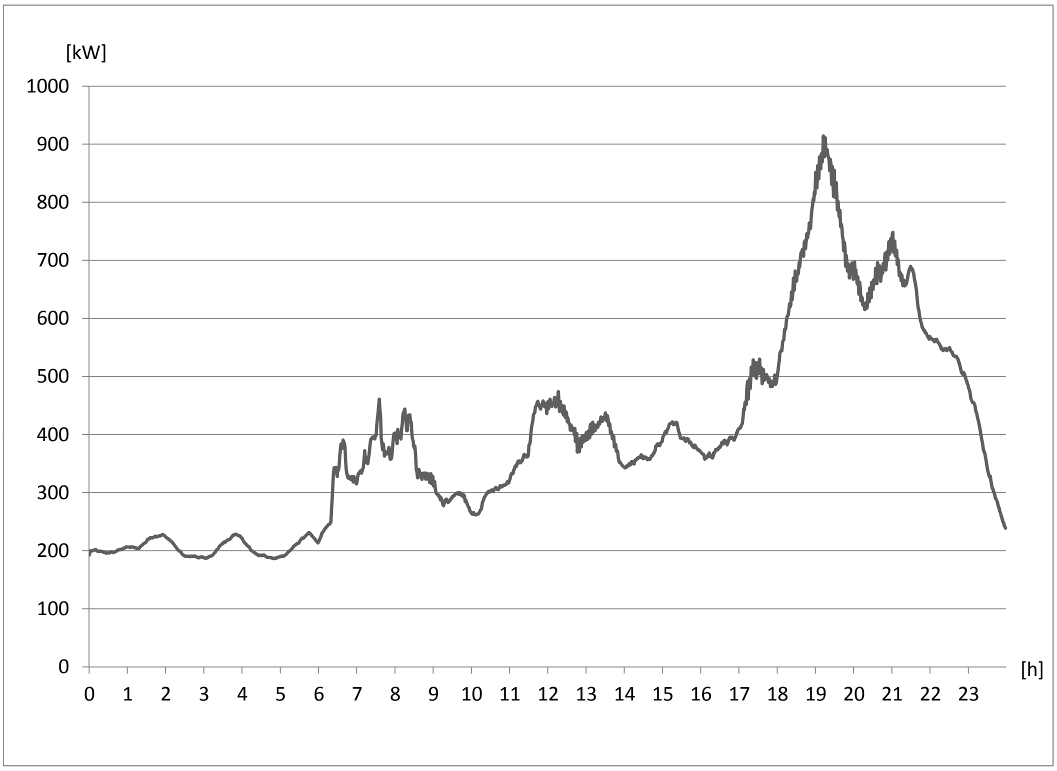


Figure 12.4: Simulated load curve of 1000 households for one day.

desktop PC for the simulation of a period of 24 hours.

The 1000 household sample is composed of a distribution of the different social groups according to real statistical data, in order to obtain a sample of households that is as close as possible to reality. In Figure 12.4, the simulated load curve for one day can be seen. The simulation is run in a high time resolution, being the time step one minute. This allows observing effects which are neglected in simulations at lower resolutions, provided by 15 minute or hourly models. Some sharp peaks can be observed, which are caused by the use of high power consuming devices in the household. A general trend to use more power during the day is clearly visible. During the night, the base load (devices that are constantly running, such as refrigerators and other permanent loads) cause a consumption that is only around a third part of the daily peak load. The configuration of the simulation can be seen in Figure 12.5.

In Figure 12.6 the curve is compared to a standard load profile of Germany for a winter weekday (according to [Bun11] and made public by [RWE11]) for

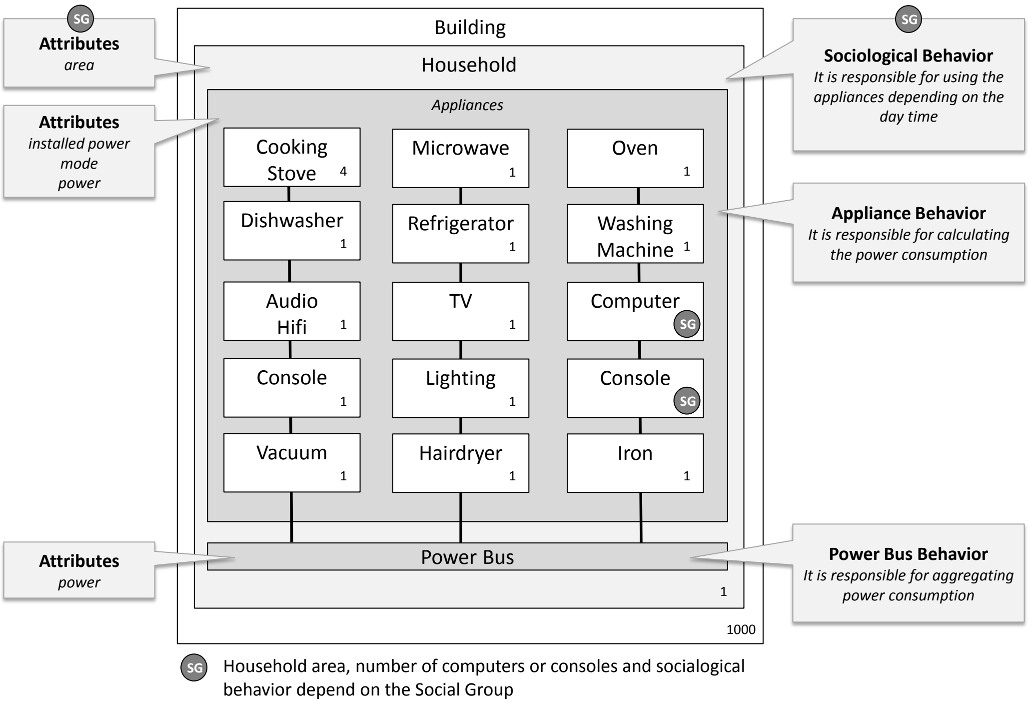


Figure 12.5: Profiler pattern. Social Group (SG) selected based on a frequency distribution.

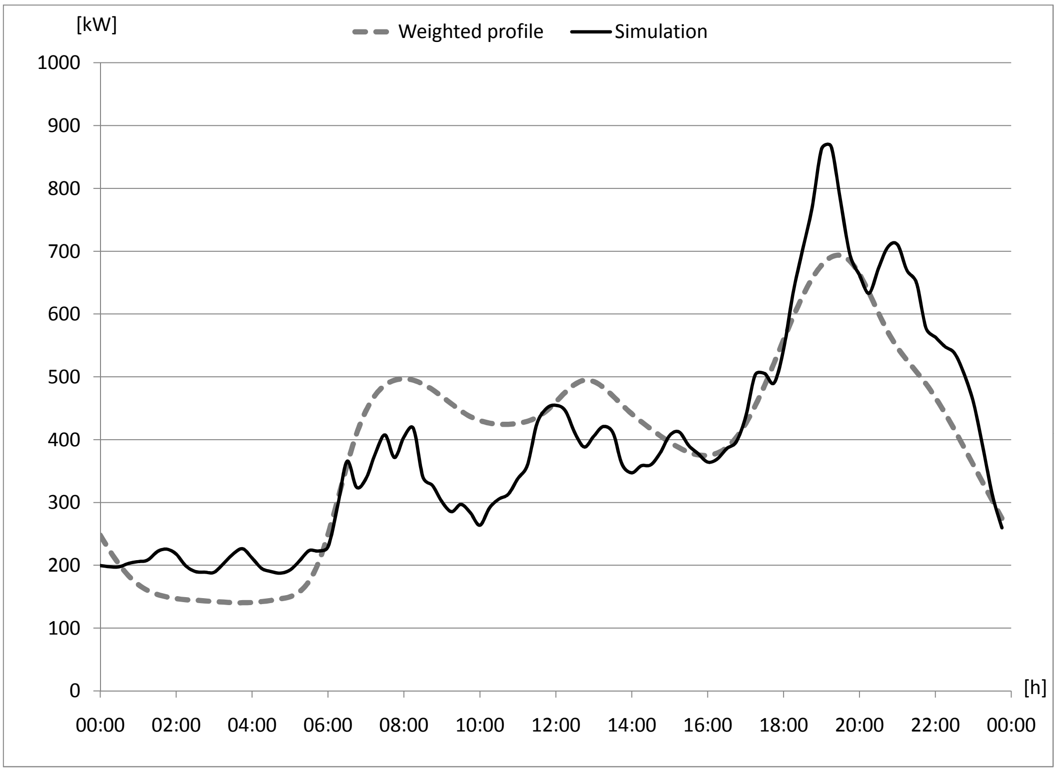


Figure 12.6: Simulated load curve vs. standardised residential load profile.

household demand. The profile is provided in a normalised form in order to be weighted with a given number of energy. In this case, and for comparison purposes, the profile was weighted with the same amount of energy as the simulation curve, i.e. that the daily consumed energy [kWh] is the same in both cases. The load profile is a smooth curve, representing an aggregated load behaviour at high levels of the electricity system for large number of consumers. They are the result of a statistical analysis based on representative samples from different consumer groups [PKZK08].

The simulated curve represents the total power consumed by a sample of 1000 households, modelled individually and with an autonomous behaviour for each of them. The curve has been adapted by averaging periods of 15 minutes in order to match the same time granularity as given by the standard load profile. The curve is more peak shaped, which is probably due to the relatively small amount of households (in comparison to the statistical samples taken to obtain a profile, which are representative) and the reduced number of behaviours (in total, only 20 different behaviours have been used). Furthermore, only 70% of the household population is modelled, neglecting other social groups which may change the curve.

Even though the selected samples in the survey are not representative for all Germany nor society as a whole (only five social groups were used), the general trend of both curves are similar. Three peaks can be observed, which are closely synchronised in time and correspond to morning, noon and evening peaks. These peaks are correlated with a large and concurrent usage of high power devices, such as cooking plates, ovens, microwaves, etc. due to alimentation habits, as well as lighting use in the evening hours. The morning and noon peaks are lower in the simulation than in the profile, whereas the evening peak is higher. The timely synchronisation of the ramps of the peaks matches quite well; this indicates that the activities (having breakfast, lunch, returning home, etc.) were modelled according to the average German user behaviour. Even though, some differences can be observed at the evening drop (21-23h), as well as a small second peak that cannot be found in the profile.

Even using a relatively small sample of households and reduced number of behaviours, a curve that represents the major characteristics (peaks and troughs, as well as their timely synchronisation) is generated. Concerning the differences in the height of the peaks, the model could be reviewed in order to check the individual power curves of each electrical device. This seems to be quite hard as almost no data is available for such a validation (at a representative sample). However, the differences could also be related to the use of only 70% of the socio-demographic population share. Further, behaviour itself is another factor to consider, as a largely simplified and almost static model was used.

Some specific characteristics of the model create peaks, which could be explained by a rather homogeneous behaviour of the groups. For example the second evening peak (around 21:30h) is possibly caused by some activities (watching TV, use of appliances, other evening activities) which start and end at similar times, because of the relatively small sample.

12.4 Discussion

The case study demonstrates that a bottom-up simulation of the residential consumption using an agent-based approach has been successful since the result curves show similarities to aggregated load curves. A comparison with a national aggregated profile shows similarities in the main characteristics of the curve.

Apart from the discussed results, another relevant result of the work is the integration of many heterogeneous models. Different types of appliances have been integrated in a simulation which contained a total amount of 40 thousand elements. Furthermore, sociological behaviours have also been integrated in the simulation in order to represent the way in which people living in households consume energy. This is, human intervention has been represented in a simulation by using information gathered on consumption habits through a survey.

From now, the simulation that has been developed will allow us to experiment and study new algorithms and strategies for electrical system management. New scenarios, new problems and new challenges will arise in the near future with the introduction of RES and a distributed production in the electrical system. Simulation is particularly necessary to design a new management approach.

The simulation will allow us with further work to study the integration of DSM strategies. Strategies, such as adaptive or reactive technologies, incentives or campaigns, can be addressed for studying their impact at load curve level.

Social behaviours need to be validated and improved. This validation is necessary for studying the emergent behaviours and for identifying the local actions of agents which provokes the desired emergent behaviour. Moreover, social behaviours must be improved in order to introduce a higher degree of heterogeneity to the models. Due to the high resolution of the household model, individual actions such as changing specific parameters on a device can be performed.

Furthermore, the model developed could be expanded in order to simulate not only the demand side, by including distributed generation or other injections to the grid (like storage), which could interact with the existing elements. This is already contemplated within Tafat, and would allow a disaggregated analysis of generation-demand balance and the possibility to estimate the impact of those measures at a system level.

CHAPTER

13

Modelling lifestyle aspects influencing the residential load-curve

This experiment, published in [HEK12] [1], is partially supported by the previous one. In this one, the focus is not only made in the disaggregation of the electrical system until the device level, but also in using a more accurate sociological model. This sociological model, as in the previous experiment, is devoted to switch on/off appliances within households.

To develop this sociological model, a big survey has been performed which allowed the extraction of information concerning consumption habits in the residential sector. Based on this information, a behaviour for Tafat that emulates people living in a household has been modelled.

[1]People that participated in this experiment: Wolfgang Hauser, José Évora and Enrique Kremers.

13.1 Description

13.1.1 Residential consumption

Residential consumption of electricity is influenced by a multitude of variables and shows big variance between households, even within the same society and geographic region: [LB08] report differences of up to factor 40 between the measured electricity demands of 1 627 households in a Northern Californian sample [MH11]. Furthermore, electrical consumption for single household tasks vary greatly between households: [AEE08] shows that electricity used for cooling devices differs by factor 10 between different households, the same applies to electricity used per person for dish-washers.

The rise of decentralized power supply raises the need for electricity demand forecasts of smaller areas. Simulations of household electricity demand are mostly based on mean values of the whole population (e.g. [PL06]); for specific areas of interest this approach results in an ecological fallacy, because different kinds of households are not equally distributed in space: In Stuttgart the average number of persons per household differs from 1.56 to 2.18 for different city quarters. To reduce the ecological fallacy it is necessary to identify determinants of residential electricity consumption, which can be linked to geographic data or building types, e.g. the number of persons living in a household or lifestyle typologies.

For the planning of power grids, not only the overall quantity of electricity consumed is of importance, but it is also important to know at what time of day the electricity is demanded: it is the load curve that matters (a load curve visualizes the use of electrical energy over time, showing watts on the y-axis and time on the x-axis). Nevertheless, measured data about electricity demand on a household level is very hard to find, especially when looking for a random sample. To simulate load curves for different types of households, we connected a simulation converting weekly or daily probabilities of energy relevant household tasks, into start times of events with simulations of appliances' load curves.

13.1.2 Methodology

Based on the lifestyle typology developed by [Ott05] usage rates and probabilities for the possession of different electrical appliances are compared between different groups of society, using data collected in a postal survey. In order to evaluate household energy consumption, the survey focused specifically on daily routines and appliance ownership, because direct data collection of household energy consumption through surveys faces a big problem regarding missing values (25-60% in most studies), which can not be expected to be distributed randomly. Usage rates, habits and probabilities for the possession of household equipment are transformed into household-agents with specific daily routines and specific household equipment, then the resulting load curves for each household are simulated in an agent based model and aggregated in different ways.

13.1.3 Data

We get information about weekly and daily usage rates for different appliances from a postal survey conducted in the run of the European Centers and Laboratories for Energy Efficiency Research (ECLEER) Ph.D program. In Germany, 4 000 Stuttgart based households had been asked to fill out a questionnaire about household usage rates, socio-demographics and lifestyle issues; 769 filled out questionnaires have been sent back, equalling a return rate of 19.2 %. Results of the survey can be compared with a survey employing the same typology, conducted in 2008 by the *Statistisches Amt* of the city of Stuttgart, with a bigger sample and a better return rate [Sch10]. Our results are more or less in line with this survey, with the exception of a higher age average resulting in smaller proportions of modern lifestyles. The biggest differences are with regard to the proportion of *Conventionalists* and *Hedonists*. It seems plausible that younger people, spending less time at home, were less willing to take the time to answer our questionnaire, as it was substantially longer than the one used by [Sch10], which also explains the lower overall return rate. A comparison of the results of both surveys is presented in Table 13.1.

13.2 Case study

After the survey analysis, the information regarding the appliances usage for every social group is available. Now, the challenge is to build up a simulation where this information is integrated in order to calibrate an agent-based model. However, this task concerns external aspects which provide the agents environment. The envir-

Table 13.1: Otte lifestyle groups in Stuttgart

Standard of consumption			
high	**Conservative well-off**	**Liberal well-off**	**Reflexives**
	4.82 (3) %	16.31 (15) %	4.68 (10) %
	Ø 66 (62) years	Ø 55 (50)years	Ø 45 (39) years
	Ø 2.12 pers.	Ø 2.46 pers.	Ø 2.28 pers.
medium	**Conventionalists**	**Success seekers**	**Hedonists**
	14.61(7) %	30.21 (27) %	6.52 (14) %
	Ø 64 (65) years	Ø 52 (48) years	Ø 42 (36) years
	Ø 2.21 pers.	Ø 2.43 pers.	Ø 2.09 pers.
low	**Traditional workers**	**Home-centered**	**Entertainment seekers**
	8.09 (7) %	12.20 (14) %	2.55 (5) %
	Ø 63 (65) years	Ø 50 (46) years	Ø 37 (33) years
	Ø 2.22 pers.	Ø 2.33 pers.	Ø 2.22 pers.
Modernity	low	medium	high

Results of the survey by [Sch10] shown in parentheses for comparison

onment is composed of two sides: the global and the local. The global regards the agent location within the world. This information arises when the agent is related to a concrete household of the scene and this is, in turn, related to a building. The first relation is due to a link and the second one is due to a contain relation. Finally, the building indicates the exact coordinates to locate itself in the world. On the other hand, the local environment consists of the appliances within the household as they are the elements with which the agent interacts.

The simulation model describes agents, appliances, households and buildings. As described in [EKM+11], the agents can be considered as intentional models and the appliances as design models. The agent model is described in section Household Model and represents parts of the behavior of a household, which have a big impact on consumption of electrical energy. The local environment of the agent is composed of many kinds of appliances which are switched by the agent. The task to develop the local environment consists of modelling the many appliances that the agent uses. The appliances models consists in the description of how they work. For instance, the washing machine model works producing a non-consumption when switched off. When the agent turns the washing machine on, a three-cycle working mode starts up producing a non-zero consumption.

This model is represented using a bottom-up approach where the electrical grid consumption comes from the agents acts since they generate the appliances consumptions at switching them. The tool that we have used to model this experiment is Tafat. This framework was used because, among other reasons, it allowed us developing the model fast since the appliances models reuse was possible. Another reason for choosing this framework was the performance as the intention is to model a huge scene where thousands of agents are running at the same time with their respective local and global environments.

13.2.1 Household Model

Each household is represented as an agent, having control over his electrical appliances; the probability to own a certain kind of appliance is derived from the distributions in the survey data and differs between the lifestyle groups (see section Survey Results). Each instance of a household draws randomly from the respective distribution, in order to determine if he owns a tumbler, washing machine, etc and what kind of cooling devices are to be found. In the same fashion, each household is assigned rates of using these appliances, as well as times of inactivity (sleep hours) and absence. Probabilities of preparing a warm lunch or dinner and the time when these take place are also taken from the survey and differ between groups. Of course, the agent can not perform the household tasks in his time of absence or sleep. It is, however, able to start multiple devices at the same time, which will run for a predefined cycle (washing machine) or for a time that is, again, drawn from the distribution of the lifestyle group he belongs to. Lights are turned on automatically between 18:00 and 7:00 if the agent is not absent or sleeping; however, we added a normally distributed error component to the start and stop time, in order to prevent an artificial peak to the aggregated load-curve. Cooling devices are running on a regular pulse. To generate a load-curve that averages the behavior of the households of interest and is robust towards random variation, 1 000 households are set up for each simulation run. The agent behavior of the households is connected to its environment through Tafat, controlling the start and stop times for the appliance models stored therein.

13.3 Results

13.3.1 Survey Results

Regarding the ownership of electrical appliances and daily routines, there are significant differences between lifestyle groups. *Conservative well-off* and *Entertainment seekers* appear to be specially suited for a direct comparison, as they are on the opposing ends of the lifestyle dimensions and show the biggest differences in appliance possession and usage rates. For this reason, these groups have been chosen for the simulation runs.

In regard to the number of cooling devices to be found in households, *Conservative well-off* own on average 1.9 devices, meaning that almost every household owns two devices, while only one in three *Entertainment seekers* households has a second cooling device. A Welch t-test between the group shows that this difference between the two groups is significant on the 99.9%-level. Furthermore, we find significant differences regarding the percentage of household owning a of dish-washer (85 % vs. 39 %) or a tumble dryer (53 % vs. 11 %). Regarding daily routines, the two groups differ significantly in the number of times they are cooking per week (8.0 vs. 5.5) and in regard to their absence from home (see Figure 13.1).

13.3.2 Simulation Runs

Data about measured load-curves of different household is very sparse and not publicly available. This was one of the motivations for our approach. Unfortunately, this prevents a good verification of our results with measured data. Therefore, we compare the simulated load-curves (cumulatedSIM) with a standard household load-profile scaled to the electrical energy consumption of the simulation (weightedSLP). To evaluate the effect of the different lifestyles on the residential load-curve, we present results of three simulation runs: one with the share of the different lifestyle groups as found in the survey, one where all the households are set to the behavior and appliance ownership of the group of the *conservative*

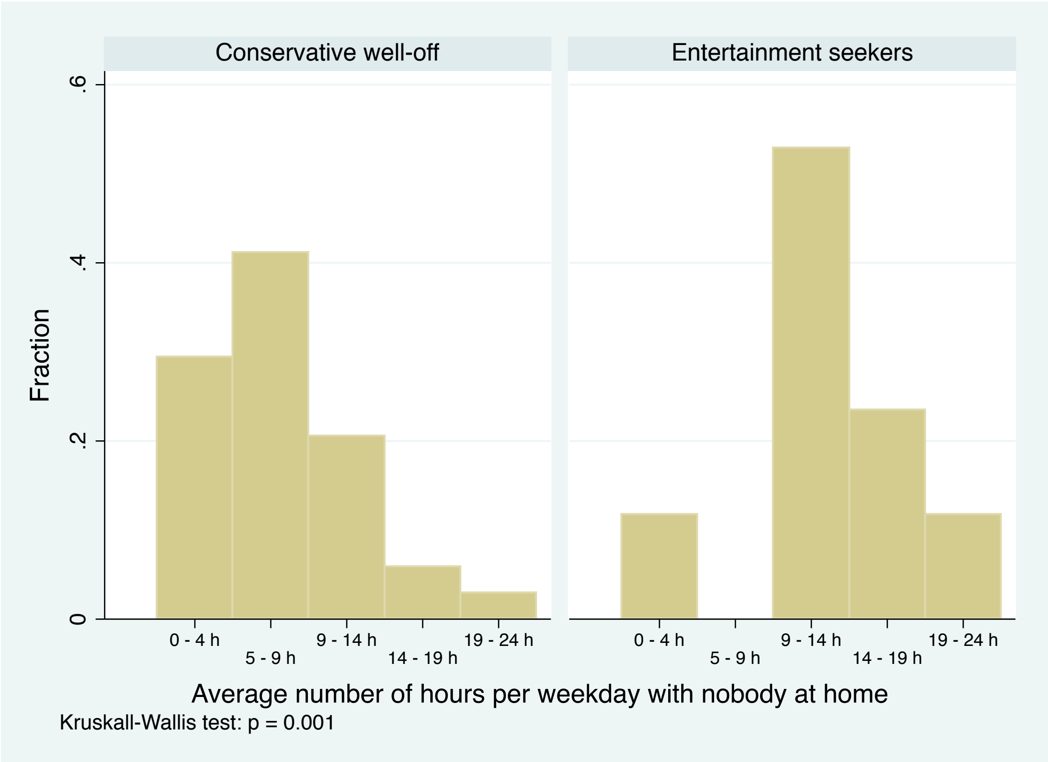


Figure 13.1: Absence from home on weekdays.

well-off and a last one with all households set to values found for the group of the *entertainment-seekers*.

With lifestyles represented in the simulation with the same shares as in our sample (Figure 13.2), we see that the noon and evening peaks are more accentuated than in the standard household load profile (SLP). One reason for this seems to be, that we model only a relatively small number of appliances by now and therefore the effect of stoves and ovens could be overestimated. Another reason - which applies especially to the evening peak - might be that employment figures in Stuttgart are very high, so that more consumption than on average in Germany is shifted towards the evening hours. A third reason might be, that we underestimated the variation in lunch and dinner preparation and should introduce a bigger error component to this decision. Another difference in comparison to the SLP is that the simulated load curve does not decline as deep in the night hours and rises less sharply in the morning. With the small number of devices which can be

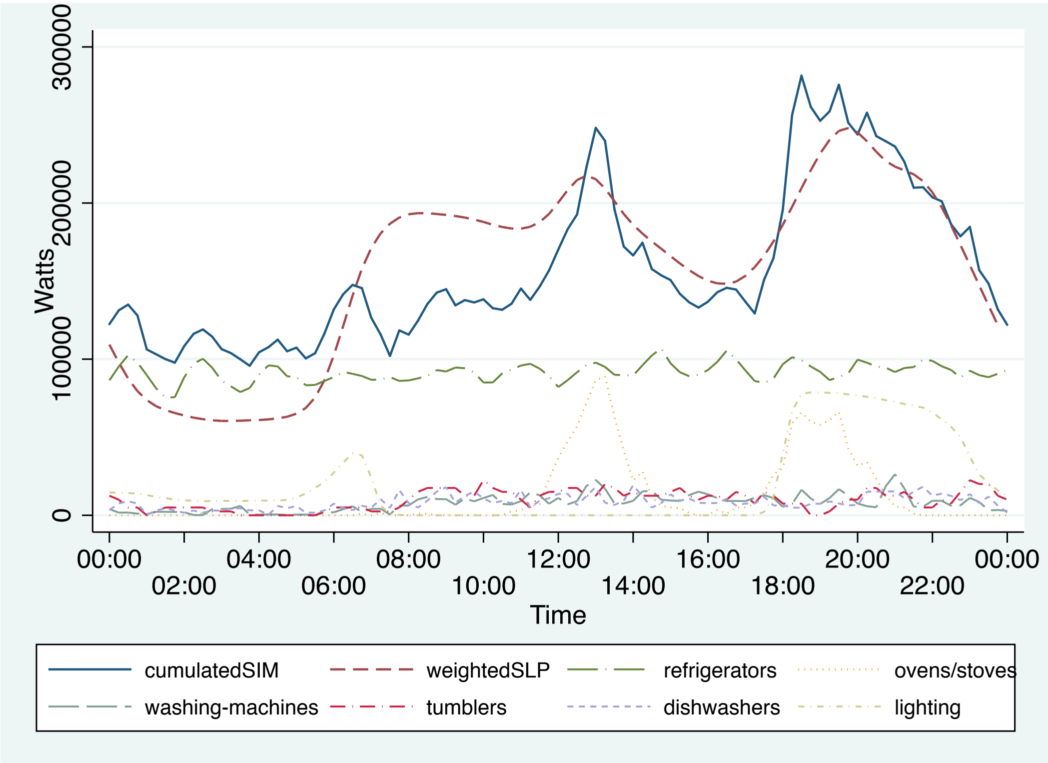


Figure 13.2: Simulated load-curve for 1 000 households with lifestyle groups distributed according to their share in the survey.

simulated at the moment, cooling devices - which produce a continuous pulse - make up a big part of the electricity consumption in our simulation; even though refrigerators and freezers have a significant impact on the electricity consumption of a household, their smoothing effect on the simulated load curve might be exaggerated here. However, our aim was to present the opportunities of a simulation model based solely on survey data and theoretical considerations, before extensive calibration and the simulated load curve shows a clear resemblance towards the SPL.

When we compare the load-curve of the *conservative well-off* (Figure 13.3) with the standard load profile or with the load curves of the other simulation runs, the most striking difference is the very pronounced peak at the middle of the day. In this group we find more elderly people than in the other groups, which explains why they spend more time at home, giving them the opportunity to consume more

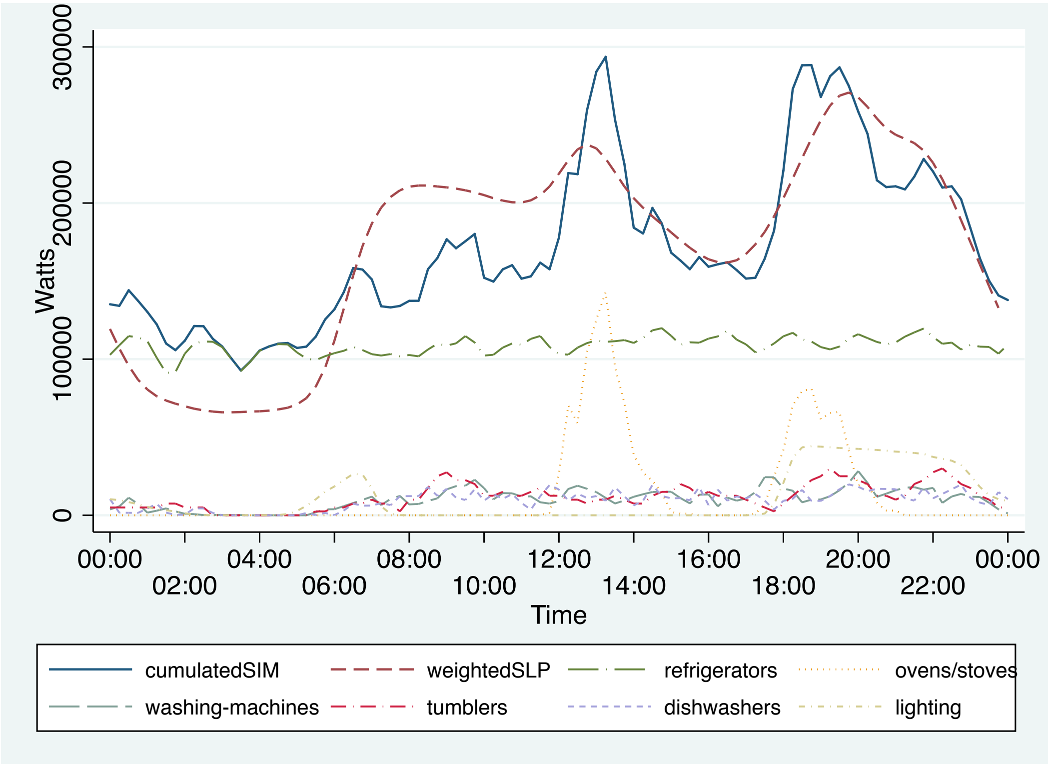


Figure 13.3: Simulated load-curve for 1 000 households with the profile of the *conservative well-off*.

electricity during work hours. The share of people having a warm lunch at home is bigger than in the other groups which raises the peak at the middle of the day to a higher level than in any other group. At the same time, the percentage of people having a cold meal for dinner is rather high in this group, which explains why the evening peak is relatively small here. Overall their load-curve is on a higher level than the others, because they live in more spacious apartments/houses and own more electrical appliances than i.e. the entertainment-seekers (see section Survey Results). Their ownership of many cooling devices and the habit of having a warm meal for lunch reported clearly shows in the simulated load curve.

The load-curve of the *entertainment-seekers* (Figure 13.4) is characterized by the absence of the peak in the middle of the day and by a lower level than the other load-curves. Households belonging to this group are rather young and there are less retired persons to be found in this group than amongst the conservative

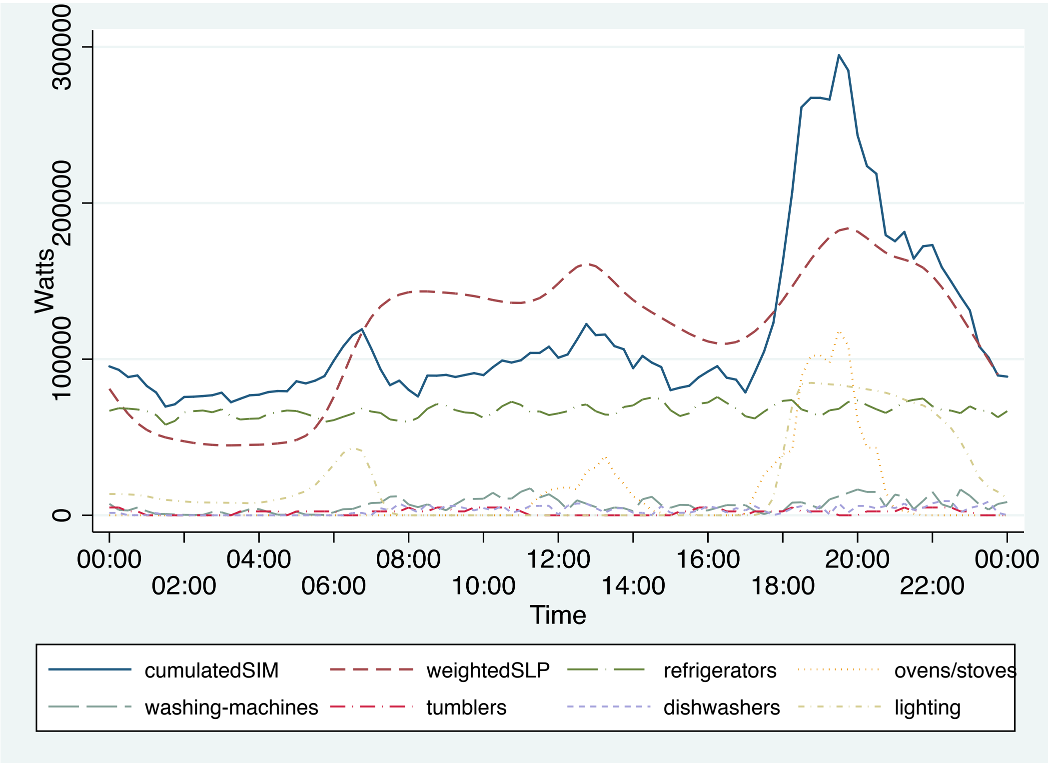


Figure 13.4: Simulated load-curve for 1 000 households with the profile of the *entertainment seekers*.

well-off, which is one reason why they spend more time out of home than other groups. Their consumption is therefore shifted to the evening hours, where they create a larger peak than the other groups. They have relatively few household appliances like dishwashers, tumble dryers and washing machines which decreases their energy consumption significantly.

13.4 Discussion

The simulation shows that the different lifestyles clearly produce different load-curves, which correspond to their behavior and ownership of appliances, even though not all the survey results and all appliances to be found in a household could be modelled by now. We expect the differences between the lifestyle groups to rise with the progressing inclusion of further appliances, e.g. water heaters and further entertainment electronics. As lifestyles are not equally distributed in space, this fact is important when planning power grids, especially when thinking about decentralized power supply. The smaller the area which is supplied by a single power source, the more important these differences become. The ecological fallacy of assuming a standard average load-profile for all households inside a specific region could lead to a wrong estimation not only of the level of electricity demand, but also of the shape of the load-curve.

CHAPTER

14

A Multi-Objective Particle Swarm Optimisation method for Direct Load Control in Smart Grid

This experiment, submitted to "Natural Computing" [1] on January 2014, is aimed at evaluating a DLC policy to lead the device control regarding the grid state which is represented by the differences between offer and demand throughout the day. DLC is implemented through a hierarchical structure where the uppermost control unit receives the grid state managed by the system operator. It receives the power demand and a reference power at each instant t. This reference power is calculated by a system which makes decisions based on the production capabilities. This reference is transmitted among the hierarchy arriving to the control units that are in the lowest level.

At this level, an optimisation algorithm based on swarming intelligence is used to convert the reference into specific orders which are sent to the related devices

[1] People that participated in this experiment: José Évora, Mario Hernández and José Juan Hernández.

accomplishing, thereby, the power constraint. These devices are within a residential sector which have been modelled as in previous experiments.

14.1 Description

14.1.1 Proposed method

In this work, MOPSO is applied to reduce loads in a Power Grid when power constraints must be satisfied. A specific requirement of this problem is that the method must provide results in fractions of a second. Otherwise the grid stability would be affected.

From Kennedy and Eberhart [KKE01], the computational complexity of PSO is in the order of $O(mn)$, where m be the number of initial particles and n the number of iterations to reach the global optimum. Particles are MOPSO mechanisms that are used to explore the search space. So, the larger the grid is, the larger the search space is, and therefore, more particles are needed to explore this space involving a longer execution time.

The proposed method addresses the MOPSO execution time requirement on power grids. The idea consists in dividing the power grid into small clusters (e.g. neighbourhoods) and executing in parallel MOPSO algorithms at each one. In this way, the execution time can be adjusted to accomplish this requirement.

Each cluster is managed by a load control unit where the MOPSO is executed. All control units are hierarchically managed by dispatching units (Figure: 14.1). These dispatching units are responsible for conveying power constraints to control units. In this way, a global optimal solution emerges from local solutions that are accomplishing local power constraints at each cluster.

Supported by this structure, a service oriented system has been developed to accomplish power constraints. In this system, the uppermost dispatching unit (named "Commander"), that is "on-line" connected to the system operator, receives power constraints that should be accomplished. Then, this unit balances these constraints among the intermediate dispatching units (named "Intermediary"). These intermediary units receive the constraints defined by the Commander, and in the same way send new constraints to the control units. These control units are located at the low level voltage transformers and are connected to neighbourhoods. At this level, con-

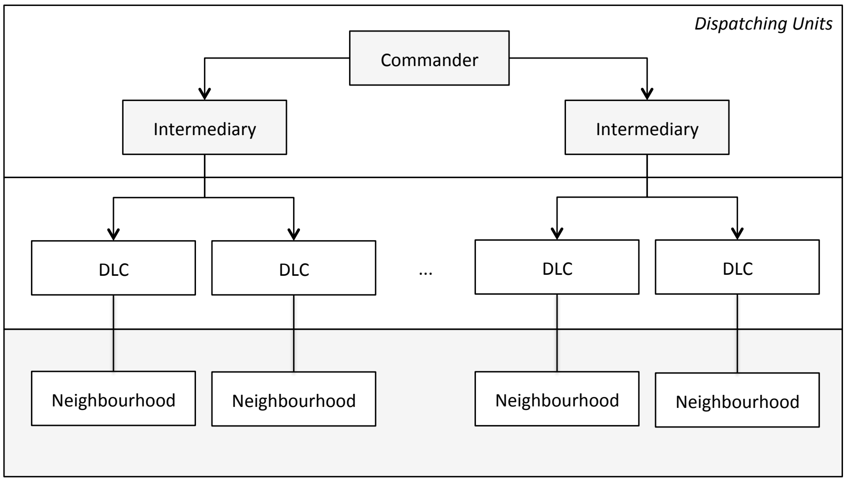


Figure 14.1: DLC structure.

straints are accomplished by acting on controlled devices using a MOPSO based algorithm. The algorithm is applied to optimise the consumption reduction acting on appliances state (e.g. ON/OFF).

This structure can be scaled according to the size of the grid. So, this distributed structure is not only addressing the execution time concerns, but also the grid scale concerns.

This method is continuously executed while there are power constraints from the system operator. This means that the optimal solution is recalculated considering the previous one.

From a computational point of view this is an interesting property since recalculations are faster. However, if a power constraint lasts very long a specific refrigerator could be switched off during all this period. Obviously, this must be avoided. So, the method must preserve the quality of service and guarantee that a device is not switched off for long periods of time.

This section is structured as follows. First, how power constraints are dispatched by the Commander and Intermediary Units is explained. Second, how these constraints are satisfied by the Control Units applying the MOPSO algorithm.

14.1.2 Power constraints dispatching

In this work, power constraints dispatching is made according to the power consumption that units at lower levels are monitoring. The underlying hypothesis is that the ability to modify the consumption is higher where the demand is higher. However, other hypotheses could be defined and evaluated in further works.

This dispatching is performed by Commander and Intermediary Units. If $R(t)$ is the power constraint sent by the system operator, the corresponding amount of power to be modified by the unit i at the lower level is:

$$R_i(t) = r_i(t) * R(t) \tag{14.1}$$

Where r_i is the ratio from the whole restriction that the unit i must achieve and is defined as:

$$r_i(t) = P_i(t-1)/P(t-1) \tag{14.2}$$

Where P_i is the power consumption that the unit i is monitoring and P is the power consumption of the whole scenario.

14.1.3 Optimisation on device controls

The idea to accomplish the power constraints among the devices at the neighbourhood level consists in executing an optimisation algorithm in the control units. The optimisation algorithm used is MOPSO [CL02, SC05]. MOPSO is a Multi-Objective version of the PSO algorithm [KE95].

This algorithm required adapting to the power grid domain. These adaptations are related to specific implementations of the algorithm concepts (particles and objectives, among others) and, on the other hand, improvements to enhance its time execution. Our version of this algorithm has been adapted from the Opt4J implementation [LGRT11]. In the next sections it is explained the modifications that have been made in order to adapt the implementation of MOPSO made in Opt4J.

14.1.3.1 Particle

Each particle (or individual) contains a feasible configuration of all controllable devices of a neighbourhood that is controlled by the corresponding Control Unit.

Device kind	Real configuration	MOPSO configuration
Lighting	On with intensity 100%	> 0.5
	On with intensity 70%	<= 0.5
Refrigerator	On	> 0.5
	Off	<= 0.5
Freezer	On	< 0.5
	Off	>= 0.5

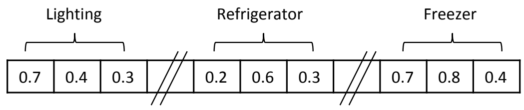


Figure 14.2: Left: real and MOPSO configuration matching. Right: particle configuration example.

That is, in the initial configuration, each particle is generated randomly and, then, its feasibility is tested with respect to the restrictions that are exposed in the next section. Each particle corresponds to a configuration of the complete set of controllable devices of the neighbourhood.

The goal of the optimisation process is to obtain the optimal solution starting from this initial configuration. The device configuration differs in relation to the type (Figure: 14.2). Therefore, the lighting control concerns the intensity modification. However, freezers and refrigerators are controlled by turning them on/off.

MOPSO generates the population at the beginning and makes it evolve every iteration. This evolution of the particles may cause them to move from a valid position to an invalid one making their configurations infeasible. Because of this, the constraints are checked at every iteration before calculating the particle fitness.

14.1.3.2 **Restrictions**

The restrictions check whether the particle's configuration is valid for a concrete device regarding several aspects. These restrictions are described in the list below:

- *Maximum time under control.* This is the time that a device can be remotely controlled by the DLC method. This time varies depending on the device kind. For instance, fridges and freezers cannot be disconnected for long periods. However, lighting may be under control for longer periods of time.
- *Minimum time before control.* After being controlled, devices are recovering. For instance, if a fridge was disconnected for 10 minutes, then it needs sometime to recover its internal temperature. Then, this time watches over the minimum time that is necessary to take before being again under control.

This time ensures that devices can operate well and also varies depending on the device kind.

14.1.3.3 **Objectives**

The device configuration that every particle has is evaluated with respect to three objectives. The first of them is related to the power constraint application while the other two concern quality of service. Below, a detailed description of the three objective functions will be described.

a) First objective

This objective looks out for particles whose configuration minimises the difference between the power consumption to be achieved and the estimated power consumption based on particle configuration. Every particle configuration is processed to obtain an estimation of the power change that this would produce if applied (Equation: 14.3).

The way in which the consumption of the device is estimated depends on the device kind. In the case of the lights, the consumption modification can be previously calculated since the current consumption and intensity is known by the control unit. Therefore, the modification in the consumption of a specific light when changing its intensity can be perfectly estimated.

In the case of the thermal loads, their consumption can also be estimated when they are proposed to stop consuming as their current consumption is known. Nevertheless, when they are proposed to stop being controlled, it is not possible to know whether they will start consuming energy or not in the same time step. This depends on the thermal load needs for cooling or freezing. In this case, based on a previously experimental study made over these appliances, this power is estimated based on the ratio of thermal loads that start consuming energy after being controlled.

$PCA_{ijl}(t)$ corresponds to the power consumption of the controllable device l of the households j and the neighbourhood i at time t. The estimation of the consumption for the controllable device l of household j in the neighbourhood i for the next t is $PCA_{ijl}(t{+}1)$ according to the configuration of the particle that is being analysed. If in the neighbourhood i there are h_i households and at each household there are a_{ij} devices, the power change expected $PE_i(t+1)$ in neighbourhood i is calculated as:

$$PE_i(t+1,\vec{x}) = \sum_{j=1}^{h_i}\sum_{l=1}^{a_{ij}}[PCA_{ijl}(t+1,\vec{x}) - PCA_{ijl}(t)] \tag{14.3}$$

That is the sum of the differences between the consumption in t of each device of the neighbourhood i and its corresponding estimated consumption in $t+1$ according to the particle that is being analysed.

In this sense, the first objective function $O_i^1(t+1)$ tries to minimise the difference between $PE_i(t+1)$, that is proposed by each particle $\vec{x_m}$, and the part of the power restriction $R_i(t)$ which is assigned to the neighbourhood i, coming from the grid operator. Where each particle $\vec{x_m}$ corresponds to a configuration of all the controllable devices in the neighbourhood i.

$$O_i^1(t+1) = \min_{\vec{x_m}} |PE_i(t+1,\vec{x}) - R_i(t)| \tag{14.4}$$

b) Second objective

This objective prioritises the control of devices that have been controlled few times. It avoids the situation where some devices are controlled many times whereas others are controlled few times. This objective concerns the times that the devices have been controlled (Equation: 14.5).

In order to address this objective, $CAP_{ijl}(t)$ is defined as the number of times that the device ijl has been controlled by the DLC method. The device that has been controlled the most, $MP_{ijl}(t)$, is calculated as the maximum of the previous calculated values of $CAP_{ijl}(t)$. Based on the previous calculated values, the objective function O_i^2 is defined as:

$$O_i^2 = \max_{\vec{x_m}} \left(\sum_{j=1}^{h_i}\sum_{l=1}^{a_{ij}}[CAC_{ijl}(t) * (MP(t) - CAP_{ijl}(t))^{\alpha}] \right) \tag{14.5}$$

Where CAC is a boolean value which is set to 1 when the device is proposed to be controlled and 0 otherwise. $\alpha > 1$ is a parameter that is intended to magnify the effect of the differences between MP and CAP_{ijl}. In our case, α has been set to 5. Then, the objective 2 shows higher values when devices controlled few times are proposed to be controlled.

c) Third objective

This objective attempts to continue controlling devices that are already under control before starting to control new ones. The aim is to ensure client comfort by avoiding many switches in short periods of time. This objective is addressed by calculating two values:

- Number of devices that can continue being controlled but are proposed to stop being controlled in neighbourhood i ($CPN_i(t)$) and,
- number of devices that are not being controlled but are proposed to start being controlled in neighbourhood i ($NCPN_i(t)$).

Ideally, a device should not start being controlled when there are others that are already being controlled and can still continue. Thereby, this objective minimises the smallest of both values previously defined.

$$O_i^3 = \min_{\vec{x_m}}(minimum[CPN_i(t), NCPN_i(t)]) \tag{14.6}$$

The reason why one of them is minimised is because some flexibility is necessary in order to accomplish the cases where the modification of the restriction is big and, therefore, the amount of devices that are necessary to attend the constraint is higher.

14.1.3.4 **Archive**

The original MOPSO *archive* contains the particles that are Pareto-optimal. This archive is inspired in the Pareto Archive Evolution Strategy (PAES) [KC00]. This means that there may be particles with one good value for a specific objective and not so good values for other objectives. Besides, this algorithm includes a crowding process that is used to establish a second discrimination criterion (additional to Pareto dominance) [CL02, SC05].

A modification has been made in this archive management procedure in order to remove particles that do not have a good value for the first objective. As the first objective is the most important one to address, a modification in this *archive* has been implemented to favour this objective. With this modification, the *archive* manager removes the particles that are far from being optimum for the first objective (when $PC - R(t) > R(t) * 0.1$). This process is made to reduce the exploration

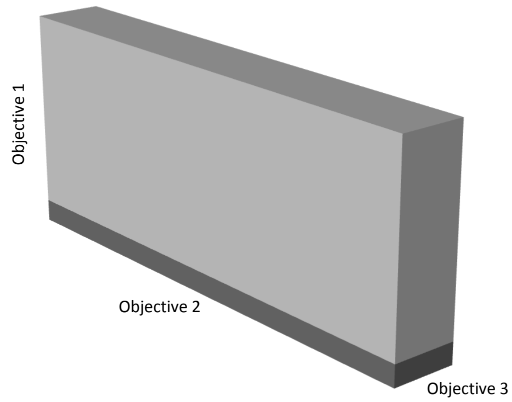


Figure 14.3: Search space reduction achieved by the new *archive*.

inside the search space (Figure: 14.3) to the really interesting region with respect to the problem.

This improvement has been done since on the one hand, the complete search space can be really huge when configuring thousands of devices and on the other hand, there is no interest in exploring the regions that are out of the limits previously defined. As a consequence of this modification, the execution time improves the original MOPSO algorithm since it reduces the number of particles needed to explore this smaller search space.

14.2 Case Study

In order to validate this method a case study has been designed. Fridges, freezers and lighting have been selected to be controlled for two main reasons. On the one hand, the consumption of these devices represent a significant portion of the whole. On the other hands, fridges and freezers are interesting since thermal inertia allows to shift ahead the consumption without affecting the quality of service. As well, light intensity can be slightly reduced in order to adapt the demand in case it is necessary.

The DLC method is evaluated simulating this case study. The simulation model consists of two main parts: control and consumer side. The control side consists of the hierarchical structure for dispatching the power constraints and controlling the consumer side. Consumer side is modelled according to a previously designed

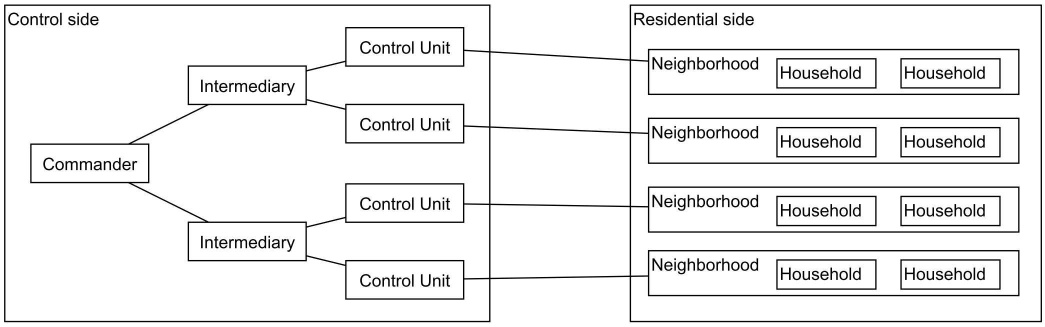


Figure 14.4: Two-sided scene: controllers and consumers.

experiment [EKM+11]. Briefly, the demand side has been modelled in detail, including all appliances at households and social behaviours.

14.2.1 Scene description

The scene consists of two main parts: controllers and consumers. On the one hand, the controller side contains: a single Commander, 3 Intermediaries and 6 Control Units. They are hierarchically structured as exposed in the figure 14.1. This scene does not contain any grid operator since the production side is not represented and, therefore, the balance between production and demand is not calculated inside. Thereby, the commander dispatching unit is provided with the power constraints to be achieved (Figure: 14.5).

On the other hand, consumers are grouped in neighbourhoods which represent a set of customers that are controlled by the same control unit. The households where the consumers live are equipped with common appliances. The appliances that a household has are related to the social group of the household inhabitants. These studies use the same information gathered in the sociological studies performed in Germany [EKM+11] and related to the social behaviour of household inhabitants in order to analyse the residential energy consumption.

14.2.1.1 Consumers

There is not a fixed number of customers that can be within a neighbourhood as is the case in most real situations. The households where the consumers live are equipped with common appliances that can be found in a household (Table: 14.1.

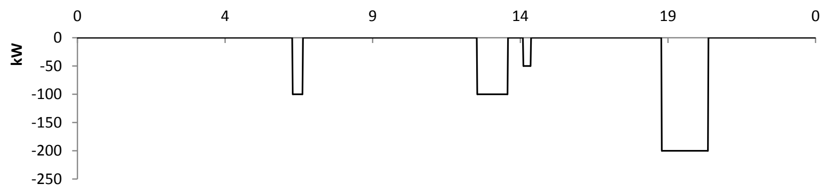


Figure 14.5: Restrictions sent by the grid operator. X-axis: time in hours. Y-axis: power constraint in kW.

The appliances that a household has are related to the social group of the household inhabitants. There are five different social groups:

- junior single,
- senior single,
- junior couple,
- senior couple and
- family with children

These groups represent 70% of the population of Germany, based on the German Socio-Economic Panel. A survey which interviewed people in Karlsruhe (Germany) provided the data about how the people use the appliances, at what times and for how long. Therefore, samples were grouped taking into account to model the social behaviour of each social group. Each sample has been modelled as a recipe [1] where the starting time and duration time of the actions contain some randomness providing different energy usages at each household.

14.2.2 Simulation parametrisation

In this section, the parametrisation of the simulation is presented. In table 14.2, the values of the most important parameters of the simulation are provided. Note that, there are MOPSO parameters that have not been changed from the original implementation made in Opt4J in which our algorithm is supported. For this reason, we refer to the library in order to find out these values and many others more that may be interesting.

[1] A list of actions to perform along the day, e.g. turn on the computer at 10 am.

Model object	**Amount**	**Remarks**
Neighbourhood	6	Several Households A control unit
Households (HH)	1350 among neighbourhoods	Several Appliances Related to an agent
Units	10	1 commander (C) 3 intermediaries (I) 6 control units (CU)
Cooking Stove	4 per HH	
Microwave	1 per HH	
Oven	1 per HH	
Dishwasher	Social group dependence	
Refrigerator	1 per HH	
Audio Hifi	1 per HH	
TV	Social group dependence	
Lighting	1 per HH	All lighting in a HH
Washing machine	1 per HH	
Computer	Social group dependence	
Waterboiler	1 per HH	
Hairdryer	1 per HH	
Vacuum	1 per HH	
Console	Social group dependence	
Agent	1 per HH	Switches devices on/off 5 kinds of agents
Power Connection	1359	1350 from HH to CUs 6 from CUs to Is 3 from Is to C
Communication lines	Dependent	From devices to CUs 6 from CUs to Is 3 from Is to C

Table 14.1: Elements existing in the model scene.

Table 14.2: MOPSO parametrisation used.

Parameters setting	**Value**
Number of Households	1350
Number of elements	$\approx 42k$
Time step	10s
Number of particles per each MOPSO	30
Number of particles per archive	100
Stop condition	30 iterations
Mutation, c1, c2, w	default values in Opt4j
Particle size	neighbourhood dependent

14.3 Results

In this section, the result analysis of the simulations performed are presented. All the simulations performed have been made under the same conditions (only the strategy was changed from non-DLC to DLC). The simulations are 1-day long in all cases and include some randomness. This randomness comes from two sources: social behaviour and MOPSO. The social behaviour includes some randomness that regards the time and duration in which the actions are executed. This is so in real situations where not all appliances of a kind are switched synchronously at all the households, but they are switched at different times with a behaviour that is modelled stochastically. On the other hand, MOPSO has a high random component to generate the initial population. Because of this, we have performed 20 simulation executions: 10 non-DLC simulations and 10 DLC simulations. Most of the result analysis made is using the averaged values for those 10 individual runs.

The results are presented following a bottom-up strategy. This is, the device level is shown in the first place and the total aggregated level at the end. The way in which the DLC strategy is evaluated is by comparing its results with the simulations where there is not any strategy working. Thereby, at the device level, we will be able to see how the devices react under the DLC control by contrasting them with the normal operating mode.

Later on, the differences between some neighbourhoods are shown where the least populated one is compared to the most populated working under both non-

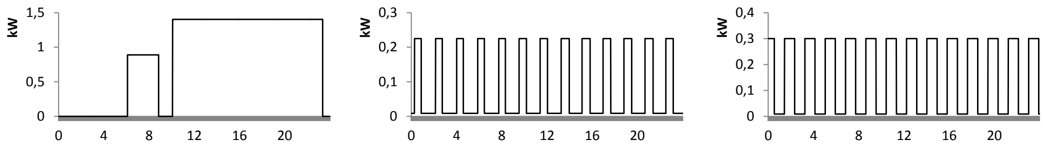


Figure 14.6: Devices without DLC policy: a light, a refrigerator and a freezer.

DLC and DLC. Then, the way in which each neighbourhood contributes to the power restriction is shown.

After finishing with the most aggregated level, a difference, device by device, is shown aggregating all of them by kinds. Then, all the lights are sorted in the lighting group, all refrigerators in the refrigerators group and all freezers in the freezers group, they are all aggregated from all the neighbourhoods of the scene. An a posteriori comparison shows the impact of each kind to accomplish the power constraints.

Finally, all the devices within the 6 neighbourhoods are aggregated showing the total consumption as a result of the emergent behaviour of the scene which is observed at the highest level.

14.3.1 **Device level**

This section presents the results at the device level (less aggregated one) where the differences between the non-DLC usage and DLC usage can be observed. The figure 14.6 shows the way in which three devices worked throughout the day of the non-DLC simulation. The lighting consumption depends on the social behaviour of the people living in the household. This is, its consumption evolves according to the behaviour of the inhabitants' social group as well as the household area. However, refrigerators and freezers have a systematic operating mode. These simplistic models represent their behaviours for cooling as a pulse train.

The figure 14.7 presents some examples of how those machines behave under a DLC strategy. In the lighting case, some small intensity changes can be seen throughout the day. This is due to the DLC order sending which modifies the lighting intensity. Refrigerators and freezers are totally switched off when they receive the order coming from the control unit. This figure shows several examples for each kind showing that they do not act synchronously. Moreover, the restriction

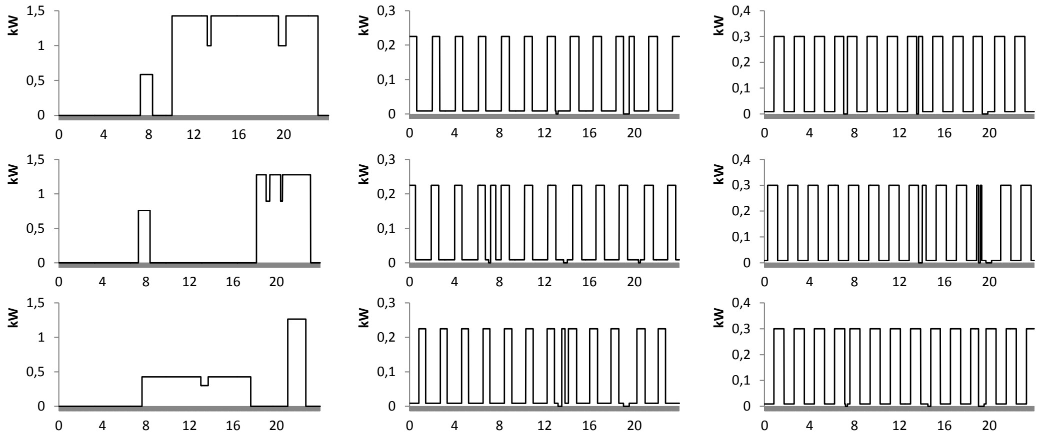


Figure 14.7: Devices with DLC policy: a light, a refrigerator and a freezer.

satisfaction to keep the quality of service is complied at every time for every appliance.

14.3.2 Neighbourhood comparison

Every neighbourhood in the scene has a different amount of inhabitants living inside. Then, the mission of the upper control units is to rightly distribute the power constraints among the neighbourhoods. In the non-DLC case, it is possible to observe the consumption difference between a small neighbourhood (50 households) and a big one (500 households) (Figure: 14.8). As the small one is less populated, the consumption is more fluctuating in the night peak due to the pulse devices such as cooking stoves. This fluctuation is lower in the more populated one since this effect is hidden by experiencing a higher consumption.

The same comparison is made for the DLC case (Figure: 14.9). The difference between both consumptions is almost the same than the non-DLC case except for the power constraints. What is important to note from this figure is the contribution each one made to each power restriction. Focusing on the last restriction, the whole scene had to reduce 200kW. Therefore, the least populated neighbourhood contribution is 8kW while the contribution of the most populated one is around 75kW.

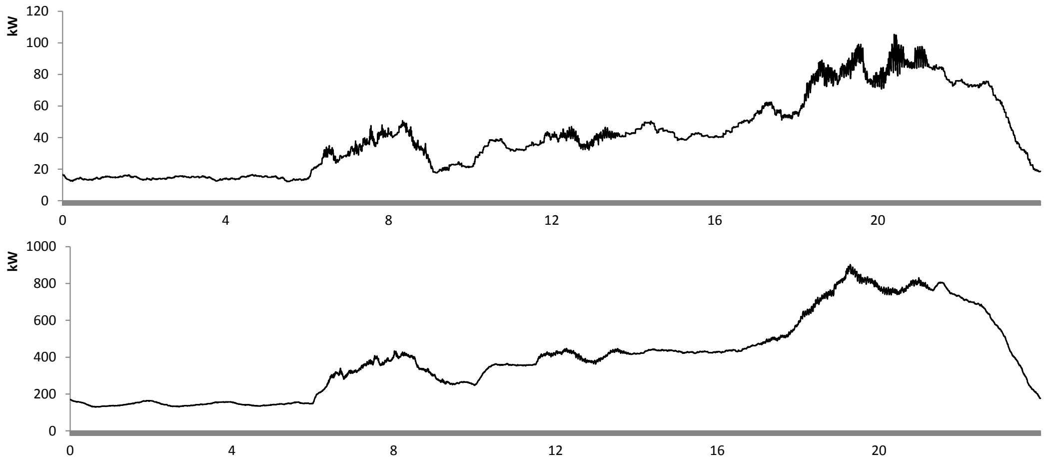


Figure 14.8: Neighbourhoods without DLC policy. Top: least populated one. Bottom: most populated one.

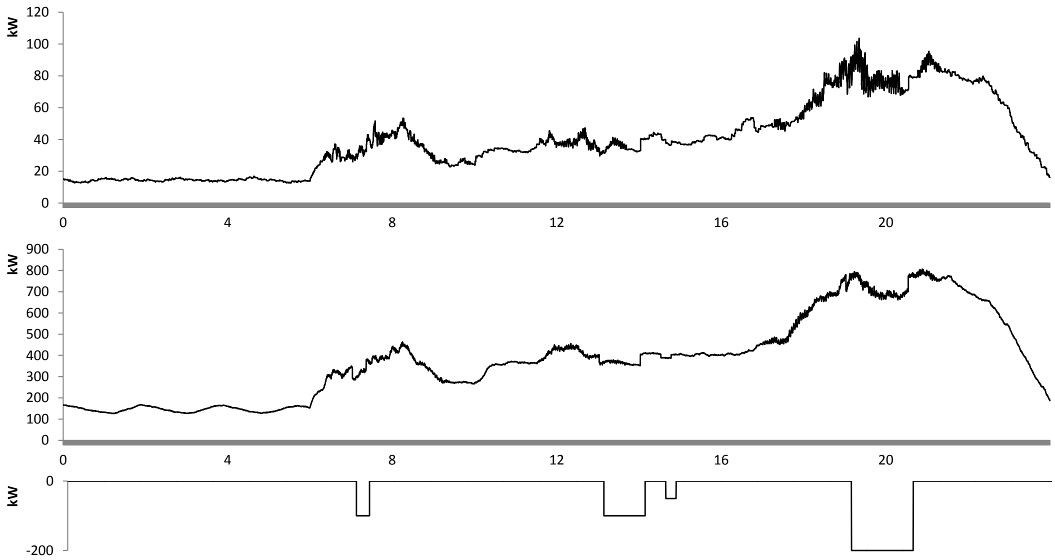


Figure 14.9: Neighbourhoods with DLC policy. Top: least populated one. Bottom: most populated one.

14.3.3 **Whole scene**

At this point, the behaviour that emerges from the whole system is presented from the point of view of power consumption. This is, all the consumption within every neighbourhood including non-controllable devices. In the non-DLC case (Figure: 14.10), we can observe what the behaviour is like when no DLC method is being

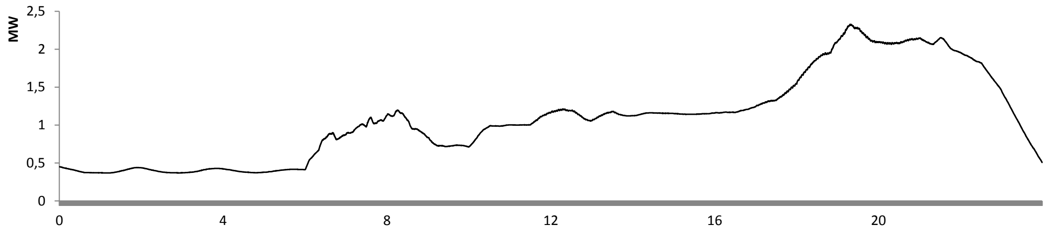


Figure 14.10: The average total consumption from ten non-DLC simulations.

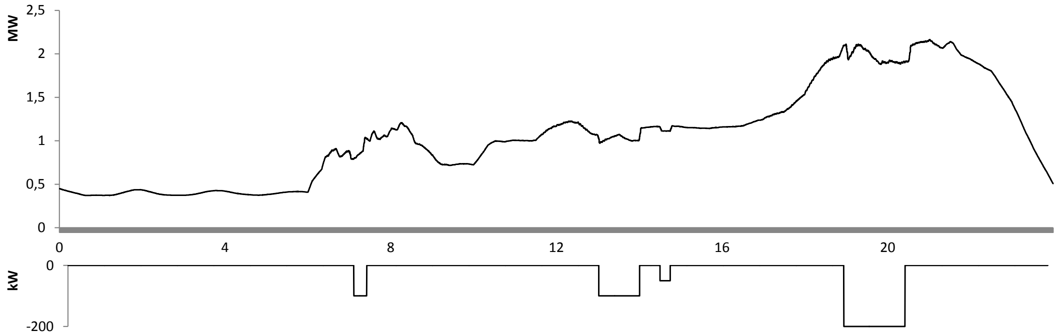


Figure 14.11: The average total consumption from ten DLC simulations.

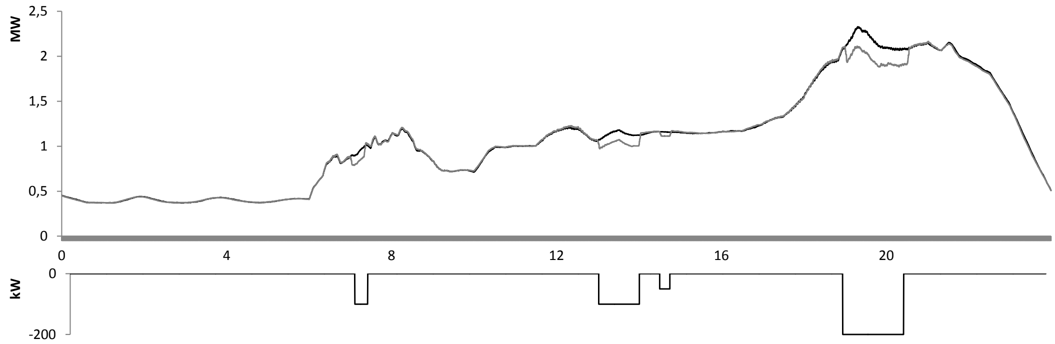


Figure 14.12: The comparison between the non-DLC total consumption (black) and DLC total consumption (grey).

applied. A more detailed discussion concerning the realistic aspects of the resulting load curves of this kind of simulations can be seen in [EKM$^+$11].

The figure 14.11 shows how the whole system acts under the orders generated by the DLC method concerning power constraints. Now the aggregation of three kinds of appliances can be seen together with the rest of the consumption. Successfully, higher power constraints imply higher reductions in the load curve.

Table 14.3: MOPSO execution times.

Results	Value
MOPSO average neighbourhood 1 (50 HHs)	212ms
MOPSO average neighbourhood 2 (500 HHs)	2193ms
MOPSO average neighbourhood 3 (100 HHs)	471ms
MOPSO average neighbourhood 4 (250 HHs)	1179ms
MOPSO average neighbourhood 5 (300 HHs)	1231ms
MOPSO average neighbourhood 6 (150 HHs)	727ms

14.3.4 Performance

One of the main concerns of this paper is to design a structure which allows to reduce the execution time of the MOPSO algorithm. As said before, the execution time that MOPSO takes can be adjusted according to the size of the clusters. For this reason, this case study has been designed to contain several neighbourhoods with different sizes so that executions time can be evaluated.

In the table 14.3, the average execution time of the MOPSO algorithm for each neighbourhood is presented. As it can be observed, the execution time linearly increases when dealing with bigger neighbourhoods. These results are not important from the point of view of their amount, since they may change if simulations are executed in a different computer. Nevertheless, they are important to be considered since it can be seen a relation between the cluster size and the execution time.

Therefore, the proposed method can be adjusted to achieve the execution times that are required by a power grid. This adjustment process must be made by clustering the power grid with the right size that allows the achievement of the desired execution time.

14.4 Discussion

A DLC method which uses a multi-objective optimisation algorithm has been proposed and evaluated through simulation. This method deals with the application of power constraints and the preservation of grid stability and customer quality of service. An interesting property of this method is that it can be scaled according to the power grid size.

In order to test it, a simulation for a case study has been developed where the demand side has been modelled in detail, including all appliances at households and social behaviours.

It has been shown that the method is able to find a global solution within the time requirements. This feature is really important since the demand management must be executed around a second. The case study shows the significant impact that this method produce in the demand-offer management of power grids. In the simulation results it has been observed how the method is able to reduce the demand up to 20% controlling three kinds of devices: refrigerators, freezers and lightings. This DSM policy over the demand increases the flexibility in the whole control of the power grid.

Future works could analyse and evaluate different dispatching strategies or even other optimisation algorithm. MOPSO has provided good performance results concerning the execution time. However, there are other optimisation algorithms that can be coupled in the proposed method. Even more, within the same structure several optimisation algorithms can be executed in different control units at the same time.

Other DLC methods should also be explored where more devices are controlled to modify the demand. For example, thermal devices such as water heaters, air conditioning or radiators could be included in future studies. This could lead to a higher control ability on the demand side, and in this way a higher flexibility.

CHAPTER

15

A large-scale electrical grid simulation for massive integration of distributed photovoltaic energy sources

This experiment, published in [EKHH12] [1], presents a case study which explores the impact of a massive PV generation introduction in a real scenario. The PV generation is integrated using a DLC policy which enables the modification of the demand in order to balance the power load.

Tests are performed "in silico" using Tafat. This case study takes places in the context of the Millener [EDFb] (Mille installations de gestion energétiques dans les îles – Thousand energy management devices on islands) project framework.

[1] People that participated in this experiment: José Évora, Enrique Kremers, Mario Hernández and José Juan Hernández.

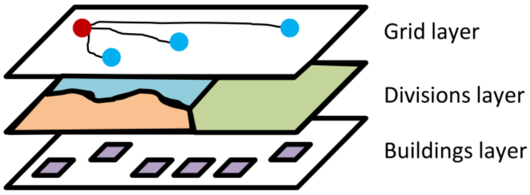


Figure 15.1: GIS layers related of the provided data.

15.1 Description

15.1.1 Modelling

The modelling process consists of two main steps: data processing and scenario modelling. Those steps correspond to the first and the second phases of the simulation's life-cycle (presented before in 8.1.5). In the first step, the goal consists of integrating miscellaneous data of the region which come from different sources. These data concern three separated aspects: buildings, administrative divisions and electrical grid (Figure 15.1). The building database refers mostly to the residential sector and provides information concerning their geo-coordinates, their floor areas and their heights. The administrative divisions information provides the geographical area as well as the information which concerns the main aspects of every division . The electrical grid information geo-localizes the substations, distribution feeders and transformers of the grid. Furthermore, it provides the consumptions at the feeder level. These data were integrated in two steps. First, the administrative divisions and buildings were related using a contention criterion. Later, the buildings were related to the closest transformer which is usually the case.

The second step is divided in two main processes: entities addition and scenario modelling. As the measure requires the PV, battery and air conditioning entities we need to disaggregate those devices at the customer level. Since the PV [Duf10] and battery [GDDBKV09] models were already developed the process only regarded their integration into the Tafat framework. On the other hand, the air conditioning behaviour does not represent the working mode of the device as we assume their consumptions are included in the consumption data. Therefore, this behaviour only represents the effects of shifting the air conditioning consumption. Finally, the solar radiation behaviour was represented using historical data of this region.

In the scenario modelling process, Profiler has been used. Therefore, a plug-

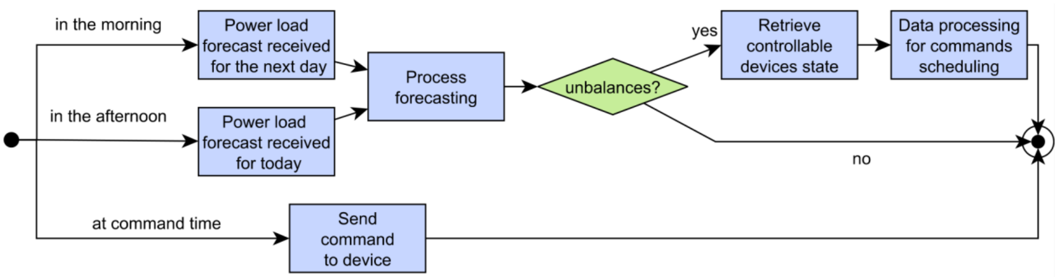


Figure 15.2: DLC strategy flowchart.

in for the Profiler was implemented to develop models according to the processed data. This plug-in parses the output data coming from the previous step and builds a model where the buildings, regions and the electrical grid are represented according to the Tafat model syntaxes.

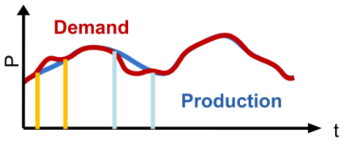
The building consumption was represented by distributing the known consumption at the feeder level considering the buildings volume. This model, named as baseline, allowed validating the disaggregation and aggregation processes by confronting the aggregated simulation results at the feeder level with the original data.

Later on, a second model was created where the devices proposed in the DLC measure (PV, battery and air conditioning) were added. As the addition of those devices is designed to be made over 250 single family houses, the information at the administrative level which regards this aspect was used to calibrate an algorithm which filtered these houses using an area and height criteria. On the control side, the Control Unit (CU) and the (Energy Box) EB were added and their behaviours developed considering the strategy description concerns.

15.1.2 **Control policy**

The proposed and evaluated solution considers global and local measures to schedule commands which will be executed in a certain time and over concrete controllable devices. This procedure is achieved by the CU through executing two main tasks: scheduling and command sending. The objective is to balance the demand and production of the overall system (Figure 15.2).

The schedule task is carried out twice. The first one is processed in the afternoon to set up the commands for the next day. This schedule task consists in comparing both the predicted production (PV generation included) and consump-



Alert\Time	0	2	4	6	8	10	12	14	16	18	20	22	24
Afternoon	0	-1	0	0	0	0	0	0	0	0	0	0	0
Morning	0	0	0	0	1	0	0	0	0	0	0	0	0

Figure 15.3: Translation process from unbalances into day-long alerts.

tion along the next day. When a difference between both is observed (alert), an iterative process to retrieve the customer devices state is started to check the actions that can be taken to bridge the gap (Figure 15.3). The use of this information over some restrictions will raise the devices which will be controlled. Joining the alert time and the selected devices the commands are created. The same procedure is made in the morning to re-schedule the commands for the same day using a new prediction gaining precision.

The second task consists in sending the generated commands according to a specific schedule to the proper customer household through the EB. In this strategy, the EB works as a bridge between the CU and the households, as well as sending the local measurements to the CU.

15.2 Case study

This case study concerns a massive integration of PV generation units over a real region. This study is performed through "in silico" simulation. The goal of the case study we present is to analyse and evaluate the impact of a DLC policy control on a pilot project proposal over a model based on a real region where 250 residential customers are equipped with PV cells and batteries. In order to provide flexibility to the DSM measure the selected customers have air conditioning equipment whose consumption can be shifted. The simulated regions are Saint Pierre and Le Tampon ($\approx$ 150,000 inhabitants) which are in La Reunion Island (France). This case study is developed in the Millener project framework whose objectives are, among others, the integration of more RES in the context of isolated grids, the improvement of the grids stability and the incentive proposal to improve the energy efficiency. Within this context a specific objective is the design, study, evaluation and implementation in real life of DSM measures.

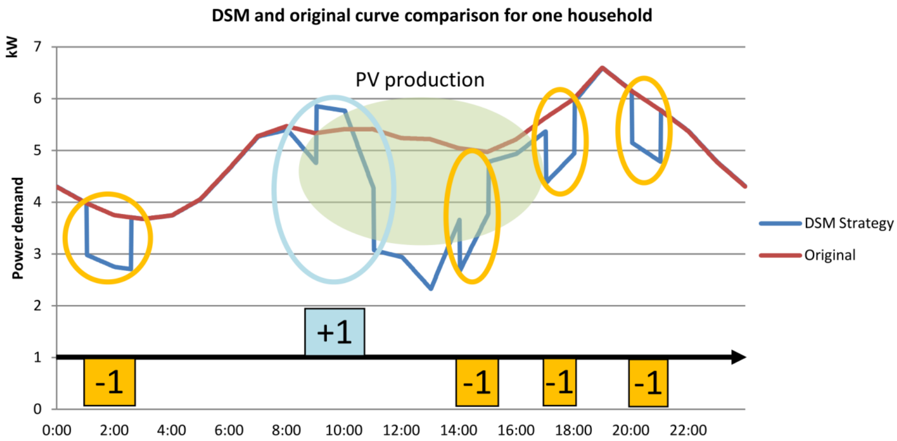


Figure 15.4: Difference between the baseline load curve and the one coming from the DSM strategy for a concrete building.

15.3 Simulation results

In this section some charts are presented to show the differences when applying the DLC policy confronted with the baseline situation (where no measures are applied) at different scales. The Figure 15.4 presents those differences for one specific customer of the region. Note the impact of the PV production from 8 to 19 approximately providing a peak of 2.5kW at noon. This PV impact is not a consequence of a command execution but is the consequence of the solar radiation. Furthermore, the commands execution can be observed at different times and they approximately achieve 1kW of increment or reduction respectively.

In Figure 15.5 the strategy is evaluated on the aggregate to all the selected customers. At this level we can observe a PV production peak at noon of approximately 650kW. As expected, the aggregation of the customers presents a reaction to the commands of approximately 250kW.

The use case results show a very flexible model which allows modifying the load curve by approximately 250 kW in both senses (reducing / increasing). This amount is variable since the selected devices for the deflection are not always the same depending on their states: battery state of charge, last time the air conditioning was shifted, etc. Zooming in at the individual household level, we can observe significant modifications of the load curve compared with the original one. Similarly, these changes can be observed when the 250 customers load curves are ag-

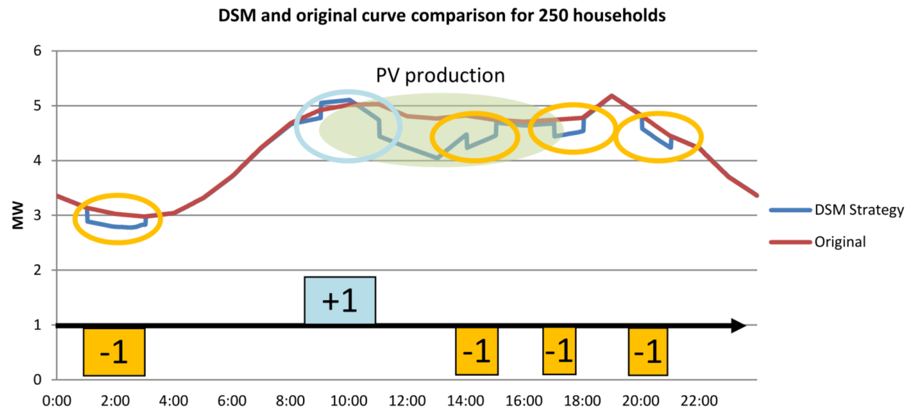


Figure 15.5: The original and the DSM strategy load curve for all the Millener customers (250).

gregated. We can expect bigger percentages when adjusting the parameters and also improving the disaggregation (the current consumption for the customers' households is oversized). Note that this impact is provoked by intervening only 250 customers of the region. Higher impacts could be expected when intervening more customers. On the other hand, the aggregated load curve can be fitted as desired allowing its flattening.

15.4 Discussion

The study and evaluation of DSM measures over a real scenario where PV is massively introduced has been fulfilled. The measure presented success in the demand modification regarding electrical grid conveniences. The usage of Tafat framework has facilitated the model creation and simulation and will allow including more requirements as the strategies require. Therefore, we can focus on several topics including the design and testing of more DSM procedures, the improvement of the regions representation, the addition of more RES technologies and the full production representation.

These simulations performed at the time in which this publication was released contained thousands of elements. Nowadays, simulations in the context of this project concern the representation of the whole island including the whole power grid, buildings, devices and social behaviours. These simulations have consisted of up to 3.5 million elements.

CHAPTER

16

Criticality in complex sociotechnical systems, an empirical approach

In this experiment, published in [VKE$^+$13] and submitted to "Applied Energy" [1] on July 2014, the criticality of sociotechnical systems is evaluated. The system that is evaluated is a Power Grid which contains many smart refrigerators. These refrigerators are able to react according to the grid frequency by switching themselves on / off.

Therefore, the evaluation regards which parameters of this massive introduction of refrigerators influence the stability of the grid. To this end, a large corpus of simulations have been developed using Tafat which allowed the identification of identify critical configurations of the system which are accountable for instability issues in the Power Grid.

[1] People that participated in this experiment: Pablo Viejo, Enrique Kremers, José Évora, José Juan Hernández, Mario Hernández, Oscar Barambones and J.M. González de Durana.

16.1 Description

16.1.1 Phase shifts and criticality

One of the most peculiar properties of complex systems is how in certain cases, their situation changes between order and disorder, stable and unstable states. The term edge of chaos was introduced by Langton [Lan90] while analysing phase transitions on cellular automata and describes a critical point which separates order from disorder. This edge of chaos can be seen as a stationary regime itself, in which complex phenomena take place [Wol02].

The aspect of phase changes can be shown in complex networks by the Watts-Strogatz model [WS98]. The model aims to show the small world effect, in which there are short paths between apparently unrelated individuals of a complex system. It is done through an ordered graph, in which the connections (edges) are changed. We start with an ordered ring lattice graph in which we rewire edges randomly. The average path length drops quickly after an initial limited rewiring, while the clustering coefficient remains almost constant, leading to a typical small world network. If we continue rewiring more and more edges, a random graph will emerge - a complete disordered system. Here we see that the average path length is still low, but has not decreased much since our small world state. It is noteworthy that the clustering coefficient dropped - another characteristic of a disordered system.

Complexity thus resides between an ordered network and a completely random topology. Here, the characteristics of both order and disorder meet and complex effects can be observed, such as changing and emergent patterns. This example illustrates the location of complex systems and phenomena situated between ordered systems and randomness. Langton [Lan90] introduced the border where a system gets chaotic, while analysing phase transitions on cellular automata. A cellular automata is governed by simple rules, but have proven to show highly complex behaviour [MCH$^+$94, MHC93].

Criticality refers to a critical location which separates order from disorder, and plays an important role in complex system theory. Some natural systems are pushed towards the edge, which can be seen as an driver for evolution. Complex phenomena can take place there, which allow the system to change its state, which unlikely to happen in a completely ordered or chaotic environment (evolutive landscapes, red queen theorem).

In order to maintain *fitness* in a changing environment, it is necessary for the system to be capable of actively transforming itself over time [All09]. Here resides the importance of the edge. In evolving environments, the system has to be adapted continuously in order to stay in efficient conditions and these adaptations and changes are more likely to happen at the edge of chaos.

An example of this edge, which is familiar to the engineering and technical domains, is the Reynolds number. It is used in fluid mechanics to characterise different flow regimes, such as laminar or turbulent flow. A laminar flow occurs at low Reynolds numbers, where viscous forces are dominant, and is characterized by smooth, constant fluid motion; turbulent flow occurs at high Reynolds numbers and is dominated by inertial forces, which tend to produce chaotic eddies, vertices and other flow instabilities.

16.1.2 Simulation study for sociotechnical systems

Socio-technical systems are present in the real world in everywhere a technical system interacts with human behaviour. This is the case for the major part of existing technical systems. In order to model the "whole" system, human behaviour has to be added to technical models. Coupling social and technical models is therefore a requirement for modelling these large, complex systems. Agents are well suited to this approach, as by definition, they include behaviour which can be interpreted as social or also technical behaviour.

As an example of this, in [MP10], urban water systems were modelled as complex adaptive systems by using ABM. The approach supported the analysis of the system, by acquiring a fundamental understanding of the processes - e.g. emergent phenomena, relating to water safety and human behaviour, which could be represented through a micro-modelling rather than being described as a probabilistic causal model at the macro level. So, it was noted that while steady-state behaviour was achieved in most cases, occasionally very different steady-states occur, and sometimes they also had sudden shifts that occur for diverse and unexpected reasons [MP10]. Being a non-deterministic model, it allows, given a set of conditions, the prediction of patterns of system behaviour, at least in a probabilistic sense.

16.1.3 Electrical energy systems as complex systems

The energy system is a highly interconnected system of systems [KV10], which is undergoing a paradigm shift moving away from a centralised and hierarchical structure, towards a more distributed system where actors have a larger influence on it.

Electrical energy systems are composed of different networks and levels, usually differentiated by their voltage. Large producers inject at high voltage levels, medium voltage level is used for distribution and mid-size production units, as well as industrial consumers, and the low voltage level delivers electricity to the final consumers. All these levels are interconnected through transforming stations. Electrical grids are usually large scale systems, ranging over vast geographical areas, such as the European or North American power grid. The system itself is in continuous change, offering a highly dynamic behaviour in time and space. Transitions emerge over time as fundamental change of large-scale socio-technical systems (lambda-systems) [CD10] such as energy infrastructures that are the backbone of society.

The paradigm shift the energy system is going through nowadays implies that production is no longer limited to large energy providers. This is due to the entrance of small decentralised producers in the form of distributed generation to the network, which are able to inject energy at much lower voltage levels than before. This allows energy to be consumed close to where it is produced, avoiding long transmission paths. Therefore, a tendency towards a less hierarchical system can be observed, which involves energy flows that become bidirectional (not only in direction to the lower levels, but also going up sometimes). This involves technical and regulatory challenges which have to be taken into account. A better knowledge of the demand side is needed in order to tackle these points.

The electrical energy system thus can be seen as a system-of-systems with a high increase of decentralized decision processes. Also, new technologies increase the means of communications and the interactions along the network. A large number of entities, with growing heterogeneity, characterise the system as a complex one.

Current tools to simulate energy systems usually focus on one scale or level. So, behaviours on different levels are hardly being represented together, as well as their interaction across scales. Only a few high resolution studies and models exist on electric domestic demand [WF07, WW09], and they are not coupled to an energy system model. However, this is needed to understand the causes and effects across a complex system.

Few applications of complexity theory on electrical energy networks can be found so far in literature. Particularly, in [RCV07], Rosas-Casals analysed the fragility of the high voltage system in different countries toward targeted attacks, using complex networks theory. They found a correlation between topology and the dynamics of the system. Other applications can be found in energy markets, where complex modelling approaches such as ABM are used to reflect market behaviour by heterogeneous actors [SGRM07]. Another example is the modelling of Microgrids [PFR09].

Cascading events in the power grid have led to major blackouts in the past. These cascading failures are related to de-synchronisation processes in the electrical system. Being composed of production, dissipation, transmission, and consumption, the electrical system represents a dynamical problem and the power grid can be seen as an example of a system of oscillators [ADG08].

Previous works in the SG domain have shown that there can be phase shifts in these systems. In [KdD12b], a refrigerator DSM simulation has been shown to oscillate largely, which would imply catastrophic consequences on a real system. Whereas refrigerators work based on pulsating loads, these loads are usually not synchronised. By intervening the system though SG measures though, in some cases the phases of the loads can coincide, leading to such a synchronization in which the system begins to oscillate. Through empirical tests using a simulation model, these stable, partially stable and unstable regimes were analysed.

16.2 Case study

In [SIF07], Short uses a simplified refrigerator model in order to simulate a large number of refrigerators. However, in any of these approaches, the effects of the different DSM strategies at system level and with a large penetration have been analysed through a detailed model. The case study we chose here is based on the works described [KdD12b, KdD12a]. A simulation model of a domestic refrigerator (micro-level) was coupled with a simplified model of an electrical energy

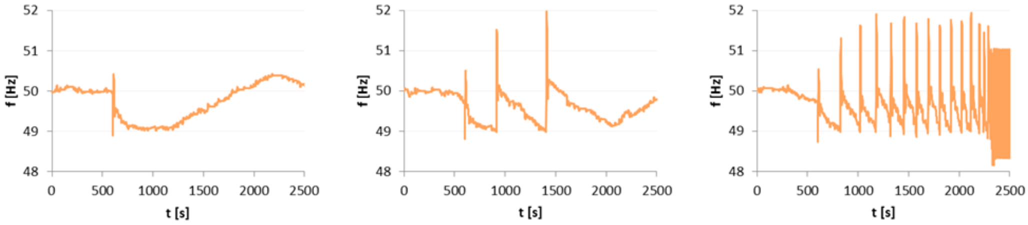


Figure 16.1: Stable (left), partially oscillating (center) and oscillating (right) regimes. Source: [Kre13].

system (macro-level). The idea behind this was testing autonomous load shedding strategies for refrigerating appliances. By using a global indicator from the electrical system, a device at individual refrigerator level decides locally whether to stop consumption, to support the reserve mechanisms of the power grid and increase the stability of the system. This mechanism seemed to work well to stabilise the system after major events like the breakdown of a production unit.

These studies also questioned that, under certain circumstances, with increasing number of controlled refrigerators, non-desired, dangerous effects like oscillations could occur. The model allowed for a first analysis by probabilistically quantifying the risk of such an event. Furthermore, besides the stable and the oscillating state, an intermediate regime was found in which it is uncertain if the system will fall towards stability or instability (Figure 16.1).

In this document, we decided to go further and analyse the critical zone in a more detailed manner, taking into account:

- an increase of the simulation entities to analyse these effects on massively replicated agents
- obtain a better understanding of the synchronisation and de-synchronisation processes
- perform a parameter analysis to detect which factors have an impact on these effects

Also, we analyse which conditions move the system towards a critical space. These steps are described in the following:

1. Analyse phase space and the external conditions of the system

2. Go towards critical space and fix conditions on the edge of chaos
3. Vary different types of parameters to detect if they are likely to shift the edge

To do so, the simulation model was transposed into a suitable simulation framework. So far, simulations were run in Anylogic. In the case described in this article, the model is implemented in Tafat [EKM^{+}11, EHH13b], which allows for better performing simulations and a larger number of agents running massively in parallel. Tafat is presented in chapter 8. For these simulations, we obtained a factor 4 of improvement in the execution time. Furthermore, Tafat was adapted in this case to perform parameter variations. simulation and result analysis.

16.3 Results

The different factors that have an impact on the system can be classified into categories. We have classified the different factors of impact on the system into

- External: related to the environment, around the system (production, failure, environmental temperature etc.)
- System relevant: related to the system that we are contemplating (Number of fridges, proportion of fridges towards the system)
- Entity internal: related to the individuals composing the system (Number of controlled fridges, parameter of individual entities of the system, door openings)

For the latter, the impact on the global system is difficult to evaluate without a complex modelling approach. Even if controlling the external and system variables, the system remains unstable, simulation can help us to determine whether if changing the internal condition can have an impact on the system. This is a typical emergent phenomenon as we have individual behaviours (conditioned by internal parameters) which can have an impact at aggregate level.

In Figure 16.2, the main simulation parameters used are presented.

The parameters of the model can be classified using these categories. All of these can be controllable or non-controllable. So, for example, the threshold setting of the load shedding algorithm and the door opening rate are both internal, but the

External or environmental			System relevant			Internal or entity related		
Description	Name	Value	Description	Name	Value	Description	Name	Value
Nominal System Power	aveProd	300MW	Number of controlled refrigerators	contrRefri	0%	Thermal parameters	Tau1-Tau9	Distribution of values according to refrigerator park. Different value for each refrigerator
Installed load (except refrigerators)	aveDemand	240MW	Scale Factor	sf	1-600	Load shedding threshold	freqOff	49 Hz
Relation refrigerators / total load	refriShare	20%						
Nominal grid frequency	freqNominal	50Hz						
Number of refrigerators	numberFridges	250-400.000						
Room temperature	roomTemp	uniform(20,30)						

Figure 16.2: Parameters and their classification.

first one is controllable and the latter not (it depends on social behaviour which is only likely to be influenced indirectly).

Not controllable parameters should be simulated as well, even if they cannot be modified in the real system. However they give an idea how a singular factor affects the system, and if this factor changes (external conditions), we can know the effect on the system.

Furthermore, we can classify the parameters as fixed or varying. Those fixed are usually similar for all the entities of the system. Random or varying parameters describe variations on the characteristics among entities of the same types. So, the refrigerators have, for example, different installed powers according to a distribution, which is taken from survey data.

In the first case, we analyse the phase space of the system. As said, given the described load shedding algorithm in [KdD12b], the system can remain in three states:

- stable
- partially stable
- oscillating

We try to avoid the oscillating state and any risks of partial oscillation, as this can lead to catastrophic oscillations between frequency and loads on the system, which would cause an immediate blackout or damage to the system.

Metrics are used to detect different phases of the system. A metric has to be adapted to the effect to be studied, in this case to detecting oscillations. These are detected by counting the number of simulation steps in which the grid frequency undergoes a threshold value, and clustering these results, to obtain the three regimes:

- stable (only one under passing at the production breakdown)
- partially stable (some oscillations, less than 20, but stabilisation after)
- oscillating state (many oscillations, more than 20 but usually several hundred until the end of the simulation).

As we are dealing with non-deterministic simulations, an identical configuration of the simulation can lead to slightly different results. So, a Monte Carlo experiment with 100 runs per configuration set was done. For each configuration set, the probability of being in a determined regime was calculated.

In a first stage, the external parameters, such as proportion of refrigerator loads vs. the total system load were adjusted (refriShare). Further, a scale factor was used, as although tested, for massive Montecarlo simulations (several thousands), simulating 400.000 refrigerators was not feasible. Before, it was shown that an aggregated behaviour was already achieved with 2000-5000 refrigerators.

16.3.1 Scenario 1: Variation of the share of controlled refrigerators

After finding a state where the system is between order and disorder, the proportion of controlled refrigerators was increased in steps of 10%. We found that the system remained completely stable below 60%. Therefore, focused on the interval 60-100% where the partial regime, the state which represents the edge of chaos, is located. The proportion was here varied in steps of 5% to have a higher precision.

In Figure 16.3, we can see a phase shift in which the partial regime represents the edge of chaos. At this stage, with fixed conditions of the system, it is not predictable whether the system will oscillate or remain stable. Through the simulation

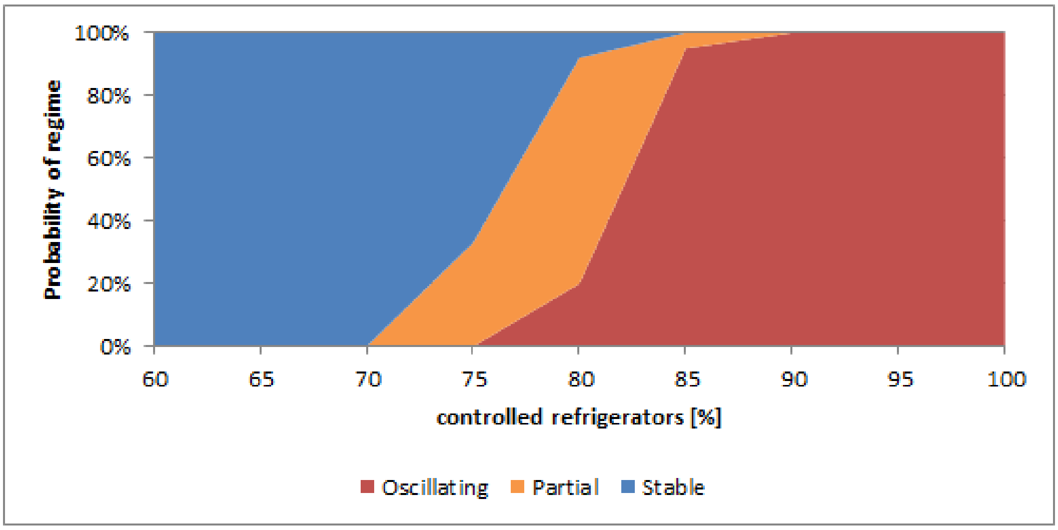


Figure 16.3: Stability of the system vs. controlled refrigerator share.

analysis however, a probability of this occurring could be calculated based on an exemplary case. In this case, we see that the variation of a system parameter can have an effect on its stability, and using the described method for a concrete system, a maximum value for integration of a certain technology could be obtained, assuring that the system will be always far from the critical zones.

Figure 16.4 shows the number of oscillation for the different shares, ordered sequentially. We see the steps in the curve which represent stable systems (2 oscillations), partial (some oscillations in between) and oscillating regimes (several hundred).

16.3.2 Scenario 2: Variation of door opening rate

In this case, we wanted to test the effect of an internal parameter, which is not controllable. We chose the door opening rate, as it is a characteristic variable for a sociotechnical systems, in which human behaviour gets into interaction with a technical system. Even if it cannot be controlled, door opening is likely to vary due to external factors, which can be predicted in some cases. This means, that without controlling the variable, we can nevertheless know how a system will react to it, if the external conditions are changing. For example, door opening probability can be increased due to extreme events, such as the world cup final or the superbowl,

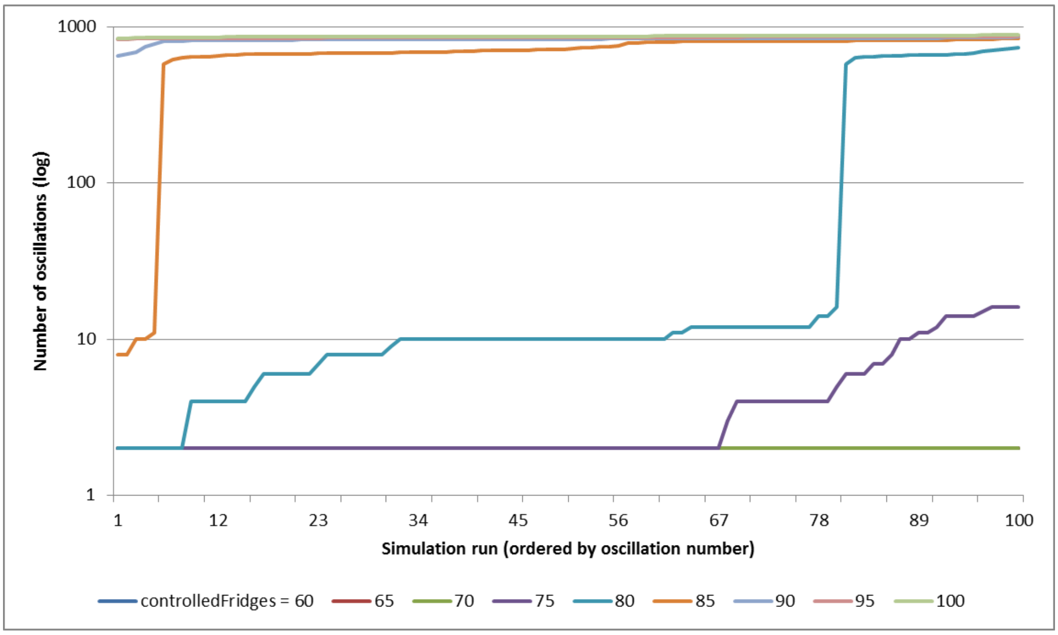


Figure 16.4: Ordered plot of 100 simulation runs for different shares of controlled refrigerators.

where a large amount of people may synchronise their behaviour. In this case, we fixed the controlled refrigerator share at 80%, which in the previous experiment showed to be at the edge of chaos. Now, the average door opening rate per day was changed and plotted in a similar manner.

We can see (Figure 16.5) that the door opening has no clear effect on shifting the edge, even if a small tendency can be seen towards a slightly more stable system when there are more door openings. The uncertainty of the partial regime is marked by a high probability (around 70-80% in all cases). Due to the probabilistic nature of the model, a larger number of simulations should be done to analyse it further. In the given range, no significant impact on the stability can be observed.

16.3.3 Scenario 3: Variation of the controlled share and frequency threshold

In this case, we are varying two parameters at the same time, to see if the threshold frequency at which the load shedding is activated has an impact on stability, and if it is dependent on the controlled refrigerator share. This simulation was run with

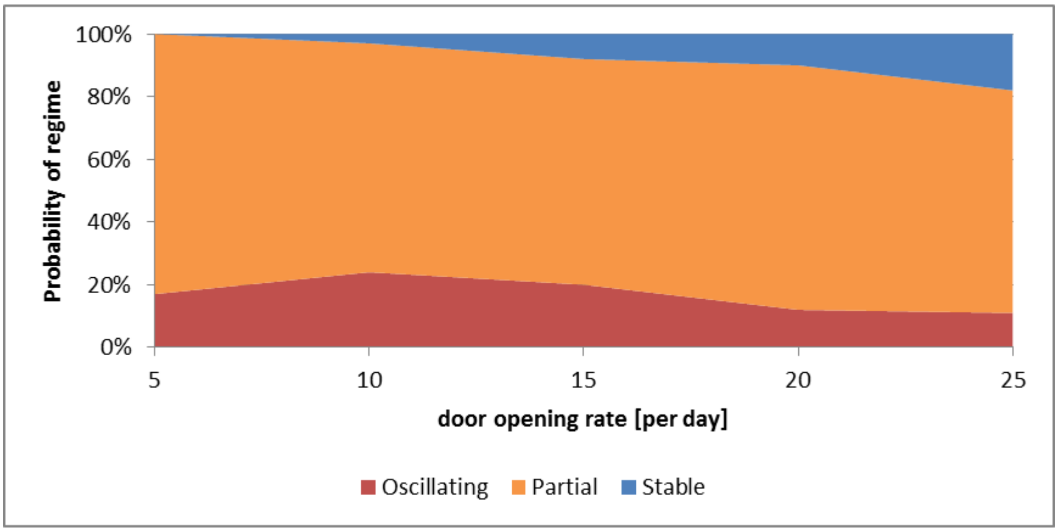


Figure 16.5: Stability of the system vs. door opening rate, for fixed controlled refrigerator share at 80%.

the same parameters as above, running again a Montecarlo experiment by doing 100 runs for each configuration setting. In this case, we have a large amount of simulations as 8 x 9 parameters were varied, which results in a total of 7200 simulation runs in the Montecarlo mode. The simulation was run for 2000 refrigerators with an adapted scale factor.

As we see in Figure 16.6, as expected with higher shares of controlled refrigerators there is a higher risk of oscillations. However, a clear impact of the disconnection frequency threshold (freqOff) cannot be recognised. A slight tendency in which at lower thresholds (< 49Hz) the oscillations occur at earlier stages, might appear, but should be analysed with more extensive Montecarlo runs.

For the partial regime (Figure 16.7) a similar conclusion as in scenario 1 can be made. At around 85% of controlled refrigerators, this regime is most likely to happen. Below and above, the probability is lower as the system is either stable, or oscillating.

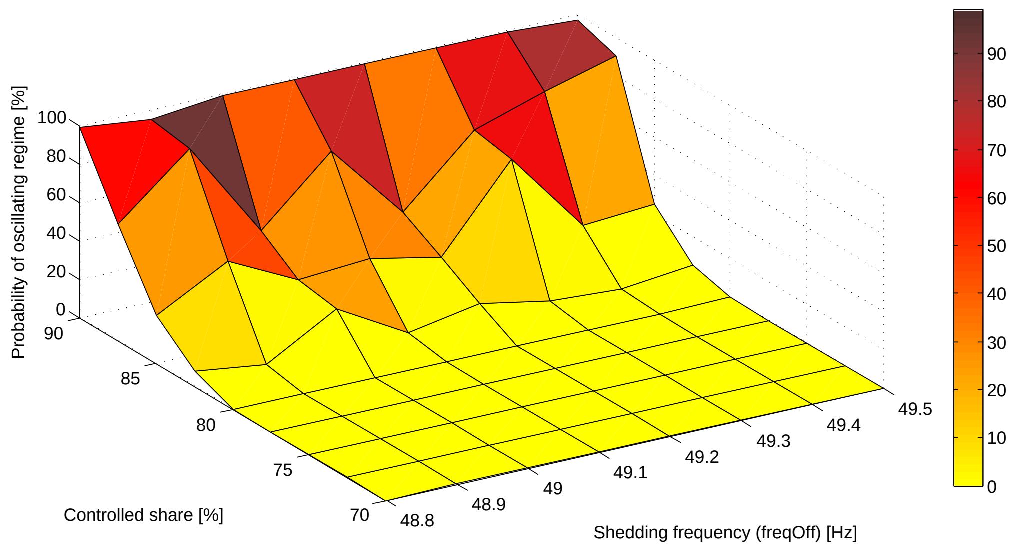


Figure 16.6: Probability of being in the oscillating regime, in function of contrRefri and freqOff.

16.3.4 **Scenario 4: Variation of the controlled share and external temperature**

In this case, we are varying two parameters using two different configurations for the refrigerators. In the first case, the initial configuration parameters of each refrigerator were recalculated (drawn from the probability range) for each new external temperature (random population). In the second case, the parameters configuration remained the same for all the variations of the others two parameters (persistent population).

This evolution from a heterogeneous set of refrigerators to a homogeneous one was made to increase the comparability of the results. Both experiments show that the parameters that define the individual refrigerator characteristics have an important influence on the global performance of the system, by shifting the solution towards the edge of chaos.

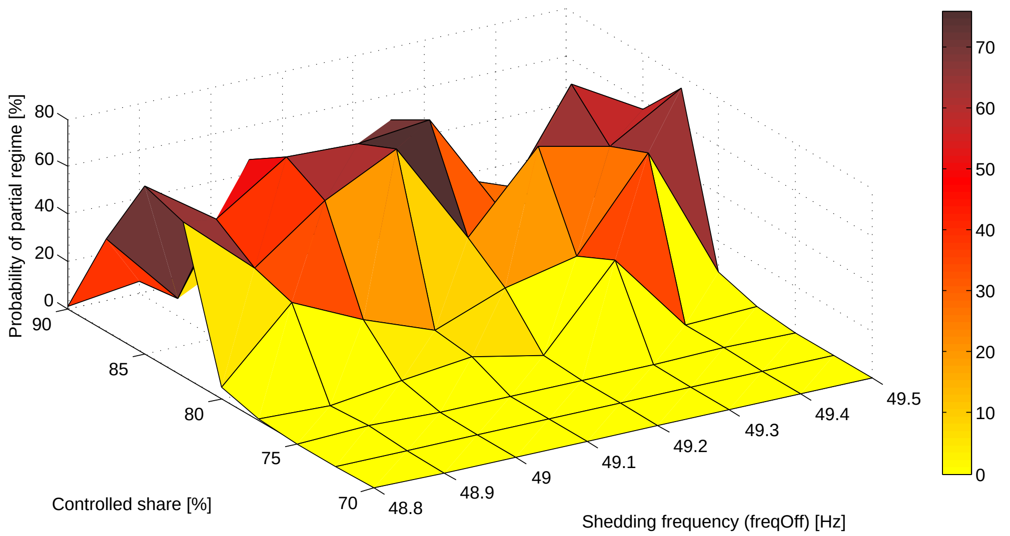


Figure 16.7: Probability of being in the partial regime, in function of contrRefri and freqOff, which represent the phase space at the edge of chaos.

16.3.4.1 Scenario 4-a: random population

This scenario, in which different configurations of the refrigerators population have been set for all parameter variations, has been simulated 1.200 times per configuration. This makes a total amount of 67200 simulations. For each experiment, 4000 refrigerators were simulated. Every refrigerator has been set with a defined external temperature, drawn from a random, uniform distribution of a fixed range. The ranges are: 20-25; 21-26; 21,5-26,5; 22-27; 22,5-27,5; 23-28; 24-29; 25-30; 26-31; 27-32 °C.

In Figure 16.8, it can be seen that there is a slight correlation between the share of controlled refrigerators and the possibilities of falling into a totally oscillating regime (completely synchronised, unstable state). These possibilities increase when the controlled share is higher - which already has been identified in Scenario 1. We can also see, that the external temperature has a similar effect. The higher the external temperature is, the higher are the possibilities of falling into an oscillating regime. However, we see that the increase of being in this regime is not monotonous, even showing some strong reduction for higher shares (between 21,5-23 °C).

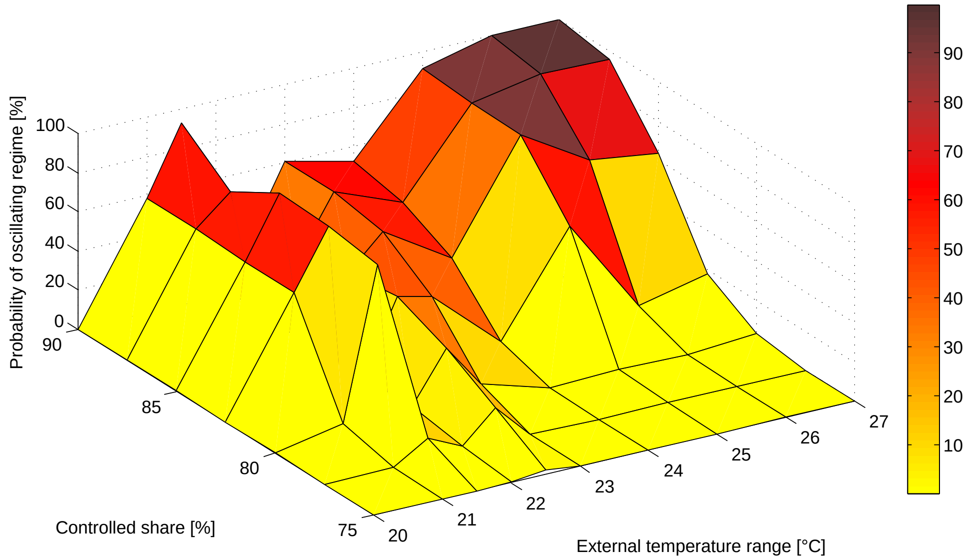


Figure 16.8: Probability of being in the oscillating regime, in function of contrRefri and external temperature.

To further explore these variations in the tendency, we included the 21,5-26,5 and 22,5-27,5 ranges. These ranges also highlighted these irregularities.

Finally we detected that the reason why the trend for growing temperatures was not monotonously increasing was in fact not due to the temperature configurations, but rather due to the initial parameters of the characteristics of the individual refrigerators (thermal coefficients and initial temperatures). These describe the refrigerators population which are differently defined at each simulation configuration. These different configurations cause the location of the edge of chaos to be shifted making the results non-comparable.

In the case of the partial regime (Figure: 16.9), a similar conclusion as in the oscillating regime case can be made. For both parameters, a slight correlation can be seen in the way they increase and how the probabilities to fall in a oscillating regime increases. In high controlled shares and external temperatures, the number of partial issues decreases since oscillating issues are happening more often.

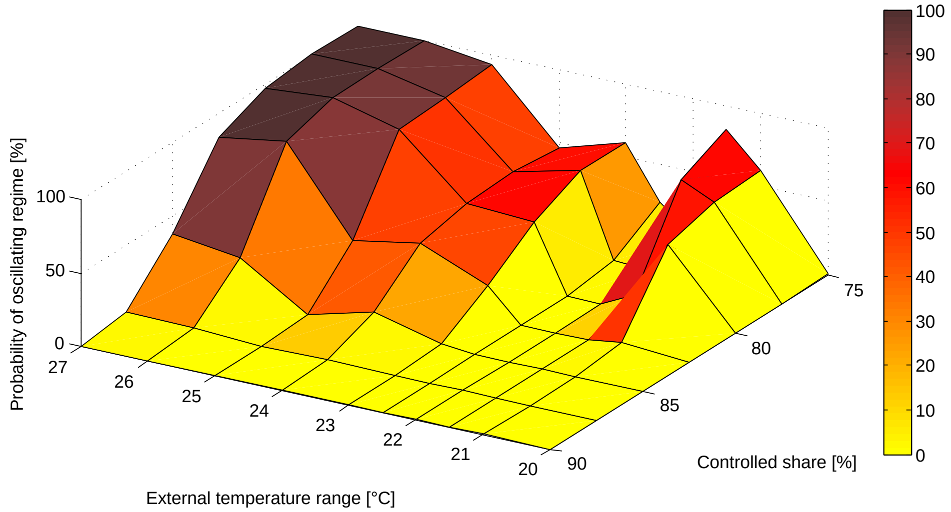


Figure 16.9: Probability of being in the partial regime, in function of contrRefri and external temperature.

16.3.4.2 **Scenario 4-b: persistent population**

This scenario, in which we have set the same initial configuration parameters for all simulations, has been simulated 1.200 times per variation. This makes a total amount of 67200 simulations. At each simulation, 4000 refrigerators were used. Every refrigerator has been set with a concrete external temperature according to the temperature ranges presented before.

In Figure 16.10, as expected the higher the share of controlled refrigerators is, the higher the probabilities of having an oscillating regime is. Furthermore, it can be seen that the external temperature has an important influence in the oscillating issues. Even for low levels of controlled share oscillating issues are achieved with high temperature ranges.

In the case of the partial (Figure: 16.11) regime a similar conclusion as in the oscillating regime case can be made. When both parameters increase the probability of falling in an partial regime increases as well. In high controlled shares and external temperatures, the number of partial issues decreases since oscillating issues are happening more often.

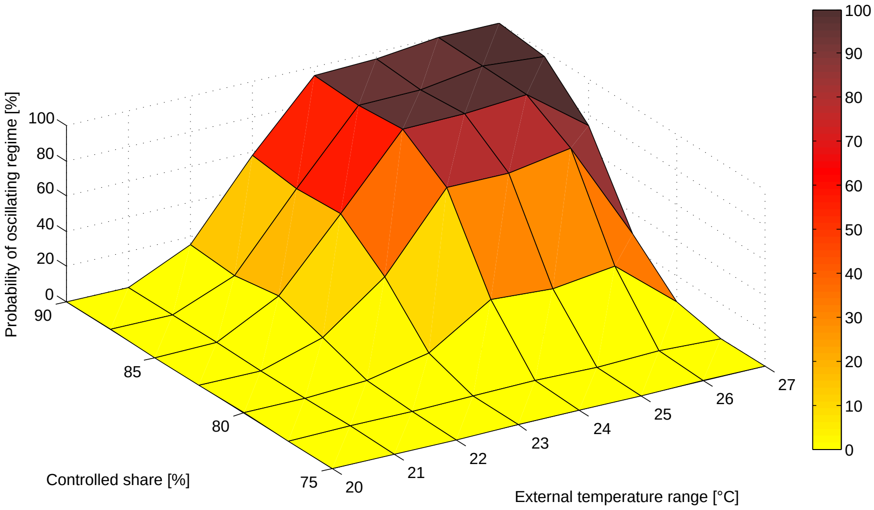


Figure 16.10: Probability of being in the oscillating regime, in function of contrRefri and external temperature.

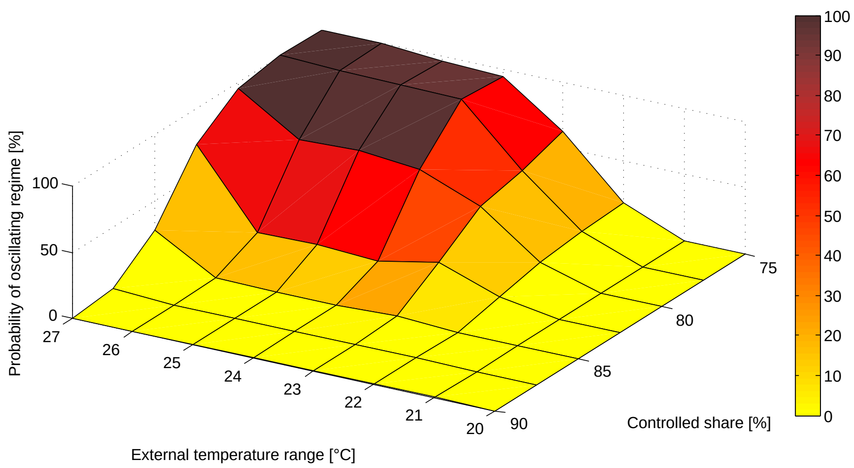


Figure 16.11: Probability of being in the partial regime, in function of contrRefri and external temperature.

16.4 Discussion

Sociotechnical systems are in constant evolution, which makes them dynamic in time and structure. This has been made clear through an example in the energy system. Using distributed SG measures which are supposed to improve grid stability, oscillating problems were detected, as the system is driven towards the edge of chaos. This criticality has been analysed empirically through simulation by using a complexity based modelling approach.

The case study shows that a border can exist, usually known as the edge of chaos, at which in only some cases critical situations occur. This border can be explored through massive simulation, allowing for evaluation of different situations and a statistical analysis of the massive simulations (Montecarlo).

For an exemplary case, the penetration rate of a DSM device was varied, observing the amount of times that the system enters a stable, partially stable or oscillating regime. This allows for quantification at which penetration rates systems remain in a safe state, and allow the identification at which point critical situations can occur. These situations occur as a phase shift, but the edge that has been studied can have a range at which only in some cases critical situations are caused. Through massive simulation these cases were found and quantified.

Furthermore, different parameters were varied to analyse if they have an effect, pushing the system towards the stable or chaotic sides. It has been shown that some parameters have no remarkable effects (door opening), and other have low effects (threshold variation), while others have significant effects (external temperature).

Simulation, and especially massive simulation on a bottom-up based approach has proven to provide a valuable tool to analyse and explore the behaviour of complex sociotechnical systems, such as the electrical grid, showing both engineering and complexity science aspects in the same model. These allow the quantification of risk in a given situation, determining cause-effect chains through a detailed observation of the systems trajectory in particular cases.

Simulation is though considered as a means to empirically explore the edge of chaos, based on a real socio-technical system which itself would be too risky to perform the experiments on.

As there is a large range of possible parameter combinations, for further studies, more of the parameters will be varied and their effects analysed. Furthermore, the sensitivity to initial conditions can be explored, which was especially high.

Following up this work, some indicators to derive the stability of the system will be designed. These indicators can give a better idea of how to quantify the criticality of the system, besides the given studies.

CHAPTER

17

Vehicle to Grid

This experiment, presented as final work of degree [1], concerns a strategy for scheduling EVs for parking lots is presented. This strategy is devoted to minimise the peaks of consumption that can be produced by a massive connection of EVs. Implemented using an optimisation algorithm, the strategy will be evaluated in a hypothetical case study where parking lots are located at the University of Las Palmas de Gran Canaria (ULPGC). The optimisation algorithm is based on PSO and watches over the grid stability by minimising the peak of demand that is produced in the morning when both workers and students plug in their vehicles. This minimisation is carried out by injecting the energy of vehicles when necessary and taking into account the users' habits. The results obtained show a considerable reduction of the peak.

17.1 Description

17.1.1 Strategy description

The objective of the strategy is to minimise the energy consumption peak that is produced when both workers and students of the ULPGC have their vehicles

[1]People that participated in this experiment: Francisco Marzabal, Mario Hernández, José Juan Hernández and José Évora.

plugged into the grid. This strategy is then compared with respect to not using a strategy. When there is no strategy, vehicles start charging their batteries as soon as they are plugged into the grid and they will stop charging whenever their State of Charge (SoC) is 100%.

What is proposed to face this issue and, thus, minimise the maximum peaks that are provoked by the massive connection of vehicles at the ULPGC is to schedule intelligently the charging process of vehicles. To this end, our strategy considers the following items:

- Vehicles do not need to start charging as soon as they are plugged in. This protects the grid from rush hours when many people arrive at the ULPGC.
- Vehicles do not need to charge until their SoC is 100%. This way, the grid of the university will not require so much energy.
- Vehicles arrive at the ULPGC with an unknown SoC.

A requirement of the strategy is to handle information regarding the users such as timetables and distances. This information is necessary for the optimiser in order to minimise the peaks. It is used to determine:

1. the SoC that every vehicle requires to go back from where they came and,
2. the time in which the vehicle must be ready with the SoC previously calculated.

This way, two important goals are achieved: minimising the peak provoked by rush hours and reducing the amount of energy consumed by the ULPGC. However, this strategy must be subjected to several constraints so that the Quality of Service (QoS) does not affect vehicle users:

- The SoC of incoming cars is not known, as mentioned before
- When a vehicle is set up to charge, it will charge until it reaches a maximum of 100% of SoC (it is not mandatory to reach this level, but it cannot be surpassed). Therefore, a vehicle with a 100% of SoC will not be able to obtain energy from the grid. This is a physical constraint.
- When a vehicle is set up to discharge (inject energy on the grid), it will discharge until reaching a minimum of 60%. (no mandatory reaching this level, but it cannot go under this one). Therefore, a vehicle with a 60% of SoC will not be able available to be used for injecting energy.

- A charge or injection can be made to a vehicle a minimum of zero and a maximum of one time per hour.

17.1.2 PSO customisation

The version of the PSO that has been used to address the implementation of the strategy exposed above is known as Binary Particle Swarm Optimisation (BPSO). This version has been selected because the states of EVs are discrete: disconnected, idle, charging, discharging. Since there are four states, every vehicle state is described with two bits so as to represent four different states.

- 00. Disconnected
- 01. Idle
- 10. Charging
- 11. Discharging

Vehicle discharging may be considered as something that can be disadvantageous from the point of view of battery performance. For this reason, this proposal considers energy transferring as a market where users who are charging the battery are paying for energy and those who are discharging their vehicles in order to help the grid are getting paid. According to this, an objective function is required to define whether a vehicle must keep its idle, charging or discharging state or when their state must be switched to another. In order to support the calculation of this objective function, a data structure was designed to represent the current state of the vehicle:

- kWhObj: amount of energy that the vehicle must have before leaving the parking lot expressed in kWh
- SoC: state of charge. From 0 to 1, it represents the amount of energy that the vehicle currently has with respect to the maximum the vehicle can have
- Mode: disconnected (00), charging (01), idle (10), discharging (11)
- DepartureTime: time in which the vehicle is scheduled to departure and, therefore, the time by which the vehicle must be charged with kWhObj.
- ID: vehicle identifier

17.1.3 Description of individuals

Considering a parking lot, each individual of the PSO algorithm is configured as an array in which, at each position, there are two bits which represent the state of a vehicle. Therefore, the size of the array is as big as the number of vehicles related to the parking lot. Every position of this array is associated with the information presented in the section above which describes the vehicle state. An example of an individual is presented below:

- Individual 1
  - Id: vehicle001
  - State: 00

  - Id: vehicle002
  - State: 01

  - Id: vehicle003
  - State: 11

Thanks to these individual structures, the proposed states of vehicles for each individual can be confronted to the data structure information in order to check whether that proposal is convenient or not.

17.1.4 Evaluation of individuals

Each individual exposes a proposal for the states of all vehicles within a parking lot. The way in which the fitness function is calculated is according to the next instructions that are evaluated for every vehicle state that is proposed:

- If charging is proposed
  - At this point, the difference between the current energy and kWhObj is calculated. If this difference is lower than 0, the decision is weighted positively. Otherwise, the decision is weighted negatively. Outline cases, as vehicles which have just a 20% of SoC are especially considered. If this happens, the decision is weighted with the maximum possible value indicating that it's almost mandatory to apply a charge for that vehicle.

  - The departure time is also confronted with the time that is necessary to reach the kWhObj. If the time to charge the vehicle is not too long, this decision will be weighted more positively.

- If discharging is proposed
  - Again, the difference between current amount of energy and the objective one is calculated. According to the result, if the difference is higher than 0, it will be positively weighted. Otherwise, it will be negatively weighted. If the current energy is too low with respect to the objective one, this will be very negatively weighted.

- If idle is proposed
  - If the current energy of the battery is close to the objective one (in a range of 300Wh) then this decision is positively weighted. Otherwise, it will be negatively weighted.

Therefore, the evaluation of every vehicle's proposal state is aggregated by summing, obtaining a value that summarises the fitness of the individual. Apart from this, another factor that influences the fitness of the individual considers the amount of energy that is necessary to be extracted from the grid. In this sense, the parking lot will try to minimize the amount of energy that is coming from the grid. How is this possible? The parking lot will try first to use the energy stored on the parked vehicles before requesting energy from the power grid. The way in which this is approached is by weighting positively values of energy demanded to the grid that are closer to 0W. In this sense, the fitness function that is used to evaluate the quality of an individual is formulated as follows:

$$f_i(t) = \alpha * calculus1_i(t) + (1 - \alpha) * calculus2_i(t) \tag{17.1}$$

Where i is the individual that is being evaluated. α is a parameter that weights which calculation has a greater impact on the fitness function. Calculation1 is the one that obtains a fitness value depending on the vehicles' state configuration. Calculation2 concerns the minimisation of the energy that is demanded to the power grid.

The way in which individuals evolve according to the PSO algorithm is made by using two dimensions of velocity and position. This is due to the fact that there are two bits that codify the state of the vehicles. This calculation is made as follows:

$$v_{id}(t+1) = \omega * v_{id}(t) + \phi * rand_1 * (pBest_{id} + x_{id}(t)) + \phi * rand_2 * (gBest_{id} - x_{id}(t)) \tag{17.2}$$

$$x_{id}(t+1) = x_{id}(t) + v_{id}(t) \tag{17.3}$$

Where v and x are the velocity and position of the individual i. d is the bit that is being updated (first or second bit of the vehicles' state).

17.2 Case study

In this section, the strategy that has been previously described is going to be tested. To this end, a scenario has been designed containing several parking lots that are located at the University of Las Palmas de Gran Canaria (ULPGC). This Spanish university is located in the Canary Islands. This place has been chosen for two main reasons:

- It is the working place of the people that have been involved in this study
- It is a good place for investing on EVs since the distances are short on this island and there are good weather conditions for renewable generation of energy. This energy can be more efficiently managed if vehicles are introduced on the power grid as they would become a flexible load.

17.2.1 Model elements

The scenario is populated with the following kinds of entities: electric vehicles, parking lots, charging stations and political areas. Before getting into the details of how these elements are modelled, the relation between them must be explained. EVs will move from one political area to another. A political area can represent an urban node in which vehicles are. A movement from one political area to another will convey a consumption of the batteries' energy. Therefore, the place where vehicles are coming from to the university is considered a political area and the university itself too. Within the university, several parking lots will be placed according to the existing facilities that the university has for this purpose. At each parking lot, several charging stations will be placed in accordance with the real amount of available parking spaces.

Since political area and parking lots are just aggregators of vehicles, they are not further described. They only contain a feature which is the location in which they are. This location allows us to calculate the distances that exist among political areas and parking lots which is used to calculate the battery drain that trips convey. The next sections will describe how EVs and charging stations are modelled.

17.2.1.1 **Electrical vehicle**

A vehicle is mainly described by its battery, which indicates the range that a vehicle can cover according to the SoC. For this reason, the main features that describe a vehicle are:

- Capacity: maximum capacity of the battery kWh
- ConsumptionPer100Km: amount of kWh drained every 100kms
- Max Input: maximum amount in Watts that the vehicle is able to receive for charging
- Max Output: maximum amount in Watts that the vehicle is able to inject when discharging
- Current In/Output: power that the vehicle is obtaining/injecting from/into the grid in Watts at a concrete time
- Current load: it is the SoC of the vehicle at a concrete time.

From the point of view of dynamics, vehicles execute two different processes at each time. The first one concerns their movement. At each time step, the position of the vehicle is updated. Secondly, the battery SoC is calculated in accordance to this movement in order to emulate the battery draining.

17.2.2 **Charging station**

Charging stations are placed in the parking lots and there are as many units as parking spaces available in each parking lot. Each charging station is the entry point of the Electrical Vehicle (EV) to the power grid. The charging stations are described in the scenario as follows:

- Input Voltage
- Input Current
- Frequency
- Phase System
- Output Voltage
- Output Current
- Output Power
- Mode. Any of:
  - OFF
  - IDLE
  - ON DISCHARGE
  - ON CHARGE

17.2.3 Scenario

The scenario is composed by the model elements that have been previously described. Therefore, using them, the ULPGC will be modelled in order to address the issue that is being approached: the massive charging of vehicles. To do so, political areas, parking lots, EVs and charging stations must be especifically defined for this scenario.

To sum up, the amount of entities that are going to be present during the simulation is detailed in the list below:

- EVs: 2000
- Political areas: 6
  - 4 in Las Palmas de Gran Canaria
  - 1 in Telde
  - 1 at the University of Las Palmas de Gran Canaria
- Parking lots: 11
- Charging stations: 1513

17.2.3.1 **Political areas**

Apart from the university political area, the areas where vehicles are coming from to the university must be defined. In order to do so, a small survey has been made. This survey revealed that vehicles are mainly coming from two main places: the cities of Las Palmas de Gran Canaria and Telde. Therefore, this can be modeled by using two political areas whose location is the city centre of both cities. This can work perfectly for Telde since it's not a big city, but this does not work so well for Las Palmas de Gran Canaria as it is a larger city. In this sense, Las Palmas de Gran Canaria has been split into four different zones so that the distances that vehicles have to cover to go to the university are more accurate. The list below provides the concrete locations of the political areas that are used within the model:

- Las Palmas de G.C. (Zone 1): 28.136410, -15.437508
- Las Palmas de G.C. (Zone 2): 28.125132, -15.430641
- Las Palmas de G.C. (Zone 3): 28.125132, -15.430641
- Las Palmas de G.C. (Zone 4): 28.115972, -15.437078
- Telde: 27.996220,-15.416737

Since the university political area is composed by the parking lots that are inside, its location is not given since, in this political area, the locations of parking lots will be used. These locations are provided in the next section.

17.2.3.2 **Parking lots**

At the university there are many parking lots scattered across the campus. However, in this study, eleven of them have been used in order to do this study. These eleven parking lots have been chosen as they had the highest amount of parking spaces. Note that private parking lots are only accessible by professors and administrative staff of the University. The list below describes these parking lots:

1. Law Faculty (340 spaces): 28.080609,-15.450591
2. Technology Park (100 spaces): 28.079937,-15.451664
3. Technology Park 2 (43 spaces): 28.080041,-15.452759
4. Law Faculty 2 (166 spaces): 28.079265,-15.449256
5. Economics Faculty (126 spaces): 28.077712,-15.450146

6. Economics Faculty 2 (20 spaces): 28.076472,-15.450481
7. Computer Science faculty (217 spaces): 28.072037,-15.451278
8. Computer Science faculty 2 (91 spaces): 28.073571,-15.452303
9. Telecommunications faculty (134 spaces): 28.072108,-15.453494
10. Engineering faculty (132 spaces): 28.071213,-15.454856
11. Engineering faculty 2 (144 spaces): 28.070532,-15.456221

17.2.3.3 **Electrical vehicles**

In order to be more realistic, real EVs have been modelled in this scenario. The vehicles that have been modelled are the ones that are the most popular at the moment. The table below details their main features, which have been translated into model parameters for the simulation:

Brand	**Model**	**Capacity**	**Range**	**Charging time**
Think	City	25kWh	200km	8h
Mitsubishi	i-Miev	16kWh	130km	7h
Citroen	C-Zero	16kWh	130km	7h
Renault	Fluence-ZE	22kWh	160km	6-8h
Reva	L-Ion	11kWh	120km	6h
Nissan	Leaf	24kWh	160km	8h

Based on this information, EVs have been modelled using two main features: battery capacity and consumption per 100kms. Based on these two factors, calculations for the SoC of the battery can be made.

These vehicles will emulate the flows of vehicles within the different political areas. This emulation is done by assigning a concrete profile of driver to each vehicle. There are four different profiles of drivers that have been considered in this study: student, professor, administrative staff and visitor. According to the profile, these actors will behave differently. The next paragraphs provide a description of their behaviours.

Every student is randomly associated with a faculty. This faculty will imply a reference to a preferred public parking lot for a concrete student. Apart from this, a political area is assigned to the student which will represent the place in which the student lives. This choice is made following a random criteria. Moreover, it is necessary to set the times of arrival and departure. Firstly, it is determined whether the student will go to the university in the mornings or evenings. If it is in the mornings, the time to arrive is determined following a Gauss distribution centered on 08.30 with a deviation of 2 hours. The time to leave follows the same criteria being centred at 13.30 with a deviation of 2 hours. In the case of students coming in the evening, the same criteria is used to arrive having 16.30 as centre with 2 hours of deviation and leaving at 20.30 with 2 hours of deviation. The ratio of students coming in the mornings is 70%.

Professors are set according to exactly the same criteria as students. This is, the working place is chosen (referring in this way to the closest private parking lot) and then the time of arrival and departure are set following exactly the same Gauss distributions as exposed above. The ratio of professors coming in the mornings is also 70%.

Administrative staff use private parking lots and have a different timetable. Their time of arrival is centred at 08.00 with a deviation of 15 minutes while the time of departure is 14.00 with a deviation of 15 minutes as well.

Visitors use public parking lots and their arrival times are in a range that goes from 10.30 until 13.30 with a deviation of 2 hours. Note that this profile does not come to the university in the evenings.

17.2.3.4 **Charging station**

Charging stations are located within parking lots and there are as many as parking spaces available within every parking lot. They are responsible for connecting EVs to the power grid. PSO algorithm will work on these units to command them in order to achieve the energy objectives. Each parking, which consists of several charging stations, will execute the PSO to configure the batteries charging. The parametrisation of the PSO that is executed at each parking is exposed in the table 17.1

17.3 Results

This section discusses the main results based on the outputs provided by simulations of the scenario described above. It must be pointed out that part of the results of this study are not only the graphics themselves, but also the simulation

Table 17.1: PSO parametrisation used.

Parameters setting	**Value**
Number of vehicles	From 0 to max number of spaces
Time step	5m
Number of particles per each PSO	30
Stop condition	30 iterations
Mutation, c1, c2, w	25%, 1, 1, 1
Particle size	Vehicles plugged to charging stations

framework which makes it possible to keep working on more new policies for EV management.

To compare the benefits of the conceived strategy, it will be compared to the use of no strategy. To this end, two simulations have been executed with a time interval of 24 hours. On the one hand, the first simulation runs the scenario considering the strategy that every vehicle starts charging as soon as they arrive and always until getting a SoC of 100% (as long as they do not leave before it is reached). On the other hand, the second simulation runs the same scenario but using the strategy that has been proposed in this study.

In figure 17.1, the sum of the power that is being consumed by all the charging stations within the university is shown along a day. The light gray line regards the case in which the strategy consists in starting to charge vehicles as soon as they come without any optimisation. Behavioural patterns of the different agents' profile can be perceived from the shape of the line. The maximum peak of the line is reached at 08:30 approximately with a value of almost 1.4 MW. Apart from provoking this significant peak, this policy causes a large valley at noon since, at this time, most of the vehicles either have left the university or are fully charged (100% of SoC). Another interesting detail is the second peak that appears in the morning which is due to the visitors who come to the university. The energy that is consumed in this simulation is 3.298MWh.

In figure 17.1, the results of our strategy in terms of power consumption are shown in dark gray. As an optimisation method is being used to control the charging of vehicles, behavioural patterns of the different agents' profile cannot be perceived as clearly as in the other strategy. The maximum peak of the line is reached at

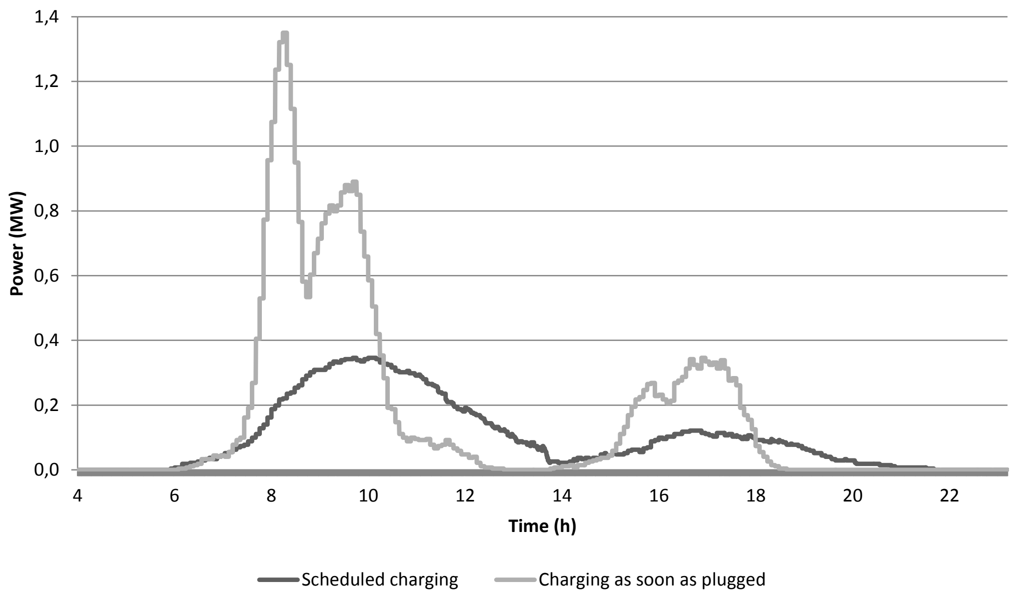


Figure 17.1: This chart shows the total power consumption of all the charging stations of the scenario for both strategies: scheduled charging and as soon as plugged charging.

10:00 approximately with a value close to 0.35 MW. The distance between peaks and valleys in this strategy is small. In this strategy, the total amount of energy that was demanded was 2MWh.

Apart from this, the energy consumption of the university's parking lots has also decreased from 3.298MWh to 2MWh (40% of reduction). This reduction is also important since this way, consumption is being shifted to better time intervals when there are not usually peaks in a power grid. The way in which this goal has been achieved has two main sources:

- Vehicles do not charge until 100% as in the other strategy. They charge until the amount of required energy is reached based on the knowledge that parking lots gather from the users.
- Vehicles that arrive with more energy than the objective energy they need to leave are used to discharge. This is, if a vehicle is plugged with a 90% of SoC and its objective SoC to leave is 70%, this vehicle can be discharged until 70% without having to be charged during its stay.

17.4 Discussion

The strategy proposed in this study means a significant improvement on power grid management. It has been experimentally evaluated, through simulations, and its results are quite interesting. Concerning the reduction of peaks, this strategy has managed to reduce the maximum peak 75%. This is an important achievement since this involves money that can be saved from the investments on power grid facilities.

The use of PSO to develop this strategy has been core. In a system where there are many agents each of which is self-interested, SI techniques are helpful to handle all these constraints. This approach has provided enough flexibility to introduce all required constraints offering good global and individual results.

CHAPTER

18

Frequency Management with Swarm Intelligence: a case study in Smart Grids

This experiment, that will be sent for publishing [1] test autonomous load shedding strategies using smart appliances. Since, the grid frequency can be measured from any power outlet, a smart appliance connected to the grid could know the current imbalance on the grid. This means that it is possible to design a smart appliance that can disconnect or reconnect itself to the grid in order to improve the stability.

18.1 Description

Based on the grid frequency measure, individually and locally, each smart appliance decides whether to stop consuming, to support the reserve mechanisms of the power grid and increase the stability of the system. When frequency reaches normal operating margins, appliances decide to restart working in their normal mode.

[1]People that participated in this experiment: José Évora, José Juan Hernández, Mario Hernández and Enrique Kremers.

This mechanism seemed to work well to stabilise the system after major events like the breakdown of a production unit.

The selected case study is based on the works described in [KdD12b, KdD12a]. A simulation model of a domestic refrigerator (micro-level) is coupled with a simplified model of an electrical energy system (macro-level). Smart refrigerators are present in many households as appliances which are connected to the power grid.

However, the studies also detected that, under certain circumstances, when increasing the number of controlled refrigerators, non-desired effects like oscillations could occur. The model allowed for a first analysis by probabilistically quantifying the risk of such an event.

In these simulated experiments, different policies that may be autonomously executed in each refrigerator are analysed. These policies make smart refrigerators react according to significant imbalances locally detected through the frequency.

Since it is a collective behaviour, we expect to have global emergent behaviours as a result of dealing with the frequency based on these simplistic local policies. The resultant implementation is a SI system in which the frequency management is decentralised.

18.2 Case study

The scenario that is used for evaluating the policies consists of 4,000 households with a refrigerator. 4,000 households is a representative sample that can be easily simulated. The results of simulations over this scenario are not substantially different from simulations in which more refrigerators were used. Households has been set with an external temperature that has been set on an uniform distribution between 20-30$\circ$ Celsius.

75% of these refrigerators are smart appliances, that is, controllable. This is, 3,000 refrigerators will be monitoring the grid frequency and making decisions on its own consumption.

The simulation implements a generation unit failure. This failure causes an abrupt drop in the frequency from 50 Hz making the power grid unstable. This drop is due to the production fall while the consumption remains the same after the failure. In figure 18.1, the effect of this failure in the frequency is presented. This is the reference scenario where there are no refrigerators that can disconnect

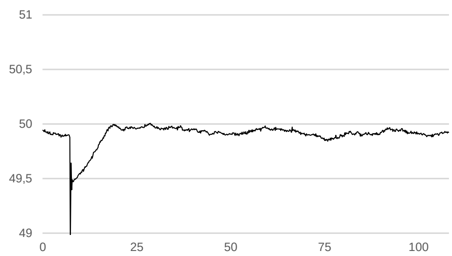


Figure 18.1: Frequency reaction facing a generation unit failure.

themselves to adjust the demand.

However, the hypothesis is that demand side may help to improve frequency recovery. If consumption were able to follow this production fall, the frequency could recover faster to the nominal value at 50 Hz. This can be done by Smart refrigerators that are able to adapt their consumption disconnecting/reconnecting when frequency drops.

In the subsections, different demand side policies are presented and experimentally evaluated. The faster they recover this production fall, the better the policy.

To evaluate the performance of the policies, three indicatesators have been defined. These indicators are oriented to allow for a quantitative comparison among the different policies. In figure 18.2, indicators named as I_i show which information is used to calculated them. T_H and T_L are thresholds that delimit what we consider the normal operating values of the frequency. We consider that the system becomes *unstable* whenever frequency exceeds these thresholds. The following list shows how these indicators are calculated:

- I_1: is the "total instability". This is calculated as the area of the frequency curve over or under the thresholds. It is the pattern filled area in the figure 18.2.
- I_2: is the "instability length". It is calculated as the average duration of unstable periods.
- I_3: is the "unstable time". This is, the duration of unstable periods.

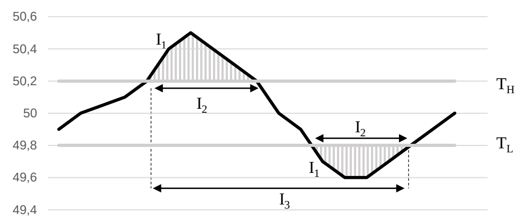


Figure 18.2: Indicators that are used to evaluate policies.

18.3 Results

18.3.1 Policy #1

This first policy consists in monitoring the power grid frequency and react according to detected events. This policy considers an event whenever the monitored frequency is lower than 49Hz.

If this occurs, the policy switches off the refrigerator stopping its consumption instantaneously. However, it keeps monitoring in order to reestablish the refrigerators state as soon as the frequency gets higher than 49Hz. This policy is the same as that used in [VKE$^+$13, KdD12a].

In order to test this policy a simulation was performed in which all controllable refrigerators of the previously described scenario were programmed with this policy. Fig. 18.3 presents the obtained results of the frequency. As it can be observed there is a first big oscillation followed by smaller ones. It can be considered that the system reaches the stability in the second 16,000. Comparing with the base situation in which no refrigerator is being disconnected, this policy clearly produces worse results than not reacting to the event (Figure: 18.3).

18.3.2 Policy #2

The second policy is based on the same idea, but the way in which it reacts changes. Here, the policy keeps monitoring the frequency grid. Whenever the frequency is lower than 49Hz, it makes a decision.

When this event happens, each refrigerator will get a random integer number in a variable range uniformly distributed. This number represents the sleep time before making the decision. Once this time is out, the refrigerator will check the

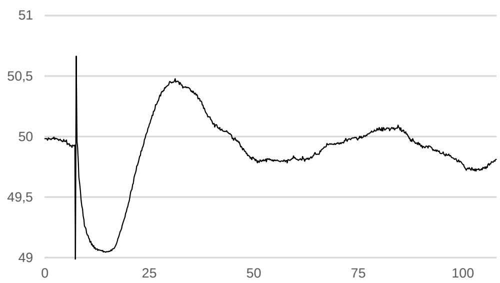


Figure 18.3: Frequency reaction facing the failure and the massive disconnection and reconnection of refrigerators. Policy #1. Axis Y: frequency (Hz). Axis X: time (seconds).

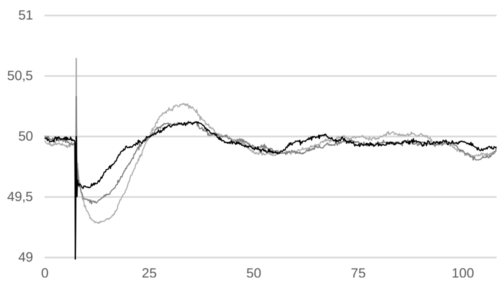


Figure 18.4: Frequency reaction facing the failure and the use of policy #2 with different configurations.

frequency state and act accordingly. If the frequency is lower, the refrigerator will stop consuming, if not, it will keep working in its normal operation mode.

The idea behind this policy is to avoid every refrigerator reacting at the same time to the frequency event as this provokes a higher impact in the power grid [KdD12b]. In fig. 18.4, frequency curves are presented for three different parametrization of this policy. In the first one, the range is between 0 and 10 seconds (light grey). The second (dark grey), between 0 and 20 and the third on between 0 and 30 seconds (black).

It can be observed that the longer the range is, the faster the frequency recovers its normal operation mode. Furthermore, oscillations are also softer with longer ranges. This policy clearly improves the first one but it is still worse than the reference scenario.

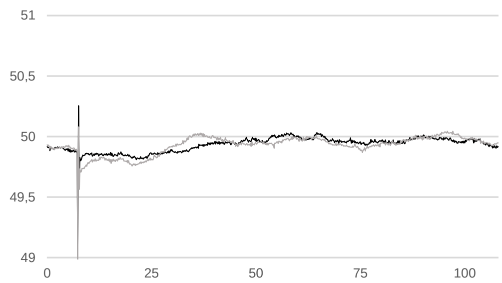


Figure 18.5: Frequency reaction facing the failure and the policy #3 with different configurations.

18.3.3 **Policy #3**

Following the same idea of the second policy, in this third policy a second delay is established. Thus, this policy uses two different delays: the disconnection delay and the reconnection delay. In the previous policy an oscillatory effect can be observed. This effect happens because all refrigerators are reconnected at the same time (when frequency recovers). This results in that all refrigerators having a synchronized behaviour, whereby they consume and stop consuming at the same time. This synchronization effect takes time to disappear.

This third policy not only applies the same delay in the disconnection, but another in the reconnection. As before, when a frequency event is observed by the refrigerator, it calculates a delay between 0 and a parametrized number. When that time is reached, frequency is checked again in order to decide whether it is necessary to stop consuming or not.

The second delay works in the same way. If the refrigerators is disconnected and frequency reaches 49,8Hz, a delay is calculated between 0 and a second parametrized number following a uniform distribution. When the delay time is reached, the frequency is checked again to see if it is still over 49,8Hz. In case it is not, the refrigerator will wait until the frequency reaches that level and the delay will be recalculated again. In figure 18.5, two different configurations of this policy are presented.

As it can be seen, oscillatory effects have almost disappeared and in a short period of time frequency gets stabilized. The first configuration seems to be better as frequency corrects faster in a good level over 49,8Hz.

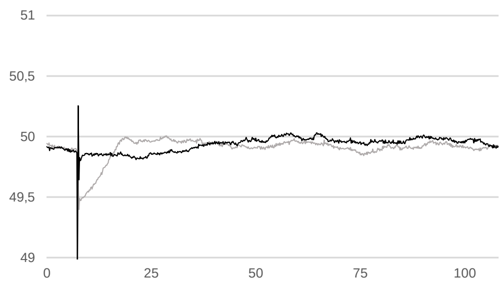


Figure 18.6: Comparison between best policy (black) and base case (grey) in which no refrigerator reacts to the frequency event.

Table 18.1: Indicators for both cases.

Indicator	Base case	Policy #3
I_1	91.15	22.23
I_2	452	23
I_3	452	23

18.3.4 Evaluation

In order to quantitatively evaluate these policies, indicators of stability and resilience are used.

In this section, we discuss the best solution of the policies presented against the base case. On the one hand, in the base case, no refrigerator is used to help frequency to improve. On the other hand, the third policy using a disconnection delay of 40 seconds and a reconnection delay of 600 seconds for the refrigerators. In fig. 18.6, it can be observed that this policy produces more stable results at the beginning of the frequency event. Frequency gets a good operating point faster than in the base case.

In table 18.1, the values of the indicators for both cases are presented. The smaller the value for the indicator, the better. As can be observed, the area outside thresholds is bigger for the base case than the policy #3 case (Indicator 1). Other indicators also reveal that the recovery time is higher for the base case.

18.4 Discussion

In this study, a hypothesis of decentralized control for the frequency management in power grids has been proposed and tested. Many smart devices on the demand side are observing the grid frequency and individually acting by disconnecting or reconnecting itself to the grid.

From the point of view of the system behaviour, this decentralised control can be analysed as a SI system. The emergent behaviour that comes out from individual behaviour should exhibit a stability and resilience improvement. The engineering of these two properties is complex since synchronized effects should be avoided when many agents coincide in the decisions they make.

Many experiments have been executed and it can be observed that small variations in device behaviour involve meaningful effects in the grid stability and resilience. Experimentally, it has been observed that Swarm Intelligent policies may improve the base case in which there are no smart devices that help to stabilize the grid. However, using a Swarm Intelligent policy is not always a guarantee for achieving an improvement. This is the case of the first presented policy.

Policies should be defined and adjusted to obtain the desired improvements. In this sense, it has been seen that deterministic policies involve synchronization among devices and accumulative effects that cause undesired instabilities. Even more, it was experimentally observed that policies that implement random disconnection and reconnection times give better results than deterministic ones.

However, these experiments has been carried out applying policies over one type of smart device, in this case refrigerators. It could be interesting to make further experiments with larger scenarios and different device types. Furthermore, other policies regarding cooperative behaviours can be further explored.

CHAPTER

19

Agent-based modelling for designing an EV charging distribution systems: a case study in Salvador of Bahia

This experiment, published in [EHBS14] [1], concerns a strategy for scheduling EVs charging for Salvador de Bahia city. This strategy is devoted to minimise the peaks of consumption that can be produced by a massive connection of EVs. Implementing a dynamic pricing, the strategy will be evaluated in a hypothetical case study with a predicted population of EVs for 2030 in Salvador de Bahía. The results obtained show a considerable reduction of the peak.

19.1 Description

The development of this model has required to use data of different natures. The next paragraphs below summarise the data that has been used and its justification.

[1] People that participated in this experiment: José Évora, José Juan Hernández, Daniel Barbosa and Paulo Sampaio.

Demography. Current number of inhabitants in Salvador of Bahia and its projection until 2030 (IBGE[Fun10]). This is relevant information for calibrating the original consumption load curve and calculate the number of vehicles.

Vehicle ownership. This is a future projection from now until 2030, calculating the number of vehicles per 1,000 inhabitants. In [DGS07], this projection is estimated for many countries, including Brazil.

Electrical Vehicle penetration. This projection provides the ratio of vehicles that is foreseen to be in coming years. In [SSS09], the author provides some penetration rates concerning whether public institutions fund the purchase of electrical vehicles or not. In our model, we decided to take the average of all the scenarios the author proposes. It must be noted that the objective of this paper is not predicting the number of EVs in Bahia, but analysing their impact.

Vehicle usage. This is an statistical data concerns the number of kilometres that are driven by the inhabitants in Salvador of Bahia from IBGE[Fun10].

Consumption. This information concerns historical data for electrical power that is consumed in a portion of Salvador of Bahia City serving approximately 142,000 inhabitants. This information will be used to compare the impact that the EVs will have when they are massively introduced into the power grid.

Due to the difficulty in obtaining complete information from the electrical company, this study uses data from one typical district in Salvador with residential, commercial and industrial sectors, including schools and faculties, malls, etc. The substation in this district consists of two 69kV/11.9 kV and 20 MVA three phase power transformers in operation and 10 distribution feeders that it supplies 58,308 consumer units in low and medium voltages.

Since the predicted penetration of EVs from 2014 to 2029 is not significant, the model has been made for a prospective scenario in 2030. Using the predictive studies, the EV fleet has been sized for the modelled district in 2030. According to the prospective studies in [SSS09], the penetration rate of EVs in 2030 is about 10.4% when averaging the most optimistic and pessimistic perspectives. For the population of the portion under study, this means 5,253 vehicles. For each of these vehicles, an agent and a charging station are attached.

Table 19.1: Vehicles

Brand	**Model**	**Capacity**	**Range**
Mega	e-City	9kWh	100km
Reva	L-Ion	11kWh	120km
Think	City	25kWh	200km
Mitsubishi	i-Miev	16kWh	130km
Citroen	C-Zero	16kWh	130km
Renault	Fluence-ZE	22kWh	160km
Nissan	Leaf	24kWh	160km
Tesla	Roadster 42	42kWh	257km
Tesla	Roadster 70	70kWh	483km

19.2 Case study

This section is devoted to explain how models have been developed. The EVs have been modelled to represent any EV in the market. It is easily parameterised by simply providing the battery capacity (in kWh), the autonomy (in km) and the starting state of charge (from 0% to 100%). In this simulation, vehicles in table 19.1 have been taken into account.

These vehicles are discharged based on the mileage they cover. This means that the energy that has to be discharged from the battery is calculated based on the kilometres that have been driven and the autonomy of the vehicle. Vehicles are charged according to the standard which is at 3,700 watts.

This electrical vehicle model needs to be commanded by an agent model in order to perform the trips (e.g. going to work). Therefore, an agent has been developed to implement the behaviour of the driver who needs transportation. This agent has been implemented in accordance to ordinary lifestyles, since it was not possible to obtain information of usage patterns for vehicles. These agents execute four activities which concern the use of the vehicle: going to work, returning from work, going to shops, returning from shops. An example of how these agents are scheduled is presented below:

- Activity 1. Id: go work. Time: 7:00. Deviation: 500s. Pattern: Monday, Tuesday, Wednesday, Thursday, Friday.

- Activity 2. Id: back from work. Time: 15:00. Deviation: 500s. Pattern: Monday, Tuesday, Wednesday, Thursday, Friday.
- Activity 3. Id: go shop. Time: 18:00. Deviation: 500s. Pattern: Tuesday, Thursday, Sunday.
- Activity 4. Id: back from shop. Time: 21:00. Deviation: 500s. Pattern: Tuesday, Thursday, Sunday.

In addition to this information, agents are also provided with the distance to work and to the shops. These distances are calculated based on a normal distribution centred in the average mileage driven by people in Salvador de Bahia. After each activity, the agents will plug in the EVs so that they can start charging their batteries. However, depending on the strategy, the connection of the vehicles to the grid will not necessarily involve its charging, as this will be decided by the strategy. This means that in the simulation in which there is not a strategy, EVs will start to charge as soon as they are plugged in. In the simulation in which there is a strategy, EV charging will start in accordance to the policies of the strategy.

19.3 Results

Keeping the same distribution infrastructures, the charging potential of the substation for electrical vehicles can be defined as the difference between the maximum deliverable power to vehicles connected to charging stations and the current energy demand: $Pp(t) = Pm(t) - Pd(t)$. Obviously, this potential is not the same in all power grid substations, since both factors, Pm and Pd, may vary from one substation to another as well as seasonally.

In the proposed charging model that this paper explores, one or more substation feeders will be enabled to allow the management of all vehicles based on these two factors. Ideally, the energy used for their charging should be the cheapest. Moreover, system operation should be facilitated so that generators are switched on and off as little as possible. To this end, the objective should pursue the flattening of the demand curve.

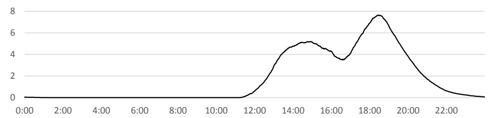


Figure 19.1: EVs consumption when they start charging as soon as they are plugged

19.3.1 No DSM policy

In this experiment, the model previously presented will be simulated considering that EVs start charging as soon as they are plugged in. In the figure 19.1, the consumption of EVs for a weekday is presented. The highest peak of this load curve can be found at 7pm approximately. This peak is due to the fact that almost everyone is at home at this time, with their EV plugged in, with the result that that their charging reaches the greatest level of overlapping. This overlapping effect is also high in the morning and afternoon, coinciding with commuter “rush hours”.

In the figure 19.2, two load curves are presented: the original load curve for 2014 (grey) and for 2030 in which EVs are included (black). In this chart, it can be seen the increment that involves the EVs in the total consumption. As a remainder, this significant impact is caused by 5,253 EVs in a power grid serving an area of 142,000 (in 2014) and 149,000 (predicted for 2030) inhabitants. If any of the three predictions that are used in this study increases (population, vehicle ownership rate or EVs penetration rate), would result in the consumption increment being higher. It can be observed that the impact of introducing 5,253 EVs can reach 8 MW, which is significant considering that the maximum consumption of this area for the analysed day is 103 MW. This approximately means an 8% increment. Along with this increment, we must also consider how the rest of the energy demand not provoked by the EVs introduction has increased as well. This leads to a total increment of approximately 15MW. This would mean that this substation would have to be re-sized.

19.3.2 RTP-based DSM policy

Knowing the valley existent in the early morning, RTP policy has been programmed to output a very cheap price for the energy that is consumed from 1am to 8am. In these experiments, all vehicles have been tuned to only accept this price. In the

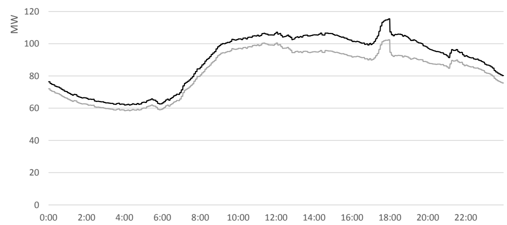


Figure 19.2: Comparison of the original load curve for 2014 (grey) with the total consumption including EVs charging for 2030 (black)

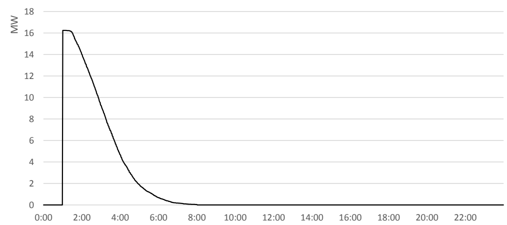


Figure 19.3: EVs consumption according to the RTP based policy which makes cheaper charging from 1am to 8am

figure 19.3, the consumption of the EVs is presented under this new policy. As it can be observed, since many vehicles have not been charged during the previous days, most of them start charging their batteries at 1am, thus producing a high demand peak of 16MW. This consumption becomes 0 at 7am approximately.

In the figure 19.4, again, two load curves are presented: the original load curve for 2014 (grey) and for 2030 in which EVs are included (black). This time, the peak that already existed in the load curve for 2014 has only increased as consequence of the population increase for 2030. This time, EV charging has been shifted to an off-peak period which is in the early morning. In this new interval, EVs charging does not increase at all the overall maximum peak of the demand allowing for a greater introduction of EVs if it was necessary. Notwithstanding, there is an abrupt increase of the consumption at 1am since almost all EVs start charging at that time due to the low prices. This is a non-desired effect for the grid stability and should be further studied in order to smooth their charging starting process.

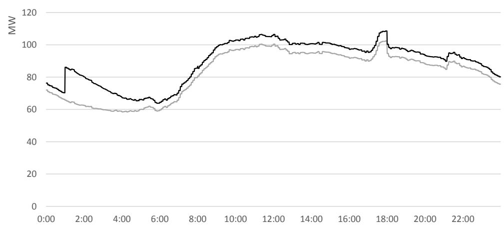


Figure 19.4: Comparison between original load curve for 2014 (grey) with the total consumption including EVs charging (black) using the RTP-based policy for 2030

19.4 Discussion

This work has been oriented to develop the model of Salvador of Bahia's power grid using an Agent-based modelling approach. This work has allowed to validate the ability of this modelling approach to:

1. Model individual decision making.
2. Consider local constraints.
3. Model power grids as social technical system.
4. Study emergent synchronisation and coupling effects when many agents coincide.

The major advantage is that this model is able to simulate human decisions and actions that would affect the functioning of the power grid. In addition, such changes in the grid would in turn influence human decisions and actions. Once this model is developed, future projects will be oriented to design and assess DSM policies that would help reduce investments on grid infrastructure. On the other hand, the agent decision model could also be improved. For example, agents could make decisions based on responses to changes in the system, which will in turn change the context for future decisions; or agents could behave in a heterogeneous way, thus maximising a certain profit, either a full charging of the vehicle or save money.

Part V

Conclusions

CHAPTER

20

Results

Research work has been conducted to deal with the problems of designing and evaluating management policies related to the SG paradigm. These policies are aimed at different stakeholders to benefit their interests. For instance, final customer's interests is the reduction of the electricity costs and the improvement of the energy efficiency. The interests of operators and retailers is to improve their profit by reducing operation and production costs. Globally, people want to have a more efficient power grid in which CO_2 emissions are reduced with a higher RES penetration.

The motivation for conducting research in the field of Power Grids is due to its proposed evolution to SGs. In this evolution, IT seems to have an important role. This evolution is proposed because there is a significant need for the improvement of power grids, such as: reduction of the dependency on fossil-fuel based energy production; market concerns such as fuel prices volatility; and reduction of GHG emissions.

SGs aim to face the next challenges: higher introduction of RES; more efficiency and flexibility; and less dependency on fossil fuels. Moreover, they are characterised by the introduction of automation at distribution level. It is expected that this evolution will overcome the challenge of the massive introduction of electrical vehicles, distributed generation and RES [HHM11]. The SG paradigm is an even more important challenge for isolated power grids as they have limited resources on the production side to balance offer and demand.

This is the case of the Canary Islands. In this territory, there are excellent environmental conditions for increasing the exploitation of RES. The Canary Islands have 2,500-3,000 hours/year of solar radiation producing an average of 5-6kWh/m^2 per day (source: Canary Institute of Technology). Furthermore, there are 3,000-4,500 hours/year of wind with a speed average of 7-8m/s. This produces 625MWh per day having installed 75MW of wind energy. In this sense, the results of this research can be directly applied to improving power grids in the Canary Islands.

The study of these smart management policies requires models representing the system at the lowest level of detail allowing to analyse side-effects. Aggregated models are not able to represent events on the demand side that should be considered in designing these policies. Therefore, Power Grid models should be disaggregated to represent all loads that can be managed. In these models, complexity arises since there are many heterogeneous components which are interacting at different scales.

The emergent behaviour of the system cannot be inferred with traditional models, because complexity cannot be represented. SGs cannot be studied using traditional approaches since these are not able to capture all of the system's properties. The best way to study complex models is through Software tools. However, if SGs are not supported by the required formalisms and methodologies, software technologies by themselves are not sufficient to perform these studies. This means, the problem is not choosing the software tool, but applying the right formalisms and methodologies. For these reasons, several authors have suggested the use of complex system based formalisms to represent SGs. A Complex Systems approach facilitates the representation of SGs at the lowest level of detail.

Under the complex system approach, only software tools that support this formalism can be considered to study SGs. Nevertheless, there are structural concerns that must be taken into account to support the engineering of large-scale complex systems. In the case of SGs, models may have millions of components. To model them, several complex system simulation software tools have been reviewed in this document. In all these tools, a lack of a methodological orientation to face the modelling of large-scale complex systems was found.

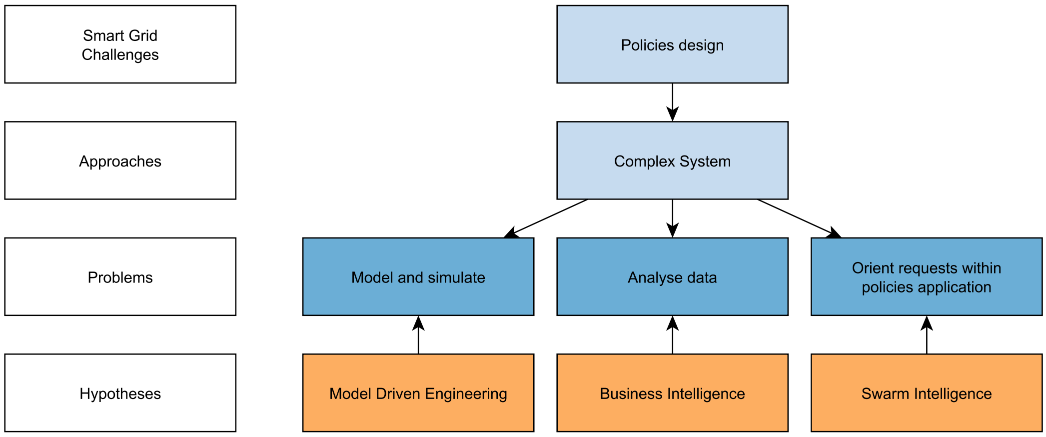


Figure 20.1: Problems within the design of policies that are encountered in this research through hypotheses.

The problem researched is how complex based formalism should be applied to build models of large scale systems. This is to say, the main issue is how methodologies may help to apply complex based formalism in such systems. Therefore, three hypotheses have been stated to support three different processes present in the study of SG policies (Figure: 20.1). These hypotheses are related to: modelling and simulating large-scale scenarios with MDE; analysing huge amounts of data coming from simulations with BI methodology; and distributing resources among many power consumers with Swarming Intelligence formalisms.

These hypotheses were researched in an exploratory approach on real SG projects performing some case studies. In this chapter, the results of this research are presented. First, we expose how the methodologies have been applied and what tools have been developed. Second, an interpretation on how these methodologies may contribute to the case studies is provided. And finally, the transfer of the research results to industry demonstrates the validity of the hypotheses.

20.1 About Smart Grid modelling

In this research, MDE has been applied to deal with the modelling and simulation of SGs. As a result of this research, a framework, named Tafat, has been developed. Tafat is a successful approach to the use of MDE in the field of SG simulations. Through Tafat, it is possible to: express a SG model by using a metamodel; and to simulate it by executing code generators that generate a simulator by interpreting the model.

This framework supports the development of complex system simulations and has been designed and developed for this research. This framework has been addressed using an MDE approach. Among others, the use of this methodology presents the following advantages:

- Fast development of complex system simulations from scratch.
- Usage of metamodels which describe the complex system elements enabling a full compatibility between simulation models that have been developed under a concrete metamodel. This means that many modellers may be working with the same metamodel and can integrate their models easily.
- Separation of concerns: the description of elements is fixed at the metamodel level while their way of acting (behavioural aspects) is variable, so that there may be different versions in which a behaviour of an element may be represented.
- Modular development which allows the improvement of the engine without affecting previously developed models.
- A high performance engine whose optimisation is constantly being improved.
- A complete set of tools which assist the different processes in creating a new simulator.
- Easy access to the modification of experiments which only requires a change in the definition of the simulation model.

20.2 About data analysis

In this research, BI methodologies have been applied to the analysis of huge quantities of data coming from SG simulations. As a result of this research, a framework, named Sumus, has been developed. This framework has been designed to overcome the issues of analysing data coming from huge simulations.

The use of a BI based approach has been useful to discover relevant information which was hidden in data results. Information visualisation studies have definitely been useful to detect relational effects amongst variables. These effects can be further studied so that modifications on the execution of different projects take them into account. Another important strategy to extract hidden or not evident information is through data mining.

This framework can be used in many other fields since they are meant to facilitate decision supporting at management positions. We have identified heterogeneous fields such as metrics to program code, product selling and public information systems. In other simulation fields, this method of data analysis can be really interesting, e.g. a set of simulations which run different configurations using the same scenario. Indeed, using an OLAP solution would make it possible to compare all of these configurations with each other.

20.3 About strategy design

In this research, SI techniques have been applied to conduct strategies of SG policies. As result of this research a framework, named Bandada, has been developed. Bandada is a laboratory supported by Tafat aimed at experimenting DSM policies. The framework has been designed to have a decoupled software architecture which makes it possible to change modules: how the demand behaves, what is the behaviour of the decision makers, which are the parameters to consider when optimising, etc.

Until now, this laboratory has integrated optimisation algorithms based on SI such as PSO or MOPSO. However, it is open to modifications so that different algorithms to support the decision making can be implemented.

Thanks to this laboratory, some experiments have been executed shortening the time to carry them out. This is possible since some of the research was previously made and, therefore, it could be reused. For instance, the experiment described in chapter 14 used all the demand side modelling that was developed in chapter 12.

20.4 Empirical hypotheses validation

All of the experiments that concerned simulations posses a high level of complexity containing thousands to millions of elements. The difficulty in developing these simulations comes from different sources. These difficulties are addressed by the methodologies proposed. The following list describes how these methodologies have been helpful in dealing with the different problems that arise when developing these experiments:

- *Modelling and simulation*
  - *Elements heterogeneity*. All of the experiments required the modelling of many different components (e.g. photovoltaic cells, electrical vehicles, etc.). Every different component has been developed independently from the others and often by different people. Thanks to the use of the metamodel, all these elements have been connected in order to perform simulations that supported the experiments described before. The use of a metamodel that provided semantic coherence for all of these models has been indispensable in allowing the use of all of them within the same simulation.
  - *Model development*. In all of the experiments described, there is a huge number of components being simulated. This task would have been almost impossible without using Profiler, the tool for automatically generating models based on statistical data. Therefore, it was only necessary to define input data for each experiment and define the way in which that data is used to generate the scenario. Based on this, Profiler generates the models for as many homes as desired concerning the data provided.
  - *Simulator development*. After developing the simulation model, the simulator was easily created. By only using the Simulator Generator tool, a simulator was created which included all required simulation elements and behaviours. This tool retrieved all the components needed to execute a simulation from both the metamodel and the repository.
- *Data analysis*. The quantity of results coming from these experiments is huge. Each of those thousands or millions of elements provides data each time step. However, it was necessary to have a way to analyse, not only if the emergent behaviour of the consumption was aligned with the expected one but also if individual ones were correct as well. In this sense, Sumus, and its approach based on BI, has been really helpful, first, to debug the models and, second, to check the emergence of the complex system under study.
- *Swarm intelligence*. The design of DLC policies is not an easy task as there are many different actors which are self-interested. Each individual interest

that every actor has must be taken into account in order to obtain a successful DLC policy. In this sense, it has been demonstrated, through experiments developed to this purpose, that swarming intelligence techniques are a promising approach for developing this kind of policy.

Therefore, the three hypotheses have been validated through experiments that have been described. First, MDE is useful for developing large-scale models with a high level of disaggregation. Second, results analysis is supported by Sumus which facilitated the comparison and data exploitation on the aggregated level. Finally, the use of swarming intelligence techniques, like MOPSO, has been successfully applied in the experiments.

20.5 Transference

The results of this research has allowed collaboration with institutions in which the knowledge have been shared. This collaboration has resulted in: the development of joint projects; participation in some workshops; and a way to experimentally validate the hypotheses proposed in this research.

Different projects have been developed in the last four years with EIFER (European Institute For Energy Research, Karlsruhe, Germany). This collaboration has increased the opportunity to work jointly with EDF (Électricité de France) R&D group through different public and private projects, such as Millener (Section: 20.6). This collaboration has been fruitful as it was possible to experimentally validate the hypotheses with real problems and real data. This is an important concern in most empirical software engineering research as it facilitates the transfer to industrial applications [SAA$^+$02]. Furthermore, through the collaboration with EIFER, it has been possible to work with the Universities of Stuttgart and the Basque Country.

In addition, participation in workshops has made it possible to exchange ideas with other institutions. For instance, it has been possible to exchange ideas with institutions as for example Supelec (École supérieure d'électricité) or Inria (Institut National de Recherche en Informatique et en Automatique) in France.

20.6 Projects

The different pieces of work that have been done during the development of this research have made possible participation in several R&D projects. The list below provides the name of the projects and our participation in each one of them:

- Framework para la simulación de la gestión de mercado y técnica de redes eléctricas insulares basado en agentes inteligentes. Caso de la red eléctrica de Gran Canaria (Framework to simulate the market and technical management of insular electricity grids based on intelligent agents. Case study: Gran Canaria electricity grid)
  - Source: ACIISI (Agencia Canaria de Investigación, Innovación y Ciencias de la Información)
  - Participation: research and development of Tafat
- Millener (Mille installations de gestion energétiques dans les îles – Thousand energy management devices on islands)
  - Source: French Government
  - Participation: research and development of case studies for La Réunion island using Tafat

20.7 Publications

Several publications have been written and published in different congresses and journals. A list containing the publications released in conferences is presented below:

- Agent-Based Modelling of Electrical Load at Household Level
  - Authors: Évora, J., Kremers, E., Morales, S., Hernández, M., Hernández, J. J., & Viejo, P.
  - Congress: In ECAL (European Conference on Artificial Life) 2011: CoSMoS - Proceedings of the 2011 Workshop on Complex Systems Modelling and Simulation
  - Place: Paris, France
  - Date: August, 2011

- Towards an interdisciplinary approach for the simulation of future SG architectures from a complex systems science point of view
  - Authors: Viejo, P., Kremers, E., Hernández, M., Hernández, J., Évora, J., Langlois, P., Daude, E., González de Durana, J.M., & Barambones, O.
  - Congress: European Conference on Complex Systems 2011
  - Place: Vienna, Austria
  - Date: September, 2011
- A large-scale electrical grid simulation for massive integration of distributed photovoltaic energy sources
  - Authors: Évora, J., Kremers, E., Hernández, M., & Hernández, J. J.
  - Congress: 6th European Conference on PV-Hybrids and Mini-Grids (OTTI)
  - Place: Chambery, France
  - Date: April, 2012
- Modelling lifestyle aspects influencing the residential load-curve
  - Authors: Hauser, W., Évora, J., & Kremers, E.
  - Congress: 26th European Conference on Modelling and simulation (ECMS 2012)
  - Place: Koblenz, Germany
  - Date: May, 2012
- Asynchronous Smart Grid Simulations
  - Authors: Évora, J., Hernández, J. J., & Hernández, M.
  - Congress: UCNC'12 (Unconventional Computation and Natural Computation): CoSMoS - Proceedings of the 2012 Workshop on Complex Systems Modelling and Simulation
  - Place: Orleans, France
  - Date: September, 2012
- Decision support for Complex Systems: a Smart Grid case
  - Authors: Évora, J., Hernández, J. J., & Hernández, M.

- Congress: UCNC'13 (Unconventional Computation and Natural Computation): CoSMoS - Proceedings of the 2013 Workshop on Complex Systems Modelling and Simulation
- Place: Milan, Italy
- Date: July, 2013

• Criticality in complex sociotechnical systems: an empirical approach
  - Authors: Viejo, P., Kremers, E., Évora, J., Hernández, J. J., Hernández, M., Barambones, O., & González de Durana, J. M.
  - Congress: European Conference on Complex Systems 2013 (ECCS'13)
  - Place: Barcelona, Spain
  - Date: September, 2013

• Asynchronous approach to simulations in Smart Grid
  - Authors: Évora, J., Hernández, J. J., & Hernández, M.
  - Congress: European Simulation and Modelling Conference 2013 (ESM'13)
  - Place: Lancaster, England
  - Date: October, 2013

• Tafat: A framework for developing simulators based on Model Driven Engineering
  - Authors: Évora, J., Hernández, J. J., & Hernández, M.
  - Congress: European Simulation and Modelling Conference 2013 (ESM'13)
  - Place: Lancaster, England
  - Date: October, 2013

• Model Driven Engineering for data miners simulation
  - Authors: Évora, J., Hernández, J. J., & Hernández, M.
  - Congress: IEEE International Conference on Data Mining 2013 (ICDM'13) - International Workshop on Domain Driven Data Mining
  - Place: Dallas, United States of America
  - Date: December, 2013

• Agent-based modelling for designing an EV charging distribution systems: a case study in Salvador of Bahia
  - Authors: Évora, J., Hernández, J. J., Barbosa, D. & Nazareno, P.

  - Congress: European Simulation and Modelling Conference 2014 (ESM'14)
  - Place: Porto, Portugal
  - Date: October, 2014

Next list contains the publications released in journals:

- Advantages of Model Driven Engineering for studying complex systems
  - Authors: Évora, J., Hernández, J. J., & Hernández, M.
  - Journal: Natural Computing
  - DOI: 10.1007/s11047-014-9469-y
- Criticality in complex sociotechnical systems: an empirical approach to electrical grids [1]
  - Authors: Viejo, P., Kremers, E., Évora, J., Hernández, J.J., Hernández, M., Barambones, O, González de Durana, J.M.
  - Journal: Applied energy

[1]This paper has been submitted and now it is awaiting for the acceptance

CHAPTER

21

Discussion

In this chapter the scientific and technical contributions of this research are described. "Scientific contributions" refer to the creation of knowledge whereas technical contributions refer to the creation of new processes to deal with specific problems.

The purpose of this research is to explore and describe the application of already existing methodologies to identify problems in SGs. This exploration has also allowed an outlook on new research targets that could be studied in future research.

21.1 Contributions

21.1.1 On Applying the complex system approach in Smart Grids

According to the well-honoured philosopher of science Thomas Kuhn, in the scientific development of a discipline there are three main stages [Kuh12]. These three main stages are: "prescience", "normal science" and "revolutionary science" (Figure: 21.1). The "prescience" stage is characterised by having numerous incompatible and incomplete theories. The consensus of a prescientific community in terms of methods, terminologies, experiments, etc. leads to the "normal science". In this stage, new paradigms can be conceived to deal with anomalies reaching the "revolutionary science" stage.

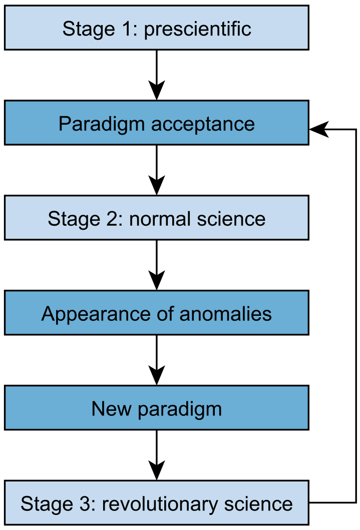


Figure 21.1: Kuhn's cycle.

In the case of the SG topic, it can be considered to be currently in the second stage: "normal science". Within the study of this normal science, several limitations have been found when SGs are being modelled (appearance of anomalies). The complex system approach has been proposed to overcome them (new paradigms).

There are several studies in the documentation that assert that power grids cannot be studied using traditional formalisms. In these studies, the hypothesis states that complex system approach could overcome the limitation of the traditional formalisms. This hypothesis cannot be absolutely verified since it cannot be proved to carry out all possible case studies. However, what the hypothesis states can be reproduced and refuted.

In this research, this hypothesis has been reproduced through the experimentation of case studies. These case studies provide additional evidence that this formalism is valid [CGLP94, PKZK08, EKM$^+$11]. Several case studies have been carried out to analyse SGs following complex system approach. It has been found that this formalism is useful for studying SG policies with a maximum level of disaggregation. When a new policy is being proposed, the behaviour of a power grid is not known. In these cases, the system can be modelled through its components' behaviours and simulated to obtain the emergent behaviour. A maximum level of disaggregation also allows to re-aggregate the components' behaviours and analyse them at different levels (e.g. appliance, household, district, country level).

Nevertheless, in this research, it has been recognised that complex system approach by itself may not be enough to deal with the study of SGs. Other issues appear when using this formalism: complexity for modelling systems, complexity for analysing data and complexity for designing strategies. An important contribution is the identification of complementary methodologies or formalisms that could be useful for overcoming these issues. In particular, the application of MDE, BI and SI should be considered.

The validity of this contribution has been demonstrated through the execution of case studies. As in the case of complex system approach, these hypotheses cannot be verified. In order to verify them absolutely, these hypotheses should be applied to perform all SG case studies, which is impossible. Nevertheless, in this research, several case studies have been carried out to provide evidence of its validity.

21.1.2 On Modelling Smart Grids

This research has found that the modelling of SGs requires the coexistence of different natures in their components. When SGs are being modelled, it is necessary to describe components that may have different natures such as: electrical, thermodynamical, meteorological or sociological, among others.

Obviously, electrical components must be represented since their behaviour is closely related to the human behaviours (e.g. when these components are switched on and off). Actually, the way in which the humans use these electrical components will determine the system's emergent behaviour. The interaction they do over this components involve consumptions that affect how the system works. Thus, many electrical models must be modelled in order to represent production, distribution, demand, etc. with the degree of detail that is necessary for testing the SG implementation.

In addition, thermodynamical components must also be modelled since it is necessary to represent the thermal transferences between thermal electrical components (e.g. refrigerators, domestic water heaters, etc.). The modelling of these electrical components must include thermodynamical behaviour that calculates the thermal gains derivative from their consumptions. Moreover, infrastructure as households may include a thermal behaviour as well. This behaviour would calculate the infrastructure temperature based on the external temperature and the thermal gains derivative from the thermal electrical loads. In this manner, thermal electrical loads can calculate their internal temperature based on their consumption and the household temperature.

The calculation of the internal temperature of a household depends on the external temperature, as previously stated. Therefore, it is necessary to represent the meteorological conditions of the environment where the household is located. Thus, another kind of behaviour which is necessary to model power grids is identified. Furthermore, meteorological conditions must be represented if there are RES in the power grid under study. Renewable energy technologies, such as PV cells and wind turbines, require the representation of environmental conditions. These conditions are used by the models of these types of technologies to calculate the energy produced.

Furthermore, the representation of SGs also use the modelling of human interactions. Human beings are users of Power Grids and the way they use energy must be represented. This is a very important consideration because of the significant impact human interaction has on the power grid. As mentioned previously, in state of the art, power demand varies geographically, seasonally, culturally, sociologically, etc. All these variations have to do with the way in which humans interact with the power grid.

In conclusion, the study of the application of SG policies requires the modelling of many components with heterogeneous natures. The behaviours that have been previously mentioned are not only different because they have different natures: electrical, thermal, social, etc. but also because they require different execution paradigms: continuous (e.g. thermal) or discrete (e.g. washing machine).

21.1.3 On Developing Complex Models

In this research, MDE has been found to be a complementary methodology for modelling SGs with a complex system approach. When it is necessary to integrate different modelling techniques, structural guidelines are needed to deal with this complexity. In these cases, MDE is useful for managing the model definition and the simulator construction.

In this research, the application of MDE has been validated by carrying out the case studies. Moreover, this approach is being used by other researchers. Currently, EIFER and EDF are developing their own case studies with this methodology. From a scientific point of view, it is known that other laboratories are researching in the hypothesis. To this day, the hypothesis has not been refuted.

The most important implication of applying MDE in this field is technical: MDE helps to build up and maintain large-scale simulation models. Modellers are assisted in the construction process by separating concerns: the scenario description (simulation model) and the description of individual components (behaviours). In this way, a set of steps are established to guide the modeller to the goal:

- Developing the simulation model.
- Describing a component in the metamodel if it has not been previously defined there.
- Developing necessary behaviours not yet modelled.
- Building a simulator based on this simulation model and simulate it.

Another MDE implication is its adaptability to changes. A single component or behaviour can be easily replaced or added to an existing model scenario. Flexibility is an important feature for engineering SGs. In the design of a new policy, it is necessary to test it in different scenarios. The flexibility provided by MDE makes this possible. These scenarios may not only use different behaviours for the components but may also organise components in a different way. In addition, the same scenario can be used for testing different policies.

Since MDE requires the definition of a metamodel, a technical contribution of this research is the development of a metamodel for SGs. In this sense, a metamodel can be understood as the creation of an agreement contract between many modellers in order to make collaboration easier. The metamodel has several important implications:

- Modellers can share their models whenever they have been built under the same metamodel.
- Modellers can contribute to the metamodel by extending its definition or creating new behaviours that can be used by others.
- Modellers do not need to describe the components of the reality every time they start a new simulation since they take advantage of the already defined metamodel.

- Modellers can use previous work developed by other modellers.
- Modellers can develop their models faster thanks to all of these above-mentioned factors.

In this way, large simulation models can be built by developing the simulation model; and, in case it is needed, developing new components and behaviours. In this sense, modellers are guided completely so they know what they have to do to develop simulation models. Therefore, the complexity of the scenario only affects the simulation model and not the overall framework.

Furthermore, tools have been developed to deal with the complexity of the representation of large-scale system in this simulation model. The tool named Profiler was a response to the need of many modellers had for the construction of large-scale scenarios. In this sense, this tool handles the complexity of building these simulation models. Based on input data provided by the modeller, this tool creates the scenario. Another important tool is the Simulator Builder which builds a simulator based on a model that describes the scenario.

21.1.4 **On Analysing Simulation Results**

This research demonstrates that the application of BI methodologies is helpful. These methodologies have been used to analyse data coming from the simulation of a complex system.

When SGs are studied as complex systems, the huge quantity of results coming from these simulations can be managed using BI methodologies. BI methodologies can be used to discover laws and regularities that provide a better knowledge of the emergent behaviour in SGs. These methodologies allow science to be done over simulated experiments. Based on the output available from the simulations, it is possible to find rules, laws, regularities, etc. automatically.

In this research, BI methodologies are used on the output of the case studies presented. The use of these methodologies is really important in order to know how the emergent behaviour in SGs is reached; and how it can be influenced.

After performing all of the case studies presented in this document, it has been shown that BI can be used in this context with naturalness. The mechanisms provided by BI are sufficient to represent data coming from SG simulations and allow its exploitation.

Another important implication of using BI methodologies is that they allow science to be performed over simulated experiments. Instead of establishing a hypothesis and, later on, validating, the procedure in this case would be different. Since all data is available (as they are the output of simulations) rules, laws, regularities, etc. can be found automatically.

21.1.5 On Designing Strategies For Smart Grids

In SG engineering, there are some objectives that must be covered by the policies designed. It has been found that the accomplishment of these policies requires the support of strategical concerns. Strategies can be defined to coordinate the behaviour of all of the participants influencing the power grid.

When SG policies are designed, the way in which their strategical concerns are implemented can be done using SI techniques. SI techniques are able to optimally coordinate all participants in such a way that accomplishes the objectives.

This research has validated that SI techniques are useful for executing resource management strategies. In SGs, energy resources must be managed efficiently maintaining a quality criteria. To this end, it has been demonstrated through experiments that SI techniques can help to distribute these resources.

An important implication of this is that the use of SI techniques permits quality management in SGs. Quality management ensures that the products sold to customers are consistent. In the case of SGs, it must be ensured that all customers are supplied with a minimum of quality. It has been found that SI techniques are useful for describing these quality issues and to applying them correctly.

A natural criteria for quality establishment in power grids can be based on the stability of the grid. In this sense, all components are involved in maintaining the stability of the grid, as in chapter 16. In this context, the power grid can be seen as a huge organisation with many participants. All these participants are responsible for executing the guidelines coming from the Quality Manager.

21.2 Other remarks

21.2.1 On Implementing Smart Grids

Through the case studies, it has been found that many participants may be involved in maintaining the stability of the power grid. For this reason, the implementation of SG policies should not be massively introduced in the early stages.

In a technological implementation process, the traditional methodology consists in a build-measure-learn cycle. In this cycle, small experiments in a small part of the system are built to validate the benefits of introducing a SG policy. Then, these experiments must be evaluated by using objective measures that can be used to determine how good the policy is.

Thanks to this process, new knowledge is acquired which can be used in the next iterations of the procedure. In this way, experiments can develop in accordance to reality. This feedback from the information of early stages makes it possible for experiments to become large-scale sustainable implementations.

This approach may also be useful in a context in which there are already different mechanisms working to stabilise the power grid. Therefore, the new policy can coexist with those mechanisms already established for stabilising the process.

Nevertheless, this implementation of new technology in an environment where there is already operative technology. Whereas the well-functioning of new technology can be tested prior to its massive deployment, previous technology may not work correctly with the introduction of the new technology. In this sense, another observation is that "smart" devices might overcome this problem if they can be remotely updated.

21.2.2 On Simplifying complex systems

The main reason for using the complex system approach for studying SGs is to obtain knowledge. The emergent behaviour that SGs may have as a consequence of applying certain policies is not possible to be inferred. This thought has led researchers to deploy complex system simulations to study SGs.

Nevertheless, another interesting question is that, once the emergent behaviour is known, would it be possible to use simplified models? If this were possible, many computational resources may be released to perform other calculations. For instance, if the consequences of applying a specific policy are known, would it be possible to represent these consequences in a simplified model and use it in other complex simulations? At first glance, this seems possible.

Obviously, this depends on the features of the emergent behaviour. In some cases, once the emergent behaviour is known, it may be possible to predict what it would be like in other circumstances. In these cases, it should be possible to have a simplified representation of such behaviour. In other cases, there are emergent behaviours whose reaction to unstudied events cannot be predicted. In these cases, the use of simplified models may not be sufficient to represent the complex model and therefore, the model cannot be replaced.

21.2.3 On Studying Smart Grids

In the study of SGs as complex systems, the need to carry out several processes has been identified. Normally, the study of complex systems deal with how such systems have to be modelled and simulated. When SGs policies are studied using the complex system approach, it is necessary to carry out simulations, data analysis and strategy designs.

The main goal of SG studies is to find the emergent behaviour and how it is reached. In addition, an important goal is to design strategies that allow achieving a specific emergent behaviour. Regarding either goal, the complex system approach by itself may not be enough. In the following list both goals are described with a greater degree of detail:

- *Understanding the emergent behaviour.* In the analysis of complex systems, this can be a difficult task if it is not guided. The emergence that a system reaches depends on the behaviour of the components of the system and their interaction. However, it is not easy to infer how such behaviour has been reached by only observing the emergent behaviour output of a simulation. For this reason, in this research, BI methodologies have been used for a better understanding of the emergent behaviour.
- *Reaching a specific emergent behaviour.* A common need is to design a strategy that can reach a specific emergent behaviour. This is the case of DLC policies. These policies must provide a specific emergent behaviour in

accordance to power production limitations that may arise. In order to deal with this, in this research, SI techniques have been used.

21.2.4 On Professional Profiles

IT can be considered to support the decision making prior to the implementation of SG policies. IT provides the possibility of analysing the consequences of applying these policies in a real power grid.

In this research, this application of the IT has been explored: several IT frameworks have been used to address these types of issues; and new approaches have been used to deal with the problems of complexity previously mentioned. This research has also been fed by all experiments that have been conducted in collaboration with EIFER and EDF. This collaboration has been very important to the results of this research.

In this sense, it is possible that, in a near future, professional software developers will be required in this field. This may be an important professional profile that will help to define and execute case studies. The design and evaluation of projects in SGs may require the cooperation of people with different skills including software engineers.

Therefore, new professionals must be trained to combine two major skills for the study of SG policies: electrical and software engineering. As said before, these kinds of professionals may also be important to support IT in the monitoring and controlling network proposed in SGs.

21.2.5 On Methodological Research

This research has been carried out keeping in mind its transfer to society. The hypotheses evaluated in this document are aimed at supplying requirements that have been identified in the study of SG policies.

As a consequence of carrying out real case studies jointly with EIFER and EDF, requirements have been identified and dealt with by this research. In this kind of work, the research procedures focus more on the transfer than on providing a broad perspective.

An important concern that should be highlighted is that in this research, several case studies have been designed to tackle real problems. It is common that case studies are aimed at validating hypotheses. However, when case studies are designed to validate hypotheses, it may be that they do not encounter real problems. For this reason, the case studies performed in this research have a particular relevance to the validation of hypotheses.

In this research, case studies have been designed to attend to real problems. Some of them have been designed in a collaboration with EIFER and EDF. In these cases, real power grids were modelled and this provided stronger evidence of validity than synthetic problems.

21.3 Future work

21.3.1 On Smart Cities

One of the main challenges that can be addressed on the MDE side is its application in Smart Cities. A natural evolution of the SG simulations could lead to what is known as "Smart Cities". In [GPM07], a city is considered "smart" when the following are present: investments in human and social capital and, traditional (e.g. transport) and modern (e.g. information and communications technology) communication infrastructure; sustainable economic growth; a high quality of life; and a wise management of natural resources. In the context of Smart Cities, modelling and simulation tools may be helpful for decision makers. Through simulations, decision makers can evaluate the consequences of applying a specific policy before its implementation.

The simulation of Smart Cities is an interesting challenge, since it may involve many heterogeneous elements from different domains which are somehow linked. For instance, promoting electrical vehicles in a territory may require the study of traffic, power grids, road networks, etc.

This involves an evolution in the metamodel since it must be able to support the definition of more components. This evolution would not only consist in describing more elements than in the SGs metamodel. It will also be necessary to define a new metalanguage to describe the Smart Cities metamodel. In this context, it can be seen that there are components that may affect different aspects of a Smart City: electricity, transportation, waste, water, etc. For this reason, the current metalanguage that is valid for representing SGs is not enough for Smart Cities.

Currently, the research is focused on the design of a metalanguage that provides enough capability to describe the reality of Smart Cities. Moreover, the analysis of Smart Cities involves new challenges that can be addressed through BI methodologies and SI techniques.

The application of BI methodologies in the field of Smart Cities has to face challenges in the design of new indicators, taxonomies and visualisations. Indicators used in SGs may no longer be sufficient for analysing the performance of a whole Smart City. In Smart Cities, more characteristics than the electrical one have to be considered and new indicators should be designed. Similarly, in the case of taxonomies, new categories are used to define the categorisations that will allow the filtering of information. The design of new visual representations will be required in order to better represent the information coming from Smart Cities simulations.

The same thing happens in the application of SI techniques. New strategies must be defined in the context of Smart Cities beyond the SGs. These new strategies may require the use of different algorithms that can be used to manage components in different aspects: transportation, electricity, facilities, waste, etc.

21.3.2 On Developing Additional Case Studies

Developing additional case studies is an interesting way to continue this research. In this work, several case studies demonstrated that the use of these methodologies is helpful. However, they account for a very small portion of the whole universe of possible case studies that related to SG policies.

In this sense, researchers can reproduce the hypotheses to provide more evidence for their validity or to refute them. As previously said, these hypotheses are valid as they are reproducible and refutable. Any researcher can try to apply MDE in order to carry out experiments of SG policies.

Moreover, new case studies can be proposed in which new challenges have to be dealt with through the application of MDE: for instance, case studies in which it is necessary to represent scenarios that evolve. This could be an interesting feature for long-term (e.g. several years) simulations. That is to say, components can evolve during the execution of the simulation so that new ones can be created or destroyed. For example, it may be interesting how a specific policy may react to an evolutionary scenario in which: new buildings are created; electrical vehicles are progressively introduced; or RES quota increases. Furthermore, case studies in which both scenarios and policies can evolve can be considered.

These kind of scenarios may be important for testing the applicability of SG policies in the long-term. In order to support them, new mechanisms must be designed within this ecosystem of methodologies and formalisms.

21.3.3 On Discovering Patterns

As said before, BI methodologies can be used to find patterns in the simulations of SGs. The discovery of patterns may have many different usages: obtaining a better knowledge of how emergence is reached, simplifying a complex model, feeding back into the design of SG policies, etc.

In this field, much work can be carried out as a continuation of this research. Data mining techniques can be explored in order to automatically detect patterns in series of data. However, this is not an easy task, as there is a main underlying concern that must be considered when exploring these techniques: the huge quantity of data.

As stated in the hypotheses part of the document, an important concern of data management is the velocity. Huge quantity of data is contradictory to the velocity. This is especially tricky when dealing with data mining techniques that require a thorough analysis of the data. For this reason, this research should consist in: investigating data mining procedures to find patterns; and finding procedures to accelerate the process.

One of the applications that this research may have is the simplification of complex models. This would open another interesting future project that could consist in automating the model simplification process. That means, the use of methodologies, formalisms or techniques to identify behavioural features in complex models that can characterise simplified models.

To this end, results coming from the discovery of patterns can be used in this research. After, software tools could be able to parse data coming from complex models in order to create simplified models. The behaviours provided by the simplified model must be certified to be working at certain intervals that ensure measurable errors. For instance, these simplified models could be used as submodels within bigger simulations. Therefore, smaller complex simulations can be used to engineer bigger ones in which the computational costs can be reduced.

21.3.4 On Other Fields

In the "hypotheses part" of this document, the definition of external validity in the context of this research was discussed. However, a hypothesis has external validity if it applies to several case studies of SG policies. This external validity has been checked in this document as these methodologies have been successfully applied in all presented case studies.

Nevertheless, the second definition provided for the external validity is not dealt within this research. This definition concerned the use of these methodologies in other complex systems apart from SGs. This is actually an extension of the scope of the hypotheses defined in this work.

It is reasonable to believe that if their application in the field of SGs has been useful, their application in other fields can be helpful as well. However, this should be checked and is proposed for future research. In this sense, the hypothesis to be researched could be: "MDE, BI and SI can be helpful methodologies for engineering complex systems of any nature". As in the previous hypotheses, this cannot be fully verified because it is not possible to perform all of the case studies in all complex systems. However, a subset of them can be carried out, which could provide new ideas or requirements.

21.3.5 On Experimenting with Overview, Design concepts and Details protocol

Another line that can be explored is the integration of the Overview, Design concepts and Detail (ODD) protocol in our approach to model agent-based models with MDE. This protocol is a proposed standard for describing agent-based models [GBB+06]. Therefore, our approach could include the description of models that are compliant with ODD.

This protocol consists of three blocks (Overview, Design concepts and Details), which are subdivided into seven elements: Purpose, State Variables and Scales, Process Overview and scheduling, Design concepts, Initialization, Input, and Submodels. Therefore, agent-based models can be described according to these seven elements. The next list summarises the contents to place at each of these elements:

- Purpose: description of the scope of the model. Clear, concise and specific formulation of the model's purpose. Explanation of the need for building a complex model.
- State variables and scales: state variables refer to low-level variables that characterise low-level entities of the model. The scales refer to the simulation configuration (e.g. length of time steps, size of habitat cells).
- Process Overview and scheduling: the process overview describes environmental and individual processes that are built into the model (e.g. food production, growth, movement). The schedule has to do with how these processes occur and their order.
- Design concepts: these concepts provide a common framework for designing and communicating individual-based models.
- Initialization: defines how the environment and the individuals have to be created.
- Input: conditions of the environment that change over space and time.
- Submodel: presentation and explanation of all the processes listed above in a detailed manner, including the parameterisation of the model.

Bibliography

[ADG08] A Arenas and A Diaz-Guilera. Synchornization in complex networks. *Physics Reports*, 469:93–153, 2008. 191

[ADHM] Robert Arritt, Roger Dugan, Robert Henry, and Tom McDermott. OpenDSS. 35, 37

[AEE08] ADEME, ENERTECH, and EDF. Campagne de mesures des appareils de production de froid et des appareils de lavage dans 100 logements. Technical report, ADEME and ENERTECH and EDF, 2008. 150

[AER] ADEME, EDF, and Région Réunion. Millener, une longueur d'avance pour la Réunion. 9

[AES07] M.H. Albadi and EF El-Saadany. Demand response in electricity markets: An overview. In *Power Engineering Society General Meeting, 2007. IEEE*, pages 1–5. IEEE, 2007. 30

[AG12] Hazem Ahmed and Janice Glasgow. Swarm intelligence: Concepts, models and applications. *School Of Computing, Queens University Technical Report*, 2012. 47

[All09] P Allen. *Complexity, Innovation and Organizational Evolution*. 2009. 189

[Ami01] Massoud Amin. Toward self-healing energy infrastructure systems. *Computer Applications in Power, IEEE*, 14(1):20–28, 2001. 4, 19

[And09] David L Anderson. *An evaluation of current and future costs for lithium-ion batteries for use in electrified vehicle powertrains*. PhD thesis, Duke University, 2009. 5

[AP96] Timothy Ward Athan and Panos Y Papalambros. A note on weighted criteria methods for compromise solutions in multi-objective optimization. *Engineering Optimization*, 27(2):155–176, 1996. 118

[B+09] Albert-László Barabási et al. Scale-free networks: a decade and beyond. *science*, 325(5939):412, 2009. 40

[BA03] Robert Blumberg and Shaku Atre. The problem with unstructured data. *DM REVIEW*, 13:42–49, 2003. 62

[BB02] Tim M Blackwell and Peter Bentley. Improvised music with swarms. In *Computational Intelligence, Proceedings of the World on Congress on*, volume 2, pages 1462–1467. IEEE, 2002. 48

[BBS+09] Prasanna Balaprakash, Mauro Birattari, Thomas Stützle, Zhi Yuan, and Marco Dorigo. Estimation-based ant colony optimization and local search for the probabilistic traveling salesman problem. *Swarm Intelligence*, 3(3):223–242, 2009. 47

[BC05] C Bueno and J A Carta. Technical–economic analysis of wind-powered pumped hydrostorage systems. Part II: model application to the island of El Hierro. *Solar energy*, 78(3):396–405, 2005. 5

[BF04] A Borshchev and A Filippov. Anylogic-multi-paradigm simulation for business, engineering and research. In *The 6th IIE Annual Simulation Solutions Conference*, volume 150, 2004. 45

[BI04] J.P. Barton and D.G. Infield. Energy storage and its use with intermittent renewable energy. *Energy Conversion, IEEE Transactions on*, 19(2):441–448, June 2004. 28

[Blu05] Christian Blum. Beam-aco-hybridizing ant colony optimization with beam search: An application to open shop scheduling. *Computers & Operations Research*, 32(6):1565–1591, 2005. 47

[Blu08] Christian Blum. Beam-aco for simple assembly line balancing. *INFORMS Journal on Computing*, 20(4):618–627, 2008. 47

[BLW05] Paul Baker, Shiou Loh, and Frank Weil. Model-driven engineering in a large industrial context-motorola case study. In *Model Driven Engineering Languages and Systems*, pages 476–491. Springer, 2005. 55

[BM01] Eric Bonabeau and Christopher Meyer. Swarm intelligence. *Harvard Business Review*, 79(5):106–114, 2001. 46

[BMW10] PE Baker, C Mitchell, and B Woodman. Electricity market design for a low-carbon future. *London, UKERC*, 2010. 24

[Boh03] Marko Bohanec. *Decision support*. Springer, 2003. 57

[BPSK02] Michael Butler, Luigia Petre, Kaisa Sere, and Stuart Kent. Model Driven Engineering. In *Integrated formal methods*, volume 2335 of *Lecture Notes in Computer Science*, pages 286–298. Springer Berlin / Heidelberg, 2002. 50, 52

[Bre00] Marilynn B Brewer. Research design and issues of validity. *Handbook of research methods in social and personality psychology*, pages 3–16, 2000. 72

[BRS$^+$10] Steve Bou Ghosn, Prakash Ranganathan, Saeed Salem, Jingpeng Tang, Davin Loegering, and Kendall E. Nygard. Agent-Oriented Designs for a Self Healing Smart Grid. In *2010 First IEEE International Conference on Smart Grid Communications*, pages 461–466. IEEE, October 2010. 36

[Bun11] Bundesverband der Energie- und Wasserwirtschaft. BDEW - Standardlastprofile (SLP), 2011. 143

[BVB08] Christian Blum, Mateu Yábar Vallès, and Maria J Blesa. An ant colony optimization algorithm for dna sequencing by hybridization. *Computers & Operations Research*, 35(11):3620–3635, 2008. 47

[BW89] Gerardo Beni and Jing Wang. Swarm Intelligence in Cellular Robotic Systems, Proceed. In *NATO Advanced Workshop on Robots and Biological Systems*, Tuscany, Italy, 1989. 45

[BW93] Gerardo Beni and Jing Wang. Swarm intelligence in cellular robotic systems. In *Robots and Biological Systems: Towards a New Bionics?*, pages 703–712. Springer, 1993. 45, 46

[BY97] Yaneer Bar-Yam. *Dynamics of complex systems*, volume 213. Addison-Wesley Reading, MA, 1997. 39

[C+94] CDIF Technical Committee et al. Cdif framework for modeling and extensibility. *Electronic Industries Assocaiation, EIA/IS-107*, 1994. 52

[C+06] European Commission et al. European technology platform smart grids: vision and strategy for europe's electricity networks of the future. *Office for Official Publications of the European Communities*, 2006. 22

[Cao04] Junwei Cao. Self-organizing agents for grid load balancing. In *Proceedings of the 5th IEEE/ACM International Workshop on Grid Computing*, pages 388–395. IEEE Computer Society, 2004. 46

[CCS93] Edgar F Codd, Sharon B Codd, and Clynch T Salley. Providing olap (on-line analytical processing) to user-analysts: An it mandate. *Codd and Date*, 32, 1993. 60

[CD97] Surajit Chaudhuri and Umeshwar Dayal. An Overview of Data Warehousing and OLAP Technology. *ACM Sigmod Record*, 26(1), 1997. 103

[CD10] E Chappin and G.P. Dijkema. Agent-based modelling of energy infraestructure transitions. *International journal of Critical Infraestructures*, 6:106–130, 2010. 190

[CDMT94] Alberto Colorni, Marco Dorigo, Vittorio Maniezzo, and Marco Trubian. Ant system for job-shop scheduling. *Belgian Journal of Operations Research, Statistics and Computer Science*, 34(1):39–53, 1994. 47

[CGLP94] A. Capasso, W. Grattieri, R. Lamedica, and A. Prudenzi. A bottom-up approach to residential load modeling. *Power Systems, IEEE Transactions on*, 9(2):957–964, 1994. 8, 32, 33, 250

[CL02] C A Coello Coello and M S Lechuga. MOPSO: a proposal for multiple objective particle swarm optimization. In *Evolutionary Computation, 2002. CEC '02. Proceedings of the 2002 Congress on*, volume 2, pages 1051–1056, 2002. 119, 164, 168

[Cla11] Cédric Clastres. Smart grids: Another step towards competition, energy security and climate change objectives. *Energy Policy*, 39(9):5399–5408, September 2011. 3, 20

[CMB10] A.J. J. Conejo, J.M. M. Morales, and L. Baringo. Real-Time Demand Response Model. *Smart Grid, IEEE Transactions on*, 1(3):236–242, 2010. 33

[CMG11] Timothy Cooper, Gillian Mead, and Phil Grünewald. The Future of Energy Storage: Stakeholder Perspectives and Policy Implications. Technical report, UK Energy Research Centre, Imperial College, 2011. 23

[CnA00] C.A. Cañizares and F.L. Alvarado. UWPFLOW Program, University of Waterloo, 2000. 35, 36

[Coa99] Meta Data Coalition. Open information model. *At URL: http://www. mdcinfo. com/OIM/MDCOIM10. pdf*, 1999. 52

[Coe98] Carlos A. Coello Coello. A comprehensive survey of evolutionary-based multiobjective optimization techniques. *Knowledge and Information Systems*, 1:269–308, 1998. 118

[Col03] Nick Collier. Repast: An extensible framework for agent simulation. *The University of Chicago's Social Science Research*, 36:2003, 2003. 44

[CSMJ13] Trinidad Chardin-Segui, Sebastien Mathiou, and Yves Jeanjean. The Millener project: Five hundred RES & one thousand fifty DSM installations for households in French islands, aggregated for grid support. In *2013 IEEE Grenoble Conference*, pages 1–6. IEEE, June 2013. 9

[Dep] Deparment of Energy, USA. http://energy.gov/oe/technology-development/smart-grid. 21

[DGS07] Joyce Dargay, Dermot Gately, and Martin Sommer. Vehicle ownership and income growth, worldwide: 1960-2030. *The Energy Journal*, pages 143–170, 2007. 228

[DMCM91] Marco Dorigo, Vittorio Maniezzo, Alberto Colorni, and Vittorio Maniezzo. Positive feedback as a search strategy, 1991. 47

[Dor92] Marco Dorigo. Optimization, learning and natural algorithms. *Ph. D. Thesis, Politecnico di Milano, Italy*, 1992. 47

[Dou06] Douglas C. Schmidt. Guest Editor's Introduction: Model-Driven Engineering. *Computer*, 39:25–31, February 2006. 49, 50, 51

[Duf10] D Duffert. *Lastmodellierung zur Untersuchung von Lastmanagement - Maßnahmen in Elektrizitätsnetzen*. PhD thesis, Final Project Degree, Karlsruhe Institute of Technology, 2010. 182

[DW07] Patrick Devine-Wright. Energy citizenship: psychological aspects of evolution in sustainable energy technologies. *People centred eco-design: consumer adoption of low and zero carbon products and systems*, page 63, 2007. 24

[Ede96] Herb Edelstein. Technology how to: Mining data warehouses, 1996. 63

[edfa] EDF webpage. `http://www.edf.com/`. 72, 139

[EDFb] EDF. LE PROJET ”MILLENER”. 9, 181

[EE10] EDSO ENTSO-E. The european electricity grid initiative (eegi)-roadmap 2010-18 and detailed implementation plan 2010-12, 2010. 22

[EH99] Russell C Eberhart and Xiaohui Hu. Human tremor analysis using particle swarm optimization. In *Evolutionary Computation, 1999. CEC 99. Proceedings of the 1999 Congress on*, volume 3. IEEE, 1999. 48

[EHBS14] Jose Evora, Jose Juan Hernandez, Daniel Barbosa, and Paulo Sampaio. Agent-based modelling for designing an EV charging distribution systems: a case study in Salvador of Bahia. In *European Simulation and Modelling Conference 2014 (ESM'14)*, 2014. 227

[EHH13a] Jose Evora, Jose Juan Hernandez, and Mario Hernandez. Asynchronous approach to simulations in Smart Grid. In *European Simulation and Modelling Conference 2013 (ESM'13)*, 2013. 132

[EHH13b] Jose Evora, Jose Juan Hernandez, and Mario Hernandez. Decision support for Complex Systems: a Smart Grid case. In *UCNC'13 - CoSMoS Workshop*, 2013. 193

[EHH13c] Jose Evora, Jose Juan Hernandez, and Mario Hernandez. Tafat: A framework for developing simulators based on Model Driven Engineering. In *European Simulation and Modelling Conference 2013 (ESM'13)2*, 2013. 75, 124

[EHHK12] Jose Evora, Jose Juan Hernandez, Mario Hernandez, and Enrique Kremers. Asynchronous Smart Grid Simulations. In *UCNC'12 - CoSMoS Workshop*, 2012. 75

[eif] EIFER webpage. http://www.eifer.uni-karlsruhe.de/. 72, 139

[EK95] Russ C Eberhart and James Kennedy. A new optimizer using particle swarm theory. In *Proceedings of the sixth international symposium on micro machine and human science*, volume 1, pages 39–43. New York, NY, 1995. 47, 116

[EKHH12] Jose Evora, Enrique Kremers, Mario Hernandez, and Jose Juan Hernandez. A large-scale electrical grid simulation for massive integration of distributed photovoltaic energy sources. In *6th European Conference on PV-Hybrids and Mini-Grids*, pages 239–246, 2012. 181

[EKM+11] Jose Evora, Enrique Kremers, Susana Morales, Mario Hernandez, Jose Juan Hernandez, and Pablo Viejo. Agent-Based Modelling of Electrical Load at Household Level. In *ECAL 2011: CoSMoS - Proceedings of the 2011 Workshop on Complex Systems Modelling and Simulation*, page 12, 2011. 8, 32, 75, 86, 137, 152, 170, 177, 193, 250

[EN08] Boris Evelson and N Norman. Topic overview: business intelligence. *Forrester Research*, 2008. 61

[ES01] Russell C Eberhart and Yuhui Shi. Tracking and optimizing dynamic systems with particle swarms. In *Evolutionary Computation, 2001. Proceedings of the 2001 Congress on*, volume 1, pages 94–100. IEEE, 2001. 48

[Eur07] European Technology Platform SmartGrids. Strategic research agenda for Europe's electricity networks of the future, 2007. 17, 18

[Eur10] European Regulators Group for Electricity and Gas. Position Paper on Smart Grids: An ERGEG Conclusions Paper, 2010. 23

[exp] ExpertChoice. http://expertchoice.com/. 60

[FJ69] Robert L Ferguson and Curtis H Jones. A computer aided decision system. *Management Science*, 15(10):B–550, 1969. 60

[FMXY12] Xi Fang, Satyajayant Misra, Guoliang Xue, and Dejun Yang. Smart grid - the new and improved power grid: A survey. *IEEE Communications Surveys & Tutorials*, 14(4):944–980, 2012. 21

[Fre08] Christoph W Frei. What if...? utility vision 2020. *Energy Policy*, 36(10):3640–3645, 2008. 21

[Fun10] Fundação Instituto Brasileiro de Geografia e Estatística - IBGE. 2010 population census: Synopsis - Salvador. *2010 Population Census*, 2010. 228

[GBB$^+$06] Volker Grimm, Uta Berger, Finn Bastiansen, Sigrunn Eliassen, Vincent Ginot, Jarl Giske, John Goss-Custard, Tamara Grand, Simone K Heinz, Geir Huse, and Others. A standard protocol for describing individual-based and agent-based models. *Ecological modelling*, 198(1):115–126, 2006. 262

[GD00] Luca Maria Gambardella and Marco Dorigo. An ant colony system hybridized with a new local search for the sequential ordering problem. *INFORMS Journal on Computing*, 12(3):237–255, 2000. 47

[GDDBKV09] J.M. González-De-Durana, O. Barambones, Enrique Kremers, and Pablo Viejo. Complete Agent based Simulation of Mini-Grids. In *Power and Energy Systems, EuroPES*, 2009. 182

[Gho84] Jay B Ghosh. Asynchronous simulation of some discrete time models. In *Proceedings of the 16th conference on Winter simulation*, WSC '84, pages 466–469, Piscataway, NJ, USA, 1984. IEEE Press. 86

[Gla02] Jean-Michel Glachant. Why regulate deregulated network industries. *J. Network Ind.*, 3:297, 2002. 18

[Gla12] Jean-Michel Glachant. Regulating networks in the new economy. *Review of Economics & Institutions/Economia, Societa e Istituzioni*, 3(1), 2012. 18

[GMR+03] A. Gabaldon, A. Molina, C. Roldan, J.A. Fuentes, E. Gomez, IJ Ramirez-Rosado, P. Lara, JA Dominguez, E. Garcia-Garrido, and E. Tarancon. Assessment and simulation of demand-side management potential in urban power distribution networks. In *Power Tech Conference Proceedings, 2003 IEEE Bologna*, volume 4. IEEE, 2003. 5, 27, 28

[GPM07] Rudolf Giffinger and Nataša Pichler-Milanović. *Smart cities: Ranking of European medium-sized cities*. Centre of Regional Science, Vienna University of Technology, 2007. 259

[GR05] Volker Grimm and Steven F Railsback. *Individual-based modeling and ecology*. Princeton university press, 2005. 43

[GRB+05] Volker Grimm, Eloy Revilla, Uta Berger, Florian Jeltsch, Wolf M Mooij, Steven F Railsback, Hans-Hermann Thulke, Jacob Weiner, Thorsten Wiegand, and Donald L DeAngelis. Pattern-oriented modeling of agent-based complex systems: lessons from ecology. *science*, 310(5750):987–991, 2005. 43

[Gup97] Himanshu Gupta. Selection of views to materialize in a data warehouse. In *Database Theory-ICDT'97*, pages 98–112. Springer, 1997. 58

[GWD+11] Nikhil Gudi, Lingfeng Wang, Vijay Devabhaktuni, Soma Shekara, and Sreenadh Reddy. A Demand-Side Management Simulation Platform Incorporating Optimal Management of Distributed Renewable Resources. *Interface*, pages 1–7, 2011. 31, 33

[Hay64] Friedrich A Hayek. The theory of complex phenomena. *The critical approach to science and philosophy*, pages 332–349, 1964. 67

[HCS06] Mike Holcombe, Simon Coakley, and Rod Smallwood. A general framework for agent-based modelling of complex systems. In *Proceedings of the 2006 European conference on complex systems*. European Complex Systems Society Paris, France, 2006. 45

[HDSQCCM07] Alfredo G Hernández-Díaz, Luis V Santana-Quintero, Carlos A Coello Coello, and Julián Molina. Pareto-adaptive ε-dominance. *Evolutionary Computation*, 15(4):493–517, 2007. 118

[HEK12] Wolfgang Hauser, Jose Evora, and Enrique Kremers. Modelling lifestyle aspects influencing the residential load-curve. In *26th European Conference on Modelling and simulation (ECMS 2012)*, 2012. 149

[Her09] Jose Juan Hernandez. *Implatación del eBusiness en pequeñas organizaciones con una orientación al modelado y la interoperabilidad.* ULPGC Scientific Publishing, Las Palmas de GC, Spain, 2009. 51, 55

[HHM11] M. Hashmi, S. Hanninen, and K. Maki. Survey of smart grid concepts, architectures, and technological demonstrations worldwide. In *2011 IEEE PES CONFERENCE ON INNOVATIVE SMART GRID TECHNOLOGIES LATIN AMERICA (ISGT LA)*, pages 1–7. IEEE, October 2011. 3, 22, 237

[HHS10] Sekyung Han, Soohee Han, and Kaoru Sezaki. Development of an optimal vehicle-to-grid aggregator for frequency regulation. *Smart Grid, IEEE Transactions on*, 1(1):65–72, 2010. 19

[Hol92] John H Holland. Complex adaptive systems. *Daedalus*, 121(1):17–30, 1992. 42

[Hor95] John Horgan. From complexity to perplexity. *Scientific American*, 272(6):104–109, 1995. 68

[Hor11] Gale Horst. Consumer engagement: Facts, myths & motivations, 2011. 24

[How12] Kerry E Howell. *An introduction to the philosophy of methodology*. Sage, 2012. 71

[HPV+08] H. Hagras, I. Packharn, Y. Vanderstockt, N. McNulty, A. Vadher, and F. Doctor. An intelligent agent based approach for energy management in commercial buildings. In *Fuzzy Systems, 2008. FUZZ-IEEE 2008.(IEEE World Congress on Computational Intelligence). IEEE International Conference on*, pages 156–162, 2008. 33

[HW05] Bernhard Humm and Frank Wietek. Architektur von data warehouses und business intelligence systemen. *Informatik-Spektrum*, 28(1):3–14, 2005. 57, 58

[IEA] IEA. About smart grids. `http://www.iea.org/topics/electricity/subtopics/smartgrids/`. 4, 20

[IK93] W Inmon and C Kelley. Developing the data warehouse, 1993. 58

[Inm92] William H. Inmon. *Building the Data Warehouse*. John Wiley & Sons, Inc., New York, NY, USA, 1992. 63

[Int09] International Energy Agency (IEA) Demand Side Management (DSM) Programme, Task XVII. Integration of Demand Side Management, Distributed Generation, Renewable Energy, Sources and Energy Storages - State of the art report, 2009. 22

[Int10] International Energy Agency. World Energy Outlook, 2010. 4

[Jac11] Jerry Jackson. Evaluating smart grid investments at us cooperative and municipal utilities. *Metering International*, 2011. 24

[Jan05] Marco A Janssen. Agent-based modelling. *Modelling in ecological economics*, pages 155–172, 2005. 42, 43

[JG12] Bryson John and Patrick D Gallagher. NIST framework and roadmap for smart grid interoperability standards, release 2.0.

National Institute of Standards and Technology, page 33, 2012. 21, 25

[Joh07] Neil F Johnson. *Two's company, three is complexity: a simple guide to the science of all sciences*. Oneworld Pubns Ltd, 2007. 67

[Kau87] Fred J Kaudel. A literature survey on distributed discrete event simulation. *SIGSIM Simul. Dig.*, 18(2):11–21, 1987. 86

[KC00] Joshua D Knowles and David W Corne. Approximating the nondominated front using the pareto archived evolution strategy. *Evolutionary computation*, 8(2):149–172, 2000. 168

[KdD12a] Enrique Kremers and J.M. Gonzalez de Durana. Emergent synchronisation properties of a refrigerator demand side management system. *Applied Energy*, 2012. 191, 220, 222

[KdD12b] Enrique Kremers and J.M. Gonzalez de Durana. Towards complex system design and management in the engineering domain - The Smart grid Challenge. In *European Conference on Complex Systems 2012*, 2012. 191, 194, 220, 223

[KdH09] S. Karnouskos and T.N. de Holanda. Simulation of a smart grid city with software agents. In *Computer Modeling and Simulation, 2009. EMS '09. Third UKSim European Symposium on*, pages 424–429, Nov 2009. 39

[KE95] J Kennedy and R Eberhart. Particle swarm optimization. In *Neural Networks, 1995. Proceedings., IEEE International Conference on*, volume 4, pages 1942 –1948 vol.4, 1995. 47, 116, 164

[KHTS09] P. Kordik, M. Hasaj, M. Tvrdy, and M. Snorek. Building Automation Simulator and Control Strategy for Intelligent and Energy Efficient Home. In *Computer Modeling and Simulation, 2009. EMS'09. Third UKSim European Symposium on*, pages 154–159, 2009. 33

[kit] KIT web. `http://www.kit.edu/`. 139

[KKE01] James F Kennedy, James Kennedy, and Russell C Eberhart. *Swarm intelligence*. Morgan Kaufmann, 2001. 162

[KR02] Ralph Kimball and Margy Ross. *The Data Warehouse Toolkit: The Complete Guide to Dimensional Modeling*. John Wiley & Sons, Inc., New York, NY, USA, 2nd edition, 2002. 63

[Kre13] Enrique Alberto Kremers. *Modelling and Simulation of Electrical Energy Systems through a Complex Systems Approach using Agent-Based Models*. KIT Scientific Publishing, Karlsruhe, 2013. 38, 40, 192

[KRH+10] Mariam Kiran, Paul Richmond, Mike Holcombe, Lee Shawn Chin, David Worth, and Chris Greenough. Flame: simulating large populations of agents on parallel hardware architectures. In *Proceedings of the 9th International Conference on Autonomous Agents and Multiagent Systems: volume 1-Volume 1*, pages 1633–1636. International Foundation for Autonomous Agents and Multiagent Systems, 2010. 45

[KSE07] Oliver Korb, Thomas Stützle, and Thomas E Exner. An ant colony optimization approach to flexible protein–ligand docking. *Swarm Intelligence*, 1(2):115–134, 2007. 47

[Kuh12] Thomas S Kuhn. *The structure of scientific revolutions*. University of Chicago press, 2012. 249

[KV10] E Kremers and Pablo Viejo. A complex system modelling approach for descentralized simulation of electrical microgirds. In *15th IEEE International Conference on Enginerring of Complex Computer Systems*, 2010. 190

[Lan90] C.G. Langton. Computation at the edge of chaos: Pahse transitions and emergent computation. *Physica D: Nonlinear Phrenomena*, 42:12–37, 1990. 188

[Lan01] Doug Laney. 3d data management: Controlling data volume, velocity and variety. *META Group Research Note*, 6, 2001. 70

[LB08] Lorenz Lutzenhiser and Sylvia Bender. The "Average American" unmasked: Social structure and differences in household energy use and carbon emissions. In *2008 ACEEE Summer Study on Energy Efficiency in Buildings*, 2008. 150

[LC11] Ferry Lienert and Matthew Carson. Smart meter roll-out for the domestic sector (GB), 2011. 23

[LCRPS04] Sean Luke, Claudio Cioffi-Revilla, Liviu Panait, and Keith Sullivan. Mason: A new multi-agent simulation toolkit. In *Proceedings of the 2004 SwarmFest Workshop*, volume 8, 2004. 44

[Lev98] Simon A Levin. Ecosystems and the biosphere as complex adaptive systems. *Ecosystems*, 1(5):431–436, 1998. 67

[LG03] Ya'an Liu and Xiaohong Guan. Purchase allocation and demand bidding in electric power markets. *Power Systems, IEEE Transactions on*, 18(1):106–112, Feb 2003. 31

[LGRT11] Martin Lukasiewycz, Michael Glaß, Felix Reimann, and Jürgen Teich. Opt4J - A Modular Framework for Meta-heuristic Optimization. In *Proceedings of the Genetic and Evolutionary Computing Conference (GECCO 2011)*, pages 1723–1730, Dublin, Ireland, 2011. 164

[LR04] Seok Won Lee and David C Rine. Case study methodology designed research in software engineering methodology validation. In *SEKE*, pages 117–122, 2004. 72

[LTDZ02] Marco Laumanns, Lothar Thiele, Kalyanmoy Deb, and Eckart Zitzler. Combining convergence and diversity in evolutionary multiobjective optimization. *Evolutionary computation*, 10(3):263–282, 2002. 118

[LV03] Justin Langseth and Nithi Vivatrat. Why proactive business intelligence is a hallmark of the real-time enterprise: Outward bound. *Intelligent Enterprise*, 5(18):34–41, 2003. 61

[LVC09] MD Lemmon, G Venkataramanan, and P Chapman. Using microgrids as a path towards smart grids. In *Position paper presented at NSF workshop in cyber physical energy systems, Baltimore, MD*, 2009. 36

[MA04] R.T. Marler and J.S. Arora. Survey of multi-objective optimization methods for engineering. *Struct Multidisc Optim*, 26, 2004. 118

[Mas13] Gilbert M Masters. *Renewable and efficient electric power systems*. John Wiley & Sons, 2013. 5, 28

[MBLA96] Nelson Minar, Roger Burkhart, Chris Langton, and Manor Askenazi. The swarm simulation system: A toolkit for building multi-agent simulations. Technical report, Santa Fe Institute Santa Fe, 1996. 44

[MCH+94] Melanie Mitchell, James P. Crutchfield, Peter T. Hraber, In G. Cowan, D. Pines, D. Melzner (editors, and Complexity Metaphors. Dynamics, computation, and the “edge of chaos”: A re-examination. In *Complexity:Metaphors, Models, and Reality*, pages 497–513. Addison-Wesley, 1994. 188

[ME04] Leon Messerschmidt and Andries Petrus Engelbrecht. Learning to play games using a pso-based competitive learning approach. *Evolutionary Computation, IEEE Transactions on*, 8(3):280–288, 2004. 48

[ME10] Anthony R Metke and Randy L Ekl. Security technology for smart grid networks. *Smart Grid, IEEE Transactions on*, 1(1):99–107, 2010. 19

[Mey03] Niels I Meyer. European schemes for promoting renewables in liberalised markets. *Energy policy*, 31(7):665–676, 2003. 18

[MH11] Janine Morley and Mike Hazas. The significance of difference: Understanding variation in household energy consumption. In *ECEEE 2011 summer study. Energy efficiency first: the foundation of a low carbon society*, 2011. 150

[MHC93] Melanie Mitchell, Peter T. Hraber, and James P. Crutchfield. Revisiting the edge of chaos: Evolving cellular automata to perform computations. *Complex Systems*, 7:89–130, 1993. 188

[Mik01] Donald C Mikulecky. The emergence of complexity: science coming of age or science growing old? *Computers & chemistry*, 25(4):341–348, 2001. 68

[Mil05] F. Milano. An Open Source Power System Analysis Toolbox. *IEEE Transaction on Power System, vol. 20, no.3*, 2005. 35, 37

[Mis86] Jayadev Misra. Distributed discrete-event simulation. *ACM Comput. Surv.*, 18(1):39–65, 1986. 86

[MJ12] Mark L Mitchell and Janina M Jolley. *Research design explained*. Cengage Learning, 2012. 72

[MOF02] OMG MOF. Omg meta object facility (mof) specification v1. 4, 2002. 52

[Mom09] James A Momoh. Smart grid design for efficient and flexible power networks operation and control. In *Power Systems Conference and Exposition, 2009. PSCE'09. IEEE/PES*, pages 1–8. IEEE, 2009. 18, 19

[MP10] M Moglia and P Perez. Modelling an urban water system on the edge of chaos. *Environmental Modelling and Software*, 25:1528–1538, 2010. 189

[MPBL10] A. Monti, F. Ponci, A. Benigni, and J. Liu. Distributed intelligence for smart grid control. In *Nonsinusoidal Currents and Compensation (ISNCC), 2010 International School on*, pages 46–58, 2010. 32

[MS84] James L McKenney and Morton M Scott. *Management decision systems: computer-based support for decision making*. Harvard Business School Press, 1984. 60

[Neg04] Solomon Negash. Business intelligence. *Communications of the Association for Information Systems*, 13(1):177–195, 2004. 61, 62

[NN12] Gregoire Nicolis and Cathy Nicolis. *Foundations of Complex Systems: Emergence, Information and Predicition*. World Scientific, 2012. 41

[NNdGW09] DT Nguyen, M. Negnevitsky, M. de Groot, and C. Wang. Demand response in the retail market: Benefits and challenges. In *Power Engineering Conference, 2009. AUPEC 2009. Australasian Universities*, pages 1–6. IEEE, 2009. 5, 30

[Nwa02] C. Nwankpa. Voltage Stability Toolbox, version 2, Center for Electric Power Engineering, Drexel University, 2002. 35, 37

[Nyl99] A Nylund. Tracing the bi family tree. *Knowledge Management*, 1999. 60

[OEC] OECD. OECD Green Growth Studies: Energy. `http://www.oecd.org/greengrowth/greening-energy/49157219.pdf`. 4, 20

[Off] Office of Electricity Delivery and Energy Reliability. Smart Grid. `http://energy.gov/oe/services/technology-development/smart-grid`. 4, 20

[OLA97] OLAP Council. OLAP Definitions, 1997. 63

[Org11] Organisation for Economic Co-operation and Development. Green Growth Studies: Energy, 2011. 3

[Ott05] Gunnar Otte. Entwicklung und Test einer integrativen Typologie der Lebensführung für die Bundesrepublik Deutschland. Construction and test of an integrative lifestyle-typology for germany. *Zeitschrift für Soziologie*, 34(6):442–467, December 2005. 150

[PAB12] Coralie Petermann, Soufian Ben Amor, and Alain Bui. A complex system approach for a reliable smart grid modeling. In *KES*, pages 149–158, 2012. 7

[Pac] Pacific Northwest National Laboratory. GridLAB-D. 38

[Par06] V Pareto. Manuale di economica politica, societa editrice libraria. *Manual of political economy*, 1971, 1906. 118

[PD11] P. Palensky and D. Dietrich. Demand side management: Demand response, intelligent energy systems, and smart loads. *Industrial Informatics, IEEE Transactions on*, 7(3):381–388, Aug 2011. 29, 31, 33

[PFR09] M. Pipattanasomporn, H. Feroze, and S. Rahman. Multi-agent systems in a distributed smart grid: Design and implementation. In *2009 IEEE/PES Power Systems Conference and Exposition*, pages 1–8. IEEE, March 2009. 8, 191

[PHS+08] Fiona A. C. Polack, Tim Hoverd, Adam T. Sampson, Susan Stepney, and Jon Timmis. Complex systems models: Engineering simulations. In *ALife XI, Winchester, UK, August 2008*, pages 482–489. MIT Press, 2008. 33, 75

[PKZK08] P. Palensky, F. Kupzog, A. A. Zaidi, and Zhou Kai. Modeling domestic housing loads for demand response. In *Industrial Electronics, 2008. IECON 2008. 34th Annual Conference of IEEE*, pages 2742–2747, 2008. 8, 32, 38, 42, 145, 250

[PL06] Jukka Paatero and Peter Lund. A model for generating household electricity load profiles. *International Journal of Energy Research*, 30(5):273–290, 2006. 150

[PM98] Katherine N Probst and Michael H McGovern. *Long-term stewardship and the nuclear weapons complex: The challenge ahead*. Resources for the Future, 1998. 19

[Pol08] Michael G Pollitt. The future of electricity (and gas) regulation in a low-carbon policy world. *Energy Journal*, 2008. 18

[Poo87] R J Pooley. *An introduction to programming in SIMULA*. Blackwell Scientific Publications, Ltd., Oxford, UK, UK, 1987. 86

[Poo01] John D. Poole. Model-Driven Architecture: Vision, Standards and Emerging Technologies. In *Workshop on Metamodeling and Adaptive Object Models, ECOOP*, 2001. 50

[Pow] Powertech. TSAT - Transient Security Assessment Tool. 37

[Pow99] D.J. Power. Decision Support Systems Glossary. DSS Resources, 1999. 63

[Pow02] Daniel Power. *Decision support systems: concepts and resources for managers*. Greenwood Publishing Group, 2002. 60, 61

[Pow07] Daniel J Power. A brief history of decision support systems. *DSSResources. COM, World Wide Web, http://DSSResources. COM/history/dsshistory. html, version*, 4, 2007. 60, 61

[PRS08] D Pudjianto, C Ramsay, and G Strbac. Microgrids and virtual power plants: concepts to support the integration of distributed energy resources. *Proceedings of the Institution of Mechanical Engineers, Part A: Journal of Power and Energy*, 222(7):731–741, 2008. 24

[RBK11] Grzegorz Rozenberg, Thomas Bck, and Joost N Kok. *Handbook of natural computing*. Springer Publishing Company, Incorporated, 2011. 12

[RC03] Ken Rudin and David Cressy. Will the real analytic application please stand up? *DM REVIEW*, 13:30–41, 2003. 62

[RCO09] N. Ruiz, I. Cobelo, and J. Oyarzabal. A direct load control model for virtual power plant management. *Power Systems, IEEE Transactions on*, 24(2):959–966, May 2009. 33

[RCV07] M Rosas-Casals and S Valverde. Topological vulnerability of the European power grid under errors and attacks. *International Journal of Bifurcations and Chaos*, 17:2465–2475, 2007. 191

[REE] Red Eléctrica de España. `http://www.ree.es/es/`. 29

[RH97] Wendy Robson and Prentice Hall. Strategic management and information systems: an integrated approach, 1997. 71

[RKU11] Anand Rajamaran, Jure Keskovec, and Jeffrey D. Ullman. *Mining of Massive Datasets*. Cambridge University Press, 2011. 110

[RLJ06] S. F. Railsback, S. L. Lytinen, and S. K. Jackson. Agent-based Simulation Platforms: Review and Development Recommendations. *SIMULATION*, 82(9):609–623, September 2006. 43, 44

[RMMG08] Sylvain Rougemaille, Frédéric Migeon, Christine Maurel, and Marie-Pierre Gleizes. Model driven engineering for designing adaptive multi-agents systems. In *Engineering Societies in the Agents World VIII*, pages 318–332. Springer, 2008. 56

[Ros77] Robert Rosen. Complexity as a system property. *International Journal Of General System*, 3(4):227–232, 1977. 68

[RPT07] Saifur Rahman, Manisa Pipattanasomporn, and Yonael Teklu. Intelligent distributed autonomous power systems (idaps). In *Power Engineering Society General Meeting, 2007. IEEE*, pages 1–8. IEEE, 2007. 21

[RVRJ11] Sarvapali Ramchurn, Perukrishnen Vytelingum, Alex Rogers, and Nick Jennings. Agent-Based Control for Decentralised Demand Side Management in the Smart Grid. In *The 10th International Conference on Autonomous Agents and Multiagent Systems-Volume 1*, pages 5–12, Taipei, Taiwan, 2011. 8, 32, 33

[RWE11] RWE Rhein-Ruhr Verteilnetz. Lastprofile, 2011. 143

[SAA$^+$02] Dag IK Sjoberg, Bente Anda, Erik Arisholm, Tore Dyba, Magne Jorgensen, Amela Karahasanovic, Espen Frimann Koren, and

Marek Vokác. Conducting realistic experiments in software engineering. In *Empirical Software Engineering, 2002. Proceedings. 2002 International Symposium n*, pages 17–26. IEEE, 2002. 72, 243

[SC05] Margarita Reyes Sierra and Carlos A Coello Coello. Improving PSO-based multi-objective optimization using crowding, mutation and ? -dominance. In *In EMO?2005, pages 505?519. LNCS 3410*, pages 505–519. Springer-Verlag, 2005. 119, 164, 168

[SCC02] William R.. Shadish, Thomas D Cook, and Donald Thomas Campbell. *Experimental and quasi-experimental designs for generalized causal inference*. Wadsworth Cengage learning, 2002. 72

[SCF+06] Jean-Sébastien Sottet, Gaëlle Calvary, Jean-Marie Favre, Joëlle Coutaz, Alexandre Demeure, and Lionel Balme. Towards model driven engineering of plastic user interfaces. In *Satellite Events at the MoDELS 2005 Conference*, pages 191–200. Springer, 2006. 56

[Sch10] Thomas Schwarz. Lebensstile und Wählerverhalten in Stuttgart. Ergebnisse der Stuttgarter Lebensstilbefragung 2008. *Statistik und Informationsmanagement*, July 2010. 151, 152

[SD06] Mohsen Amini Salehi and Hossein Deldari. Grid load balancing using an echo system of intelligent ants. In *Parallel and Distributed Computing and Networks*, pages 47–52, 2006. 46

[Sei03] Edwin Seidewitz. What models mean. *IEEE software*, 20(5):26–32, 2003. 51

[SGRM07] Frank Sensuß, Massimo Genoese, Mario Ragwitz, and Dominik Möst. Agent-based simulation of electricity markets-a literature review. *Energy Studies Review*, 15(2):2, 2007. 36, 191

[SH05] Alena Shmygelska and Holger H Hoos. An ant colony optimisation algorithm for the 2d and 3d hydrophobic polar protein folding problem. *BMC bioinformatics*, 6(1):30, 2005. 47

[Shi04] Y Shi. Feature article on particle swarm optimization. *IEEE Neural Network Society, Feature Article*, pages 8–13, 2004. 48

[SIF07] J. A. Short, D. G. Infield, and L. L. Freris. Stabilization of grid frequency through dynamic demand control. *Power Systems, IEEE Transactions on*, 22(3):1284–1293, 2007. 191

[Sim91] Herbert A Simon. *The architecture of complexity*. Springer, 1991. 8, 12

[Spe10] Alexa Spence. *Public Perceptions of Climate Change and Energy Futures in Britain: Summary Findings of a Survey Conducted from January to March 2010*. Cardiff University, 2010. 24

[Spr80] Ralph H Sprague. A framework for the development of decisoin support systems. *MIS quarterly*, 4(4):1–26, 1980. 58

[SPT06] S. Stepney, F. A. C. Polack, and H. R. Turner. Engineering emergence. In *Engineering of Complex Computer Systems, 2006. ICECCS 2006. 11th IEEE International Conference on*, page 9 pp., 2006. 33, 41

[SS03] Kwang Mong Sim and Weng Hong Sun. Ant colony optimization for routing and load-balancing: survey and new directions. *Systems, Man and Cybernetics, Part A: Systems and Humans, IEEE Transactions on*, 33(5):560–572, 2003. 46

[SSS09] John L Sullivan, IT Salmeen, and CP Simon. Phev marketplace penetration: An agent based simulation. 2009. 228

[Ste89] Ralph E Steuer. *Multiple criteria optimization: theory, computation, and application*. Krieger Malabar, 1989. 118

[Str08] Goran Strbac. Demand side management: Benefits and challenges. *Energy Policy*, 36(12):4419–4426, December 2008. 5, 27, 28, 29, 31

[Swa74] E Burton Swanson. Management information systems: appreciation and involvement. *Management Science*, 21(2):178–188, 1974. 58

[SWC^{+}02] J.P. Shim, Merrill Warkentin, James F. Courtney, Daniel J. Power, Ramesh Sharda, and Christer Carlsson. Past, present, and future of decision support technology. *Decision Support Systems*, 33(2):111 – 126, 2002. ¡ce:title¿Decision Support System: Directions for the Nest Decade¡/ce:title¿. 59

[Tesa] Leigh Tesfatsion. AMES Market Package. 35, 37

[Tesb] Leigh Tesfatsion. DCOPFJ Package. 35, 37

[Tes03] Leigh Tesfatsion. Agent-based computational economics: modeling economies as complex adaptive systems. *Information Sciences*, 149(4):262–268, 2003. 67

[TM11] Marc Tritschler and William Mackay. Uk smart grid cyber security. *KEMA, June 25th*, 2011. 24

[tpftenotf] European technology platform for the electricity networks of the future. Smart Grid definition. 21

[TW04] Seth Tisue and Uri Wilensky. Netlogo: A simple environment for modeling complexity. In *International Conference on Complex Systems*, pages 16–21, 2004. 44

[TWDS09] Chris Thompson, Jules White, Brian Dougherty, and Douglas C Schmidt. Optimizing mobile application performance with model–driven engineering. In *Software Technologies for Embedded and Ubiquitous Systems*, pages 36–46. Springer, 2009. 56

[Uni00] University of Waterloo. TEFTS Program, University of Waterloo, 2000. 35, 37

[US] Climate Change Division US EPA. Greenhouse Gas Emissions: Transportation Sector Emissions. 3, 19

[Vac12] John R Vacca. *Computer and information security handbook*. Newnes, 2012. 71

[vis] VisiCalc information. `https://archive.org/details/VisiCalc_1979_SoftwareArts`. 60

[VKE+13] Pablo Viejo, Enrique Kremers, Jose Evora, Jose Juan Hernandez, Mario Hernandez, O. Barambones, and J.M. Gonzalez de Durana. Criticality in complex sociotechnical systems, an empirical approach. In *European Conference on Complex Systems 2013 (ECCS'13)*, 2013. 7, 187, 222

[War12] Colin Ware. *Information Visualization: Perception for Design*. Morgan Kaufmann Publishers Inc, 2012. 63, 111

[WF07] Andrew Wright and Steven Firth. The nature of domestic electricity-loads and effects of time averaging on statistics and on-site generation calculations. *Applied Energy*, 84(4):389–403, 2007. 191

[Wid95] Jennifer Widom. Research problems in data warehousing. In *Proceedings of the fourth international conference on Information and knowledge management*, pages 25–30. ACM, 1995. 58

[Wil02] C Willen. Airborne opportunities. *Intelligent Enterprise*, 5(2):11–12, 2002. 62

[Wol02] S Wolfram. *A new kind of service*. 2002. 188

[Wol12] Maarten Wolsink. The research agenda on social acceptance of distributed generation in smart grids: Renewable as common pool resources. *Renewable and Sustainable Energy Reviews*, 16(1):822–835, 2012. 24

[Woo02] Michael J. Wooldridge. *An Introduction to Multiagent Systems*. John Wiley, Chichester, West Sussex, United Kingdom, 2002. 32, 43

[Woo09] Michael Wooldridge. *An Introduction to MultiAgent Systems - Second Edition*. John Wiley And Sons, 2009. 80

[Wor10] World Scientific Publishing Co. Advances in complex systems, 2010. 38

[WS98] D.J. Watts and S.H. Strogatz. Collective dynamics of small world networks. *Nature*, 393:440–442, 1998. 188

[WSZ+04] Mark P Wachowiak, Renata Smolíková, Yufeng Zheng, Jacek M Zurada, and Adel Said Elmaghraby. An approach to multimodal biomedical image registration utilizing particle swarm optimization. *Evolutionary Computation, IEEE Transactions on*, 8(3):289–301, 2004. 48

[WW09] Joakim Widén and Ewa Wäckelgard. A high-resolution stochastic model of domestic activity patterns and electricity demand. *Applied Energy*, 87(6):1880–1892, 2009. 191

[XABO+14] Dimitrios Xenias, Colin Axon, Nazmiye Balta-Ozkan, Liana Cipcigan, Peter Connor, Davidson Davidson, Alexa Spence, Gary Taylor, and Lorraine Whitmarsh. Scenarios for the Development of Smart Grids in the UK: Literature Review. Technical report, UK Energy Research Centre, 2014. 20, 23, 24

[YKF+00] Hirotaka Yoshida, Kenichi Kawata, Yoshikazu Fukuyama, Shinichi Takayama, and Yosuke Nakanishi. A particle swarm optimization for reactive power and voltage control considering voltage security assessment. *Power Systems, IEEE Transactions on*, 15(4):1232–1239, 2000. 46, 48

[ZG97] R.D. Zimmerman and D. Gan. Matpower, Documentation for Version 2, Power System Engineering Research Center, Cornell University, 1997. 35, 37

[Zho] M. Zhou. InterPSS. 35, 37

[ZLB07] Günther Zauner, Daniel Leitner, and Felix Breitenecker. Modeling structural-dynamics systems in modelica/dymola, modelica/mosilab and anylogic. In *EOOLT*, pages 99–110, 2007. 45

[ZSL00] Chunkai Zhang, Huihe Shao, and Yu Li. Particle swarm optimisation for evolving artificial neural network. In *Systems, Man, and Cybernetics, 2000 IEEE International Conference on*, volume 4, pages 2487–2490. IEEE, 2000. 48

[ZW97] Marvin V Zelkowitz and Dolores Wallace. Experimental validation in software engineering. *Information and Software Technology*, 39(11):735–743, 1997. 72

[ZZGW05] Fahua Zhu, Youfei Zheng, Xulin Guo, and Sheng Wang. Environmental impacts and benefits of regional power grid interconnections for china. *Energy Policy*, 33(14):1797–1805, 2005. 19, 20

Declaration

I herewith declare that I have produced this work without the prohibited assistance of third parties and without making use of aids other than those specified; notions taken over directly or indirectly from other sources have been identified as such. This work has not previously been presented in identical or similar form to any examination board.

The dissertation work was conducted from 2010 to 2014 under the supervision of Francisco Mario Hernandez Tejera and Jose Juan Hernandez Cabrera at the University of Las Palmas de Gran Canaria.

Las Palmas de Gran Canaria,

INSTITUTO UNIVERSITARIO
SIANI
INGENIERIA COMPUTACIONAL

UNIVERSIDAD DE LAS PALMAS DE GRAN CANARIA

Instituto Universitario de Sistemas Inteligentes y Aplicaciones Numéricas en Ingeniería
Edificio Central del Parque Tecnológico
Campus Universitario de Tafira
35017 Las Palmas de Gran Canaria
www.siani.es