Self-Forecasting Energy Load Stakeholders for Smart Grids

In Fulfilment of the Requirements for the Degree of

Doctor of Engineering

Faculty of Informatics
Karlsruhe Institute of Technology (KIT)

Approved

Dissertation

by

Dejan Ilić

Born on 31. August 1983 in Belgrade, Serbia

Examination Date:	11 July 2014
Primary Examiner:	Prof. Dr.-Ing. Michael Beigl
Secondary Examiner:	Prof. Dr. Orestis Terzidis

For my parents,
who have selflessly invested their time and energy
to get me here.

Acknowledgement

First of all I would like to thank my primary adviser Prof. Dr.-Ing. Michael Beigl for this amazing opportunity. Not only I need to thank him for his support, but I also need to thank to his entire team at TecO. The given freedom and scientific support of the TecO group helped me in following my own creativity in achieving the goal. Secondly, I would like to thank my second adviser Prof. Dr. Orestis Terzidis, who encouraged my ideas and beliefs to deliver this contribution. His suggestions and observations significantly improved my understanding of the value behind this work. Additionally, I need to thank Prof. Dr. Alberto Bartoli who took the big decision on supplementing the knowledge I gained in my academical path.

It was never easy, but in simple words there is one important thing I learned from my SAP supervisor Stamatis Karnouskos: "It does not matter how slowly you go as long as you do not stop". Surely there is no way to thank him for dedicating countless hours of patience and expertise to help me rise the value of my work from all possible aspects. I also dedicate these acknowledgements to Per Goncalves da Silva, who invested his time to discuss and advance my ideas. I would also like to thank Dr. Axel Schröpfer for his continuous scientific and, most importantly, friendly support from my very first day at SAP. Additionally I would like to thank my former SAP colleagues, in particular for clarifications of Prof. Dr. Anke Weidlich and efforts of my former manager Dr. Zoltan Nochta. I also need to thank for the contributions of the students I supervised, and particularly to Malte Jacobi, Martin Griesemer and Martin Wilhelm.

Although I'm not really sure, but my instinct tells me that every single moment of my life plays a role here. Therefore I have to thank many of my friends that we're always on my side. I first have to thank Miljan Kastratović, who picked up my hand and took me to my first academic steps. From that day on, I knew where to walk. Without friends on my side I'm sure I wouldn't make it. So I give out my most sincere gratitude (by last name ordering) to Mauro Drusković, Goran Dželajlija, Dr. Dawud Gordon, Stevan Kastratović, Andrija Mesterović, Igor Alberto Molina, Nemanja Nikitović, Tirdad Rahmani, Dr. Kay-Uwe Schmidt, Pasqualino Stancanelli, Miroslav Stojković, Nenad Vasiljević and many others who always motivated and supported me in taking this step. Most importantly I will always remember the great moments we had together. Thanks!

Finally, I cannot thank enough to my parents Vladimir and Snežana, to whom I dedicate this work, as well as to my family members. Words here won't help much, but I have to express my deepest gratitude to all of them. They simply deserved it. Same I have to give out to Cathaysa, my wife, whose encouraging words appeared impossible to me. Yet, she was right.

Abstract

Predictability of energy loads is a big challenge for electricity grids. As the consumption loads are forecasted, the system must stay in balance even when a forecast error occurs. These errors, or imbalances, are simply pushed upstream to the parties responsible for balancing them. With reliable sources of production in place such techniques have been successfully applied for a century now. However, the penetration of unreliable energy supply from renewable sources will completely change operation of energy industry. By adopting the renewable sources of energy, today even the traditional consumers became producers, or the so called "prosumers". As such, not only energy is intermittently produced, it will also come from distributed resources. This takes complexity one step further, where forecasted consumer loads are powered by unpredictable and distributed resources. Thereby the value of reliability will significantly grow.

Although many events cannot be predicted, such as natural disasters damaging power lines, a significant portion of unpredictability comes directly from consumers. In order to improve system reliability, the emerging business models and roles in Smart Grids call for active participation by traditionally passive consumers. Such opportunities include for instance active involvement in grid operations, participation in local energy markets, or demand response programs etc. To participate in such programs, an accurate self-forecast of energy loads is a prerequisite of key importance. If a prosumer could achieve determinism in his energy signature, via highly accurate load forecast and potentially control over the deviations from that forecast, he could act as a resource that can reliably support needs of other stakeholders. Still, not all stakeholders can achieve it, but for those who do (by any means) additional benefits are expected.

This work uses Smart Grids as foundation to build a solution that enables active contribution of the traditionally passive consumers. The challenges are to (1) enable an efficient communication in between stakeholders, (2) reach sufficient forecast accuracy of an individual or a small group of consumers, and to (3) build a system that enables active involvement of the traditionally passive consumers. The main contribution of this dissertation consists of:

- Design and (real world) evaluation of an enterprise integration and energy management system – including scalability and performance issues

- Assessment of the forecast accuracy impact on small scale aggregations and relevance of energy storage solutions to absorb the forecast errors

- System proposal for enabling the deterministic behaviour of traditionally passive consumers – evaluated on a real world case

Following the vision of Smart Grids, this work proposes an enterprise integration and energy management system as the foundation for efficient communication between stakeholders. Their awareness is raised by the accessibility of the energy services designed and evaluated in this work. Key performance points of their scalability are also investigated to support a large number of smart meters that will stream their energy readings at high resolution e.g. 15 minutes. Even though the data can be collected, many services are highly time dependent and on-demand near real-time data processing must be in place as well. Great amounts of continuously streaming data challenge such systems. An evaluation of the entire infrastructure is made in a real world trial with 5000 smart meters, as well as the actual implementation of an application built on top of the platform's energy services.

Traditionally, an accurate energy forecast is achieved by large scales of customer aggregation. However, many added-value services of Smart Grids are envisioned for smaller scales, or even individuals, thus a question if a sufficient accuracy can be achieved by them is raised. This work contributes by investigating how accurate smaller scales of aggregation can be. Results show that small scales, e.g. of 150–200 residential stakeholders, or even individuals, e.g. commercial building, can already achieve a significant accuracy. This accuracy is still lower than what retailers of today would achieve (in an aggregation of tens of thousands), and static storage solutions are investigated for further improvement. The results show the potential to address the forecast errors with capacities of 6–10% of stakeholder's daily consumption. Still, the static solutions bear costs and this work investigates potential of available assets to replace them. Electric vehicles were identified as a promising alternative. Although their behaviour is dynamic, the simulation results show their huge potential in absorbing the errors.

If an accurate self-forecast of a stakeholder (or group of them) is achieved by absorbing the errors locally, an external stakeholder cannot be aware of it. Hereby the same infrastructure of smart metering is proposed to be used for continuous reporting of the self-forecasted intervals. Still, a smart energy system needs to be in place to autonomously support stakeholders in respecting their reported load. With this system in place, deterministic behaviour is achieved and new opportunities for many Smart Grid stakeholders are expected. Since the stakeholder's determinism can be measured, self-forecasting stakeholders can benefit from the flexibility based on the state of their storage. This work proposes an architecture that is used for system design that is evaluated for one of the proposed strategies. The evaluation results showed, in a real world case, that combined contribution of this thesis will lead us to existence of self-forecasting energy load stakeholders.

Deutsche Zusammenfassung

Die Vorhersagbarkeit von Energieverbräuchen ist eine große Herausforderung für Stromnetze. Obgleich Verbräuche vorhergesagt werden können, muss das Gesamtsystem ausgeglichen sein, auch wenn die Vorhersage einen Fehler beinhaltet. Diese Fehler werden an die darüberliegenden Parteien weitergeleitet, welche für einen entsprechenden Ausgleich verantwortlich sind. Dank verlässlicher Produktionsquellen konnten derartige Techniken ein Jahrhundert lang erfolgreich eingesetzt werden. Die Durchdringung mit unzuverlässiger Energie aus erneuerbaren Quellen wird den Betrieb der Energieindustrie jedoch vollständig verändern. Durch die Einbringung erneuerbarer Energiequellen wurden herkömmliche Konsumenten zu Produzenten, sogenannte Prosumenten. Somit wird Energie nicht nur periodisch erzeugt, sie stammt auch von unterschiedlichen verteilten Ressourcen. Dies erhöht den Grad der Komplexität, indem vorhergesagte Verbräuche durch unvorhersagbare verteilte Ressourcen bedient werden. Der Wert der Vorhersage wird deshalb signifikant an Bedeutung gewinnen.

Obgleich viele Ereignisse nicht verhindert werden können, wie etwa durch Naturkatastrophen beschädigte Stromleitungen, stammt ein signifikanter Anteil der Unvorhersagbarkeit unmittelbar vom Konsumenten. Um die Systemverlässlichkeit zu erhöhen fordern aufkommende Geschäftsmodelle in Smart Grids die aktive Teilnahme von traditionell passiven Konsumenten. Derartige Möglichkeiten umfassen beispielsweise die aktive Einbindung in den Netzbetrieb, Teilname an lokalen Energiemärkten sowie Programmen zu Angebot und Nachfrage. Für die Teilnahme an solchen Programmen ist eine genaue Vorhersage des eigenen Energieverbrauchs eine maßgebliche Notwendigkeit. Sollte es dem Prosumenten gelingen, seine Energiesignatur durch höchstgenaue Verbrauchsprognosen und eigene Kontrolle in deren Abweichung zu bestimmen, könnte er verlässlicher anderer Teilnehmer unterstützen. Wenn dies auch nicht für alle Teilnehmer gilt, so werden für diejenigen, denen es gelingt, zusätzliche Anreize erwartet.

Diese Arbeit verwendet Smart Grids als Grundlage um eine Lösung zu bauen, die ein aktives Beitragen von traditionell passiven Konsumenten ermöglicht. Die Herausforderungen sind (1) das Ermöglichen einer effizienten Kommunikation zwischen den Teilnehmern, (2) das Erreichen einer hinreichend genauen Vorhersage individueller Konsumenten oder kleiner Gruppen von Konsumenten und (3) der Aufbau eines Systems, welches einen aktiven Einbezug traditionell passiver Konsumenten ermöglicht. Die Hauptbeiträge dieser Dissertation bestehen in:

- Entwurf und (praktische) Evaluierung eines Unternehmensintegration- und Energieverwaltungssystems – unter Einbezug der Schwierigkeiten von Skalierbarkeit und Performanz

- Bewertung des Einflusses der Vorhersagegenauigkeit auf Aggregationen im Kleinen und Relevanz von Energiespeicherlösungen, um Vorhersagefehler zu absorbieren

- Systemvorschlag zur Ermöglichung deterministischen Verhaltens tradionell passiver Konsumenten – evaluiert in einem echten Anwendungsfall

Der Vision von Smart Grids folgend, schlägt diese Arbeit ein Unternehmensintegrations - und Energieverwaltungssystem vor als Grundlager für effiziente Kommunikation zwischen Teilnehmern. Deren Aufmerksamkeit wird durch die Benutzbarkeit der Energiedienste geweckt, welche in dieser Arbeit entworfen und evaluiert werden. Schlüsselpunkte hinsichtlich der Skalierbarkeit werden ebenfalls untersucht, um eine große Anzahl von Smart Metern zu unterstützen, welche ihre Energiewerte in großer Auflösung, etwa 15 minütig, senden. Obgleich der Möglichkeit Daten zu sammeln, sind viele Dienste sehr zeitkritisch und erfordern darüber hinaus bedarfsgesteuerte nah-echtzeit Datenverarbeitung. Große Menge kontinuierlicher Daten strapzieren solche Systeme. Es wird eine Evaluierung der gesamte Infrastruktur wird anhand 5000 Smart Metern mit echten Daten durchgeführt, sowie eine Evaluierung einer auf der Energiedienste der Plattform aufbauenden Anwendungsimplementierung.

Im herkömmlichen Ansatz wird eine genaue Vorhersage des Energieverbrauchs durch die Aggregation von großen Konsumentendatenmengen erzielt. Viele wertschöpfende Dienste von Smart Grids sehen allerdings Datenmengen kleinerer Gruppen oder Individuen vor. Daraus resultiert die Frage, welche Genauigkeit von ihnen erzielt wird. Diese Arbeit leistet einen Beitrag, indem untersucht wird, wie genau kleine Datenmengen sich diesbezüglich verhalten. Die Ergebnisse zeigen, dass auch kleine Datenmengen, etwa 150–200 Anwohner, oder sogar Individuen, beispielsweise kommerzielle Gebäude, eine signifikante Genauigkeit erzielen. Die Genauigkeit ist noch unter jener vom Einzelhandel erzielten Genauigkeit (Aggregation Zehntausender). Ferner werden statische Speicherlösungen hinsichtlich zukünftiger Verbesserungen untersucht. Die Ergebnisse zeigen, dass Potential existiert, die Vorhersagefehler mit Kapazitäten von 6–10% des täglichen Teilnehmerverbrauchs zu kompensieren. Dennoch verursachen statische Speicherlösungen Kosten, wobei diese Arbeit untersucht, wie diese durch vorhandene Anlagen ersetzt werden können. Elektronische Fahreuge wurden dafür als vielversprechende Alternative idenfiziert. Obwohl deren Verhalten dynamisch ist, zeigen Ergebnisse einer Simulation ihr großes Potential, um Fehler zu absorbieren.

Selbst wenn eine genaue Eigenvorhersage von Teilnehmern (oder Gruppen von diesen) durch Absobierung der Fehler lokal erreicht wird können externe Teilnehmer dies nicht wahrnehmen. Hier wird die gleiche Infrastruktur von Smart Metern vorgeschlagen, um selbst vorhergesagte Intervalle zu melden. Nichtsdestotrotz wird ein Smart Energy System benötigt, um autonome Teilnehmer in Hinblick auf ihre gemeldeten Werte zu unterstützen. Mit dem Vorhandensein eines derartigen Systems wird ein deterministisches Verhalten erreicht und neue

Möglichkeiten für viele Smart Grid Teilnehmer erwartet. Da der Determinismus der Teilnehmer gemessen werden kann, können selbstvorhersagende Teilnehmer von der Flexibilität basierend auf dem Zustand ihres Speichers profitieren. Diese Arbeit schlägt eine Architektur vor, die als Systementwurf dient, der für einer der vorgeschlagenen Strategien evaluiert wird. Die Evaluierung zeigt anhand eines Beispiels aus der realen Welt, dass die Kombination der Beiträge dieser Thesis zur Existenz von selbstvorhersagenden Energieverbräuchen von Teilnehmer führen wird.

Contents

1 Introduction

Technology is embedded in almost every aspect of our daily lives. Information and Communication Technologies (ICT) are now present even in small physical objects that enable them to communicate and interact, eventually forming the Internet of Things (IoT) [1]. Intelligent networked devices (such as sensors and actuators) amalgamated with everyday objects, house-hold appliances, industrial systems, etc. lead to the fusion of the physical and virtual worlds [2]. This wide availability of data acquisition and communication is the basis of a global ecosystem of interacting entities cooperating via innovative cross-domain services [3]. In the future, this trend is expected to bring more information, more detail, more speed that will finally reshape the world we know today [1].

All of this has an effect on businesses [4]. Many businesses are facing a new technology landscape, going towards mobile [4] and cloud technologies [5], but more importantly the active and real-time involvement of their end customers. Although information provided to consumers is of high relevance for keeping them informed, sometimes the information about the overall system state is of greater importance for them [6]. In fact, feedback is the critical component required to maintain stability of many (automated) systems, thus traditionally isolated business would also benefit from feedback and wide acceptance of ICT is the enabler. This is the turning point, as today's acceptance of ICT not only will enable users to be better informed of a system state [7], but in this set-up they will be able to actively influence it. This is a significant change from current passive user models, e.g. in the energy domain, where residential consumers consume without caring about the needs of other stakeholders involved. By integrating users willing to contribute the system's operation, win-win situations arise that are beneficial for many stakeholders [8].

The motivation of the research is described in section 1.1 list our major questions that arise mentioned in section 1.2. Although the same principles of this thesis could be applied in different industries, the focus here is on energy domain. The answers to research questions contributed to point out the main contribution of this work in section 1.3. Finally, the outline of the thesis structure can be found in section 1.4.

1.1 MOTIVATION

Immediate access to information changed completely behaviour of both businesses and customers [7]. In fact, some business are affected by technology to an extent that their customer delivery channels have changed. As an example, we can see this trend in retail, where ordering of the same products can be made directly from producers via Internet. Therefore, many have started to change their business models, but it is still to early to predict the final outcome of this revolution. As one can imagine, some businesses adopted the technological shift in their model and further improved it due the speed of information. Old business processes had to be adjusted as well, and some technologies resulted in great advantages for consumers. The energy industry is an example of that, as today energy is bought/sold on Internet trading platforms that attract more and more participants due its effectiveness i.e. ease of accessibility and competitive prices.

The focus of this thesis is on awareness in electricity grids to help the many problems they face due to the unpredictability of energy loads [9, 10, 11, 12]. In energy, forecasting plays a crucial role for planning and management activities [13]. Traditionally, the electricity load needs are forecasted so that delivery by central generators can be properly scheduled to address the loads in a cost-effective way. However forecasting brings uncertainties [14], and therefore even if scheduling is properly done, imbalances are to be expected due forecast errors. If managing such a system is challenging, one may imagine the complexity growth with the adoption of Distributed Energy Resources (DER). The concept of DER moved us from scheduling a few centralized power plants, to big, or even huge, number of distributed smaller plants e.g. such as solar panels on rooftop of residential consumers. Not only is their decentralized managing difficult, but many of them are expected to be Renewable Energy Sources (RES), as from sun and wind, and therefore they will produce energy intermittently [10]. This brings uncertainty even from the supply side of today's electricity model [15], while reliable resources are crucial for operation of the electricity grids.

To make this vision of power networks (adopting DER and RES) successful, the Smart Grid concept was introduced [16]. One basic functionality that has already been realised is the Advanced Metering Infrastructure (AMI) that enabled smart metering [17]. That allowed access to the energy consumption data of smaller consumers to be collected remotely and directly from an on-premise meter, and on much higher resolutions (e.g. 15 minutes) than before. Although the initial application for smart metering was billing, the same infrastructure allowed many more innovative applications to be developed [18]. However, most of such cases benefit from analysis of this "Big Data" [19], e.g. user profiling, but there are not many efforts that go beyond analytics where stakeholders actively use the Smart Grid infrastructure to make power networks and resources more efficient. Thus, many Smart Grid concepts emerged [20, 16] to enable the next-generation of electricity networks, where stakeholders of the grids are interconnected. With technology in place, one can delegate part of the reliability intelligence to the traditionally passive consumers, which now can actively contribute [21]. As stakeholders are interconnected, if the energy consumption requirements can be accurately communicated upfront, in similar way retailers do today, further optimization can be expected (as described in Appendix A). From the perspective of what is done with AMI today, this is similar to collection of smart meter readings with an offset, e.g. at 9:30 collect smart meter readings for 21:45 – 12 hours in advance. If a stakeholder is capable of forecasting accurately its own load, reporting this forecast and making sure that the load report is respected, we can view this behaviour as deterministic.

Deterministic behaviour is a prerequisite of many added-value services envisioned for Smart Grids. As an example, local energy markets were proposed [22, 23, 24], however to benefit from such services one would need sufficient forecast accuracy [25]. Another example is that some services may not be adopted by stakeholders if their load changes cannot be measured [26], such as flexibility [27]. In other words, behavioural change cannot be verified if one would not be able to measure that change [12]. If determinism can be achieved by a stakeholder, it would be possible to benefit from the added-value services. This is an another pivotal point that marks the transition from passive consumers of electricity, to active ones, which are now enabled to actively contribute needs of other stakeholders. By actively adjusting their loads [28], they can further contribute (as a DER) to help addressing the unpredictability we will face in electricity grids of future.

There are certain limitations depending upon what pre-conditions are considered to be deterministic evidence of an event, but to achieve determinism in this case one needs to trade-off the capability of resources in ownership e.g. assets that can absorb errors of a self-forecast. Traditional storage solutions, such as Battery Energy Storage System (BESS), demonstrated to be successful in improving reliability of many power systems. They helped high penetration [29] and unpredictability balancing of RES [10] today, so one can expect to have same application on the side of prosumers. Still, energy storage systems bear costs but they can be reduced, or eliminated, by maximizing usage of on-premise resources of a stakeholder, or a group of them (such as in a neighbourhood). Due the great potential of Vehicle-to-Grid (V2G) in electricity grids [30] and technology making the Electric Vehicle (EV) concept acceptable for consumers, penetration of EVs [31] need to be considered in future applications. As transportation vehicles of today are 96% idle [32], one must understand (as this work will show) that potential of using them for future storage systems cannot be omitted. The opportunity EVs bring to a stakeholder achieving the deterministic energy signature need to be closer investigated.

All these aspects indicate a significant change to the way stakeholders can interact with the electrical grid in the near future [3]. Among ongoing research and development projects, there are efforts towards better grid management, integration of smart-houses [33] and smart-buildings, accommodation of intermittent energy resources including EVs, demand-response schemes [26], local energy markets for business interactions, etc. To facilitate this interaction, new services [21] and tools [34] provide near real-time features, such as access to historical energy consumption and production, load forecasting, generation mix etc. However, even if a stakeholder is completely aware of system's needs, without being "predictable" one cannot verify its active contribution to other stakeholders. This way no benefit of a stakeholder can be measured. If an active contribution can be measured, the traditional infrastructure owners will be able to take advantage of the new stakeholder capabilities by tapping into their flexibility with respect to adjusting their energy behaviour [28]. Furthermore such a system is designed in this thesis and its autonomous operation is achieved if assets in ownership can absorb the errors of a self-forecast.

1.2 RESEARCH CHALLENGES

The Smart Grid is a complex ecosystem of heterogeneous entities that can interact via modern ICT and benefit from the plethora of information that it brings [35, 28]. Its realization will empower advanced business services, offering their stakeholders desired services [36] such as near real-time information, as well as new analytical services and applications [34]. Figure 1 gives an excellent example, illustrated by the NOBEL project [37], of how complex is interconnecting prosumers in Smart Cities – that will be evaluated later. Even with the entire infrastructure in place, for traditionally passive consumers to benefit from the added-value services envisioned, the deterministic and active behaviour is a pre-requisite. The main research question is raised here as **"how to incorporate traditionally passive stakeholders to generate revenue from the Smart Grid added-value services?"**. In order to address that, we need to answer the major research questions listed in following paragraphs.

CHALLENGE 1 – ACTIVE STAKEHOLDERS The Internet of Things envisions billions of connected devices sensing, possibly collaborating and providing real world information to and from enterprise systems [1]. Even if information can be gathered, one needs to pre-process it and finally offer it to other stakeholders to act upon it. Although this already poses many challenges [38], including other requirements such as designing enterprise services to access the information via the Internet, one need to be able to do this on scale. The research question here is **"how to enable an efficient bi-directional communication in between the stakeholders of future Smart Grids?"**. If fine-grained information can be exchanged through AMI in a timely manner, the awareness of stakeholders and their feedback can be made, thus the Smart Grid promises a new generation of innovative applications and services that can be realized [34].

CHALLENGE 2 – ACHIEVING FORECAST ACCURACY Forecasting already plays a crucial role for electricity grids [13]. It is applied to huge groups of energy

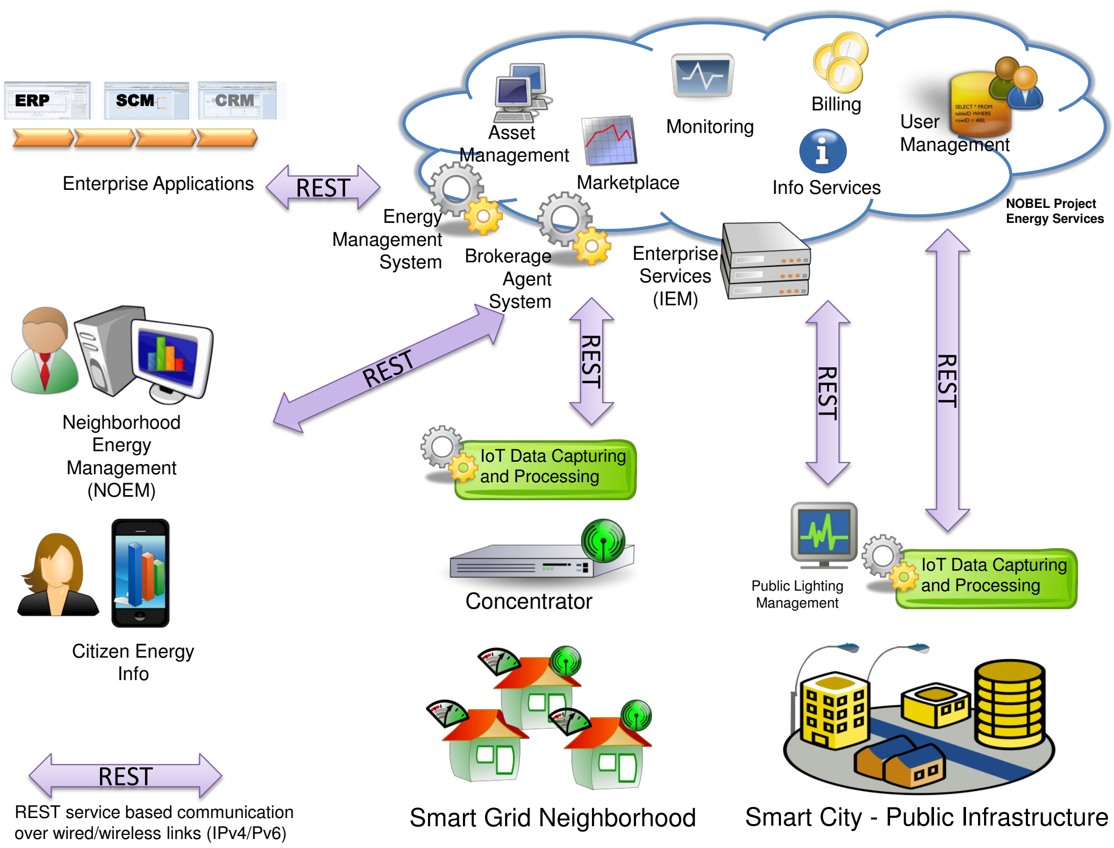


Figure 1.: Overview of a platform providing services and to interconnect stakeholders in a Smart City

consuming (or producing) entities in order to achieve a greater forecast accuracy [39]. However, as smaller scales of stakeholder aggregation tend to have lower levels of predictability, their individual forecast would not make economical sense [40]. Still, the added-value services offered to stakeholders of Smart Grids will depend on having an accurate forecast [25]. The research question here is **"how to achieve sufficient forecasting accuracy of stakeholders on lower scales?"**. Using grouping on smaller scales, BESS as well as assets in potential ownership, such as EV fleets, need to be understood. Achieving a sufficient level of accuracy is important [25], because load forecast errors may result in high penalties for stakeholders which would be a show stopper for many benefits envisioned by Smart Grids [16].

CHALLENGE 3 – DETERMINISTIC BEHAVIOUR The emergence of Smart Grids implies also new roles [18] that aim to deliver, among other things, a wide-range of better or new value-added applications and services. Some of these may not feasible for the current stakeholders – as their electricity signature is not predictable enough [25]. However, if sufficient forecast accuracy can be achieved, they can execute a self-forecast and report it in advance (to the respective stakeholders). This deterministic behaviour of a stakeholder also makes him a measurable resource

[12] and open doors to new opportunities. The research question here is **"how to enable traditionally passive stakeholders to be have a deterministic energy load in Smart Grids?"**. This, however, does not imply that actual behaviour of a stakeholder is changed, e.g. business processes, but rather he appears to have an accurate self-forecast. For those who achieve deterministic behaviour can participate in many city-wide energy management schemes such as demand-response [26], local energy markets [28], improving grid operations [41], etc.

1.3 THE CONTRIBUTION

Electricity networks are undergoing a significant change towards more adaptive, intelligent, self-managing, collaborative and information-driven systems [42]. The contributions of this dissertation are new **concepts, architectures and evaluation results** towards addressing the aforementioned challenges. Empowered by modern IT technologies [35], two key concepts are introduced and evaluated using real world data, namely **Variable Energy Storage (VES)** and **Self-Forecasting EneRgy load Stakeholder (SFERS)**. As this thesis will show, they are promising concepts for Smart Grids where they can contribute to a more versatile and intelligent network of collaborating actors. Eventually this will lead to better utilization of resources, better management, and will enable achievement other goals, such as energy efficiency.

The main contributions are:

- Enabling and assessing the visionary scenarios of the Smart Grids where stakeholders are able to interact with enterprise energy services. This involves design of an architecture and services for such system, referred to as an Integration and Energy Management system (IEM). The proposed solution was developed and integrated in a real world scenario using 5000 smart meters for several months. It is evaluated from performance perspective on different metering rates, identifying convergence at bulk size of 60 readings per message, as well as processing and providing information in near real-time for new analytic services and applications within Smart Grids. Such a futuristic application, called Neighbourhood Oriented Energy Management (NOEM) system is also presented and evaluated to have 95% of requests with delay less than 1 second on the running IEM services.

- The investigation of how forecast accuracy progresses along lower scales of aggregation of stakeholders in a group are made. A group of approximately $150 - 200$ households already benefit a significant forecast accuracy, even though off-the-shelf algorithms were used. An evaluation of the impact of a BESS on the forecast accuracy, as well is its replacement with assets that can offer a storage capacity e.g. batteries of EVs. Dynamic systems composed from the storage capacity of available assets on-premise are introduced as part of a VES. The VES is composed of a static (e.g. BESS) and a dynamic part (e.g. an EV fleet), and their impact is evaluated on different configurations e.g. scaling 20% of dynamic capacity with 80% of the static one. It was shown an energy storage capacity of size that approximates 6% of stakeholders daily consumption can take accuracy down to $2-5\%$ of Mean Absolute Percentage Error (MAPE) – what is the accuracy of energy retailers today [40]. The evaluation results are based on real world fleets without the technological barriers we face today.

- Architecture is proposed to achieve the pre-required deterministic behaviour of stakeholders . Strategies that can be applied for achieving the deterministic behaviour are proposed as well. One of the proposed strategies is further evaluated on a real world case. Using the proposed architecture, a real system is designed, where a stakeholder adopts a VES to become SFERS. To make the system evaluation relevant, an entire simulation environment was built, where all the components of the running system (especially due the dynamics of VES, where every storage unit has its individual state [43]) are evaluated over an entire year. It was shown that such system would benefit the forecast accuracy of a retailer today $(2-5\%)$ if less than 20% of stakeholder's traditional vehicles would be replaced by EVs. This was achieved with simple, but yet efficient, algorithms this work proposes for managing a VES. Most importantly, the Key Performance Indicators (KPIs) of such systems are identified.

Additional contribution is towards future existence of the proposed concepts. As such systems have to be accepted by stakeholders, the evaluation of user acceptance was made here by conducting a survey. It was found that 94% of consumers think favourably of the idea of smart and self-managed devices, while less than 50% would allow third-party direct control of their consumption devices. Furthermore, the business relevance was assessed, and knowledge gained from the discussions with many (experts in different fields of energy) is also presented in this work.

1.4 STRUCTURE OF THE DISSERTATION

A look on unpredictability in power systems is made in chapter 2 and an overview of the approaches applied towards addressing it. At the end, section 2.2 will explain how this work goes beyond the current State-of-the-Art.

As shown in Figure 1, interconnecting prosumers in Smart Cities is proposed over the cloud technologies, as also depicted in (3) on Figure 2. The chapter 3 evaluate performance (of smart meter concentrators) and scalability issues of metering platforms in section 3.1. Proposal of the Smart Grid energy services is made in section 3.2 and evaluated in section 3.3, including the evaluation of the NOEM application in section 3.3.3. As evaluation resulted in some performance issues from the monitoring services, when smart meters are observed in groups, the performance improvement is further made in section 3.4.

In (4) from Figure 2, one can see that chapter 4 focus on forecasting accuracy of many different stakeholders. Initially, section 4.1 demonstrates importance of forecast accuracy for one of the Smart Grid added-value services, more precisely, the local energy trading [23]. The accuracy convergence of different aggregation scales is investigated in section 4.2 and further accuracy improvements with BESS are evaluated in section 4.3. Experiments from section 4.4 show that the requirements of an energy storage capacity to address forecast errors will vary significantly on intraday basis. This calls for a more precise investigation of potential of assets in ownership, such as dynamic storage from an EV fleet in section 4.5.

In chapter 5 the main focus will be on SFERS – the system design and the concept evaluation. Firstly a look on where AMI stands, to support this concept, is made in section 5.1. The capability of bi-directional communication took the stakeholders a step further, where potential flexibility scenarios can be considered (as presented in section 5.2). To get these opportunities, in section 5.3 an architecture is proposed and evaluated in section 5.4. As Figure 2 depicts in (5), VES management is made directly by the intelligence of the SFERS system and self-forecast is reported to the energy services.

The conclusion ot the dissertation is made in chapter 6 and section 6.1 explains how the research challenges are addressed. A great part of the future research points are noted in section 6.3. At the end of the document, reader can find the additional contribution on business relevance (that also involve third parties from Figure 2) and results of a survey on stakeholder acceptance for systems similar to one a SFERS would run, in Appendix A and Appendix B respectively.

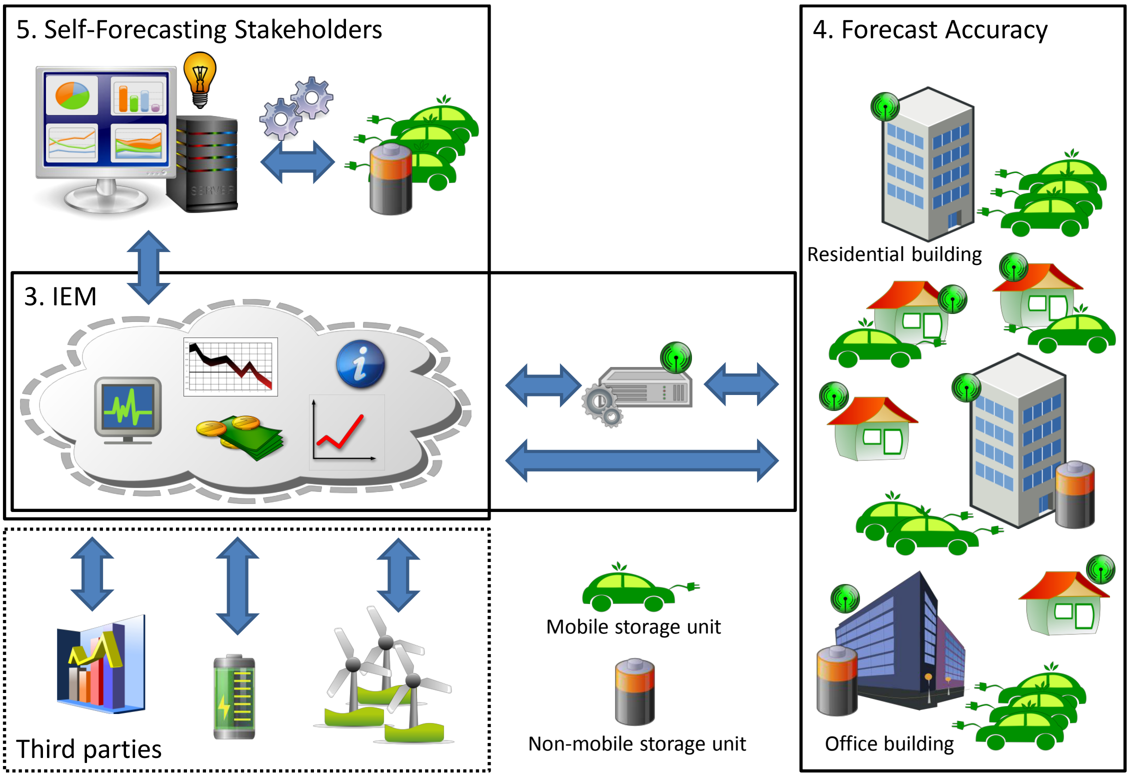


Figure 2.: Chapter structure of the thesis

2 State of the Art

To be able to follow electricity demand, power systems are build with base-load, intermediate-load and peak-load power plants. On one hand, base-load plants have a long start-up time (many hours), a high efficiency at full load and decreasing efficiency when operated at partial load. Peak-load power plants on the other hand, can start and stop very quickly (in minutes). These plants have different run times within a year, but their existence is required to keep the balance of variable demand – as this is essential for safeguarding the system security [15]. However, the full-load efficiency of peak-load power plants is much lower than that of base-load power plants, thus making them even more costly. For example, in many cases these peak-load power plants are gas turbines with efficiencies between 25% and 35%. Since the power system has to balance supply and demand – and must be able to react on disruptions – the overall system efficiency is much lower. For example, the Dutch power system efficiency is about 40%, while in almost all other countries it is even lower [44]. This has as a consequence a higher operational cost, which is propagated to the consumers. As the latter are mostly isolated in power systems of today, and can not be used actively for demand-response (except large industries), they cannot be coordinated to avoid critical fluctuations in demand [44].

As (great part of) electricity demand is not controllable while should be available, the forecasting techniques play a pivotal role in increasing the efficiency of power systems. These forecasts come together with errors [39], thus affecting maintenance of the real-time balance for a Distribution System Operator (DSO) and finally influencing a Transmission System Operator (TSO). Naturally the balancing costs are already included in the costs of the network usage. As such, one can conclude that the unpredictable operations bring costs, while stable operation is crucial for the power networks. The introduction of Distributed Generation (DG), in traditional power grids as we know them for over a century now, will completely change the game in future. Their penetration, in particular of small capacities, e.g. the introduction of prosumers [28], as well as the electrification of transportation [45], will significantly impact the complexity of scheduling the production and consumption within the power system. If penetration keeps on growing, one can expect that the complexity will reach the level where power network operations will no more be forecast-oriented, but rather information-oriented [46].

Not only the concept of DG challenges the power networks, but RES push it to the extremes. As RES are mainly used for DG [47], their (to some extent) unpredictable behaviour may affect stability of prices on intraday, but even day ahead markets [15]. Wind, for example, is almost impossible to predict some days in advance, and even hard to predict a day in advance, so sudden fluctuations in wind production can severely affect prices. This unpredictability already affects the efficiency of power systems even though they hold a small percentage in energy mixes of today. Furthermore, it was noted that Balance Responsible Partys (BRPs) will face additional problems for a correct estimation of their exact imbalance position due to the increase of unpredictable loads of RES [15]. In such situation, a BRP will know more or less an interval in which probably their imbalance is situated, but it will not know the exact position, so their costs will rise. Nevertheless, RES are there to stay and their penetration will further increase in the future [48]. As such, the unpredictability factors will further challenge the management of complexity, until it becomes a difficult issue [49].

2.1 TOWARDS POWER SYSTEM EFFICIENCY

Power requirements increase as adoption of electricity powered devices keeps on growing, e.g. adoption of EVs [31]. Even though our electric grid is the greatest engineering project of 20th century, its ageing brings costs, thus infrastructure limits need to be efficiently used for the resource allocation. This is not an easy task and intermittent DER brings further challenges, but many projects show that operations can be improved by wide acceptance of ICT [50]. Current adoption of ICT enables the vision of Smart Grids with the aim to not only deliver efficient resource allocation but also guarantee security, resilience, and responsiveness of the grid too. The only way to achieve that is to properly use the information available in Smart Grids [17], thus prefer usage of software instead of copper. This is why Smart Grid researchers introduce sophisticated concepts where traditionally passive stakeholders, in particularly the uncontrollable consumers, can actively contribute to the needs of other stakeholders [18]. This is required to keep up with rising electric demands on the grid, and if done right, costs can be from 3 to 10 times less than the current costs [51].

Great part of efficiency in power networks is addressed via sophisticated software solutions using predictive algorithms and monitoring of grid resources, e.g. the State of Charge (SoC) of distributed energy storage [52]. Due the existence of AMI and wide ICT adoption accessibility to its detailed information is possible [53]. In [54], even traditionally passive consumers were able to monitor their personal information and therefore indirectly contribute to efficiency of the grid. The future of Smart Grids reside on sophisticated multi-channel applications that will rely on such data, in a way that end-user groups can benefit from future added-value services [34]. In [37] a significant effort was invested to towards enabling multi-source data and Internet provided basic services. However, many of these services require to be continuously observed by humans, which is unwanted by end users [55]. Additionally, some applications are too complex to be observed by a human operator, or even time dependent, so human interaction cannot be considered. Instead, software solutions are used for automatized reaction on real-time [11], or near real-time operations [56].

Demand Side Management (DSM) took a relevant part in increasing efficiency [49]. Many of these solutions, however, work on DSM today to solve the problems the utilities face [57]. Still though a big part of the loads remains inflexible, due the process inflexibility and even complexity of expressing flexibility of a stakeholder. This characteristic took DSM to the point where only highly static loads are managed, or be easily predictable, whose reaction on demand is verifiable [27]. Besides that, many loads are even adopted to be controllable on demand, but many stakeholders didn't like the idea of remote parties controlling their loads [36]. With that in mind, a trend of an energy market usage in Smart Grid neighbourhoods was proposed [23, 58], where users (or their software agents [56]) can trade their energy needs. One can observe these markets as a part of DSM, where traders indirectly affect the power network efficiency [59]. It was found [25] that only stakeholders with an accurate forecast can participate in such markets. Overall, today the state of the art in DSM is the Demand Response (DR) program, which requires behavioural changes of the stakeholders and is not relied upon by the distribution grid operators.

In past, DR were closely related to the industrial facilities, but today a "guaranteed load drop" can be provided by many companies and (even) residential consumers. Doing so at scale, a significant difference can be made, e.g. companies today use even residential demand response which can equal up to 40% in the utility's service territory [57]. Therefore, companies that were once energy efficiency providers, are all now offering platforms that can help behavioural efficiency, such as involving aggregation of controllable residential thermostats in power markets [60]. This aggregation is a virtual power capacity in agreement with a utility, claiming to provide the capacity as though it were a generator [61]. Their business model is to guarantee that, when their customers need the capacity, they will provide it. This however includes an additional player in the supply chain that can be avoided by adoption of ICT [23]. Furthermore, such flexibility should not be considered just for emergencies, e.g. events involving thousands of homes for some hours, instead it should be continuous and used on a soft basis too (as described in section 2.2).

The latest trend is to enhance power networks with BESSs in order improve the infrastructure efficiency by reducing the need for copper [62]. Others applied BESS to RES that are intermittent, seasonal and non-dispatchable, so their efficiencies can be quite low due to these characteristics. Still, solar Photovoltaic (PV) and wind turbines have the highest potential to satisfy needs of tomorrow [63]. To address this, many reports have explained value of having an energy storage closer to customers [52], especially those who would like to adopt renewable energies [47]. As an example, the goal of [64] is using a BESS in conjunction with a large wind farm, in order to allow the combined output to meet an hour-ahead predicted output within 4% at 90% of the time. Thus, with BESS one can maximize utilization of RES, generation can be adapted to load profiles, and surplus wind energy can be used at peak load times, as shown in previous studies [29]. But there is always the question of storing costs. Therefore studies compared the energy efficiency of storing electricity versus curtailing it [63]. As curtailing is considered wasting, others propose to keep the energy in second-hand EV batteries [65].

Ultimately, efficiency of power systems is expected to be reached by customer engagement, with focus on mobile [66]. Today customers' experience is rather simple as at the end of the month customers get a bill, or they attempt to reach the utility only after power goes down. This linear flow relies heavily on manual communication and non-automated processes. With that in mind surveys were conducted [36], where investigation is focused on consumer desired services. To address these issues, some utilities provide websites to customers for historical consumption monitoring and even enable them paying their bills [21]. Even though their interaction is infrequent and short, this resulted in a two-dimensional relationship with the consumer, involving only billing and outages, so in response to this weak relationship customers offer no loyalty. Leading utilities envision that other processes, that require human interaction, can be powered by mobile processes and interactions, what will contrast the traditional value chain [55]. Therefore, many utility providers aspire to engage customers on social media, by mining their intelligence and behaviour data, that would help developing personalized service offers.

2.2 BEYOND STATE OF THE ART

All the aforementioned methods are trying to increase grid efficiency that will be highly impacted in future e.g. due the unpredictability of the resources [64]. Value of the predictable behaviour will rise and flexible resources should be rewarded for their load flexibility. One example would be to consume when wind is blowing. This is why this work focuses on determinism by providing a self-forecast. Realization of accurate self-forecasts for energy signature is expected to enable an active participation in new business models and roles that emerge [18]. Without having deterministic load, or so to say predictable one, the load changes cannot be verified [12]. This is crucial, as active involvement requires from stakeholders to be measurable and verifiable, in particular for grid operations [26], participation in local energy markets [23], or demand response programs [67] etc. Nevertheless, deterministic loads of the traditionally passive stakeholders will significantly impact the way grid operates today, such as reduced need for grid balancing, or frequency regulation and other challenges mentioned in Appendix A, they just need to be enabled for active participation.

Current software solutions are not designed for the traditionally passive consumers, even though consumers are always connected to their grid and affect its state. Wide acceptance of ICT will enable their accessibility to information of Smart Grids, and move from the traditionally passive end-point on the gird that energy is provided to [28]. The prevalence of energy services proposed in section 3.2 will empower both traditionally passive consumers, as well as emerging ones to slip into new roles and will provide new innovative solutions to the market [61]. Not only this work focuses on individuals, but even to small scale aggregation such as neighbourhoods or residential buildings. If they can become predictable, the unpredictability of loads in power systems will reduce, what will directly affect the overall system efficiency. In fact, [68] shows that stochastic behaviour of a resource reflects to its entire cluster, so it doesn't really matter if (for example) unpredictability is solved on consumer's or producer's side. Furthermore, the proposed IEM platform is evaluated on a real world trial in section 3.3 – in order to confirm practical feasibility and timely delivery of the futuristic Smart Grid services. This is a big step forward in engaging the passive "resources" to improve grid efficiency.

As batteries are identified to enable the second grade of convergence in forecast accuracy [69], their application will be considered even for individuals (beside the accuracy achieved with the aggregation step [40]). With energy storage in place, flexibility of SFERSs can be expressed based on their SoC, rather than offering flexibility from a complex business process. Even though the step of achieving accuracy has a significant business relevance (as discussed in Appendix A), storage solutions can bear costs. This is why results in section 4.4 are based on (potentially) owned assets that are capable to absorb errors of a self-forecast, e.g. data centres or interior/exterior lighting. Hereby the focus is on assets that will otherwise be idle for most of their time, or in other words considered as "wasted" resources. These are EVs and section 4.5 evaluates them on a real world case. The vision is primarily possible due to their penetration [32], which will help us avoid the additional costs of static storage solutions [70]. This is of main relevance, as the statistics show that we have over 1 billion vehicles on the road worldwide, resulting in 10 times as much power on the road than totally installed in power systems [44]. According to that, if stakeholders adopt the EVs [45], their on premise presence is to be considered within VES for a stakeholder, or a cluster of them.

The main contribution of this dissertation benefits by unifying the methods and solutions mentioned in section 2.1. This work combines them in order to design a system that will enable deterministic and flexible behaviour of the traditionally passive stakeholders. Once energy load activity of passive stakeholders can be verified, they can actively contribute to power system efficiency (rather than passively affecting it). Such stakeholders would join the vision of Smart Grids and exchange/get information from it to adjust their behaviour. Hereby self-forecast of a SFERS will be passed autonomously to external parties, i.e. same as smart metering with an offset, thus a system to do so is presented in section 5.3. As stakeholders do not want to continuously be involved in their energy decisions [55], the SFERS system is designed to be autonomous – using only software solutions to achieve determinism. However, decision on making a load change for benefit, as discussed in section 5.2, can be (but not necessary has to be) decided by humans e.g. accepting a critical DR event. It is important to note that the SFERS system is designed as real-time system to contribute the overall power efficiency, while data collected in Smart Grids of today is mostly used for measurements and post-analysis. In short, a SFERS smartly contributes to power efficiency by its non-stochastic behaviour, while assists in needs of other stakeholders, and this work evaluates it on a real world case in section 5.4.

3 Enterprise Integration and Energy Management System

Initially the main application for smart metering was billing, but the accessibility to the smart meter data widely opened the doors to new opportunities [18]. Today, key driving forces behind the Smart Grid efforts are the need for higher energy efficiency and better management of available resources in the electricity grid. To achieve these objectives, collecting fine-grained information inside the grid (such as energy data) is essential. In 2009 market statements for the Smart Grid era provided some hints on expected growths and business significance: Hattar, estimated that the Smart Grid network will be "100 or 1000 times larger than the Internet" [71], and similarly Sikka stated that "The next billion SAP users will be smart meters" [72]. In fact, according to the Smart Grid vision [8], any electronic device connected to the grid will be able to communicate its consumed or produced energy in almost real time [73]. Although the acquisition of smart metering information from such a large scale distributed infrastructure is challenging, not understanding components of the proposed Smart Grid architectures might result in the architectural performance bottlenecks [74]. Understanding the key performance indicators will help in designing large-scale smart metering systems and its composition of such components.

Guided by the idea to empower an entire Smart Grid city [8] and desires of its stakeholders [36], several energy services capturing the common needs of stakeholders are needed. As a result of a (potentially) common platform that offers basic energy services, rapid development of applications can be realized without the need to start from scratch. The NOBEL project [37] targeted the advanced management of Smart Grid neighbourhoods, and, as depicted in Figure 1, the aim is to use a common energy services platform named enterprise Integration and Energy Management system (IEM) [21] to interconnect the prosumers (producers and/or consumers of energy) as well as various other stakeholders, e.g. Self-Forecasting EneRgy load Stakeholders (SFERSs). Advantage of such platforms is that its services provided can be accessed via Internet through multiple channels such as web and mobile. Such platform, located in the cloud of Figure 1, is envisioned to have several energy services [21], such as: (i) Energy Monitoring, (ii) Energy Prediction, (iii) Management, (iv) Energy Optimization, (v) Billing, (vi) Energy Trading (Brokerage) and (vii) other value-added services. These services later can be mashed up in order to provide key functionalities for applications, such as an energy portal, mobile applications, and also a district monitoring and management centre [34].

The IEM has been evaluated as was extensively tested and used operationally in the second half of 2012 as part of the NOBEL project trial which took part in the city of Alginet in Spain [38]. Data in 15 minute resolution of approximately 5000 meters were streamed over the period of several months to the IEM, while the IEM services were making available several functionalities ranging from energy monitoring up to the futuristic neighbourhood energy trading [23]. As such, the energy services have not been only identified and analysed, but rather implemented in the context of a wider enterprise system architecture. Their functionality, usage and development challenges and experiences are therefore provided from the pilot and even lessons learned that affect their design and performance.

The usage of the IEM services was demonstrated in a web application for monitoring and managing a Smart Grid neighbourhood that was developed on top of the platform [34]. Being operated by a human, the application helped identifying the potential bottlenecks in service performance when energy data aggregation is done on different levels e.g. if the collected energy readings from thousands of smart meters are aggregated to be monitored in real-time. This performance points in hereby further investigated, where the traditional (row-based) database systems are compared to the emerging column-based approach [75], such as having the real-time analytics suitable in such scenarios.

This chapter considers the entire path of designing successfully such platform and demonstrate the feasibility to actively involve the Smart Grid stakeholders (or group of them) in a timely manner. In section 3.1 the limitations of Advanced Metering Infrastructure (AMI) infrastructural components are investigated, resulting in better understanding of the Key Performance Indicators (KPIs) in real-world deployments. Once data is accessible, the platform services are proposed in section 3.2, which are to be provided to the stakeholders in future smart cities. Its full implementation and evaluation in a smart city pilot was done in section 3.3, including the IERM service consumption from a neighbourhood management application of section 3.3.3. Methods shown in section 3.4 were used to achieve sufficient performance to offer a group of services to be consumed in such real-time applications. Finally, in section 3.5 a discussion on the overall lessons learned and future work takes place, and the chapter is concluded in section 3.6.

3.1 METERING DATA SYSTEM

To achieve objectives of higher energy efficiency and better management of available resources in the electricity grid, collecting fine-grained sensor information inside the grid is essential. An AMI needs to be in place to enable measuring, collecting, and analysis of data from remote meters for electricity, gas, heat, water, etc. Having such a large number of remote metering points in future, one can expect challenging collection of their sensing, thus data concentration is proposed. A concentrator is the interface between many low-speed, heterogeneous, usually asynchronous, channels and one or more high-speed, usually synchronous, channels. It acts as an interface between the smart meters and the enterprise, usually responsible for collecting the meter reading data and submitting it to an enterprise server. Figure 3 depicts how metering data is collected from various sources via strategically positioned concentrators.

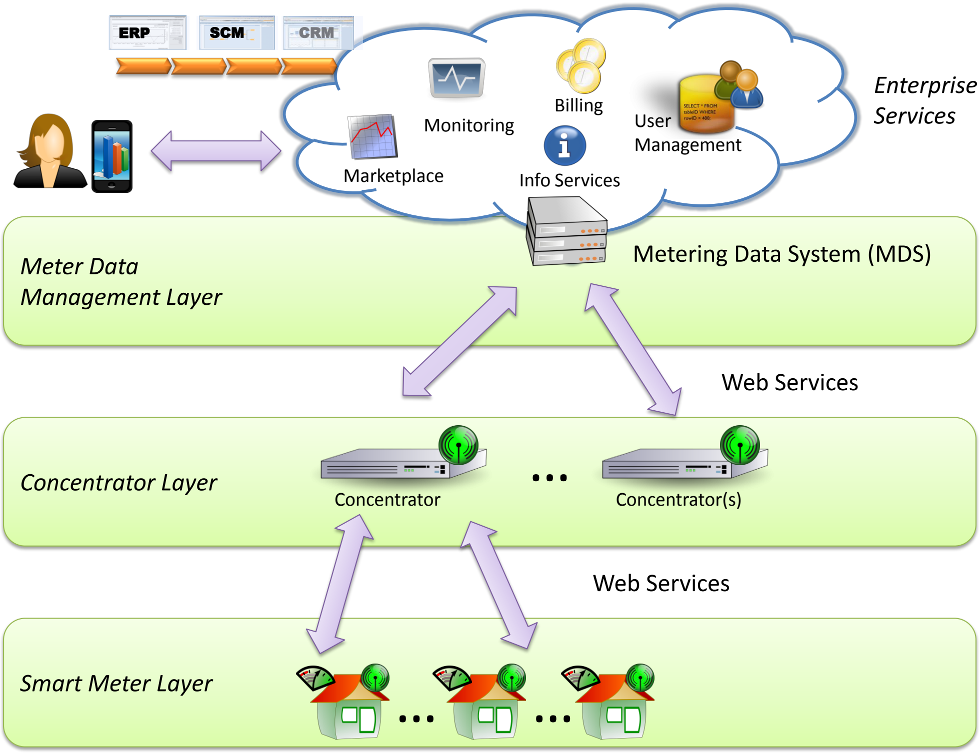


Figure 3.: AMI overview in the smart grid era.

Although smart metering is a key milestone towards realizing the smart-grid vision, not much evaluation of it is done at a detailed level. This is especially true when referring to a large number of metering points submitting data upstream. Typical evaluations refer to the number of measurements that can be achieved in 15-minute intervals. However, most real-world trials focus only on a few hundred meters. Having said that, a use case where great number of sensors transmits their high-resolution samples over the Internet enabled Smart Grid needs to be investigated [46]. To achieve this, open source software can be used for development and investigation of infrastructure limitations. As such, goals and restrictions are:

- to design a simple and scalable approach for large-scale and low-cost smart metering
- to use standardized Internet-based technologies i.e. web services between the AMI layers

- to use existing (off-the-shelf) open source software, and commonly available PCs as the hardware platform
- to simulate large number of metering points (smart meters)
- to evaluate the performance limits of the key AMI components, i.e. at concentrator and metering-server level
- to acquire hands-on experience and insight into large-scale smart metering performance

3.1.1 *Metering Architecture*

One set-up of AMI is a three-layered hierarchical architecture, similar to what is depicted in Figure 3. Bottom-up one can clearly distinguish:

- Meter Layer: the last mile, where the (residential) meters are passively tapping in and measuring the energy consumption or production of the attached devices.
- Concentrator Layer: the meters connect to this layer via various (often proprietary) protocols to report their measurements. The reported data is aggregated and submitted to the Metering Data System (MDS).
- Metering Data Management Layer: here usage data and events with respect to the infrastructure are collected for long-term data storage, analysis and management. This is typically used by enterprise services in order to empower applications such as billing, forecasting, etc.

The approach here adopted follows the same model, albeit some technological and context constraints were considered. In the context of this work, a fully IP-based three-layered service-oriented infrastructure is assumed. This implies that all messaging among the layers is done over web services. Also, in the Smart Grid context, the components depicted in the layers (such as meters, concentrators, and MDS) are designed to handle high volumes of data at high rates, hence, permanent connections with possibly high bandwidth among them might be expected. This is a clear analogy to the Internet which is composed of end-devices, routers, and servers. Similar motivation such as heterogeneity management, scalability, and performance exists in the Smart Grid.

3.1.2 *Performance Experiments*

The performance experiments conducted here were made within a partially simulated environment [74]. They were aimed to gather high volumes of metered data from the meters up to the MDS, so that enterprise applications can take advantage of the almost-real-time data. In order to achieve such goals, detailed measurements of data exchanged between the architectural components as well as in their interworking is needed. The overall performance of the architecture is measured in terms of its capacity to handle certain number of requests per second at different layers. These measurements are taken against the MDS and concentrator components in order to determine their limits, and also their reliability under heavy load.

In total two experiments have been conducted:

I. Assumption Validation

II. Concentrator MDS Performance

The evaluation of performance of each component individually and then an assessment for the whole system is accepted for methodology. In the first experiment, major assumptions are quantified and validated. In the second experiment, the performance of the MDS is measured against variable concentrator configurations to derive a practical high-throughput configuration. Details on additional experiments and the entire simulation environment is available at [74].

Metric Definition

In the hierarchical structure from Figure 3 , one can see that concentrators and MDS depend on number of sub-components (i.e. i meters or j concentrators). The key performance indicator common to all components is the *meter reading rate* r of received meter readings from meters. There are n single meter readings being submitted within a time interval t, or request of bulk size b coming from a concentrator. Thus, the rate of meter readings can be defined as:

$$r = \frac{nb}{t} \tag{1}$$

However, due to the nature of the aggregation of meter readings by the concentrator, the *request rate* q variable is introduced, that depends on the aggregation of messages in bulks of size b done at each sub-component j. Therefore, for each concentrator or MDS, the request rate q is defined as:

$$q = \sum \frac{r}{b} \tag{2}$$

Throughout this work it is assumed that each meter m_i is submitting one measurement at a time (and not aggregated meter readings), thus $b_i = 1$. This is what is expected in real world applications, at least from the always-connected meters. As a result, the request rate r for a concentrator c is $q^c = r^c$. Similarly, if a single meter reading is also propagated further not as part of a bulk (thus $b = 1$) from the concentrator c to the MDS s, then $q^s = r^s$. Furthermore, assuming minimal impact on the rates (e.g. no losses, no significant processing overhead, etc.), it could be argued that $q^s \approx q^c$.

Assumption Validation

The experiment carried out rely on an assumption, which will be assessed experimentally in order to verify it:

Meter readings can be processed at a higher rate, if they are processed in bulk, i.e. multiple meter readings at time;

The assumption comes from the fact that there is a time cost associated with the whole process. The cost for each message is associated to transmission, processing the eXtensible Mark-up Language (XML) and extracting the payload, storage etc. Most of these variables depend on a non-deterministic condition such as the network available bandwidth, the server load etc. To make it more concrete, for each measurement, a connection is established, the data is transmitted, and upon acknowledgement, the connection is closed. The server processes each request upon reception (by extracting the payload) and then stores the reading data for further processing. If this is done for one meter reading at time, cumulatively, the server will be spending a significant amount of processing time per request, leaving fewer resources for the payload processing – done by the Enterprise Java Beans (EJB). If this processing overhead between receiving the request and sending the data to the EJB could be minimized, the throughput of the MDS (i.e. meter readings ratio) would increase. Hereby is assumed there is a point from which the difference between the cost associated with a single meter reading submission vs. a bulk of them makes a significant difference on the overall performance of the MDS. If this assumption is true, the proposed approach would outperform any other approach where the meters communicate directly with the MDS.

EXPERIMENTAL VALIDATION OF BULK SIZE CONSIDERATIONS In order to validate this assumption, an experiment was conducted in which several requests were made from a single concentrator to the MDS, over a range of bulk sizes. For each bulk size, the time taken for the MDS to process the request, the server overhead, and the time taken for the EJB to process the metering data were measured. For the purposes of this experiment, a request rate was chosen so that the metering server would not be overburdened. As such, behaviour of the server could be measured under normal operating conditions.

The experiments where the selected range of the bulk size b parameter is defined as $\{b \mid 2 \leq b \leq 100, b \in \mathbb{N}_2\}$. For each b there are 1000 requests from the concentrator to the MDS. The average of meter reading rate r^s for each b is then calculated according to Equation 1. In Figure 4, the Application Server (AS) overhead, EJB and total processing time per meter reading are shown for each tested bulk size.

There is a clear correlation between bulk size and performance. It can be noted that the average time to process a single meter reading decreases as the bulk size increases. Furthermore, the rate of improvement decreases as the server converges to its processing limit. An interesting observation is that the EJB processing time

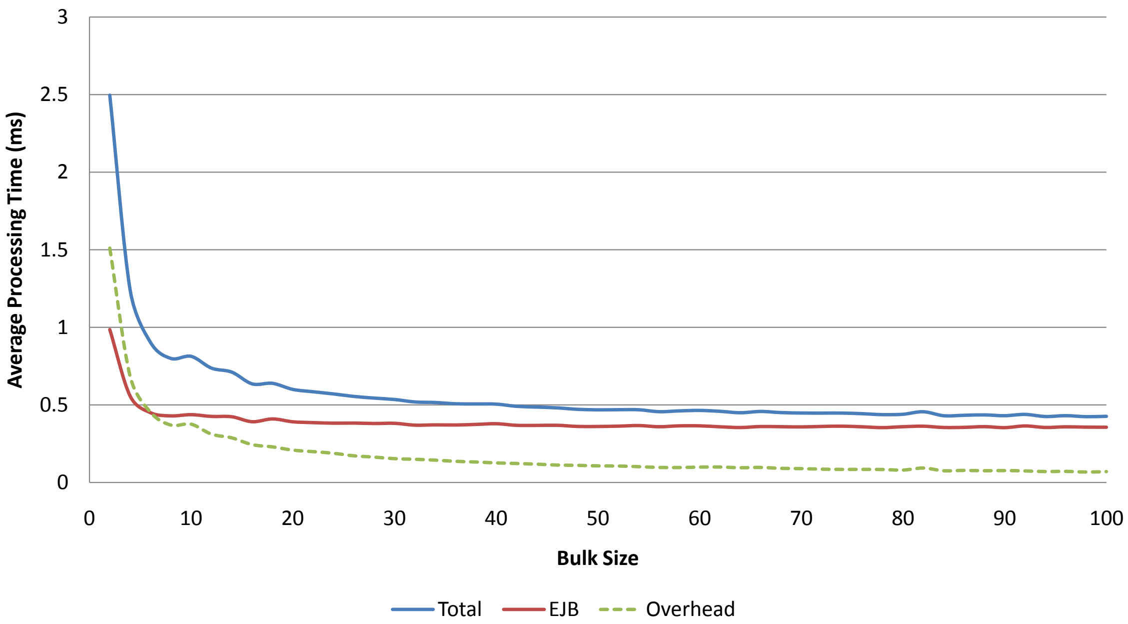


Figure 4.: Average total processing time, average EJB processing time, together with the average AS overhead per meter reading for a range of bulk sizes.

reaches its limit faster than the overhead. Thus, minimizing the AS overhead is a candidate for increasing the overall performance. This overhead occurs between the request arrival at the server and the meter reading data arrival at the EJB.

Concentrator MDS Performance

Having already verified the performance effect of bulk processing, the focus goes onto the performance of the MDS under the load of a number of j concentrators. For that reason, the average request response time for a variety of request rates and bulk sizes needs to be monitored. The main objective here is to determine the best performing bulk sizes for particular request rates experienced by the MDS. Once the boundary conditions is ascertained, predictions can be made as to how to best configure the infrastructure in order to handle certain predetermined rate of incoming meter readings. Their expected request rate q can be calculated as shown in Equation 1, out of which also the theoretical rate of meter readings arriving at the MDS can be extracted. By comparing an expected meter reading rate r_e^s and the actual (measured) rate r_a^s, for each parameter setting, the processing limits of the MDS can be determined. The processing capacity on s has to be tested for certain bulk sizes. As can be seen in Figure 5 , the practical capacity on MDS grows in-line with the parameter b. More specifically it grows fast for small bulk sizes, and for higher ones (e.g. $b = 100$) it starts to converge towards $r^s \approx 3900$.

Once the impact of b on r^s is assessed, other few scenarios are chosen to prove this finding. For this experiment, the bulk sizes used are $\{b \mid 10 \leq$

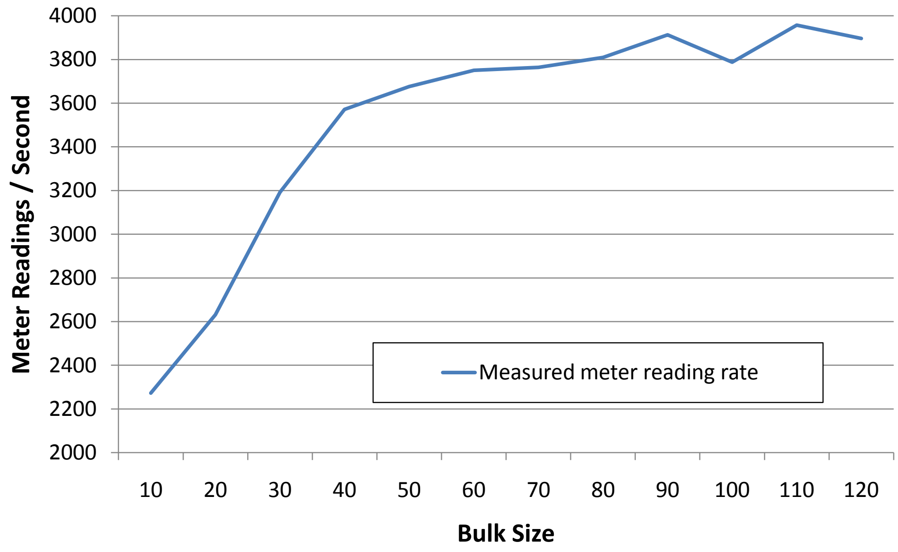


Figure 5.: Stable processing capacity of s over the dimension of b.

$b \leq 200, b \in \mathbb{N}_{10}\}$. Additionally, five MDS request rates are chosen i.e. $\{q \mid q = 100, 50, 33.33, 25, 20\}$, yielding 100 parameter combinations. The experiment results are depicted in Figure 6, where qualitative view of the effects of the request rate and bulk size on the processing rate can be seen. The expected request rate performance vs. the actual one is depicted. Here a "turning point" for each experiment is identified, where the request performance of the MDS starts to greatly deviate from the theoretical performance (were actual b can be calculated by Equation 2). All curves are pivoted around the bulk size where the threshold between normal and deviating server behaviour can be seen.

From the actual threshold bulk size value for each of the request rates (as seen in Figure 6) one can identify that the threshold bulk size increases as the request rate decreases. This is a consequence of Equation 2, as the lower that request rate is, the bigger the bulk size needs to be in order to maintain the same total meter reading ratio r^s at the MDS. As such, one can conclude that the "turning point" for the $q^s = 100$ requests per second is at $r^s = 3000$ meter readings per second, which is significantly lower than the $q^s = 20$ requests per second rate which has a "turning point" of approx 4100 meter readings per second. The result of this experiment is that the lower the request rate q^s is, the highest the meter reading rate is, where one can verify the effect of the bulk size parameter b. As such, the "turning points" described here are not entirely accurate, because the true "turning point" will be somewhere in the range $[b..b + 10]$. However, these results are more than enough to prove the importance of parameters that may variate in these scenarios. It should also be noted that the exact measured values are tied to the hardware and software configuration used for the experiments [74], still the pivotal points and relevance of the bulk transmission is expected to approximate ones on other configurations.

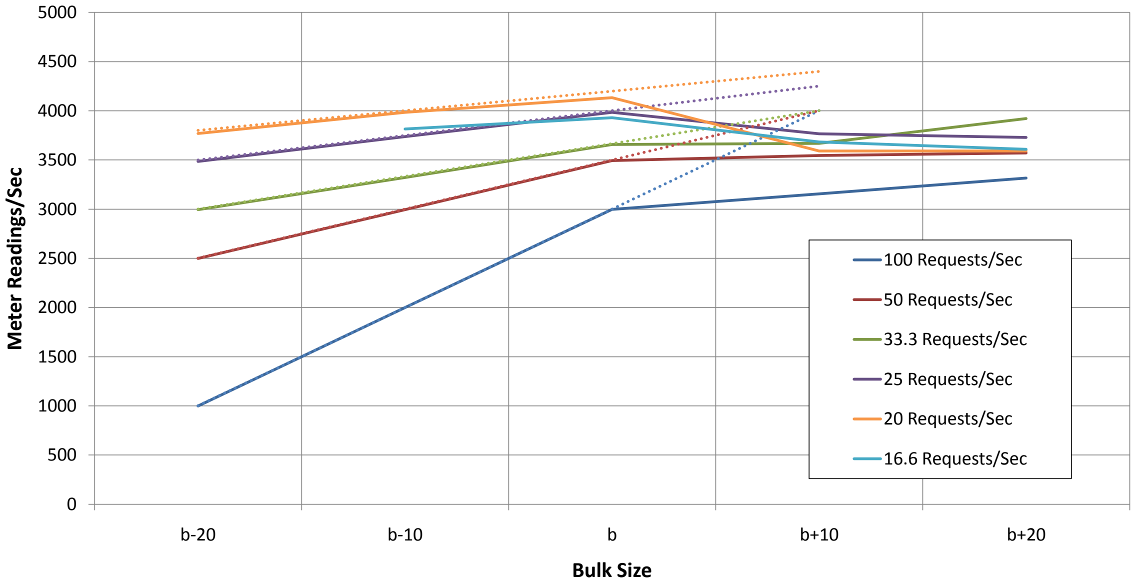


Figure 6.: Change of MDS processing capacity for different request rates and bulk sizes. The curves pivot around their threshold bulk sizes.

3.1.3 *Key Findings*

The goal of experiments was to shed more light to the performance considerations that arise when one attempts to realize the AMI envisioned by the smart grid. Since little quantitative work is in bibliography, hereby is investigated how easy it is to implement it with open source tools and made several thoughts about the possible problems that one has to deal with. The implemented testbed was used as a proof of concept. Several results have been already analysed during the experiment description sections, and it is obvious that trying to tune the whole system towards high performance (within the constraints listed in section 3.1), more complex inter-dependencies need to be considered (as done in [74]).

Two interrelated key performance indicators were identified and used to evaluate the overall performance of the proposed architecture i.e. request rate q, and meter reading processing rate r. Measuring the capacity of a component to handle different magnitudes of these rates is the first step in configuring the parameters of the architecture for high throughput. These two rates are of course related by the bulk size b, which, as witnessed (in first experiment from section 3.1.2), plays a pivotal role in the maximum capacity that a component can reach.

Through the experiments, the performance of the MDS against concentrators was measured. It was found that increasing the bulk size increased the throughput, this only happened up to a particular threshold bulk size. This behaviour can be explained from the results of the first experiment (in section 3.1.2). Increasing the bulk size increased the throughput of the server by lowering the overhead processing on the application server. However, since more payload (larger number of meter reading due to bulk size) needs to be processed as well, the request response time also increases. Since the measurements integrate the time for processing a request, and this time increases as the bulk size increases, it is clear that the bulk size cannot grow arbitrarily.

One key result that struck us was the amount of bandwidth waste. We already knew that a big percentage of the message transmitted would be devoted to the actual wrapping and envelope of the data in XML. As you can see in Figure 7 for a single HTTP request 60% of the message is occupied by the SOAP envelope while only 9% is devoted to the actual meter reading data. In case of bulks where we have aggregated meter readings the percentage of information increases but still e.g. for $b = 100$ approximately 68% is occupied by Simple Object Access Protocol (SOAP) while 30% of it is devoted to the actual meter reading data.

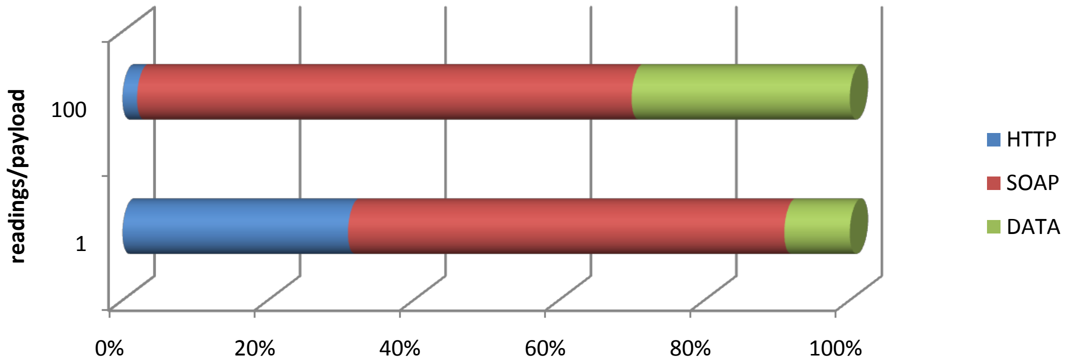


Figure 7.: Message payload efficiency.

The experiments have demonstrated the importance of the concentrator component. While the MDS and the concentrator cannot be compared exactly one to one, the MDS can cope with nearly 8 times the number of meter readings. This would suggest that another avenue for exploration would be bulk metering data being sent from the meters to the concentrators. This could further improve the capacity of the concentrator and therefore reduce the required number concentrators needed in a deployment. The only down side is that, while one could attain a higher level of granularity in the readings, the time taken for the user to access them through any enterprise system would still be limited by the period. For instance, if a meter submits a meter reading every 5 minutes, a bulk message with 5 meter readings could be sent at the same rate, thus attaining one minute granularity in the readings. This could be a cost effective strategy for increasing the reading granularity with little changes to the system.

3.2 ENERGY SERVICES FOR SMART CITIES

The smart grid vision [18] heavily depends on the increased energy data acquisition granularity for understanding better how energy is produced, consumed and where fine-grained and timely adjustments can be done. Hence, monitoring and control will increasingly play a key role for the future smart grid infrastructure and the applications that will depend on it [8]. It is also expected that the future energy monitoring and management systems will be in close cooperation with enterprise systems and heavily depend on IT technologies [35]. Well designed MDS systems (presented in section 3.1), with strategically deployed concentrators, will enable importance of timely monitoring and management of dynamic entities, such as the prosumers. Same is valid for districts and neighbourhoods, as they will also increasingly play a key role in the smart cities, as they are expected to be able to autonomously manage their energy resources e.g. a public lighting system, a shopping mall, a PV or wind farm etc. By offering a way to enable business oriented interaction among the stakeholders, one may achieve better energy management as well as enhance the procurement of energy from external providers. To achieve this, appropriate energy services must be in place, that will integrate information collected by a MDS from highly distributed smart metering points in near real-time, process it, and provide an insight upon which appropriate decisions can be taken.

In this section, a platform providing several Internet-accessible services has been designed and implemented which, in turn, can be used to create mash-up applications that deliver customized functionality and additionally let the stakeholders active in a smart city e.g. participating in energy trading [28]. Collected information at MDS has been integrated and made available in near real-time via various services. Decision making applications rely on them to provide sophisticated functionalities both on the consumer as well as the energy provider side. The NOBEL project [37] is realizing and trailing with real-users in the city of Alginet in Spain, such a set of energy services (`www.ict-nobel.eu`). This is significant for the energy domain, as is proposed to move away from the traditional heavyweight monolithic applications, towards a more dynamic mash-up application development environment.

3.2.1 *IEM Architecture*

The vision of lightweight and rapid mash-up application development, where Internet services are used as basis for creating dynamically customized applications at the end-user side is being increasingly adopted as the standard for popular large-scale on-line services e.g. in Facebook, Twitter, Amazon, etc. Adoption of such trend could indicate a paradigm change for the energy domain, where heavy-weight monolithic applications are substituted by far more dynamic, real-time and interactive applications. The designed and implemented architecture offering such energy services is depicted in FMC notation (`www.fmc-modeling.org`) in Figure 8. As it can be seen there are several layers i.e. the device layer, the middle-ware, the enterprise services and mash-up applications. The embedded devices, such as smart meters, concentrators, and generally any prosumer device are representing the lowest layer in the architecture. A middleware layer is acting as an information acquisition and processing component that connects to and aggregates functionalities near to the point of action (e.g. the meter). On the service layer we have various enterprise services as part of the Integration and Energy Management system (IEM), that put the data into business context and provide sophisticated functionalities.

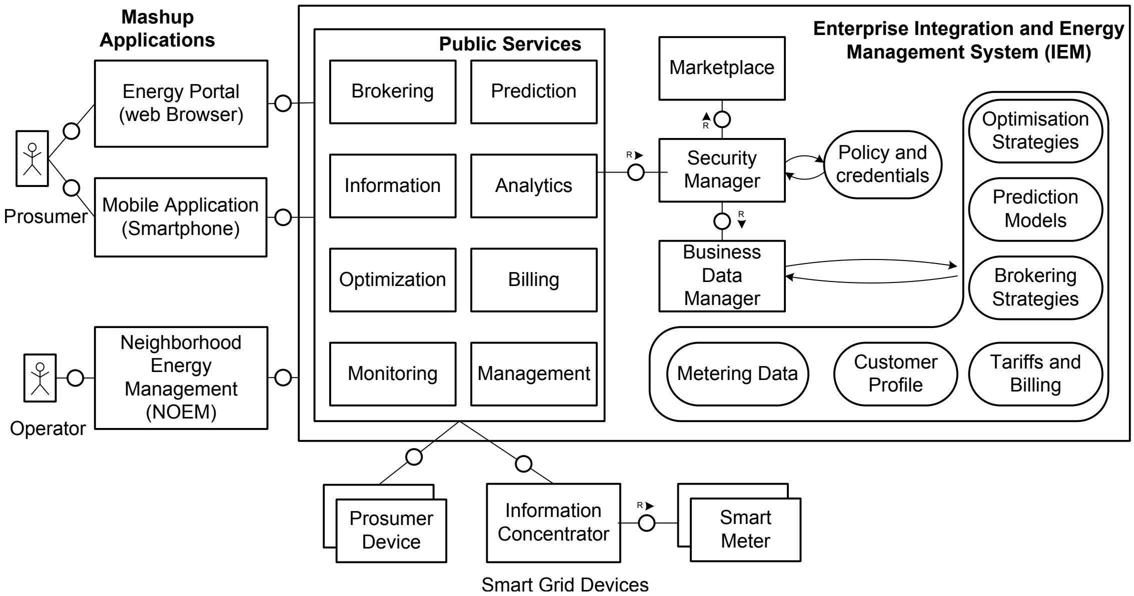


Figure 8.: Overview of the entire system architecture

The IEM is the "heart" of the entire system, which is composed of several services that provide the core of business services. They foster the Software-as-a-Service (SaaS) approach and are hosted in the "cloud". Typically, such services can be mashed up in order to provide key functionalities for applications such as an energy portal, mobile applications, and also a district monitoring and management centre. Apart from the middleware assisted communication and data processing for obvious reasons, the devices can also directly communicate with the IEM and vice versa. Once such services are accessible, it should be easy enough to use existing enterprise services as building blocks in other more complex services in a mash-up way. The diversity of the involved entities leads to increased complexity in interacting and managing the approach. Challenging issues such as the overall system management and performance expectations were considered at design phase. Adhering also to uniformity, this section proposes interfaces that can be called in same fashion from any kind of device (bound to functionality), for instance, concentrator, or end user device, while the resulting output is intended for automated friendly integration (and not really humans) in the Internet of Things [1].

3.2.2 *Realized Enterprise Services*

The IEM services have been fully implemented and tested with meters that have IPv4/IPv6 connectivity and can report their measurements to a concentrator or directly to a smart metering platform. To gain performance, it was decided not to make use of the typical heavyweight SOAP web services, as shown in Figure 7. Instead, the Representational State Transfer (REST) approach [76] is used in order simplify the API and its implementation, as well as to enable rapid application integration over the Internet. This imposes some architectural style selections e.g. client-server approach, stateless interactions, uniform interfaces and a layered system. All interactions, among the key parts of the architecture depicted in Figure 8, are also done via REST and no strong requirements on co-location exist. Additionally, REST seems to be the technology of choice for APIs directed at sharing data easily (e.g. Facebook, Twitter, Amazon, and others use it). It is expected that, due to the volume of data collected in the smart metering era, users will be given a higher degree of control and access over it. Using a RESTful approach, may open the door for energy users to leverage their data against services provided by third parties. Following service listing is relevant for this thesis work, while a complete description of the IEM services can be found in [21].

Energy Monitoring

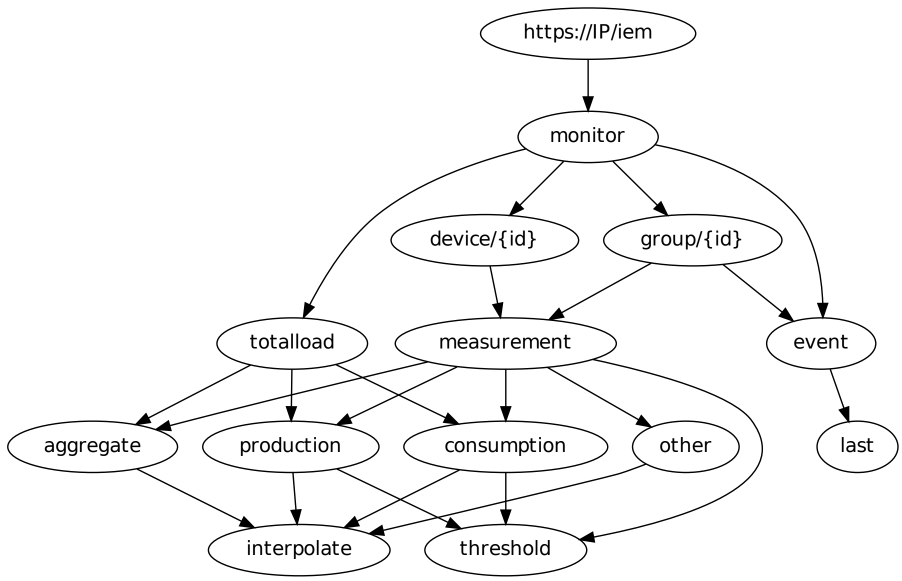


Figure 9.: Energy monitoring service overview

The *Energy Monitoring* service is responsible for acquiring and delivering data related to the energy consumption and/or production of a prosumer device. It offers a near real time view of the energy consumption/production as reported by the smart grid prosuming devices (e.g. PV, smart meter, electric car etc.). As this is a key functionality coupled with privacy concerns, it is made sure via the security framework, that only authorized users may access a subset of its functionality. For instance, users should only have access to their own production/consumption data, and group managers to the aggregated data of the group. The group functions enable monitoring of a group of devices, which provides extra flexibility and support for community-driven smart grid [42] behaviours. Figure 9 depicts a possible structure of the RESTful web service; one can derive the actual URIs of the service that can be called by following a path in the graph.

Energy Prediction

The main goal of the *Energy Prediction* service (structured as depicted in Figure 10) is to provide forecasts for energy consumption or production given a context e.g. historical information, weather prediction, prosumption device capabilities etc. Forecasts can be used by the users and operators to help with their electricity planing and trading activities; hence it would, for instance, enable operators to take advantage of opportunities, such as bidding at national energy markets, or even to comply with energy market regulations. A modular approach was taken where new algorithms can be integrated on the back-end of the service without altering its API. Additionally practice has shown that sometimes users (who have better knowledge of their future behaviour) want to be able to adjust or further customize the prediction, hence such adjustments are also possible.

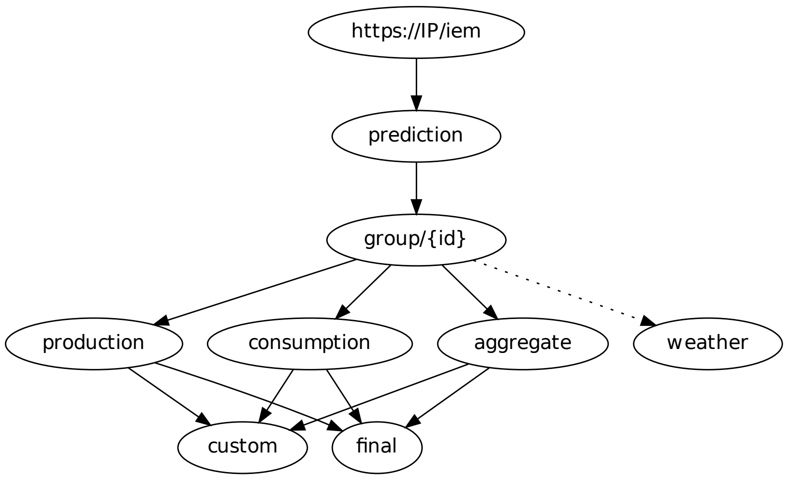


Figure 10.: Prediction service overview

Of significant importance is again the group support, where forecast accuracy and energy storage availability play a key role [68, 69]. As such, an operator could via this service get real-time prediction of the energy production/consumption for one specific neighbourhood, building or even a dynamic group of prosumers e.g. formed based on social, financial, behavioural or other criteria. As the focus is on the architecture and functionality provision, the independence of the prediction algorithms realization and the availability of a stable service interface so that new prediction models can be plugged-in without breaking the interactions with the service's clients was carefully considered. For this work, this service implementation was heavily dependent on the *R language*, as a multitude of algorithms and statistic models are already available via it.

Energy Optimization

The *energy optimization* service (from Figure 11) provides the user with load profiles that are optimized for a particular set of constraints. Given the user preferences and the available data, the optimization service aims at providing a suggested load profile that the user may want to adhere to, in order to attain more benefits, according to his goals and preferences. However, a requirement to join such a program critically depend on forecast accuracy of stakeholders, such that they can be considered as an adequate resource [77]. A case is investigated in section 5.2, where neighborhood energy is optimized by interacting with a public lighting system as a predictable stakeholder [27]. This is however not straightforward for other stakeholders, e.g. residential users, thus focus in chapter 4 is given to the achievement of sufficient forecast accuracy for such stakeholders (or a group of them acting together [61]) so they can get additional energy related benefits.

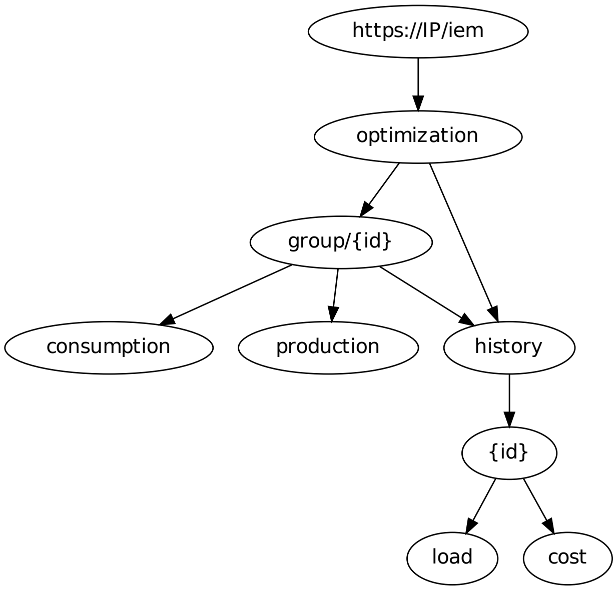


Figure 11.: Optimization service overview

Brokerage

The Brokerage service covers one key functionality envisioned for the future smart grid neighborhoods: the capability of being able to trade energy [28] in local marketplaces i.e. buy energy needed and also sell surplus. This service aims at enabling these brokering capabilities in order to allow users (or their surrogate agents [78]) to participate in the electricity marketplace [23]. As such, this service targets the operators, for managing the market, the users, for actively buying/selling electricity, and also for their automated brokerage agents. A user's brokerage agent acts as a proxy for (group of) users [56], and can be e.g. configured via a mobile device and act driven by the user's preferences, even at time when the user is unavailable. Aside from services enabling market participation, like placing an order, simple market analysis tools such as price and volume statistics, are also offered to aid the user or the automated agents in making trading decisions.

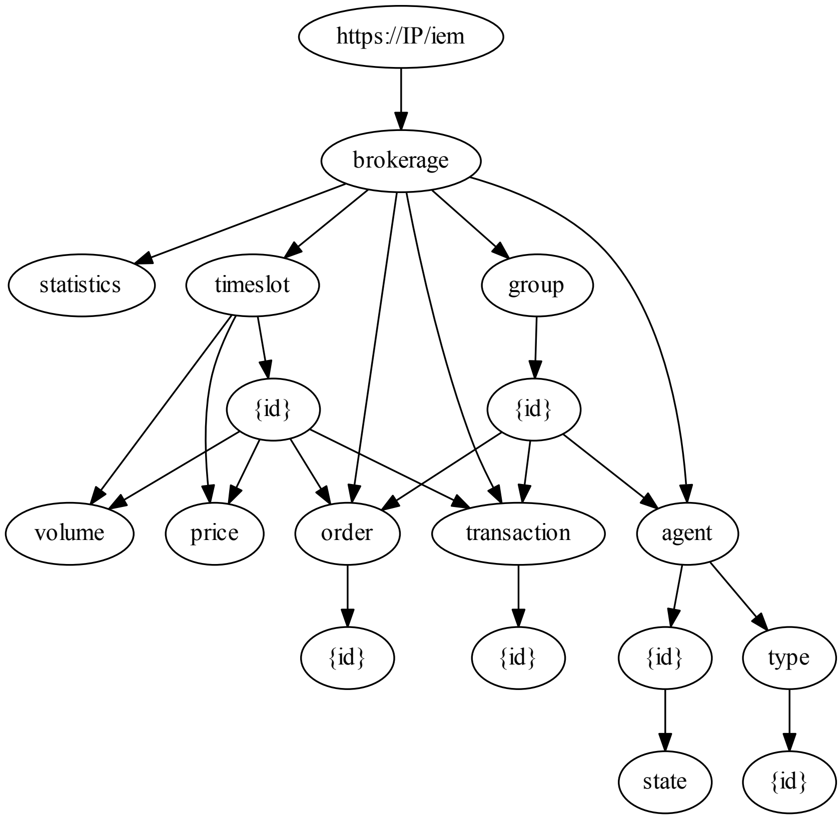


Figure 12.: Brokerage Services Overview

The structure of the brokerage service is depicted in Figure 12, where a group can be composed of one or more stakeholders. The service enables the participants to place orders, and retrieve information regarding transactions (matched orders), market orders and trading time slots. It also provides auxiliary functionalities such as statistics regarding trading price and volume, as well as the last traded price for each time slot. The `order` part of the service returns current market information on the order, as list of orders in a time frame, or in a particular time slot. Transactions are made by matching, either fully or partially, the orders in the order book. They represent the contract made by each party to consume or deliver the specified amount of energy at the agreed price. The time slot service returns information regarding each time slot, for instance, the state of the time slot (open, closed, trading, etc.), as well as the start and end of the time slot, and its energy delivery time frame. Additionally, beside the general statistics offed by the `statistics` service, one can get fine-grained information e.g. list of last traded prices for each time slot over a time frame, while the price curve can return a list of traded prices within a timeslot. Same is valid for the volume service, which returns a list of traded quantities, depending of how services are used i.e. if identification of a `timeslot` is provided.

It is absolutely critical to mention that the forecast accuracy must be sufficient in order to perform economically with the Brokerage services [25]. One can imagine that a group of stakeholders will consume Prediction service to place an order to the market, while Monitoring services will be called in order to feed smart meter readings of a group to the forecasting algorithm. Beside the effort in chapter 4 to enable stakeholders (or a group of them [61]) to achieve a sufficient forecast accuracy to benefit from such IEM services, section 3.4 will aim to deliver these services in a timely manner.

Other services

There are several other services, general ones, which are required by any management system. For instance the informational service is providing notifications to the users e.g. possible (urgent) messages from the energy provider, additional informational material such as network warnings, advertisements, news etc. The same service can be used to enable bidirectional interaction between the end-user and his provider e.g. for feedback, maintenance, communication of problems in his service etc. Such services are the indication of the active stakeholder involvement to the smart grid operation, what will lead to the main contribution of this work in chapter 5. It is also considered, that in the future, the analytics service might be of key importance. This would enable for instance enterprise users to further customize analytics engines to do data mining on the available information gathered at IEM level.

3.3 SMART CITY PILOT

Numerous embedded devices may connect directly or indirectly (e.g. via gateways) to the services provided in a smart city. The main aim of the IEM platform presented in section 3.2 was towards enabling lightweight Internet accessible energy services for thin clients over multiple channels, thus involving them in the Smart Grid activities. On the IEM service layer, one can mash up services to provide customized functionalities for various applications [21], such as an energy portal (e.g. accessible via web browsers), mobile applications, or a Neighbourhood Oriented Energy Management system (NOEM) presented later in section 3.3.3. As seen in Figure 8, there are several architecture parts e.g., the device layer, the middleware, the enterprise services and end-user mash-up applications. Furthermore, enterprise services process the collected data and provide advanced functionalities such as validation, analytics, and business context specific processing.

Since the web services offered by the IEM platform adopt the REST architecture, the implementation and integration of thin clients accessing them is simplified. As such, better performance for data exchange are expected in comparison to the

SOAP web services. In the second half of 2012, as part of the NOBEL project pilot which took part in the city of Alginet in Spain, these services have been extensively tested and used operationally. Data in 15 minute resolution of approximately 5000 meters were streamed over the period of several months to the IEM, while the IEM services were making available several functionalities ranging from traditional energy monitoring up to futuristic energy trading [34]. Results presented here is the report of the experiences during design and implementation, as well as the assessment of the pilot operation.

3.3.1 *Importance of Data Quality*

One of the key problem areas faced within the trial, was that of data quality. High quality data sets are important as they impact all the other dependent services such as energy prediction, grid problem identification, energy trading etc. Data may be validated against multiple criteria, e.g. values are expected to be within some limits, check of the syntax, correct time-stamping, duplicate detection, etc. It was suggested to take a data processing step that enables working with the data that is either already stored or is flying in (stream data). This implies data adjustment, e.g. it might be necessary to normalize data, introduce an estimate for a value that is missing, re-order incoming data by adjusting timestamps. Its significance was identified during the pilot, as the data quality (such as missing data) provoked a significant impact on the forecasting algorithms, especially as this formed the basis for those users being able to trade electricity on-line (i.e. buy or sell energy).

In Figure 13 an overview of the three month pilot against the density of data (number of meter readings) received by the IEM is depicted. Some strictly "red" areas indicate problems in the infrastructure, e.g. fallout of a concentrator, delayed or missing data etc. If this heatmap is plot in "real-time" it may assist the energy acquisition stakeholder to identify potential problem areas and initiate response mechanisms to investigate the real reasons, e.g. meter communication problems, infrastructure congestion, malformed data, security problems (such as data replay, reconfiguration) etc. One can also follow the behaviour of an individual meter or groups of meters (information concentrator) and their performance in delivering the required data in the expected quality.

Figure 14 depicts a quantitatively representation of the collected data on IEM, analysed on daily basis. Similarly to Figure 13, events like infrastructure fallout can be identified, but we also get a quick view on the overall smart metering behaviour and the load (e.g., number of smart metering events, number of meters reporting measurements etc.) on the IEM side. Being able to map real-world events to the visualized data may provide interesting correlations. For instance, a key event shown is the infrastructure fallout for some days (6-10 Oct 2012), where all meters did not report any data. Although in this case we could trace back the problem in a server failure during a weekend that was followed by a Spanish bank holiday, one can use such metrics to assess multiple aspects such as

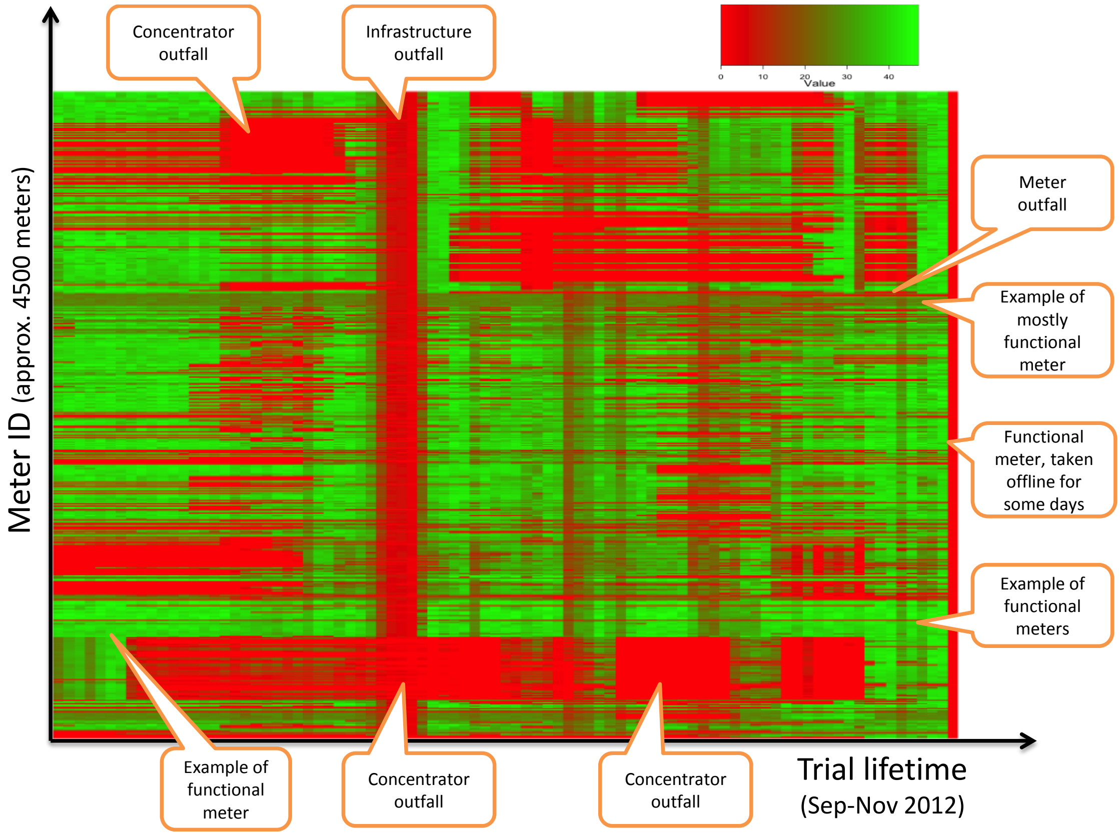


Figure 13.: Smart meter readings heatmap for the pilot period

infrastructure resilience, quality of information provided, etc., that may impact the deployment and operation of future Smart Grid services.

Figure 14 depicts the total count of received meter readings per each day, and we note that the number of received readings follows (as expected) the number of smart meters, indicating that the average number of received meter readings per smart meter, is quite stable. On the right vertical axis the number of smart meters with at least one reading per day is shown, and we realize that even though a high number of smart meters was live (and was expected to deliver meter readings), still an overall low number of them was received. From results depicted in Figure 14, in average 73% of all smart meters have delivered more than 50% of readings during a day. Still, for some days (excluding the infrastructure fallout), additional analysis on the data revealed that all meters had less than 50% of readings delivered to IEM. This was especially visible before the total infrastructure fallout (which could indicate a warning sign for such events).

As both Figure 13 and Figure 14 depict, being able to assess the quality of acquired data is key into understanding the infrastructure as well as if any future application operation could be supported or what aspects need to be enhanced to do so. In our case, many of the smart meter "failure" to deliver the expected number of meter readings could be traced back to extensive testing and reconfiguration of the infrastructure and the meters themselves. This had no impact on the real-world, as billing is the only service currently offered live in the

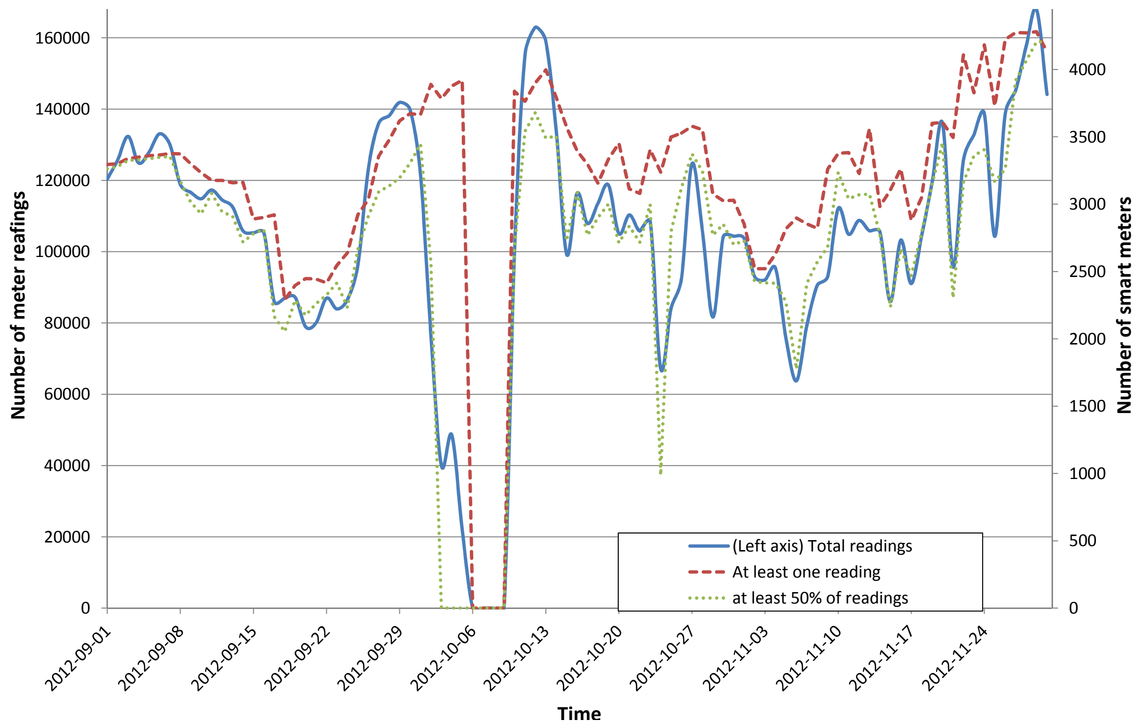


Figure 14.: Overview of the accumulated received meter readings per day

city and any subsequent and even delayed meter reading has the accumulated value of energy consumption. However, this had a significant impact on many other services presented in section 3.2, such as energy prediction and indirectly on the trading. It is clear that high quality of data and their timely assessment can provide a much accurate view on what is happening in the grid, and assist with a wide range of value added services [21].

3.3.2 *IEM Service Assessment*

As mentioned in section 3.2.1, all of the IEM services have been implemented. They were developed as *Java REST services* and deployed in a *Glassfish 3.1 Application Server* (`glassfish.java.net`) enabling their accessibility over both IPv4 and IPv6. The business data is stored in a *MySQL* DB (`www.mysql.com`). Specialized analytics and statistics are realized mostly on *R language* (`www.r-project.org`). All communication with the IEM is done over an encrypted channel i.e. *HTTPS* and a security (with role-based authorization and authentication) framework is in place based on *Apache Shiro* (`shiro.apache.org`). Additionally for performance reasons, all services interact using *Google Protocol Buffers* (`code.google.com/p/protobuf/`) which offer a highly efficient binary format. The implementation of the IEM constitutes of approximately 39,000 Source Lines Of Code (SLOC) implemented in Java.

IEM Service Request Analysis

The IEM services are implemented following the REST paradigm and hence can be accessed via the standard methods GET, POST, PUT, DELETE. Figure 15 depicts all requests made to the available services per group as these are shown in the architecture (in Figure 8), fully described in [21]. The POST method for Monitoring services (or RESTful create) was the most popular, as expected, since the metrics of smart meter data has been streamed to IEM via the monitoring services. Interestingly the Billing service had a lot of POST requests, but further analysis revealed that this was due to the contract creation and their assignment to all the customers during the begin of the pilot.

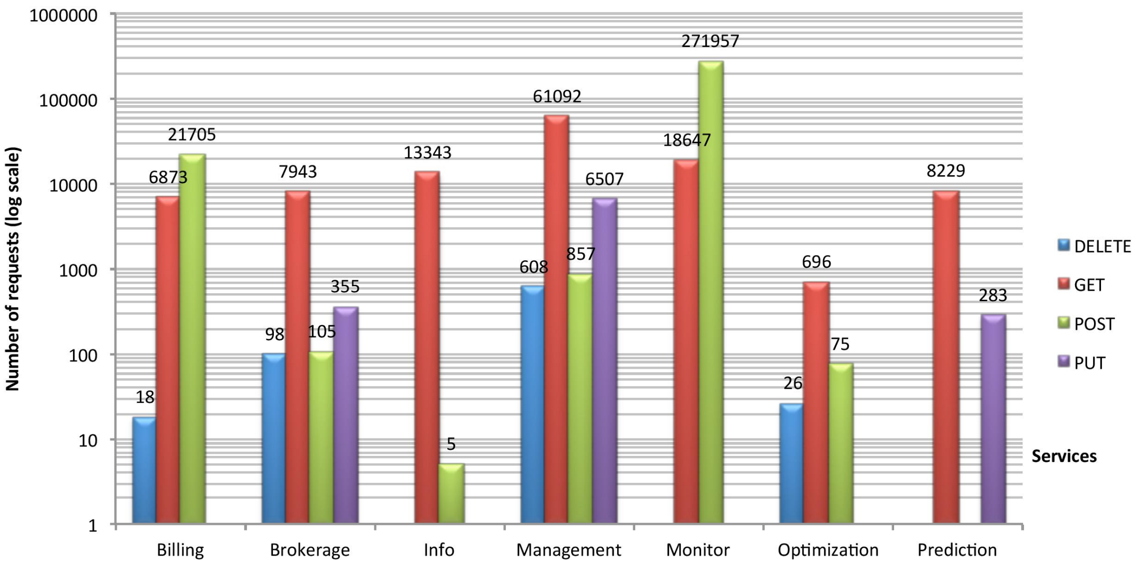


Figure 15.: Overview of service categories invocation for different methods

Requests during the pilot were made by three distinctive applications as shown in Figure 8, i.e. an energy portal and a mobile application via which mostly prosumers interacted, as well as the IEM monitoring and management application [34] used by the local utility administrators. All the service categories depicted a high number of requests for the GET method. From overall observation of Figure 15 one can conclude that Management, Brokerage, Monitor and Billing services were the most popular ones. Further analysis revealed more detailed information on similar usage patterns by the applications. For instance the increased Management requests can be traced back to the authentication process during the log-in stage of the application(s) etc. Further details of the analysis can be found in [38].

IEM Server load and DB Analysis

During the pilot, the IEM which provided the services for all applications in NOBEL was hosted in an on-line server farm (virtual machine) in Germany. The IEM server heavily relied on the MySQL database in order to hold all pilot data with a total of 6.1 GB of hard disk space spread over approximately 40 different tables. Since the pilot features more than 5000 consumers (and even more distinct smart meters), the values on Figure 16 are shown in percentage to the total DB size. Interestingly Figure 16 reveals that 98.65% of the space was dedicated to the meter readings i.e. energy reading, levels of reactive energy and other relevant readings (such as Voltage, Power, etc.). Using these percentages one can estimate DB requirements for a future large-scale solution, as well as get a notion where good design decisions are required, e.g space- and performance-wise.

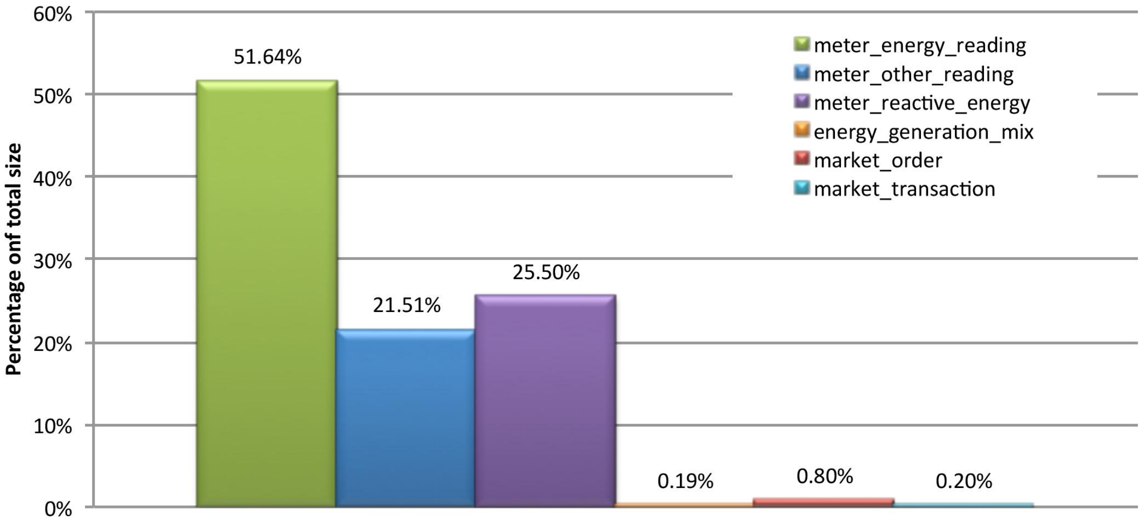


Figure 16.: Overview of the six most space consuming DB tables

For the SQL queries executed during the pilot i.e. Create, Read, Update and Delete (CRUD) operations, as expected the biggest part is devoted to storage of data, as well as acquiring information from the DB. Since the combined tables for smart meter measurements are responsible for the DB size (as shown in Figure 16), it is not surprising the resulting 53% of SQL INSERT and 44% of the SQL SELECT queries. However, the transmission of data relevant to the queries reveals interesting aspects. Approx. every SELECT query resulted in average to almost 7 times more data than INSERT; more specifically 8.58 GB was exchanged in total with an average SELECT of 4880 bytes and an average INSERT weight of 710 bytes.

From the analysis results so far, one can consider that a real-world system implementing the functions offered by IEM should be able to handle increased incoming load while the actual outgoing load depends on the end-user application request rate. However, both incoming and outgoing data rates could be estimated based for instance on the density of data metering or other information acquisition as well as functionality offered at the end-user application side. The communication part does not really offer an insight on the server load, especially when a simple service invocation might result in spikes in the server load due to massive data acquisition and analysis, while the final transferred result may be of minimum size [79]. Typical example might be the analytics over historic data that spans a custom-defined time frame of several weeks. Hence, careful design at DB level should consider the expected data flow as well as the service offering and restrictions on their functionalities. Although 5000 smart meters could not produce a great load on the server [74], in case of a high loads one should consider applying methodologies shown in section 3.1.2, rather than e.g. replacing hardware.

IEM Service Performance

The host platform where the IEM services are located play a key role on their performance. IEM has been designed to run on a distributed infrastructure and all of its components could (if wished) be installed in different systems with different computational, storage and communication capabilities that correspond to the expected load for those parts. To do so, all components of the architecture depicted in Figure 8 had to communicate strictly over REST APIs and no local dependencies were allowed. Figure 17 depicts an overview of the response times (in ms) of all services, as well as a categorization of the three applications that were accessing the IEM (as also shown in Figure 8). However, it is important to mention that hardware configuration used was moderate and no real optimization techniques have been applied, thus mostly the off-the-shelf component configurations were used.

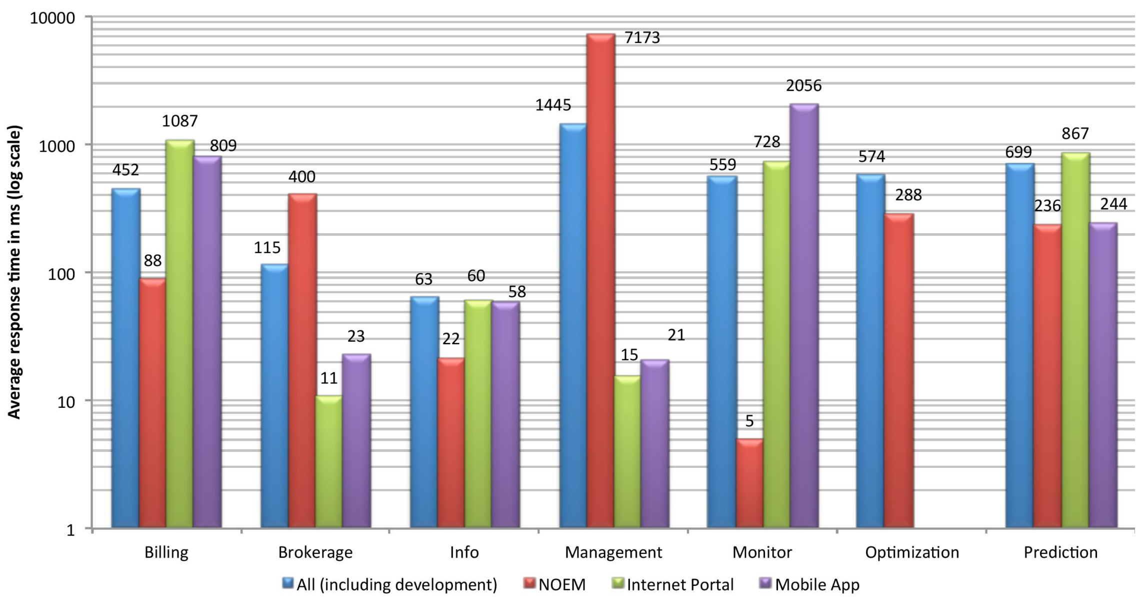


Figure 17.: Average response time per requester application for all service categories

Optimization in any aspect may result in better performance for the respective service. Of course, there is also a significant dependency on how and which services are used and the amount of processing requested on the server side as well as the data to be transferred. For instance, in Figure 17 the Management services indicate a high response time for requests coming from NOEM; however this can be fully justified as there are more the 5000 smart meters and more than 5000 customers, which would result in significantly higher payload transfer for Management services called by NOEM (that depict a system view) than the ones called by the other single-user applications where data of a single device/customer were needed. A detailed analysis identified some bottlenecks i.e. when invoking the service for all the customers (average 8235.5 ms) and all the smart meters (average 17737.54 ms), while fetching a single customer or device resulted to 8.36 ms and 7.12 ms respectively. The same logic applies to other services like the Brokerage services. The payload of these two responses may differ more than 5000 fold in our pilot. Although a complete analysis can be found in [38], already at this point is obvious that the performance of some services need to improve. This is particularly the problem for cases when multiple stakeholders are observed as a group, what will be further investigated in section 3.4.

3.3.3 *Neighbourhood Oriented Energy Management System*

One on the main consumer of the services provided by IEM is the Neighbourhood Oriented Energy Management System (NOEM) application, which will be presented in this section. Its design and functionalities developed aim to provide a tool to monitor as well as manage Smart City neighbourhoods. Some challenges faced were coupled with the need to be able to visualize real-time monitoring of various key indicators including energy production & consumption, prediction of energy usage, CO_2, energy trading volume, brokerage market management activities, energy optimization, customer communication etc. just to name a few. All the functionalities that NOEM provides depend fully on the IEM services from [21], while hereby only the relevant ones from section 3.2.2 are shown. In its final implementation [34], it acts as a demonstrator of the real IEM capabilities while it targets mostly the administrator of the envisioned neighbourhood infrastructures.

Monitoring and Management with NOEM

The NOEM is a web application being successfully implemented to mash-up the services presented in section 3.2. Complete application can be loaded by any web browser and is divided into eight functional areas: Overview, Monitoring, Management, Prediction, Brokerage, Optimization, Billing and Customer communication. Each functional area is accessible through its own tab, with the exception of the customer communication, which is accessible by clicking on the envelope icon located on top of the tabs. Most of the functional areas follow a simple pattern with an asset navigator on the left hand side, and the operational area on the right-hand side. The asset navigator allows the user to choose assets like: device, customer, group, etc. while the operational area displays the types of operations the user can effectuate on the asset. For instance, in the "Monitoring" tab, the user can select a device and view its current and historical demand/supply.

OVERVIEW The "Overview" tab gives a high level view of the energy production and consumption by aggregating all device measurements [79]. It also provides some additional information about the generation mix used to produce the electricity in the grid. The overview can also provide historical values by using the "start time" and "end time" dates located on top of the overview chart. Its usefulness is to provide a high level view for the Smart Grid neighbourhood with the main KPIs at place.

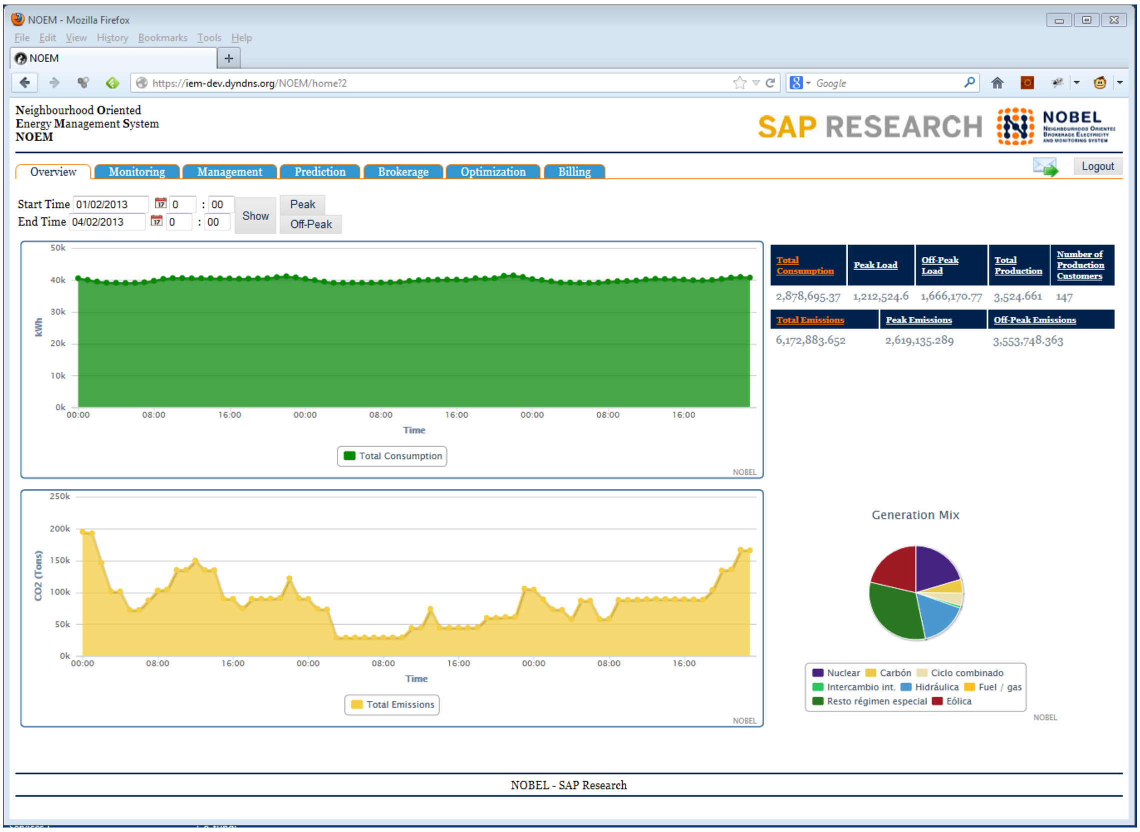


Figure 18.: The "Overview" tab showing total demand, CO_2 and energy mix

MONITORING The primary motivation of the monitoring view is to provide a high resolution report on energy production and consumption of the prosumer. The energy monitoring of NOEM also includes alarms (event monitoring) and notification capabilities. The navigator is in a tree format and can be used to select a particular group, customer, or device to be viewed; in addition search-as-you-type functionality has also been embedded to ease searching of specific customers, ids, meters etc. The NOEM application calls the relevant IEM services to present the required data. It provides the capability of monitoring the total demand reported by all the meters (already shown in Figure 18), as well as the demand and supply reported by individual meters, or groups of them [79] in a timely manner. Furthermore, additional metering data, that is, voltage, active power, reactive power, frequency, current and power factor, can be viewed. This data can be queries for specific time periods using the available controls. NOEM also allows users to set thresholds (via ”threshold” service in Figure 9) for devices, customers, and groups in the form of power (W) or energy (kWh). If the customer's, devices, or group's consumption exceed the threshold, an alarm could be sent to the user.

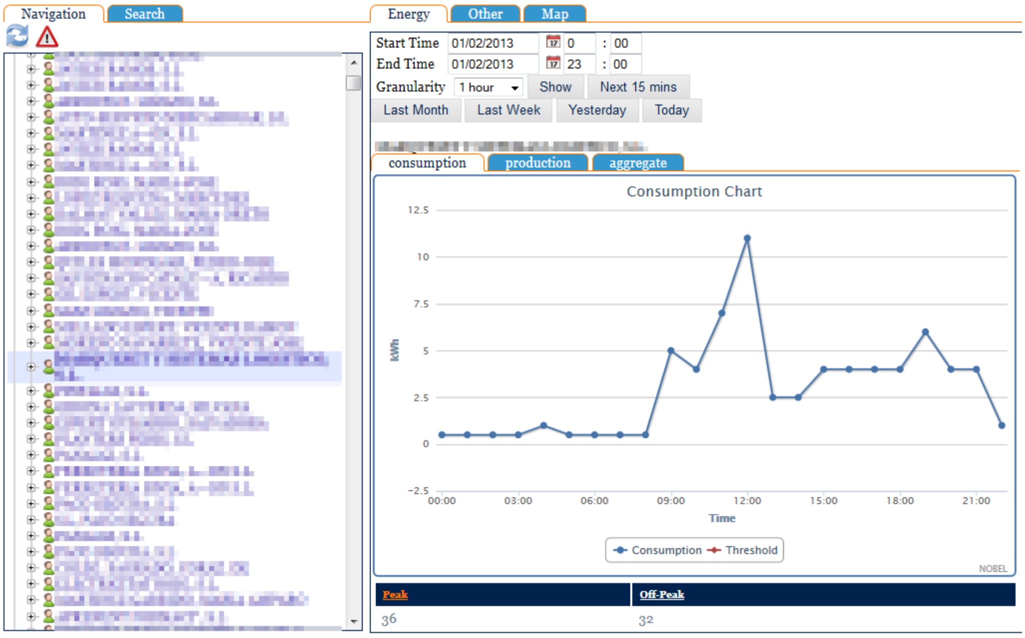


Figure 19.: The “Monitoring” tab showing a customer's daily demand profile

ENERGY BROKERAGE Energy trading is one of the key parts of NOEM as it is used also during the pilot to monitor and manage user's energy trades. The brokerage view in NOEM allows operators to manage several aspect of the participants in the NOBEL electricity marketplace [23]. Functionalities include monitoring of all market activities, overview of all market orders in a time frame, or in a particular time slot, visualization of information such as the trading price and total volume for selected time slots (in time window), as well as the last price curve and volume curve for each time slot. Operators also use this to control several aspects of the market itself e.g. the market participation for customers based on their capability of participation, or even disabling/invalidating customer activities e.g. trades. Since the NOBEL project envisioned automated brokerage agents, market operators can also manage the brokerage agents used by the participants. The “Market” view (Figure 20) allows the user to view current and historical market prices and traded volume. The “Order Book” view allows the user to inspect the order book (the current buy and sell orders) for a particular trading timeslot. Similarly the “Transactions” view enables a full list of market transactions that occurred in a timeslot including information on price, quantity, transaction time, specific participant, and specific order. The “Market Participants” view allows the user to select a customer (or multiple customers) and enrol into or remove them from the market. The "Agents" view shows a table with all of the automated trading agents in the IEM along with information about the customer on whose behalf they are trading and the management capabilities for that agent.

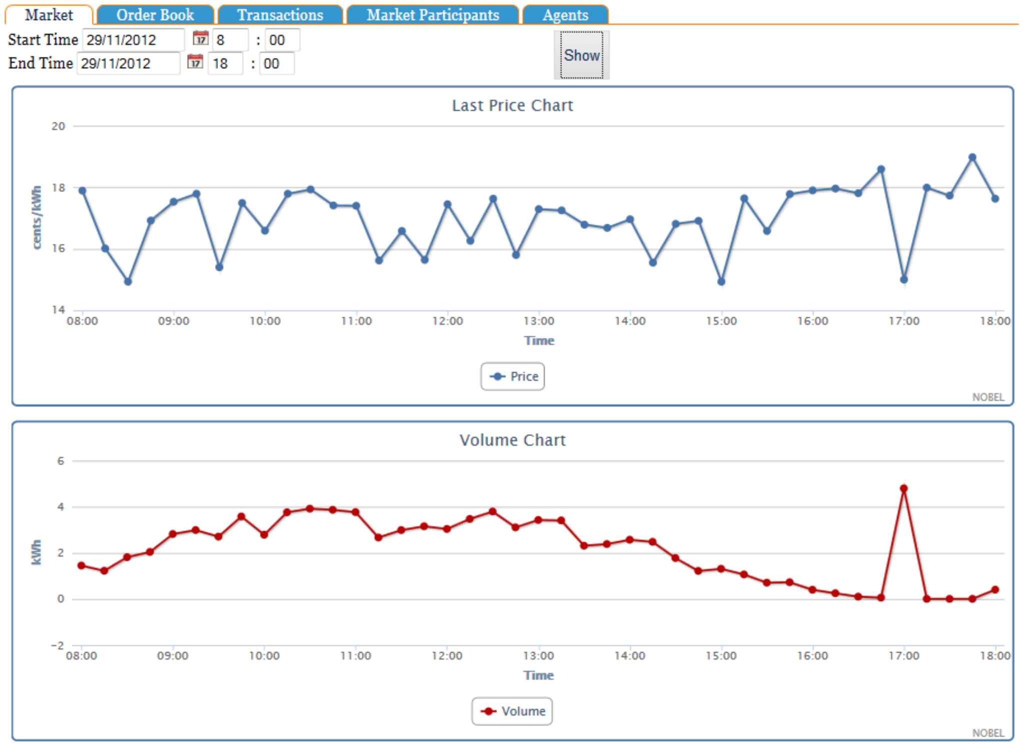


Figure 20.: The "Brokerage" tab with an overview to the "Market" depicting traded volume and prices

PREDICTION NOEM enables users to update the demand or supply prediction for individual customers, assets, or groups (as depicted in Figure 10). Prediction can be used by the customers and operators to help with their electricity planning and trading activities. The "Prediction" tab is used in a similar way as the "Monitoring" tab. A group, customer, or asset is selected and the standard time controls are used to specify the period over which to predict that asset's consumption, production, or aggregation. Once a prediction is made, the table on the right-hand (Figure 21) side can be used to correct any prediction errors. This can be done based on human intelligence via NOEM, or apply an automatized logic that will consume the IEM prediction services.

CUSTOMER COMMUNICATION The NOEM front-end application depends on the interaction with the information services from the IEM. Operators are able to directly interact with the customers and capable of messaging groups or individual customers. In our prototype the customer communication button allows the sending of text messages which are depicted on the mobile device or

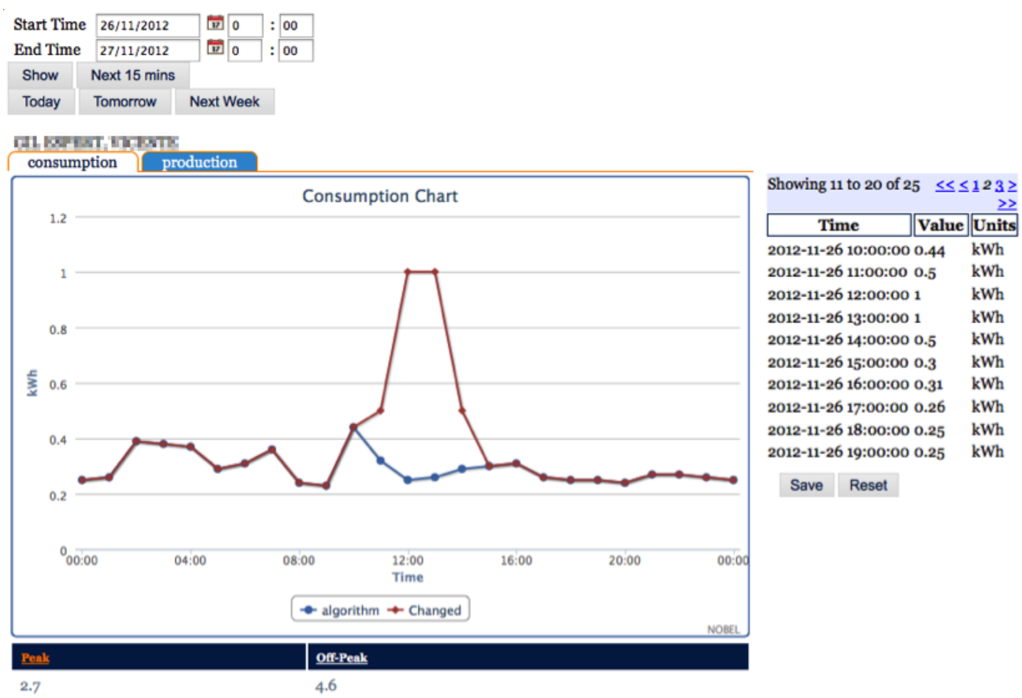


Figure 21.: Modifying the demand prediction of a customer

as e-mail. The bilateral direct communication of customers and operators has the ability to eliminate telephone center costs by enabling the user to create automatic tickets in the provider's system as well as receive customized offers.

NOEM Performance

Since NOEM was live for supporting approximately 5000 users in the city of Alginet (Spain) in 2012, the data collected allowed deeper understanding of the system performance (with more details in [34]). All data the application visualizes is retrieved from the IEM through its REST service API. The retrieved data must then be deserialized, processed and formatted for presentation on the browser. In order to evaluate the response time of the NOEM, the response times for requests directed at the NOEM were extracted from the access log file to produce a "response time duration curve" that shows the percentage of requests for which the response times were above a particular threshold. Requests to the NOEM application cover everything from retrieving data from the IEM, to browser requests for any other element responsible for presenting the retrieved data and other UI elements, for instance, images, JavaScript files, CSS files, and HyperText Markup Language (HTML) files. Some of these might maybe very quick to serve and might skew the response time duration curve, thus is depicted by the logarithm scale in Figure 22.

In general, the NOEM performance are well with roughly less than 5% of response times being above 1 second. Generally variability on response time can be expected as some of the requests can be quite resource and computationally intensive, such as requesting interpolated time-series data for long time periods

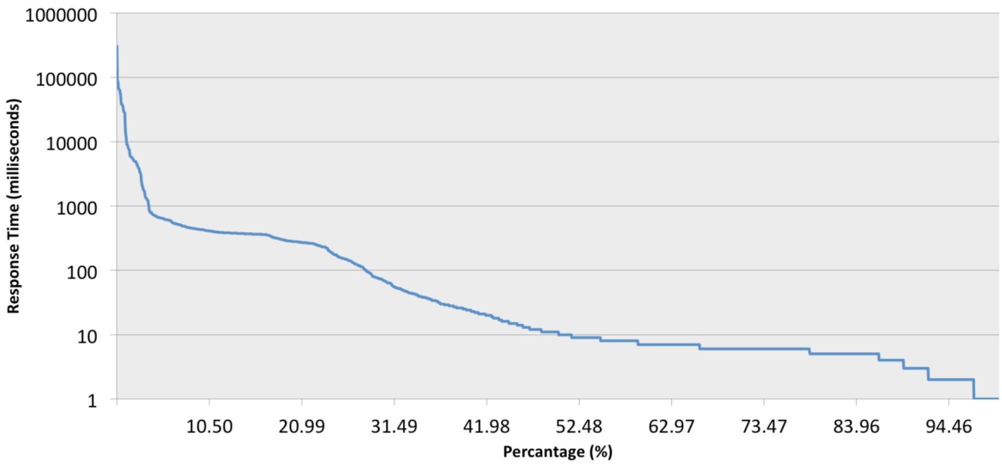


Figure 22.: Response time duration curve for NOEM requests (log scale)

(e.g., 3 months) or large number of customers in a group [79]. The longest recorded response time was of about 5 min (300975 ms). This was unusually high, with the next highest being about 1.5 min (88747 ms), and occurred while trying to login to the NOEM. Given the duration of the response time, it is likely the login attempt occurred while the server was starting up and the server probably tried to handle the request while waiting for the NOEM sub-components to start-up. Nevertheless, the application depends on the services offered by the underlying platform, thus performance shown in Figure 17 define the complete usefulness of the application. With that in mind, it was noted while running the trial that the monitoring services faced a significant performance drop as number of the metering samples increased in the database (depicted on Figure 16). To overachieve that bottleneck of large datasets, section 3.4 further investigates how data should be organized within the system.

3.4 PROCESSING REQUESTS IN REAL-TIME

Timely access to information collected from Smart Grids will allow new generation of innovative applications and services to be realized. For instance analytics on the vast amount of energy data [80] can lead to better prediction of energy customers [69] and offer new energy-related services [21] both in residential as well as in industrial environments [81]. Timely assessment and understanding may lead to qualitative better decisions and assist for instance city administrators to better run them. As an example a smart city energy cockpit providing city-wide information on energy usage and comparative analysis may enhance empower city officials to take decisions towards better energy management [34], CO_2 reduction, dynamic RES integration, EV charging, public infrastructure energy cost reduction, city investment planning, simulation of "what-if" scenarios, etc. Although data is collected and available, many services of Smart Grids are envisioned to be beneficial to users only if a sufficient accuracy can be achieved. To do so however, integration of multiple sources of data is needed, and subsequently heavenly processing should be applied e.g. consuming the Monitoring services from section 3.2 at building, neighbourhood, or even city level. The IEM services from are the perfect example, where before placing a market order via Brokerage services the Prediction service will be called to predict a load of the entire group which requires the Monitoring service. Therefore, services of such platform must be timely accessible such that groups of stakeholders in smart cities can benefit from advantages of acting together [25].

To achieve such objectives many of the enterprise systems are relying on the On-Line Transaction Processing (OLTP) for their operations; however the need for high-performance analytics has given rise to separate specialized systems delivering On-Line Analytical Processing (OLAP). If one drills down to real-time business analytics, while also taking into consideration the drastic performance improvements of in-memory systems, using an in-memory column database has some profound implications [82]. Independent of the row vs. column comparisons [83], the performance of the (in-memory) column-based solutions gained attractiveness as one is able to efficiently work in analytical as well as transactional workload environments [75]. To assess that, one can adopt several open source e.g. MonetDB (`www.monetdb.org`) [84] and commercial e.g. [85] column-based databases. In this section the focus will be on aggregation of smart metering data of various group sizes by assessing usage of a traditional DB (i.e. MySQL), including its in-memory variant, and an in-memory column-based DB (i.e. MonetDB). The aim is to assess some aspects with respect to energy measurement aggregations to improve the overall performance of IEM. This is done by using out-of-the box existing DBs without really diving deep to their tuning which could yell some additional performance benefits.

3.4.1 *Data Processing*

Several experiments were carried out measuring the aggregation performance of smart metering data. For all of them the acquired real-world dataset during the trials of the NOBEL project [37] in 2012 (`www.ict-nobel.eu`). The data has been collected by the IEM services [21] already presented in section 3.2, and contains the cumulative time series of smart meter energy readings. The meters have an energy resolution of 1 kWh sampled in between 15 minutes and one hour (depending on the meter). Since the resolution is 1 kWh many meters produced constant values and therefore were removed from the raw data set to reduce its overall size to be stored in database. This has resulted in having measurements with different time distances and are not available for every hour. As aggregation of time series data requires samples of same timestamps, this implies that the interpolation step is needed in order to provide data that fulfils this requirement. Their removal not only reduced the size of the set, but the overall precision is expected to be improved after an interpolation step (since samples in between 1 kWh can be expected).

Two different approaches for executing the data aggregation step are considered to evaluate the overall impact on the group aggregation performance i.e.,

I. Interpolating the raw smart meter data of a specific group and subsequently aggregating it: here the advantage is flexibility on the required sampling resolution, since the interpolation is done during runtime. Although, this approach leads to reduced usage of storage, the disadvantage is that the individual smart meter interpolation is done during runtime (for a selected group) that possibly can lead to lower performance.

II. Use pre-interpolated data and only execute the aggregation during runtime: the advantage may lie in skipping the individual interpolation within a group, but this approach requires much higher storage. Additionally, the flexibility is constrained by fixing the time resolution of the interpolation. However, the aggregation simplicity in runtime of the fixed resolution is expected to result in better performance.

To clearly depict the difference between traditional and emerging tools, two open source DBs i.e. MySQL to represents the traditional row-based domain and the MonetDB to represent the emerging column-based world. MySQL can store smart metering data on the hard drive and in memory, while the MonetDB by default stores all data in memory. To show the benefits of the in memory storage over the traditional solution InnoDB (hard drive) and in-memory were compared (both are part of MySQL). The selection of the InnoDB engine was due its caching algorithm of the frequently accessed data. As cached data is kept in memory, time to access the data can be reduced, but it still may require to access some data located on the hard disk. The in-memory storage engine by contrast, stores all the data completely in memory and no access on the hard disk is required.

Based on the need for interpolating the raw data set and the investigation of the smart meter grouping behaviour, a stored procedure for group interpolation is implemented for the traditional solution. This procedure is invoked by a thin client for a specified group (of numerous smart meters), the time frame and the resolution of the interpolated series. The boundary points, before the first and after the last data point, are also required in order to calculate the entire interpolation time frame. For simplicity reasons the algorithm is implemented as the linear interpolation method. As such, within the group interpolation procedure, every single device of the group is interpolated individually and subsequently all the smart meters are aggregated to a single time series. The experiments conducted in section 3.4.3 will help us understand why the interpolation stored procedure was required, and how it differs from a distinctive feature of column stores that can apply aggressive data compression.

3.4.2 *Experimental Datasets*

The original dataset consists of 5032 different smart meter devices and more than 3 million unique meter readings. For the experiments, two subsets were created from the original dataset; the first subset contains the data for only one month, whereas the second subset is complete with about six months of smart metering data. A detailed overview of the created datasets is provided in Table 1. As table indicates, the 15 minute interpolation will significantly increase the total number of points. The original sets will be referred to as A and B, while the pre-interpolated sets will be noted as A' and B'. Furthermore, the set distributions of the meter readings play a significant role in understanding the experimental results. As duplicates and constant values of energy readings were removed from the set beforehand, the final set description was calculated based on percentage of meter readings present in the set. More non constant meter readings available in the set will result in less interpolated points (as discussed in section 3.4.1). The table also shows the average number of non constant samples per each meter, which is calculated as percentage from all the samples for every smart meter in the set. [79] holds more details on the experimental environment.

	Set A (reduced)	**Set B (complete)**
Device count	4 020	4 382
Days	30 (1 month)	170 (≈ 6 months)
Time resolution	15 minutes	15 minutes
Sample count	537 604	3 365 627
Sample %	4.6%	4.7 %
Samples per device	2 880	16 332
Interpolated count	11 581 620	71 571 206

Table 1.: Overview of the two data sets used in experiments

3.4.3 *Aggregation Performance*

Aggregation of the energy readings of a group may be performed by few approaches mentioned in section 3.4.1. The first approach executes aggregation on the original measurements that contain no constant readings of energy, where each individual meter requires the interpolation step before the aggregation. The second approach is done on the pre-interpolated data, which actually requires more storage space, but the operation complexity approximates to the regular *GROUP BY* statement of SQL. In this section, performance of both approaches is evaluated by the aforementioned DBMSs on data sets of different sizes. The results of the experiments, independent of the selected set, are always referring to the time frame of aggregation (including the interpolation in first case) for exactly one month of the set, that is September 2012.

Interpolation and aggregation

If sampling of energy readings of smart meters is not made on equal frequency, or if samples are lost, the sampling frequency needs to be adjusted. The original datasets, presented in section 3.4.2, have the distorted samples of energy readings collected from the smart meters. As such, the aggregation step will require data (of interest) to be interpolated at runtime. Once raw data is collected (by submitting an SQL query to a DBMS), the individual interpolation is executed and the aggregation step is performed. The performance results of those complete operations is presented as the execution time in relation to the group size for the two different MySQL engines. However, the relevance of the overhead as explained in section 3.4.2 should be noted as can be witnessed in all experiments, and resulted in higher execution times for all small groups e.g. of less than 50 devices.

For the InnoDB case, the experiences are conducted on the reduced and complete set, respectively set A and B. A performance comparison is made by the count of rows in the table, thus the storage size required, to be processed by the MySQL DBMS instance. Their comparison will offer a better understanding of how the execution time differs, when data sets of significant different sizes are stored on a hard disk. The interpolation algorithms presented in section 3.4.1 are used for both sets and the results are shown in Figure 23. The execution time per device decreases with a higher group size for both datasets. However, one can also witness that the performance of the InnoDB engine suffers from the increase in the dataset size. This execution difference does not scale linearly with the size of the set, since the size of the set B is more than 6-fold. Still, because hard disk access is too expensive, one can immediately notice how performance suffers from the overhead for smaller group sizes. Finally, one can conclude that both scenarios continuously converge to the constant execution times per device as the group size increases.

The same experiment is conducted with the MySQL in-memory engine. In contrast to the InnoDB engine, where performance of hard disk I/O operations must be considered, the in-memory engine stores the complete data set in memory. As for InnoDB, experiments here also use the stored procedure interpolation algorithm (as explained in section 3.4.1), and Figure 24 shows the runtime results of numerous experiments. One can see the slight performance drop of the algorithm running on the complete set B. A significant convergence rate can be also noticed, in comparison to InnoDB, even for very small group sizes. However, it is interesting to see that the InnoDB engine for reduced set A performed almost equally to the in-memory engine on the complete set B. These results show that in-memory engine is not significantly affected by the dataset size, thus further analysis is conducted.

Although comparable results with in-memory and InnoDB engine for the experiments on the reduced set A can be witnessed, the experiments on the complete set resulted in performance differences. The interpolation procedure of the datasets with InnoDB engine dropped for approximately $20ms$, while in-memory engine suffered only a drop of approximately $2ms$. This difference (of few milliseconds) gets significant if an application requires thousands of devices within a group, that is further impacted by growth of a dataset. In overall,

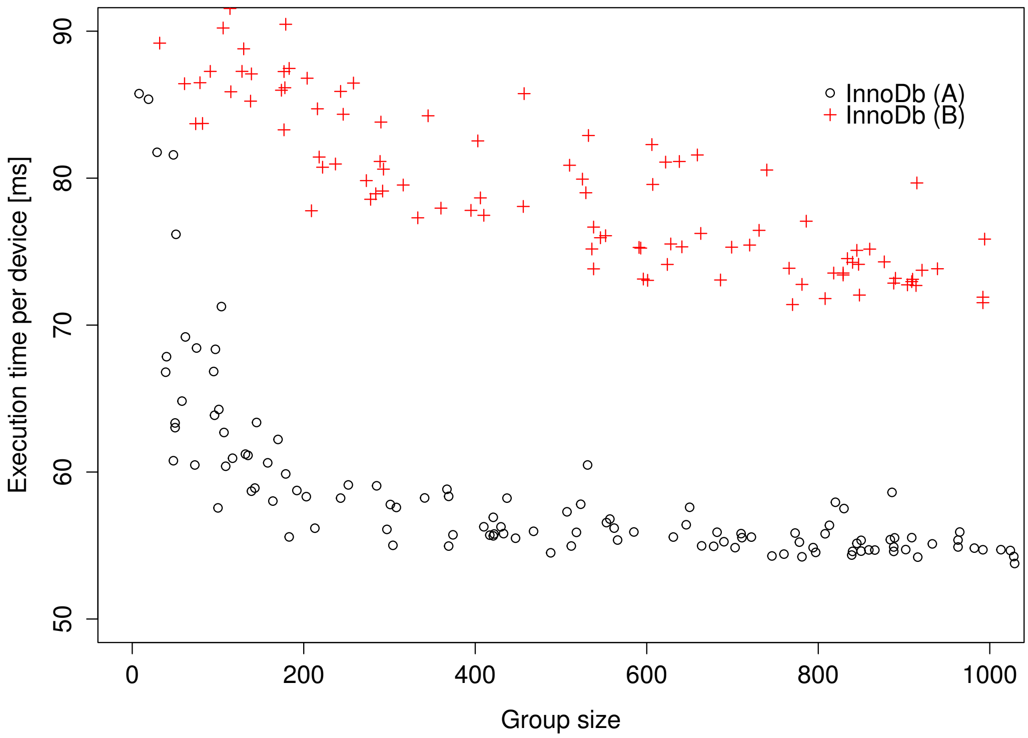


Figure 23.: Execution time for interpolation and aggregation with the MySQL InnoDB engine

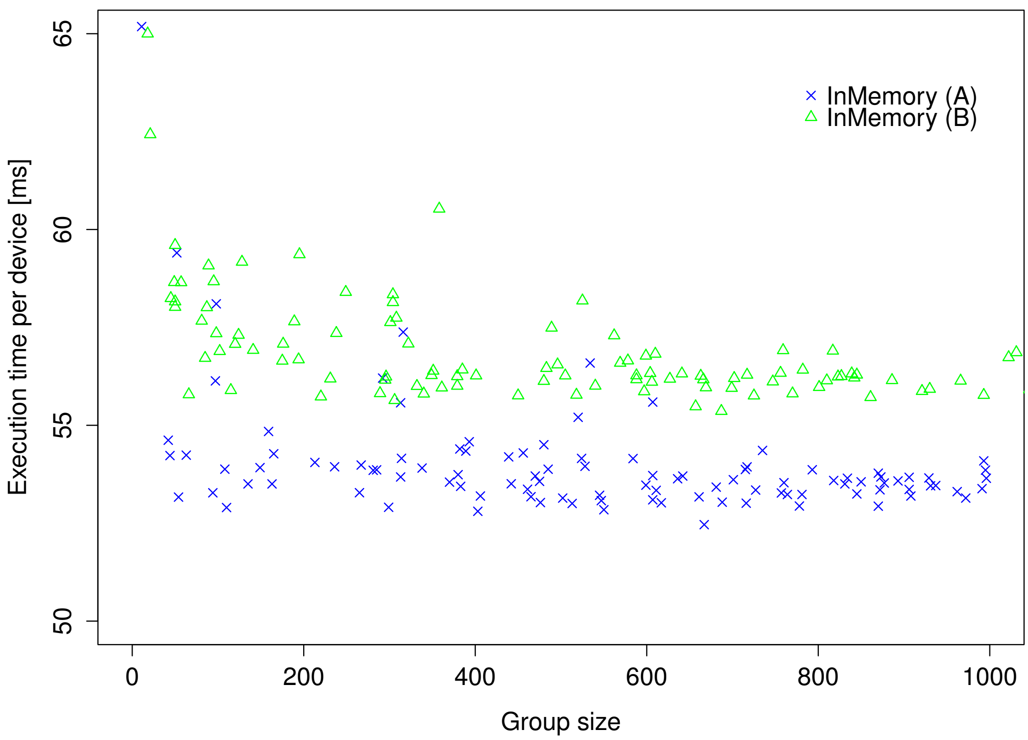


Figure 24.: Execution time for interpolation and aggregation with the MySQL in-memory engine

one can conclude that the in-memory engine performed better for all group sizes, while the InnoDB suffered approximately 10-fold more than the in-memory engine by the increase of the set size. If these results are adopted, even for the better performing engine, the aggregation will still take approx. 60 seconds for a group of 1000 devices. In real world applications such performance may not be sufficient, especially in near real-time systems eager for more efficient usage of the available resources [74], further experiments are conducted in the following section.

Aggregation on pre-interpolated data

In contrast to the previously executed experiments, the aggregation here is done using the pre-interpolated smart meter datasets. The distorted samples of energy readings within the original data sets (A and B) are pre-interpolated (A' and B') for each smart meter individually; hence now all the smart meters have the same sampling frequency (as this is stored within the DBMS). Pre-interpolation has some disadvantages such as fixing the sampling resolution and increasing the storage requirements, however it is expected to improve performance by reducing the aggregation time needed. For the experiments carried out here, the sampling resolution was fixed to 15 minutes and contains a much higher count of sampling points, as shown in Table 1. Via this preprocessing step, the interpolated data can be aggregated directly by executing a simple SQL query i.e. a *GROUP BY* statement. The performance of the complete operation is presented in form of the execution time with respect to the group size (for both MySQL and MonetDB). Similarly to the previous experiment, the relevance of the overhead explained in section 3.4.2 is included in all the experiments.

First the MySQL case is considered for both DBMS engines (InnoDB and in-memory) and compared to the reduced set A and complete set B. However, for the in-memory experiments the complete set of pre-interpolated meter readings could not fit the available physical memory of the machine. Still an assumption can be made, from the previous results of the in-memory experiments, that the execution time will not differ significantly in case of bigger data sets. Figure 25 depicts results of these experiments. Although the InnoDB shows an exponential performance improvement for higher group sizes for both sets, if compared to the results of previous experiments, a drastic drop in the performance can be immediately noticed. These results show once more that the performance of the InnoDB engine is highly penalized by the size of the dataset. The in-memory engine shows exactly the opposite behaviour, at least for the reduced pre-interpolated set A'. Interestingly the engine performed very well for all group sizes, especially if compared to the results of previous experiments depicted in Figure 24. From these experiments one can conclude that the in-memory engine performed a lot faster, because the high-cost access time to the hard disk do not incur. As an indicative example, the execution time for group sizes of 1000 with the InnoDB is approximately $60ms$, while for the in-memory engine resulted in only $1.5ms$.

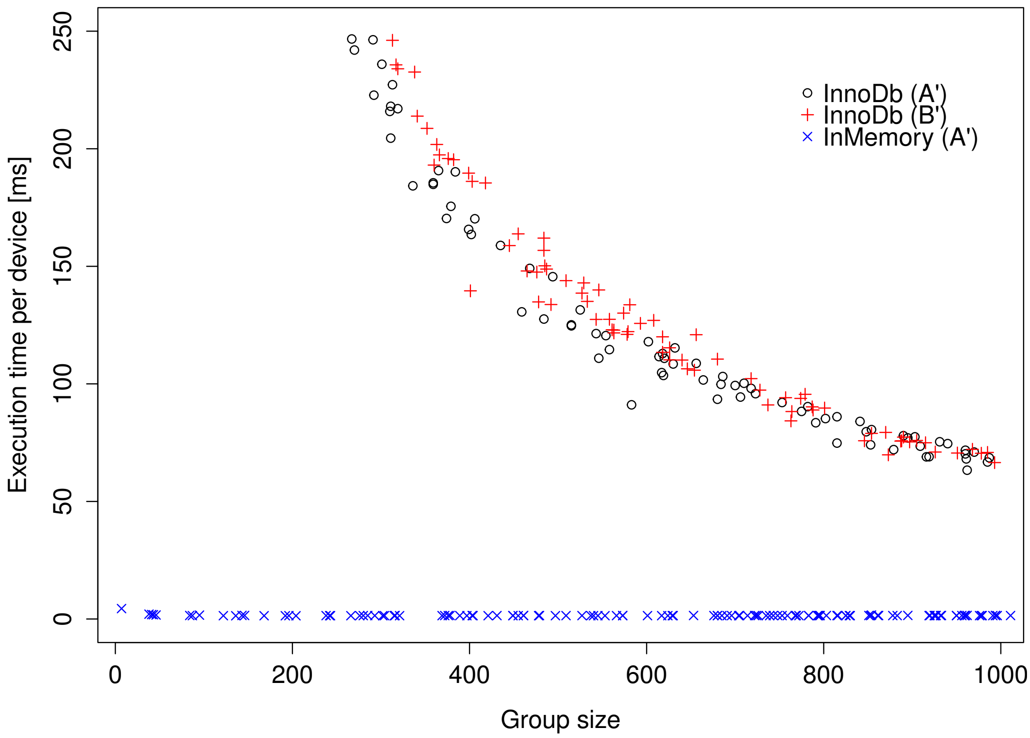


Figure 25.: Execution time for the pre-interpolated data with the MySQL

Due to the superior performance of the in-memory engine, it was decided to experiment with an in-memory column based DBMS i.e. the MonetDB, and conduct the same experiments. In contrast to the in-memory case of the MySQL experiments, the MonetDB solution has no problem storing reduced and complete datasets in memory. A distinctive feature of column stores is the application of aggressive data compression. In this way, one can use compression and some extra CPU cycles in order to fit the entire data set into the physical memory (what was not possible with the MySQL in-memory engine). The results of these experiments are depicted in Figure 26, where execution time per device of the MonetDB resulted in much higher performance when compared to the in-memory solution of the row-based MySQL. As an example, the performance of interpolation and aggregation of 1000 smart meters (on Figure 23 and Figure 24) is $\approx 60ms$, while for this experiment we can see that it is 60-fold faster ($\approx 1ms$). The size of the datasets had a minimal impact on the performance, although the size of the data set B' is more than 6 times larger than the set A'. For both sets a fast convergence rate can be seen for group sizes greater than 1000, while the performance for the group sizes less than 1000 had a certain drop in performance. It is expected that the performance improvement rate is actually the software overhead, also being affected by the decompression time (as discussed in section 3.4.2).

The acquired results are further analysed to cherry-pick the best ones. Obviously, the in-memory engine performed far better than InnoDB, thus it was selected for the overall comparison that is depicted in Figure 27. It is evident

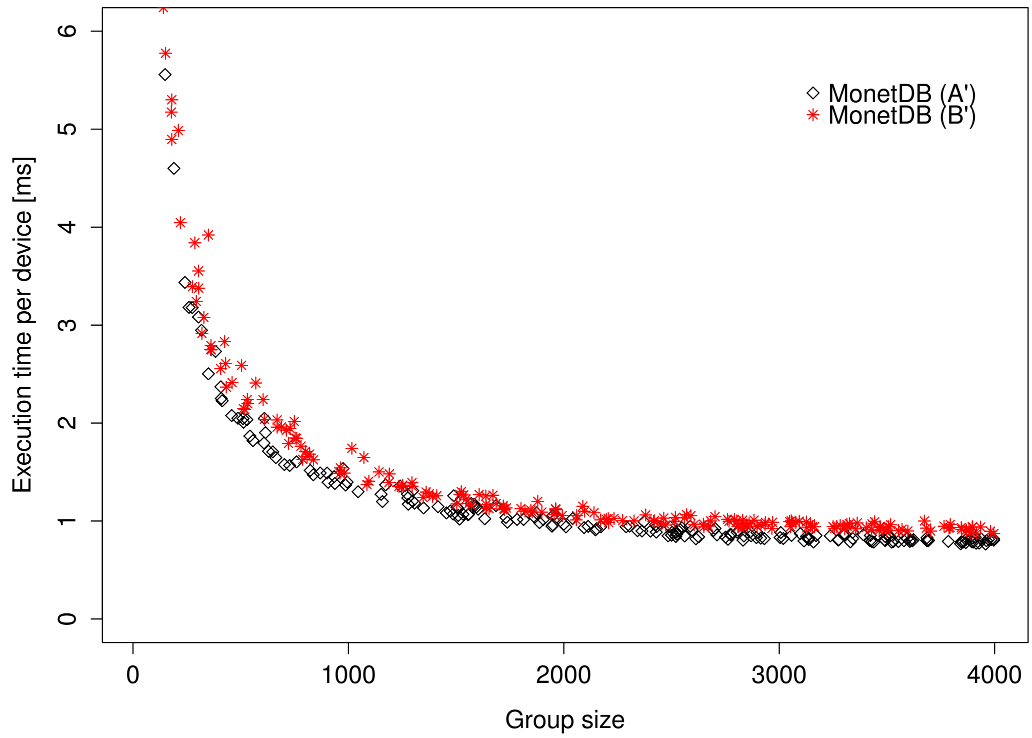


Figure 26.: Execution time for pre-interpolated data with MonetDB

that the MySQL execution time decreases up to a group size of 200 and then starts growing, while the MonetDB execution time continuously decreases and converges to the execution time of less than $1ms$ per device. These experiments witness that the MySQL in-memory engine exceeds the MonetDB performance for smaller group sizes, but limited only to the reduced pre-interpolated set A'. However, even though the MonetDB resulted in expensive execution times for smaller groups, it performs significantly faster for bigger group. For future industrial and business applications exactly these large groups are the main point of interest, and hence constitute our main focus.

Overall Comparison

The experiments conducted for both defined scenarios depict how the aggregation performance is affected by the various storage technologies as well as the potential pre-processing of data such as the pre-interpolation. The analysis of the experiments revealed that pre-interpolation of data has significant impact on the performance boost. The column based DBMS (MonetDB) proved powerful by storing reduced and the complete dataset entirely in memory, while the traditional DBMS (MySQL) had severe limitation in our experimental environment. Still, the MySQL in-memory engine over performed MonetDB for smaller groups, while MonetDB showed continuous improve even after over performing MySQL. To get a better understanding of the performance benefit of both solutions, the total execution time needs to be compared from the best of breed DBMS cases. For that purpose, the overall execution times of experiments on the reduced

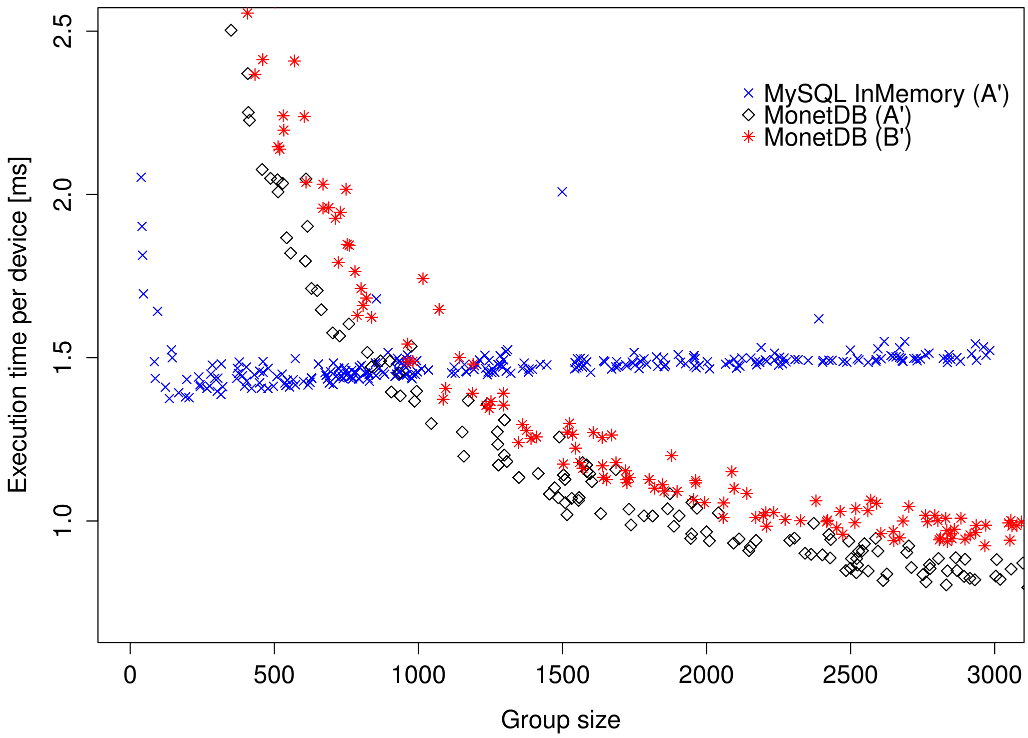


Figure 27.: Execution time comparison for the pre-interpolated data per device

pre-interpolated set A' can be calculated from Figure 27. The high performance improvement realized by the MonetDB even for bigger groups (as shown in Figure 26), leads to increase of the gap between the overall performance between the two DBMSs. Although overall MonetDB performed better for the experiments of bigger group sizes, Figure 27 indicates that total execution time increases for the bigger groups, however at a much lower rate than MySQL.

These experiments can be used as a rule of thumb towards making an informed decision for solutions running in different environments. Depending on the performance requirements and data one can select a configuration that fits its business objectives. Hence for a highly limited environment, one needs to consider the performance limitations if the interpolation is executed at runtime. However, if high performance is required (such as for group monitoring services from section 3.2), one should focus on the pre-interpolated data sets as these assist towards removing the performance penalty (time) needed for the interpolation step. Pre-interpolation of data sets can be scheduled more flexibly for historical data, as such action leaves only the real-time data to be analysed. If observed from the angle of the IEM platform located in a cloud, the server capacity allows pre-interpolation of data consumed by high performance services, thus significant improve of the results from section 3.3 can be achieved.

3.5 LESSONS LEARNED AND FUTURE WORK

The MDS experiments realized in section 3.1 were measured in a high bandwidth, single hop and unconstrained network. This, however, is not a reasonable expectation for a real world scenario (as shown in Figure 1), where timely delivery is impacted by far harsher network conditions. The challenge exists especially in between a concentrator and smart meters where a variety of heterogeneous networks is expected (e.g. residential ADSL connection, power line communications or even through existing wireless mobile phone networks). To completely understand delays in data delivery, one should experiment with conditions that are more reflective of reality. The simulation environment, on the other hand, was useful in understanding the theoretical limitations of such system [74].

From the technical viewpoint, a web service enabled infrastructure was considered and more specifically the traditional implementation of web services where SOAP is used. Although Figure 7 depicts higher efficiency for greater bulk sizes, there are more lightweight approaches out there. Using technologies that would enable several other aspects e.g. lightweightness, high performance, backward and forward compatibility etc. were considered. All services are Internet based and use the HTTP as a complete application protocol, which also defines the semantics for the service behaviour (as followed by RESTful approaches). Additionally to the REST style, the Google Protocol Buffers (as an extremely efficient binary format) for enhanced performance was used, since significant amounts of data had to traverse the network from the platform to the applications. Bulk data transfer was also used [74], instead of many smaller messages. This further improved the system efficiency and make the applications more network friendly, as resulted from experiments in section 3.1.2. In overall, the initial results for web service enabled devices such as the smart meters appear to be promising [86].

To validate the design of services in the IEM platform, multiple web applications was created composed solely from a mash-up of the IEM services [38]. During the process of creating such applications, for a variety of different stakeholders (residential end-users, utility etc.), valuable lessons, both technical as well as other related to design and social aspects, were acquired. Many functionalities are identified as "generic" that serve the majority of applications, should be hosted on the server side and include the sophisticated logic. In that line of thought, more lightweight applications can be developed, while their functionalities are decoupled from the data processing logic and intelligence of the service, which can evolve independently. Beside NOEM from section 3.3.3, similar applications can take advantage of the same services are also built for mobile devices [87]. Their common requirements were identified early and built into the platform. At this point, the initial basic set of services is being maintained, while the IEM platform is extended with additional functionalities e.g. those later needed for SFERS in chapter 5. Finally, strategies to detect service performance deterioration and handling are also needed. As many of these pose a vivid research area especially in cloud computing domain, can be assumed that aspects of service monitoring and life-cycle management will be provided by an underlying platform hosting these energy services.

The problem of extensive times in server processing occurred for some requests (due to their nature or a server overload) and in the meantime the client either timed-out or was blocked. In example of service failures within IEM pilot from section 3.3, they may occur both at client and server side even during the processing of a request. Typical examples are those of network failures, time-outs, service crash, etc. A typical case for when the group of smart meters was request in detail depiction of energy data for a long period of time e.g. 10 months. Here one can either consider asynchronous calls or publish/subscribe mechanisms. Both approaches will bring additional complexity in addressing them [73], therefore performance optimization steps were made as presented in section 3.4. These results indicate that in a real operational environment it would be expected to make use of high performance in-memory DBMS [85] in order to deliver analytics over mass data in “real-time” [79]. On the other hand, some services could not benefit from such technologies, simply because service an application inappropriately invokes services. For instance on the Brokerage service in Figure 15, observed that the service fetching the orders of a specific customer (average 217.67 ms) is slower in the response time than the service returning all the orders (average 26 ms). Further analysis showed that an inefficient use usage of the service by client application introduced these delays i.e. delivering all customer’s orders (filtered on the application side instead by a proper service parametrization).

Scalability is of key importance, especially when considering that all the services now hosted under IEM, will have different usage patterns and performance requirements. This work targeted to make the IEM platform scalable and distributed. Thus, design decision to enable service oriented RESTful interactions among IEM services and not take advantage of other intra-component calls (which may have resulted in better performance) was justified. Anyway such a design decision may impact the data quality, whom within the trial resulted as to be the critical point. Missing or delayed data was identified to have a significant impact on key functions such as prediction or analytics, and a cascading effect on decision-relevant processes depending on them, e.g. energy trading, preventive maintenance etc. Hence adequate identification of data quality issues, as well as estimation of missing or delayed values should be further investigated.

Security, trust and privacy are challenging issues that are expected to be an integral part of design, implementation and deployment of energy service platforms. All MDS experiments from section 3.1.2 did not considered the security or privacy layers. Clearly integrating any solution there will a significantly impact on the overall performance. Experimenting with WS-Security, secure channels (HTTPS) or encrypted meter readings, will give an insight to the magnitude of such impact [88]. Additionally the use of latest hardware (not dedicated though) which have the native AES support may assist in minimizing the performance drop if such layers are implemented. In case of IEM the basic HTTP authentication and authorization were used by all services provided by the platform, and secure interactions over encrypted channels i.e. all REST calls were made over HTTPS. The trust here was placed on the end-devices delivering valid data; however device authentication as well as data checks (for replay, modification of values, other sanity checks etc.) should be made in operational environments. Developing secure resilient infrastructures in the Smart Grid era is considered a grant challenge [89]. Clearly, precautions must be carefully considered for real-world deployments which might additionally include message signing and encryption of service hyperlinks.

3.6 Conclusion

Within the scope of MDS in section 3.1, the approach was demonstrated to be used as a rule of thumb when high resolution meter reading should is targeted within an AMI. The straightforward 3-layered hierarchical architecture was taken and its performance is evaluated through the component point of view. In a methodological way, potential problem areas, as well as the line of thought that should be followed in order to find possible inter-dependencies and roadblocks, was identified and discussed. By investigating each component limitations, the evaluation results in this chapter narrow down the operational ranges one could use to achieve high performance. It was demonstrated that is possible to realize a high performance AMI based on common hardware and open source tools, without tweaking was done on the hardware or general tools i.e. operating system, application server etc.

Once data is available, significant effort is invested towards creating innovative applications for the emerging Smart Grid. It is expected that the future of Smart Grid applications resides on sophisticated web applications that can rely on a multi-source data and Internet provided basic services that can be easily customized for the specific end-user groups. The approach from section 3.2 and the experiences acquired, while developing the platform for energy services that can be used to empower end-user mash-up web-based applications, have demonstrated the need for it as well as its promising potential [87]. Also the benefits of group activities [25] were assessed from the performance perspective and demonstrated by using the traditional row-based DB (MySQL) and the in-memory-column based approach (MonetDB). These efforts take us one step closer the vision of integrating real-time analytics into modern applications, such as a smart city energy cockpit in projects such as SmartKYE (`www.SmartKYE.eu`), clearly points out to the need of being able to do open and high-performance processing of the energy related data.

The evaluation delivered upon the real world aspects of such platforms and, furthermore, demonstrated via the NOEM web application is of crucial importance. A balance among real-world requirements, needs of futuristic functionalities, technology-driven decisions and use of open source technologies is achieved. As such, the trial results shown promise that ICT of today can support sophisticated end-user applications for the fully-blown vision of the Smart Grid [18], what will be one of the main requirements for SFERS. Still, without an accurate self-forecast many of the IEM services, but SFERS concept as well, won't be economical. Thus, after enabling en efficient communication in between the stakeholders, research in achievement of higher forecast accuracies need to be done.

4 Craving for Forecast Accuracy

With the dramatic increase on fine-grained data, new challenges arise as forecasting can now also be done in near real time and detailed time-series data [17], allowing new opportunities for future applications and services [21]. The role of foresting and stakeholder interaction in real-time will be pivotal in the envisioned Smart Grids [73], such that added-value services from section 3.2 can be consumed. With that in mind, an accurate forecast is becoming increasingly important, in particular if the stakeholder base expands with the traditionally passive consumers taking advantages of the new Smart Grid offerings. As an example, one may trade energy on smart city level marketplaces [23], or even benefit by offering its predictability to third parties [90]. Even today we face such as cases, e.g. demand-response programs, whose participation requires from resources to be measurable [12] and therefore predictable. As one can see, an accurate forecast will be the enabler to take advantage of such opportunities and this chapter will further investigate how traditionally passive consumers can get there.

The grid load forecasting has a prominent position as it is a crucial planning step and today is made through the highly-aggregated data [14]. As one can imagine, different scales of customer aggregation has its progress towards the accuracy of the highly-aggregated data [40]. Still, the question of impact on smaller scales of aggregation remains unanswered. Today this is beneficial for the customers, as incurred costs of retailing forecasting errors are shared homogeneously among all customers within such aggregations i.e. via their electricity bills. That way, if individuals are behaving stochastically, costs of their stochasticity propagates to the entire aggregation [39]. Although it is not always possible to eliminate or significantly limit the prediction error, clusters composed of predictable customers will have higher accuracies [68]. As such, stakeholders may cluster them as part of a virtual "predictable" group [61] in order to represent an adequate [77] or a measurable resource [12]. Such constellation is called prosumer Virtual Power Plant (pVPP) [42], where their usefulness is directly bound to understanding their potential contribution to the grid and the ability to control it.

It is not always possible to reduce forecast errors [91], not even within the highly-aggregated predictions [39]. Even retailers today report Mean Absolute Percentage Error (MAPE) of 2% – 5% [40] and these errors are passed to the electricity bills of consumers. Interestingly, Renewable Energy Sources (RES) suffer even from a higher stochasticity, but many methods can be applied to make them appear predictable [92]. Deployment of storage solutions [70], in highly volatile RES systems [10] or strategic deployment to improve general grid operations [62], as therefore identified as the step towards resource adequacy. As one can expect, same solutions can be applied to further improve the forecast accuracy achieved in grouping scales, resulting in second grade of the forecast accuracy convergence in a cluster [69]. As such, one may expect grouping and storage technology may be applied already on the level of modern buildings [93], thus enable them for the active engagement due their accurate energy signature [94]. The question of how storage units, that may already exist in smart cities [52, 70], can assist in achieving a better forecast accuracy (for groups composed of a small number of prosumers) remains unanswered.

Many practical approaches cover their excesses and shortages in forecast by static storage systems, such as Battery Energy Storage System (BESS), and their relevance can vary significantly [95]. In this chapter, an empirical approach towards understanding how errors of a stakeholder can be better addressed if average capacity needed is properly distributed on intraday intervals. This is an important observation since storage solutions are expensive [65]. Instead, one should consider involving assets that are capable of absorb the forecast errors and therefore replace the storage units (at least in certain point of time). Hereby, existent Electric Vehicles (EVs) from a fleet of a stakeholder are actively used to compose energy storage [94]. The EV presence on premises allows usage of its battery to absorb forecast errors, and hereby is empirically assessed with the smart metering data of a commercial stakeholder. The final results will show that capacity of the fleet available can greatly reduce need for a static storage solution, while its support might be beneficial for intervals of a low fleet presence.

An initial evaluation on how important is the group act for these added-value services is done in section 4.1. Such results raise further questions of how predictability is affected on different scales of aggregation and is investigated in section 4.2. Still, the accuracy obtained by grouping step may not be sufficient for some stakeholders, thus section 4.3 investigates the second grade to its improvement. Due significance assessment within the experiments of the storage technologies, section 4.4 and section 4.5 will investigate steps to reduce costs of such solutions via owned assets i.e. an EV fleet. The conclusion on the knowledge gained in this chapter is done in section 4.6.

4.1 TRADING AS A GROUP

Electricity markets are seen as the cornerstone of liberalized power systems. They provide an efficient mechanism for the allocation and pricing of the generation capacity used to meet power demand. The paradigm change in [8] will empower a second generation of innovative applications and services, and one of them, already presented in section 3.2.2, is the local energy trading. As said before, passive consumers, such as households and small businesses, are being empowered to also become producers [47]. As they are outfitted with a generation capacity, such as roof-mounted solar Photovoltaic (PV) panels, they can even take a more active role e.g. by consuming the brokerage services of Integration and Energy Management system (IEM). These services, envisioned for the smart cities, can be considered as a "soft management control" at local level, thus dramatic effects on the power grid mentioned in [96] can be properly addressed. Still, the locational and sometimes intermittent character of distributed generation will emphasize local energy management and require higher stakeholder engagement. Some suggest to improve energy management by creation of cooperatives [61], or "energy communities" [24], and local electricity markets [22].

Market models such as [97, 98], and the NOBEL market model [23] used in in this section, have been shown to be an effective method for the coordination of local consumption and production. In a local and intraday electricity market such as NOBEL, forecasting accuracy plays a key role for success of participants and market itself. However, forecasting demand requirements for small highly-dynamic entities, such as single households, can lead to higher errors and consequently to the market-related penalty costs. This might be a potential barrier for economically feasible participation and realization of such markets. Many possible solutions to improve forecast accuracy could be applied [28, 40, 69]. Using such methods, not only one can exhibit lower forecasting errors, as it will be demonstrated, but can also potentially lower the risk of market participation for their members through the internal sharing of resources, costs and benefits [90].

In this section, the positive impact of grouping to obtain higher forecasting accuracy [68] will be applied and exploited for an effective participation on the NOBEL market (accessible via the brokerage services presented in section 3.2.2). The evaluation is carried out through a simulation of the market interactions of various prosumers, consumers, or a group of them. All the simulations are based on smart metering, solar irradiation and weather data collected during the NOBEL field trial in the city of Alginet, Spain. Complete results of the evaluation can be found in [25].

4.1.1 *Evaluation Methodology*

The effects of grouping on the forecast error reduction are evaluated through a discrete simulation model of the NOBEL market in a similar setup to the one described in [99]. The simulation comprises 1897 participants trading on unique 15 minute intervals of the market for the month of September 2012. The participants are divided into two roles: consumers and prosumers. All participants have their own predicted electricity demand profile, while only the prosumers have generation capacity. The simulation advances 15 minutes per time-step, the duration of a market timeslot n. To these timeslots participants can submit market orders based on their forecasts. Data for each participant (e.g. real demand, predicted demand, quantity bought/sold) and for each timeslot (e.g. total consumption, total production, total energy traded) is collected for the evaluation.

Simulation Model and Data

The electricity consumption for each of the simulated participants is based on real smart-metering. These measurements, with a sampling resolution of 15 minutes taken during the NOBEL field trial. Although recent work demonstrated slightly better results for the Holt-Winters (HW) algorithm [68], same work also demonstrate highly accurate results for the Seasonal Naïve (SN) algorithm. Due its performance and simplicity, the SN forecasting algorithm is applied (to each smart meter individually) to predict demand for each participant.

The generation profile of the prosumers in the evaluation scenarios is simulated. The PV generation technology was chosen as it is a main player in the context of distributed generation [100], and due to its increasing growth in the residential and commercial rooftop segment [47]. Installation of any participant with generation capacity is sized so that it will produce up to 50% of the participants total demand for the simulated period. For instance, if a participant consumes 100 kWh within the simulated period, its PV installation would ideally produce 50 kWh over the same period, weather effects notwithstanding. At this level, the average self-consumption rate of prosumers will be roughly 70% of their total consumption. This was observed to be the saturation point of self-consumption for Spanish prosumers equipped with a photovoltaic system [63].

To execute trading strategies the Zero-Intelligence Plus (ZIP) agents [101] are utilized by participants to simulate their market interactions. Full description on trading strategies and their parameters can be found in [25], hereby only the relevance assessment of forecast accuracy in consuming advanced services brought by the Smart Grid era. Still, for understanding of prosumer loads, Figure 28 depicts the demand, predicted demand, and generation output of one day for one of the participants.

Evaluation Measurements

The evaluation is carried out in two cases. Firstly, the outcomes of group participation are evaluated. This is done by creating a group of participants that trade on the forecast aggregate behaviour of its members. Latter, same participants of the group are compared against their individual performance. The experiments in section 4.1.2, the evaluation is centred on four key measurements: demand imbalance, uncapitalised generation, unnecessary buys and sells. If every participant acting on a timeslot n has its individual variables, that is (actual) consumption $C[n]$, (actual) production $P[n]$, amount bought from the market $B[n]$, and amount sold to the market $S[n]$. These measurements are defined as follows:

Definition 1 *A participant can have a Demand Imbalance $\delta[n]$ on a timeslot if there was an amount of energy bought from the market that could not be used by the participant due to insufficient demand. That is, $\delta[n] = max(B[n] - C[n], 0)$.*

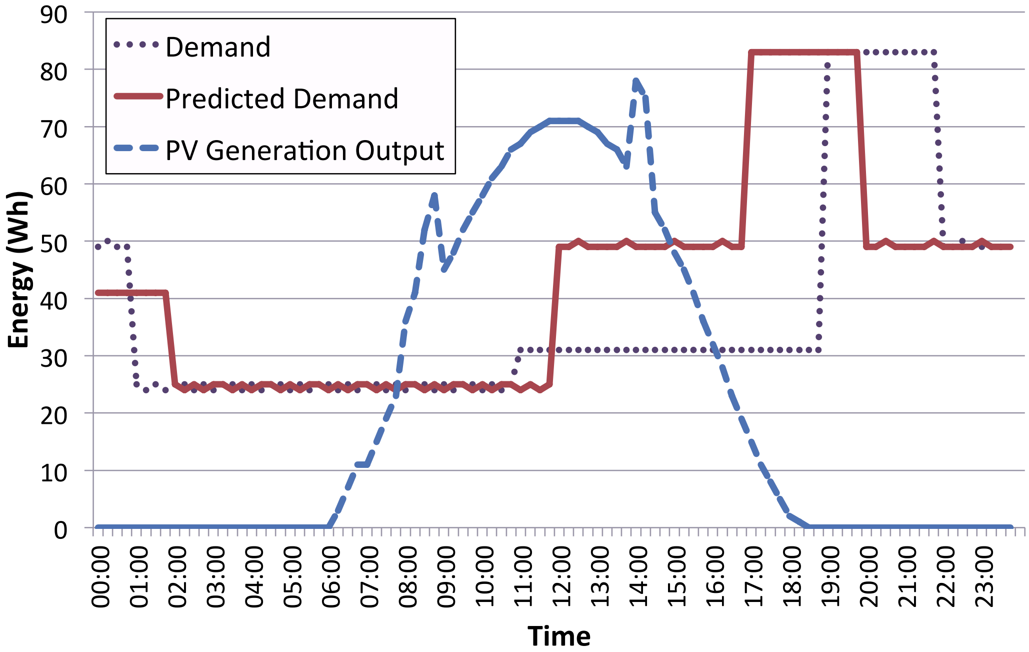


Figure 28.: Example day of actual demand, predicted demand, and generation output for one of the participants.

Definition 2 *A participant can have Uncapitalised Generation $\gamma_{p,t}$ on a timeslot if there was an amount of energy it could have produced that was not sold on the market, and could not be used to service its internal demand. This could happen due to trading inefficiencies, that is, it was unable to sell all of its excess production. Additionally, due to forecast errors, the participant might have sold less than it should have, or bought energy when it could have used its own generation. That is, $\gamma[n] = max(P[n] - S[n] - max(C[n] - B[n], 0), 0)$.*

Definition 3 *An Unnecessary buy $\beta[n]$ occurs when a prosumer, a participant with generation capacity, buys energy from the market in lieu of using its internal production. That is, $\beta[n] = max(B[n] - max(C[n] - P[n], 0), 0)$, if $P[n] > 0$. Unnecessary buys are caused exclusively by forecast errors.*

Definition 4 *An Unnecessary sell $\sigma[n]$ occurs when a prosumer sells energy to the market that could have been used to abate its internal demand. That is, $\sigma[n] = max(S[n] - max(P[n] - C[n], 0), 0)$, if $P[n] > 0$. Unnecessary sells are caused exclusively by forecast errors.*

The demand imbalance measures the amount of energy for which a participant would have to pay penalties. If it buys more energy from the market than it can use, this results in a broken contract. A "supply imbalance" can also be considered when a participant sells more than it can produce. However, because generation forecasting errors are not considered, the supply imbalance will always be zero in our case.

Uncapitalised generation measures the amount of energy a participant could not capitalize on. This happens either due to an inability to sell it on the market or through miscalculation given the demand forecasting errors, which resulted in it not selling as much as it could have. We make no assumptions as to what happens to this energy, if the participant ramps down its production to avoid possible imbalances, or if the energy is injected into the grid anyway. As such, it may or may not be penalized. In any case, it characterizes the opportunity cost of the prosumer given that it did not sell, or use the energy itself.

The unnecessary buys and sells measure the volume of erroneous trades on the market by the participant. The level of their impact is directly related to the transaction costs of the trades. For instance, in the case of an unnecessary sell, in certain circumstances it could make economical sense to sell the entire capacity on the market, rather than use it internally. This would only happen if the acquired revenue is greater than the costs and savings of using the energy. A similar point can be made about unnecessary buys.

4.1.2 *The Benefits of Forecast Accuracy*

The impact of group trading on the market is evaluated by comparing two cases: the *group case* and the *individual case*. The group case simulates trading on the market with a group. In the individual case evaluation, all participants of the same group trade individually. A probability of 60% of a participant having a PV installation is assigned in both cases. This penetration level was chosen as it was the highest level that displayed only slight levels of excess generation [99]. Hence, all of the generation can be used in the system, while any excess generation will only be a small component of the results. The group behaves like any other participant in the market; the only difference is that it trades based on its aggregated generation capacity, and on the prediction of the aggregated demand.

Participation of a group mixed of consumers and prosumers is simulated, as such their performance is evaluated in this experiment. Hereby a location-based selection is adopted, that is, the group is composed of geographically proximate participants, which can be seen as a small neighbourhood [69]. This group contains 183 participants with an average daily consumption of 1.5 MWh. Once probability of 60% of a participant having a PV installation is applied, within the 183 group participants, 108 are prosumers. Although the average individual MAPE for participants of the group is 48.53%, as a group they achieve MAPE of 10.59%. That is, the predictability of the group is nearly five times better than the individual average, resulting similar to results of other datasets [40]. For the investigated month of September, the total energy consumed by the group is approximately 38 MWh. If loads are predicted individually, the absolute prediction error results in 20 MWh (52%), while as a group it results in only 4 MWh (10.6%). The evaluation measurements are aggregated over all participants in both cases and can be observed on Figure 29. The figure indicates the performance increase when comparing the group performance and the aggregated performance of the group members in the individual case.

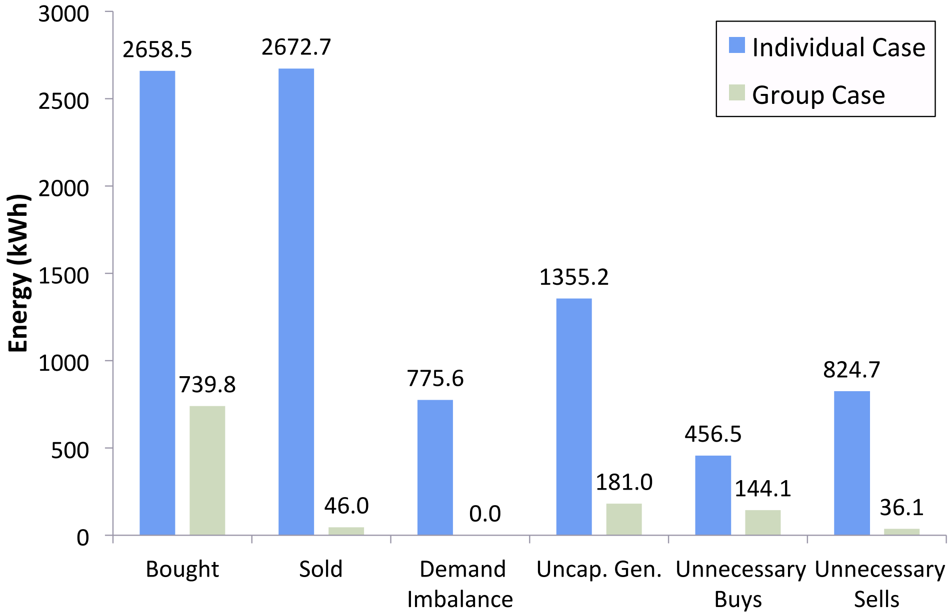


Figure 29.: A comparison of group performance vs. the aggregated performance of the individuals. The amount of energy bought and sold is also added for reference.

One can observe that the group performs far better than the aggregate of the individuals in the individual case. Obviously, this is due the improved forecast accuracy that reduced MAPE five times if same individuals act as a group. Furthermore, in a group scenario, a prosumer's surplus generation, which normally would have been placed on the market, is now shared between the members. Therefore, the decrease in uncapitalised generation (86%) and unnecessary sells (95%) is largely due to the drastic reduction in the amount of generation placed on the market. Most importantly, the increased forecasting accuracy has contributed significantly to the group reduction in unnecessary buys (68%) and in demand imbalance (100%), the latter being the major penalty component as it represents a broken contract. The performance improvement of the group also implies multi-party benefits. For instance, its future behaviour is better assessed, any penalties from erroneous behaviour are reduced in total, and depending on the cost mitigation policies of that group, this could imply enable more effective market participation for all participants.

4.1.3 *Group Trading Remarks*

A high degree of forecast accuracy by the participants will be required to ensure that participation makes financial sense [25]. It was shown that forecasting errors can lead to erroneous trading behaviour, creating uncapitalised generation and other opportunity costs and penalties. The results depict clear benefits from the overall reduction and cost sharing can lead to a more economically effective form of market participation. In a mixed group of 183 participants (including both consumers and prosumers) an overall reduction in uncapitalised generation, erroneous transactions, and imbalances was found (when compared to the aggregate performance of its individuals). These results were achieved through the use of a simple trading behaviour and forecasting methodology. This emphasizes that different scale of aggregations and more sophisticated methodologies can be applied for an effective participation, thus in chapter 4 this is further investigated.

4.2 GROUP-FORECASTING ACCURACY BEHAVIOUR

Behavioural patterns of electricity consumers may significantly differ in ease of being predictable [90]. If some customers are highly stochastic, their behaviour will propagate to the entire group of them, finally being distributed to their electricity bills. Simply, their behaviour can't be accurately predicted, that in some cases may take to extremely high balancing costs for the entire group. If one drills down to smaller groups of customers, or even individuals, the forecasting algorithms struggle even more [40], in particular when no additional information is provided [91]. For instance some appliances are highly predictable while others have completely stochastic behaviour. Consequently if a consumer is unpredictable, its incapable to benefits from added-value services of platforms, such as IEM [21], what was already presented in section 4.1.

The newly deployed Advanced Metering Infrastructure (AMI) may assist in this direction by providing a better insight on the individuals in the context of both timing and quality of information [35]. Accessibility to the smart metering data allow new dimensions in analysis e.g. predictability level assessment for any customer, or a group of them. It is important to understand that the focus does not necessarily have to be on customers as such, but virtually any grouping (based on certain criteria e.g. location, economic, social etc.) of devices or users being connected to the grid [42]. This section will investigate how predictability of prosumers improves by grouping them on much smaller scales than retailers do today. Such evaluation is possible due the accessibility of real-world smart metering data collected via AMI within the NOBEL project [37]. Finally, the investigation results of this section will lead to evaluation of the assumption that the overall performance of a group depends on predictability of its individuals.

4.2.1 *Smart Meter Grouping*

A traditional retailer's business and internal cost benefit analyses rely on the existence of large customer numbers, where individual effects are absorbed by the overall group behaviour. This is true due the mathematical behaviour of time series aggregation, especially if aggregated time series hold similar patterns. A smart meter is denoted with $m \in M$, where M is the total set of X smart meters. If n is an interval (e.g. 15 minutes), actual consumption of a m inside an interval n is denoted as $y_m[n] \geq 0$. The forecast energy load for the same interval is denoted as $\hat{y}_m[n] \geq 0$. The energy difference between forecast and actual consumption is calculated as $w_m[n] = \hat{y}_m[n] - y_m[n]$, having surplus if positive or shortage if negative.

Two types of aggregation are possible: One the one hand, if the prediction is calculated before the aggregation step, a perfect fit for aggregation of two meters a and b extracted from the set M, where $a \neq b$, if $w_a[n] - w_b[n] = 0$ or having no prediction error. A perfect example would be aggregating the forecast errors in shape of the $\sin(t)$ and its π shift $\sin(t + \pi)$ functions. On the other hand, energy of any meter $y_m[n]$ may be aggregated with any other $m \in M$ for each interval n. The resulting series can be further used for the calculation of the aggregated prediction. This step produces a subset denoted as $G \subseteq M$ of size $x \leq X$, where x represents the number of meters in the subset. The aggregation of any G for one instance results is denoted as $y^G[n] = \sum_{m \in G} y_m[n]$, that is used actually for calculating the prediction $\hat{y}^G[n]$ for consecutive l intervals.

Applied Approach

The approach used in this work can be characterized as some kind of brute-force method; the computational cycles are used to build random groups, create forecasts for these groups and measure the resulting forecast accuracy. The steps in the grouping approach rely on random numbers. The Monte Carlo method is used to build a group of randomly chosen smart meters from the original set. The probability, independently of a group size x, must be equally distributed in order to ensure comparability between all group sizes. Thus, all time series have the same probability to be chosen for a group.

For every experiment the series length l is fixed e.g. in this section $l = 96$ and represents exactly one day in 15 minute intervals. Still, every smart meter $m \in M$ contains time series $y[n]$, indexed with m, where $m \in [1, X]$. Once the size $x \leq X$ of the subset $G \subseteq M$ is determined, G gets populated by randomly drawn smart meters from M, without replacement. Finally (one or more) accuracy comparison measurements between the two time series, $y^G[n]$ and $\hat{y}^G[n]$, are stored as result of the experiment. This system is fully implemented as described in [68].

Forecasting Algorithms

Energy load forecasting is influenced by several factors, the most fundamental of which is the prediction horizon, hereby noted with l. The focus of this work is forecasting the next day load, categorized as the short-term forecast [13, 102] as its horizon is between one and seven days. Besides the forecast horizon, methods can be additionally categorized by considering seasonality. As smart meter energy readings are available, the time series forecasting methods, as they use only historical data of a variable for prediction [14], have been selected. The approach is to reveal the internal structure (e.g. seasonality, trend) by using statistical properties of the time series. Due to their robustness and implementation simplicity, time series forecasting methods are popular in short-term load forecasting. The most commonly used approaches are auto-regression [103] or exponential smoothing models [104].

For this work, the exponential smoothing forecasting method was chosen mainly for its robustness e.g. the method of HW. Exponential smoothing shows good forecast performance in empirical studies and outperforms more complex methods [105]. In order to compare experiments of forecast models, a naïve forecast method was used. Since energy load data is highly seasonal data, the SN algorithm was chosen. The principle behind the SN method is the usage of values from the previous season (e.g. day, week) as forecast value for the current season [104]. For example, the forecast value for Monday is equal to the last observed value for Monday.

Accuracy Measurements

To evaluated the forecast accuracy, the historical values (that were used to build the forecast model) compared against the predicted values. Therefore, the available historical observations are split into training and test sets. The historical series y^G are used as the training-set to fit the forecast model, later compared against its the predicted values $\hat{y}^G[n]$ for l intervals. As forecasts of different scales must be compared, the MAPE is chosen due its scale-independence. MAPE estimates the fit of a model by expressing its accuracy as a percentage, the advantage of which is that it is not fixed to a specific unit. Therefore, arbitrary models can be compared regardless of the unit of their values or their level. The MAPE is calculated as the sum of the absolute errors, normalized by the actual value [106] i.e.:

$$MAPE(G) = \frac{100\%}{|n|} \sum_{\forall n} \frac{|\hat{y}^G[n] - y^G[n]|}{|y^G[n]|}$$

where n is one interval and $|n|$ is cardinality of the discrete timeseries of the group G. The major disadvantage of this error metric is that the MAPE has no upper bound, as there is only a lower bound, which is zero. Due to this missing upper bound, extremely high values for certain time series distort the comparability of the MAPE. Especially for the case of a small denominator $y[n]$ the MAPE tends to infinity. However, this problem disappears already within the groups of few smart meters.

4.2.2 *Clustering For Accuracy*

The evaluation experiments reveal the grouping impact on the forecasting accuracy, and how the group accuracy depends on the accuracy of its individuals. An entire system of this evaluation is designed and implemented [68], in addition being feed with the real-world data from the NOBEL project [37] which runs a trial with Spanish consumers (`www.ict-nobel.eu`). This original data set is filtered in order to acquire a high number of smart meters without any invalid measurements. The resulting set had $X = 1974$ smart meters without missing, or faulty, meter readings from 03 March 2011 to 09 June 2011 (98 days in total). As metering data was collected within the project trial, it was discovered that set had 2.8% of the 15 minute resolution and 97.2% had the 1 hour resolution of the metering data of 1 kWh precision. In order to keep unique resolution, devices with the resolution of 1 hour were linearly interpolated to 15 minutes.

The process described in the section 4.2.1 is repeated using this set. In the case of the HW method, the chosen seasonality was within-week seasonality. Using a weekly season achieved the best forecast accuracy in preliminary experiments, which is also reported in [107]. For the SN forecast, a within-day seasonality was used, which means that the observations of the last day are the predicted values for the next day. This configuration depicted a superior forecast accuracy in preliminary experiments. Finally both algorithms use historical data to predict a specific date i.e. Tuesday on 7 June 2011.

Grouping Impact on Accuracy

Grouping hides the stochastic behaviours and their impact. However, today with the fine-grained smart metering data offered, one can make detailed analysis on the impact of such individual stochastic behaviours on the overall accuracy. To understand their affect, an experiment was executed 100 times for every x in the spectrum $x = [1, 180]$, which was split into 4 sub-intervals. The first subinterval was $x = [1, 25]$ by an incremental step of 1 (25 groups steps in total); the second $x = (25, 50]$ by a step of 5; the third $x = (50, 100]$ by a step of 10; and the fourth $x = (100, 180]$ by a step of 20. Figure 30 shows the result of the experiment where the average MAPE per group size is shown for both HW and SN algorithms.

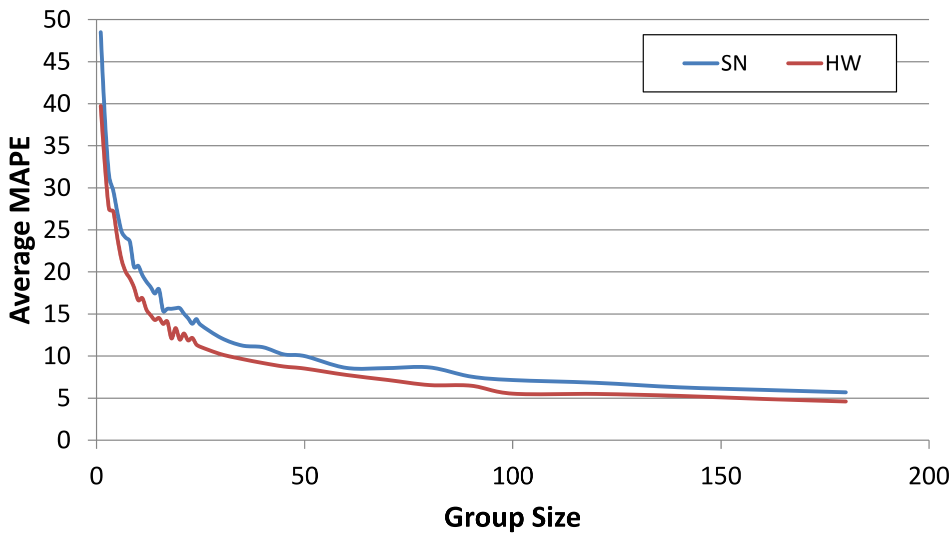


Figure 30.: Grouping effect on the prediction accuracy (MAPE)

As expected the accuracy increased with greater group sizes, showing a higher improvement rate for smaller group sizes in comparison to bigger ones. Interestingly, the results also revealed that the simple forecast method (SN) performs almost identically to the more complex one i.e. HW. However HW depicted a slightly better accuracy for all group sizes, having $\langle MAPE(G^{160})\rangle_M < 5\%$ already at the group size of $x = 160$. Further experiments conducted revealed that a lower variance of series y^G $\forall G$ is the reason of the accuracy improvement rate. The same experiment within the winter season, where variance of the meter readings is higher, depicted slightly lower improvement rate. This experiment resulted in $\langle MAPE(G^{160})\rangle_M \approx 8\%$ for the HW method. Still, both experiments resulted in a significant convergence rate. One can observe that proper parameterization of algorithms [14], and application of storage technologies [69], can lead a small group of customers to accuracies affecting the retailers of today [39].

Achieved results triggered further investigation for the competitiveness of the two selected algorithms. In order to validate the resulting behaviour, additional experiments were conducted. As shown in Figure 30, it was decided to fix x for comparison within a rolling-time window (other days of the week) to cross validate. Group size of $x = 50$ was chosen as greater x resulted in slighter accuracy improvement. Thus, one-day ahead was predicted 100 times (and therefore with different G) for every day of the week. The results for Tuesday 07 June 2011 are depicted in the Figure 30; however now a rolling-time window allow us to move along the rest of the weekdays, the results of which are shown in Figure 31.

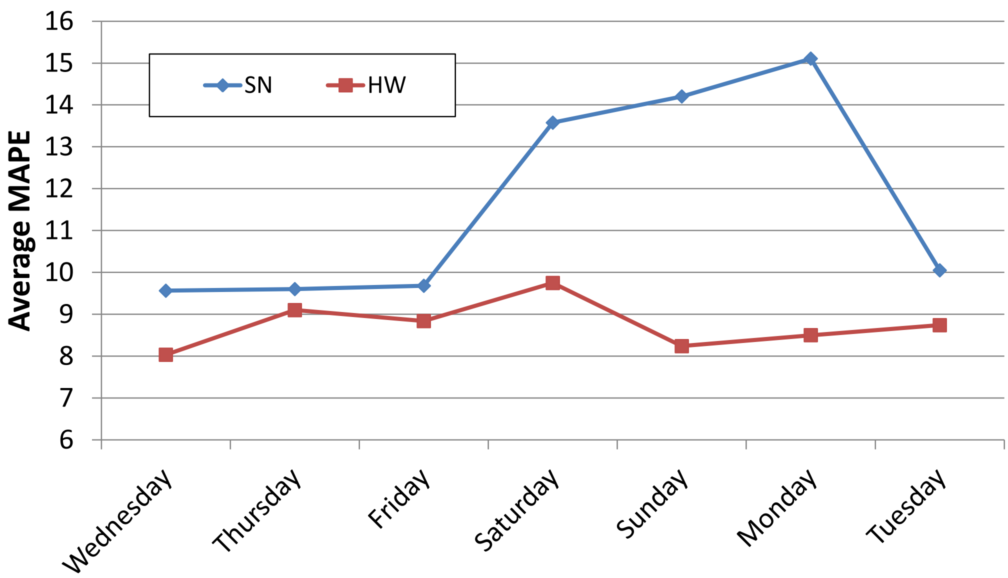


Figure 31.: Example where SN algorithm fails to predict next day ($x = 50$)

The results depicted in Figure 31 show that the HW algorithm performed better for all days, as its $\langle MAPE(G^{50})\rangle_M$ averages in between range of $8 - 10\%$. Although SN was comparable to HW from Tuesday to Friday, its forecast accuracy degraded for Saturday, Sunday and Monday. Such behaviour was expected due the fact that energy data actually contains two seasons, daily and weekly. Since SN predicts day-ahead, e.g. future 96 instances of 15 minutes, using data of one day-before, i.e. previous 96 instances, one can expect that customers (residential or commercial) behave differently on Saturdays in comparison to Fridays. As an example, load characteristics of a commercial customers usually change drastically over non-working weekend days.

The results of this experiment leads to the conclusion, that the forecast accuracy improvement by grouping is not a random effect. However, it is remarkable that the SN algorithm performed almost as good as HW for all the other weekdays (for this data set). Since the HW method depicted greater accuracy it was selected for identifying the key accuracy indicators.

Key Accuracy Indicators

It shown that the forecast accuracy improvement by grouping is not a random effect and may be represented as a function [40]. It was shown that the SN algorithm failed to predict correctly for Saturday, Sunday and Monday, even when they were aggregated within a group of 50. Obviously many devices within such group in 100 runs resulted in higher MAPE for Saturday than for Tuesday. However, one can generally expect that grouping impact is improved if every individual has a good prediction (or lower $MAPE(G^1)$) on its own. In other words, grouping two predictable smart meters will result into a lower MAPE than two unpredictable ones.

To confirm this assumption, an experiment is conducted where the MAPE for every smart meter is calculated individually using 4-weeks of historical data to predict Tuesday (07 June 2011). Figure 32 shows the cumulative density function of the HW method in dependency of the resulting MAPE values. The median of this data set resulted to a MAPE of 36.06% and is assumed to be an indicator for creation of groups with greater and smaller forecast accuracy. If the hypothesis holds true, one would be able to create "good" and "bad" groups from the individual prediction accuracy. Using the median value from Figure 32, the time series with MAPE lower than 36.06% were considered as devices in the "good" predictability set ($A \subset M$), while the remaining were considered as devices for the "bad" predictability set ($B \subset M$).

To confirm the hypothesis, the same experiments from section 4.2.2 were conducted for both sets (A and B) individually. Figure 33 shows progress of the original set M and the derived sets. Noticeably, both sets followed the exponential improvement of the original set with a slight offset. However, the results of this experiment revealed that creating groups out of good/bad time series, based on the individual forecast accuracy, proved the prior hypothesis. The set A showed greater accuracy ($\langle MAPE(G^{160}) \rangle_A \approx 4,09\%$) than the set B ($\langle MAPE(G^{160}) \rangle_B \approx 5,93\%$), concluding that the hypothesis was correct if empirical results of the depicted experiment are considered.

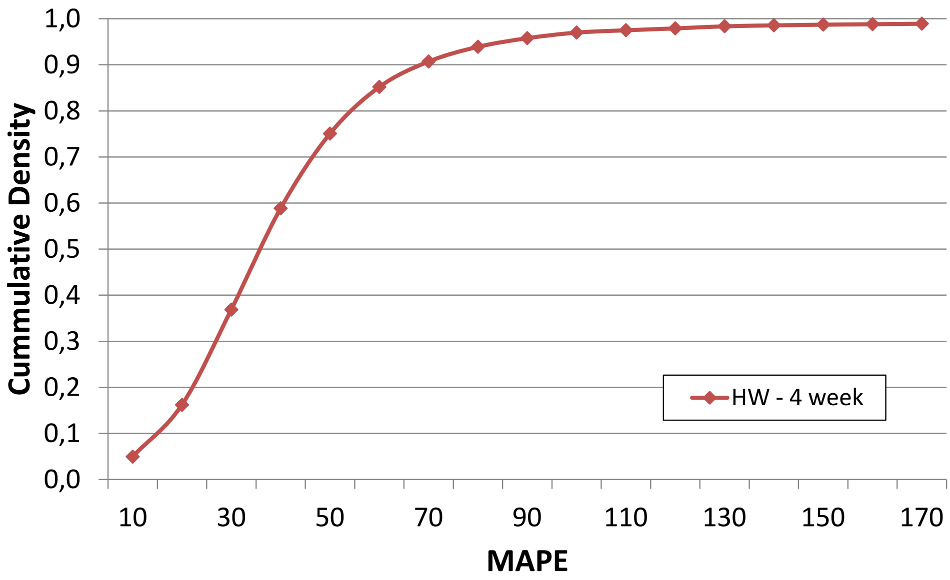


Figure 32.: Cumulative density function built of individual predictability from every device in the set

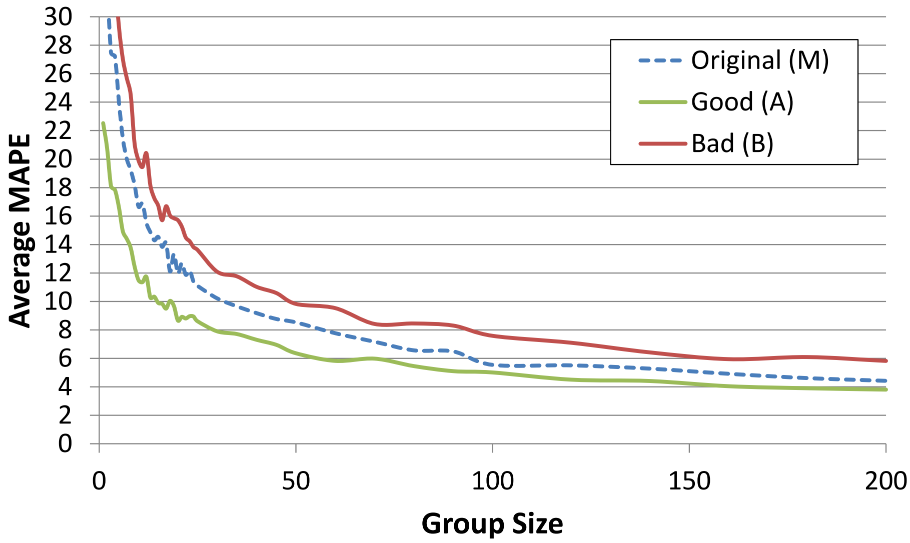


Figure 33.: Impact of grouping using good(A) and bad(B) predictable sets

Summary

Two different forecast methods were used to demonstrate and to compare the impact on the forecast accuracy by time series aggregation. Interestingly, results showed competitiveness between simple and robust algorithms. The robust algorithm HW (requiring 4 weeks of historical data) used, slightly over-performed the simple one SN (that requires only the previous day). The HW method was further used to show that groups with individuals of higher and lower predictability affect the overall predictability of a cluster. This assumption was also proved to be true independently of the used forecast algorithm.

4.3 Improving by Storage Technologies

For stakeholders to benefit from Smart Grid services envisioned in section 3.2, such as active participation in a smart city [27] or neighbourhood marketplace trading from section 4.1, consistently accurate prediction is identified as a business advantage [25]. Still, any forecasting algorithm would struggle to consistently meet the high-precision for the load-forecasting of an individual (e.g. a household). As such, forecast accuracy becomes a significant factor for the realization of prosumer Virtual Power Plants (pVPPs) from highly distributed resources [108]. The results from section 4.2 showed that clustering enables a significant convergence rate in prediction accuracy, converging even for a relatively small number of prosumers (e.g., 100 households) within a cluster. Although the prediction accuracy by clustering converges [68], it may not be sufficient to achieve accuracy required for to consuming the future Smart Grid services.

If a cluster of prosumers decide to create a pVPP, all the grid imbalances (as result of the prediction errors) occurring within the grid will lead to financial penalties [28]. In order to avoid that, all the electricity injected into– or extracted from– the electricity grid by a cluster needs to be highly predictable to become an adequate resource on the grid [77]. TO gain on predictability, significance of storage solutions was identified [52], in particularly, their importance to further improve the predictability for highly volatile RES [92, 109]. In same fashion, instead of focusing to improve forecasting algorithms as such, this section investigates advantages of a potential storage unit availability in order to improve forecast accuracy of a pVPP.

The section will focus on investigating how storage sizing impacts the predictability and affects its behaviour within a cluster. It will be shown that a significant second grade convergence, beyond the first grade by clustering [68], can be achieved with the storage mechanisms. These results are not focused on any specific aspects on the nature of the storage units, which could be composed of multiple heterogeneous resources, such as fleets of electric cars [32], residential storage, small industry storage (e.g. supermarket refrigeration units) and generally any kind of facility that could act like so for power systems [70].

4.3.1 *Forecast Accuracy in a Cluster*

Due simplicity and good results demonstrated by the SN algorithm, in comparison to be more sophisticated Holt-Winters exponential smoothing algorithm in section 4.2.2, its daily seasoning was also chosen for this section. The trial data from section 4.2 is used here as well, but on longer time frame such that SOC of a storage is kept over the entire frame. The original data set is filtered in order to acquire a high number of smart meters with the highest number of measurements (more than 50% of samples). This set M resulted with $X = 3564$ smart meters without sampling interruption from 15-Aug-2012 to 15-Sep-2012 (31 days in total). For daily SN forecasting, their individual MAPE averaged at 47.35%, what is slightly higher than the 42% resulted in evaluations with more robust algorithms [40]. Therefore, even from observing Figure 32, one can conclude that only few smart meters behaved as "predictable" individuals.

In order this poor individual predictability, different cluster sizes were analysed by the same approach from section 4.2. Aggregation is done by randomly selecting x smart meters from M into a cluster $G \subseteq M$, without repetitions. The time-series data for each smart meter is then aggregated to produce a single time-series (for measures of 15 minutes intervals) for the cluster. The experiment is repeated 50 times for every cluster of size x, denoted as G^x, with an incremental step of 20 for the spectrum $x = [20, 400]$. The box plot in Figure 34 shows MAPE and variance of M that resulted after applying the SN algorithm to clusters of different scales. The figure confirms that aggregation level increase brings rapidly M to its saturation from the SN algorithm, what was also noted for robust algorithms [40], while stochastic behaviour of the individuals diminishes in bigger clusters. Slightly higher values than Figure 30 are not a surprise, since the experiment results include 31 consecutive days, while Figure 31 depicts how SN algorithm introduces fails for three days per week (Saturday, Sunday, Monday). However, MAPE values are not critical for evaluation in the following sections and it was shown that can be improved with more sophisticated algorithms [68]. One can even make a comparison against other datasets, what is not doable only via number of smart meters, thus the power based observation is proposed in [69].

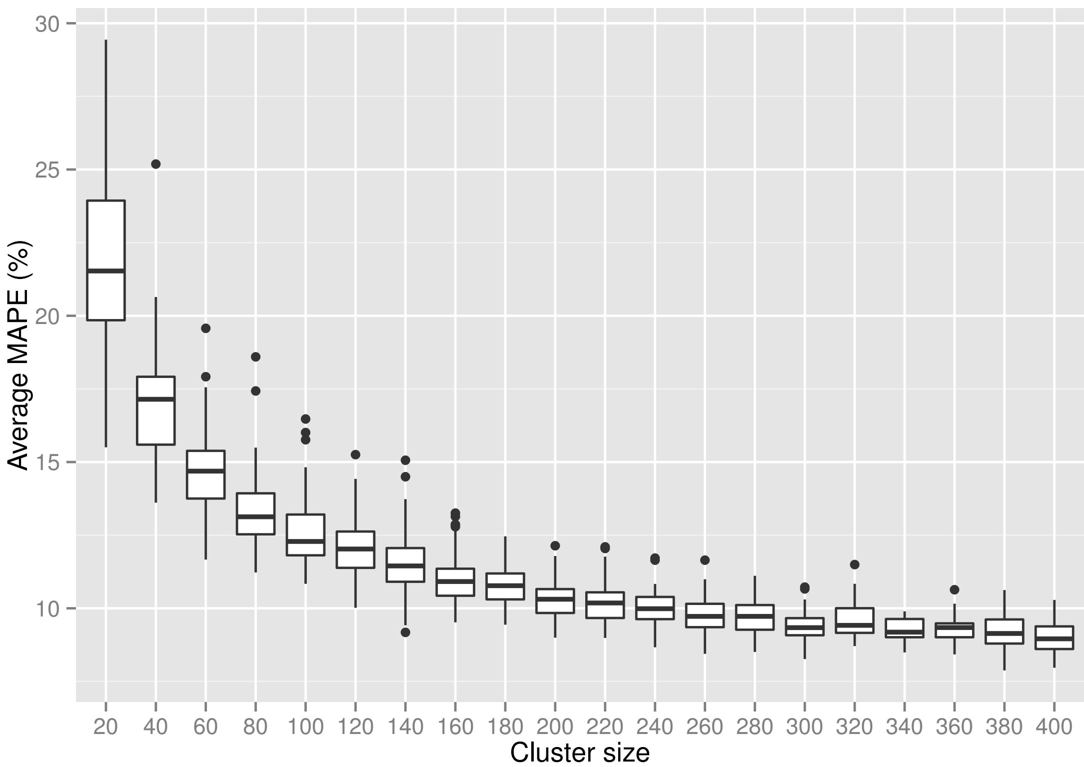


Figure 34.: Clustering effect on the prediction accuracy (MAPE)

4.3.2 *Effects of Capacity Available*

Intermittency of RES also imply the usage of different techniques to maintain their predictability e.g. wind farms can brake turbines from spinning if power exceeds a generation set-point. However, due the wind unpredictability their set-point needs to be set low, where predictability is high enough (since turbines can be stopped). Once turbine starts braking, until fully stopped, its potential to produce energy is wasted. For that reason, wind farms try to improve their business by applying different types of storage [10]. Observation of these real-world cases improving their business by applying storage technologies, the same ideology can be applied to create the predictable clusters. The forecast improvement by clustering from Figure 34 is expected to be further improved by adding a storage unit. In this section a non-variable storage unit, with its minimum E_{min} and maximum E_{max} charge limits, is simulated for different capacities to investigate the improvement potential. These experiments do not consider storage efficiency or controlling. Therefore, the actual energy load E_l, positive or negative, applied to the storage unit at an interval n is described as

$$E_l[n] = \begin{cases} 0, & \text{if } w[n] > 0 \ \& \ E_{SOC}[n+1] = E_{max} \\ 0, & \text{if } w[n] < 0 \ \& \ E_{SOC}[n+1] = E_{min} \\ E_{max} - E_{SOC}[n+1], & \text{if } w[n] \geq E_{max} - E_{SOC}[n+1] \\ E_{min} - E_{SOC}[n+1], & \text{if } w[n] \leq E_{SOC}[n+1] - E_{min} \\ w[n], & \text{otherwise,} \end{cases}$$

where $E_{SOC}[n+1]$ is its state of charge (in Wh) before an interval n. The described component is then used to carry out multiple simulations and measure the impact of different storage sizes on reduction the forecast errors w. In other words, as E_{SOC} has cumulative characteristics, every interval simulation of a cluster where $E_i[n] = |w[n]| - |E_l[n]|$ is considered as load imbalance for every $E_i[n] \neq 0$. Every simulation is repeated 50 times for every cluster size x and every storage capacity c, where storage capacity is calculated individually for each cluster G as

$$c(G, s) = \langle P^G \rangle * 24 * s,$$

where s is used for defining percentage of cluster's average daily energy usage (thus 24 hours) and $\langle P^G \rangle$ is the average power usage of the group. Finally, a cluster of size x and its storage sizing s is denoted as G_s^x. Figure 35 visualize the results of simulations for storage sizing $s = [0\%, 16\%]$ in power of 2, to show how MAPE is affected by both improvement methods.

The simulation results confirm that the storage presence resulted in faster convergence rate and further MAPE reduction. For $s > 0\%$ we see that smaller storage has a much higher impact on the predictability improvement, than bigger ones. As an example, for $x = 200$ there is a noticeable average MAPE reduction for first 2% increase by

$$MAPE(G_0^{200}) - MAPE(G_2^{200}) = 3.5\%,$$

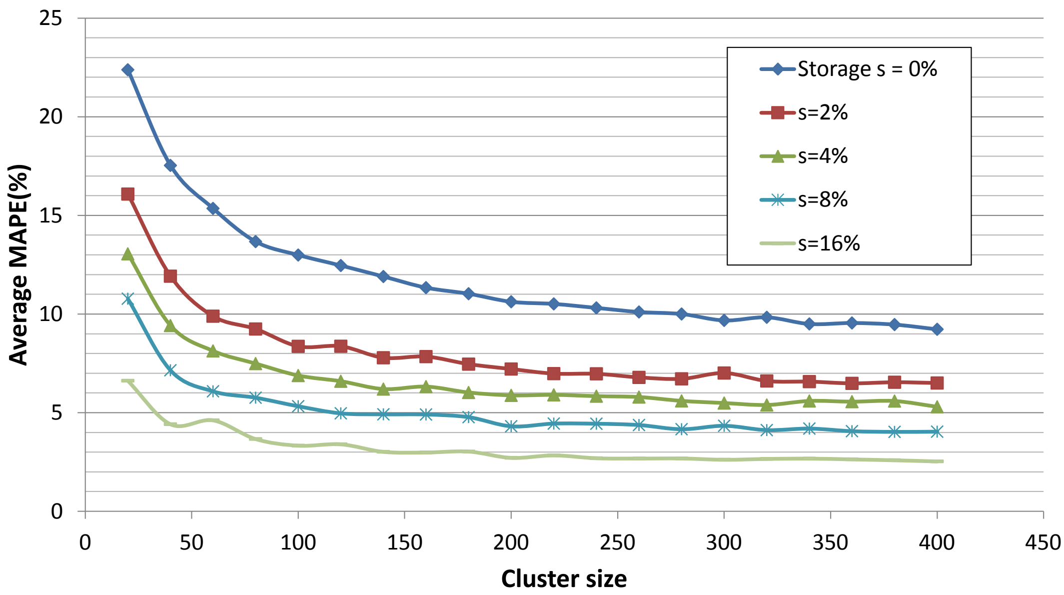


Figure 35.: Measuring impact of a storage to improve prediction accuracy of prosumer clusters

than the second increase

$$MAPE(G_2^{200}) - MAPE(G_4^{200}) = 1{,}3\%.$$

Every increase in capacity follows the same behaviour, until capacity increase seems to be of minor importance e.g., average MAPE improved only for

$$MAPE(G_{14}^{300}) - MAPE(G_{16}^{300}) = 0.25\%.$$

Finally, results confirm that a small cluster with a storage component can have the forecast accuracy of an approx. 4 times bigger cluster (without storage). For example, clusters of size $x = 60$ with $s = 2\%$ are expected to average around $MAPE(G_2^{60}) = 9{,}89\%$, while clusters of $x = 300$ with no storage average around $MAPE(G_0^{300}) = 9.67\%$. Such a small difference (0.22%) is only possible due the fact that predictability converges much faster with storage increase, while same difference without storage equals to

$$MAPE(G_0^{60}) - MAPE(G_0^{300}) = 5.67\%.$$

Understanding how different s affects the imbalance E_i of a cluster is crucial towards realization of pVPPs. Simulations of different storage sizing will lead to evaluating quantitatively the impact on the imbalance reduction, which is later relevant for an economic analysis. Still, the increase of s also has its saturation point, where further storage expansion would not make any significant improvement. Figure 35 shows how improvement for storage sizes close to $s = 16\%$ progress almost in parallel. It is important to mention that prediction algorithms play a key role here. The simplicity of the SN algorithm can assist towards understanding the importance of storage and one should expect even faster saturation if more robust algorithms are used, or if storage controlling is applied [64]. Hence the economic significance of the storage needs to be evaluated for every case individually, e.g. a neighbourhood of interest, to understand how s parameter affects its individual imbalances E_i.

4.3.3 *A Real World Use Case*

Although the pVPP concept may be appealing, it is challenging to have them as an adequate resource [77]. Furthermore, not all the clusters are equally predictable and selection of a reliable one is highly relevant for its overall forecast accuracy [68], as resulted in Figure 33.

Beside the individual predictability, one needs to understand capability of a storage unit to absorb cluster's forecast errors. To demonstrate that, a cluster of the real electrical grid prosumers is selected to represent the pVPP; in Figure 36 the polygon defines the cluster. Hereby constellation of a cluster is to include prosumers with a physical proximity to a storage unit.

Figure 36.: Clustering prosumers in a GIS-aware system by physical proximity

The selected cluster contains 186 smart meters with individual average power consumption of $\langle P \rangle = 0.34$ kW and their individual MAPE with SN is measured to be average to 49.75%. However, as a cluster the average power consumption of $\langle P^G \rangle = 63.74$ kW and its MAPE is measured to be 11.43%. The predictability of this cluster is almost 5 times better than their individual average. If measured in kWh for the same t_1 and t_2 defined in section 4.3.1, the total energy consumed from the cluster is approximately 47.4 MWh. If loads are predicted individually, the absolute prediction error results in $23.5 MWh$ (49.6%), while the cluster resulted with the error of only $E_e^G(t_1, t_2) = 5.6$ MWh (11.8%). The follow-up analysis will evaluate the impact of absorbing the forecast errors in a storage unit within the cluster.

Since the Monte Carlo approach was used in previous experiments no information on achieved imbalance reduction of a cluster is shown by enhancing it with different sizes of the storage unit. Once G is fixed, as in Figure 36, one can demonstrates not only how much reduction of energy imbalances due prediction error is achieved, but also the progress of the cluster's MAPE by expanding storage capacity. Figure 37 shows that linear increase in the storage size by the s parameter, is not followed by linear improvement of the MAPE for the selected period. Without having linear improvement of MAPE, reduction in energy deviations are expected to behave the similarly. For comparison, 2% of the cluster's daily average energy consumption is equals battery capacity of 1 or 2 EVs.

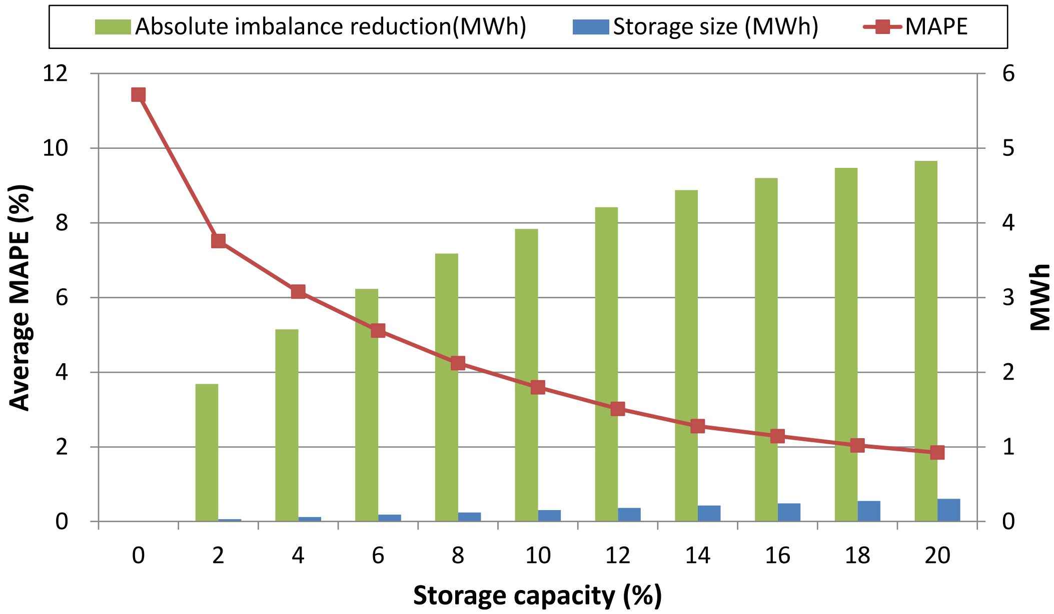


Figure 37.: Capacity increase is reducing the cluster prediction error

However, what is not easily measurable is the performance of a forecasting algorithm used for the simulation evaluation. As one can imagine, the better the algorithm, the less storage for its imbalances is required. Still energy loads in general have high correlation to the activity within the cluster, thus significance of storage availability may vary, even on intraday level [43]. Such forecasting errors need to be carefully understood in the context of the storage requirements and how they vary from interval to interval, what section 4.4 will further assess.

4.4 VARIATION OF INTRADAY STORAGE REQUIREMENT

Storage already plays a key role in future energy management scenarios [110], and limitations are already well known [70]. In this section a simple method to estimate storage capacity required to absorption the induced errors is proposed. In fact, the load behaviour of stakeholders and accuracy of their forecast plays a pivotal role in the required storage capacity that would be needed to balance the excess or shortage of energy over time. Figure 38 shows a difference in energy consumption of a commercial building, which suggests that same forecast accuracy in peak times, will simply need more storage capacity to absorb the errors. Using an on-premise available energy storage could help towards enhancing stakeholder's predictability, such that benefits as of the "resource adequacy" can be achieved [77].

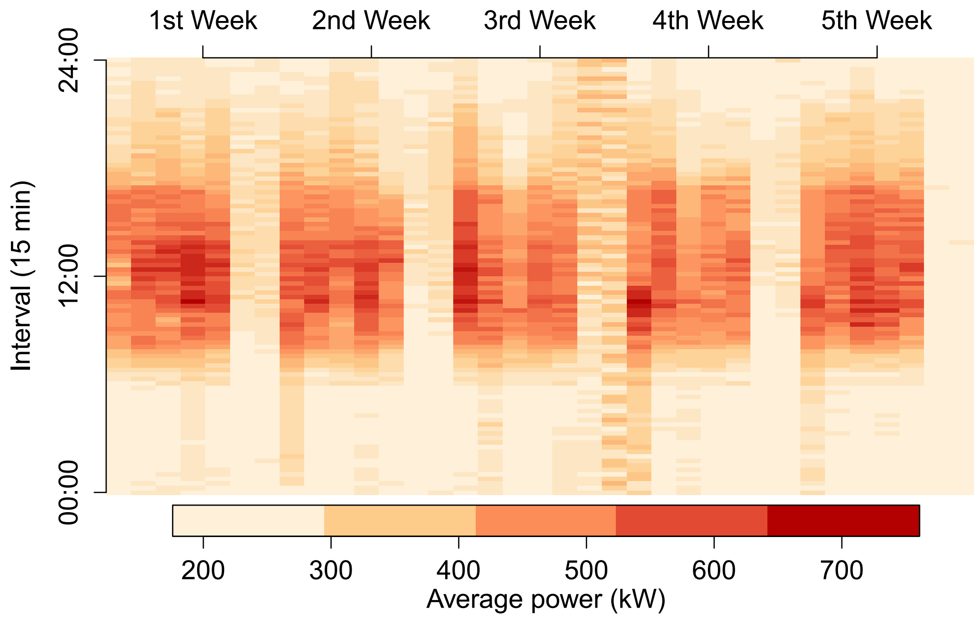


Figure 38.: Heatmap of energy consumption of a commercial stakeholder

Although this, and previous sections, doesn't really consider any specific storage technology, one needs to have in mind that many battery-based approaches are still expensive for most grid storage applications [65]. Hence this section stays at high level with respect to the actual storage characteristics, and benefit that the approach proposed can also be applied with existing assets, such as a fleet of company electric cars. The interest on the latter is that (i) will be available in high numbers in the future and are still company-owned assets whose management can be realized together with their facility goals, and (ii) they offer a variable storage capability depending on their availability and usage patterns. With that in mind, if forecast errors can be absorbed locally and if capacity needed to absorb them will significantly vary within the intraday intervals, one need to understand importance of those intervals.

Hereby the approach is empirical; it depends on smart metering data of a commercial building applied in a step-by-step way. The main aim is to understand how important the storage capacity available within the highly volatile intraday energy loads. Once the impact is understood, one may use such knowledge to propose "solution shapes" i.e. timeline of the quantity of storage needed to improve prediction of stakeholders. The evaluation will show how where different shaping of the storage can significantly reduce the overall capacity needed. The actual impacts of a specific technology as well as other side-effects on business or financial aspects are left for future work.

4.4.1 *Intraday Forecast Accuracy*

The wide availability of smart metering accompanied by the Smart Grid, enabled new approaches to analysis of generation [17], even down to individual customers [53]. The smart metering data is collected usually at a constant sampling period T, thus samples of energy consumed are represented as a discrete-time signal $y[n] \geq 0$, where n is an integer. Figure 38 depicts smart metering data sampled at $T = 15$ minutes of a commercial stakeholder. Although only 5 weeks are shown, this is a representative pattern as the variations in consumption repeat continuously through the entire year. The difference in consumption over days, led us the set split into working and nonworking days (including holidays).

If observed on the total yearly consumption, workdays resulted four times greater than the nonworking days. This is important to notice, due to the potential of improving stakeholder's predictability since storage sizing required can drastically vary in between two day types. To better understand this difference, a view of the average daily 15-minute intervals over the two created datasets is shown on Figure 39. The curve depicted shows the impact of building's processes for preparation of the workday, and the impact of employees arriving at the office as well as the actual usage of it during office hours. The drop seen is also expected after the leave of employees from the facilities and conclusion of other tasks (cleaning, maintenance etc.) which lead to an almost minimum operational level after approximately 21:00, which is where the two curves converge. The observation on the intraday load difference between the two day types indicates that the capacity required to address errors may vary.

Quantitative Observation of Forecast Errors

In this work, the consumption self-forecast is done only for a short-term horizon, for one day. Many forecast methods could be applied to the time series data produced by a smart meter [111, 103] and other indicators can further improve the forecast [91]. If an interval forecast is $\widehat{y}[n] \geq 0$, the total forecast error of that interval is $w[n] = \widehat{y}[n] - y[n]$. Since that error will be accommodated within an available storage capacity, the forecast errors will be observed quantitatively (kWh) rather than by the scale independent MAPE. This is important to understand as observing some measures by percentage, e.g. as temperature, is meaningless [106]. If X is a set of index intervals of interest, e.g. first interval of every work day in 2011, one can measure its average interval error for any

$$W[n]_X = \begin{cases} \frac{1}{|X|} \sum\limits_{k \in X} w[k-n], & 0 \leq n < \frac{\Delta}{T} \\ 0, & \text{otherwise,} \end{cases}$$

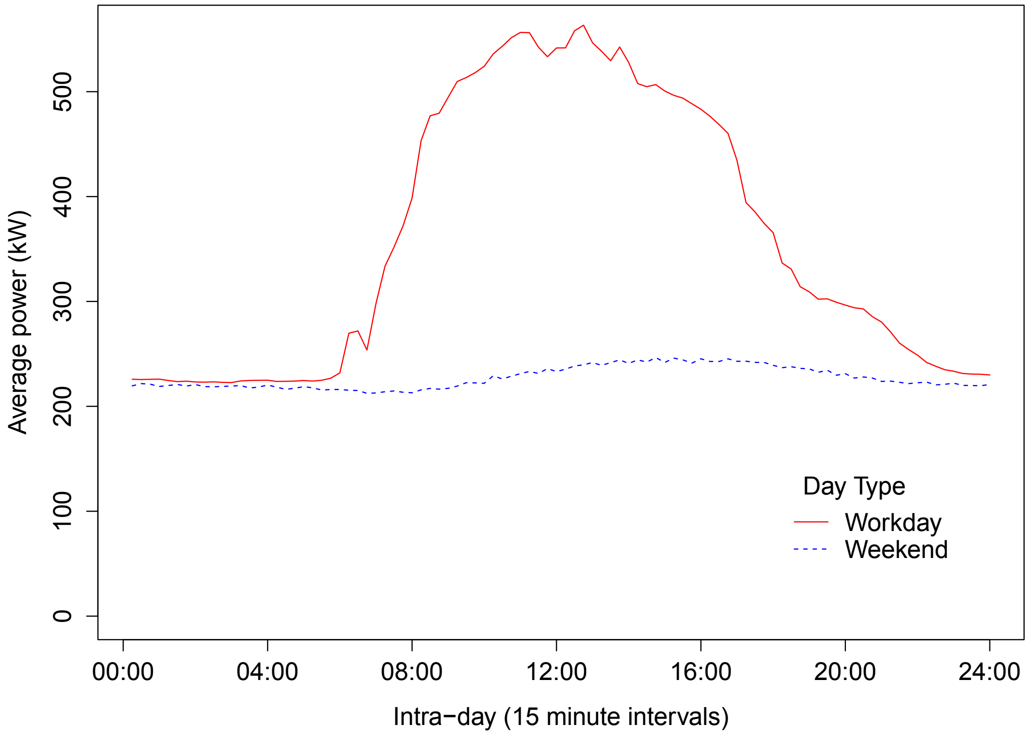


Figure 39.: Average power (kW) per interval averaged over entire year

where $|X|$ is cardinality, Δ is the season length and k represents the element of the set. Since the forecast error vary on the intraday basis, it is expected to have variations in required storage capacity to address these.

Intraday Errors

The empirical part of this work uses real-world data; hence the forecasting is done via stakeholder's smart metering data. Figure 38 depicts an example of the load produced by this commercial building with offices with 139 working places and its resulting consumption in 2011 was 2.7 GWh. As it can be seen, the building is mainly used in between 08:00–17:00 and there is a significant difference in energy load for different days of the week. The average daily power (over entire the year) approximates to 342 kW and 210 kW for working and nonworking days respectively. As such, the interval set X is divided into sets of first indices for all working and nonworking days, X_w and X_n respectively.

For the self-forecast the Seasonal AutoRegressive Integrated Moving Average (SARIMA) model was selected, as it can be used to relatively accurately predict electricity demand [112]. A forecast for next day is done on weekly seasonality and the model is trained with 4 seasons (28 days). The model training is made only with the samples known from 4 seasons up to a forecasted day and the model parameters are extracted from the same set using the off-the-shelf "auto.arima" method offered by the ARIMA libraries in the forecast package of R (`www.r-project.org`). Observations are made for average daily 15-minute intervals of $\Delta = 1$ day, and Figure 40 depicts the resulting functions of $W[n]_{X_w}$ and $W[n]_{X_n}$.

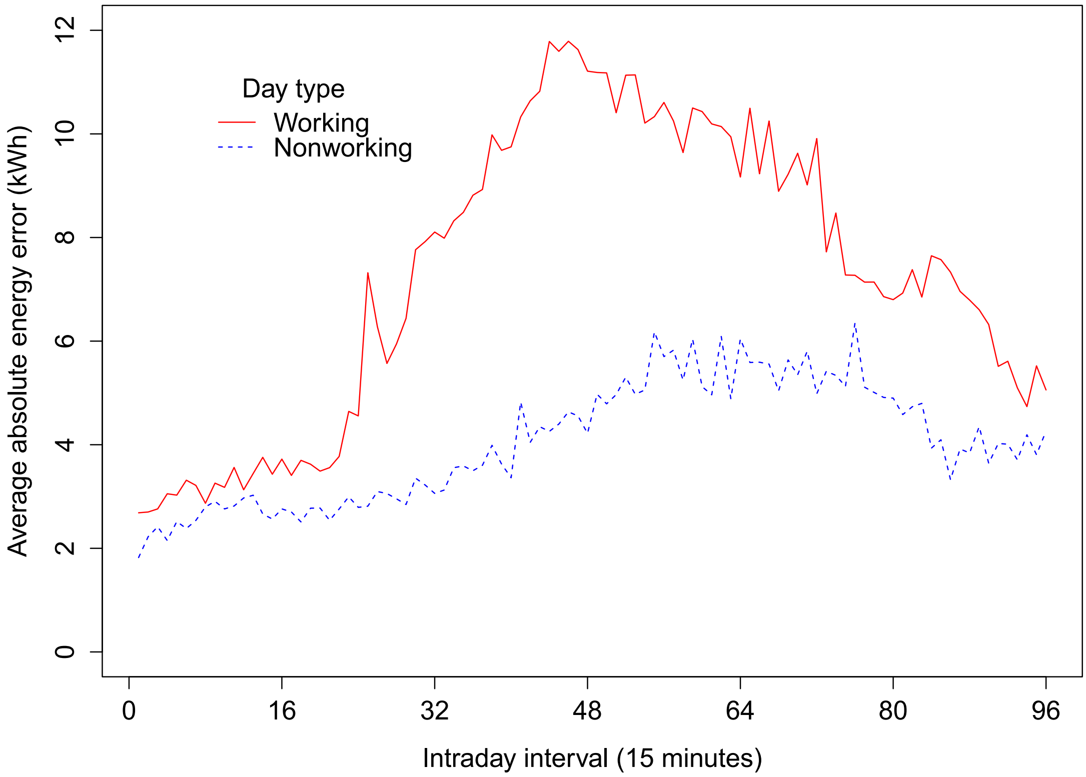


Figure 40.: Absolute forecast error averaged over intraday intervals to identify significance of their forecast error contribution

On a daily average, workdays resulted with $\text{MAPE}_{X_w} = 8.55\%$ and nonworking days with $\text{MAPE}_{X_n} = 7.56\%$. Due to the actual consumption difference for X_w and X_n, these forecast errors result in average daily error of 702 kWh and 388 kWh respectively. As Figure 40 indicates, quantitatively (and not by percentage) the errors differ significantly for the same intraday intervals. For lower values of n, working $W[n]_{X_w}$ and nonworking days $W[n]_{X_n}$ have a comparable forecast error, however the error of workdays around midday increases significantly. Although the real cause for the error is hard to pinpoint, it appears to be highly correlated with the working hours. Hence, one may expect that also other commercial stakeholders may experience a similar correlation to errors of their self-forecasts.

4.4.2 *Estimation Method*

Uncertainty of the prediction algorithm from section 4.4.1, affects the propagation of the forecast error to the storage capacity demand and need to be considered for dynamic sizing. Others identified similar behaviour [95], however some stakeholders may even have intraday variation requirements. For estimation the resulting discrete time series $w[n]$ is used, where positive value indicates a surplus and negative a shortage. Storage size is then estimated via the cumulative function from all intervals $n \in [1, l]$, extracting its extremes as the indicators. Figure 41 represents an example of estimating the storage capacity c_e as from the sum of $w[n]$ for one day. Mathematically, if a set of cumulative forecast errors is presented as $w_{cum} = \{w[1], \sum_{n=1}^{2} w[n], \ldots, \sum_{n=1}^{l} w[n]\}$ then the estimated storage size is calculated as $c_e = \max(w_{cum}) - \min(w_{cum})$. It is important to mention that this method returns the optimal sizing for each day individually, where method automatically sets their initial state of charge to the optimal position.

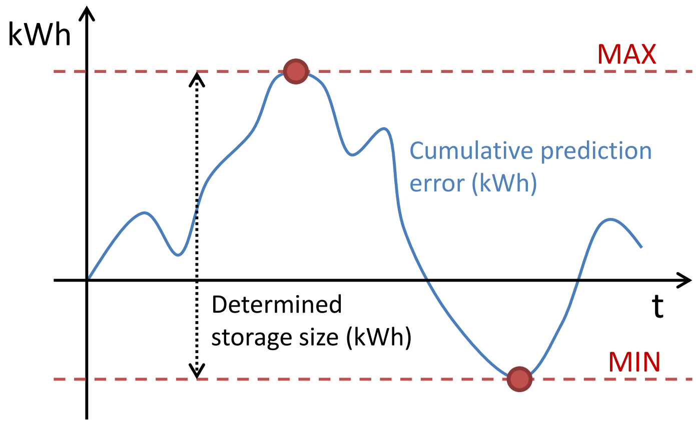


Figure 41.: Determination of an ideal storage capacity to address errors of a daily prediction

4.4.3 *Estimating Capacity Required*

The storage estimation method proposed in section 4.4.2 is hereby applied to the stakeholder data from section 4.4.1. Its metering samples are kept in same resolution, having a day represented as $l = 96$ of 15 minute intraday intervals. Same is valid for the resulting forecast, what was already discussed in section 4.4.1. Figure 42 depicts results of the estimated storage sizes (and their average), for each day from the two considered datasets. There, one can see the variation in storage needs imposed by the weekdays. As Figure 39 indicates, the nonworking dataset X_n had clearly lower consumption and understandably resulted in a lower storage requirement c_e. The average estimated storage for workdays is approx. $\langle c_e \rangle_{X_w} \approx 475$ kWh, requiring almost double the storage in comparison to $\langle c_e \rangle_{X_n} \approx 305$ kWh. For the figure, one can clearly see that the average required storage capacity $\langle c_e \rangle_{X_w}$ is covering approximately 85% of the estimated storage size $\langle c_e \rangle_{X_n}$. These results are considered as the second clue of the needed for considering a variable capacity.

For comparison to other cases, these estimations should be observed via daily energy consumed, such that one can understand sizing of its storage to improve the predictability. If every day is observed individually, their estimated storage size is covered with only 6.4% and 5.5%, for working and nonworking days respectively. Interestingly most of the forecast error is covered with storage size

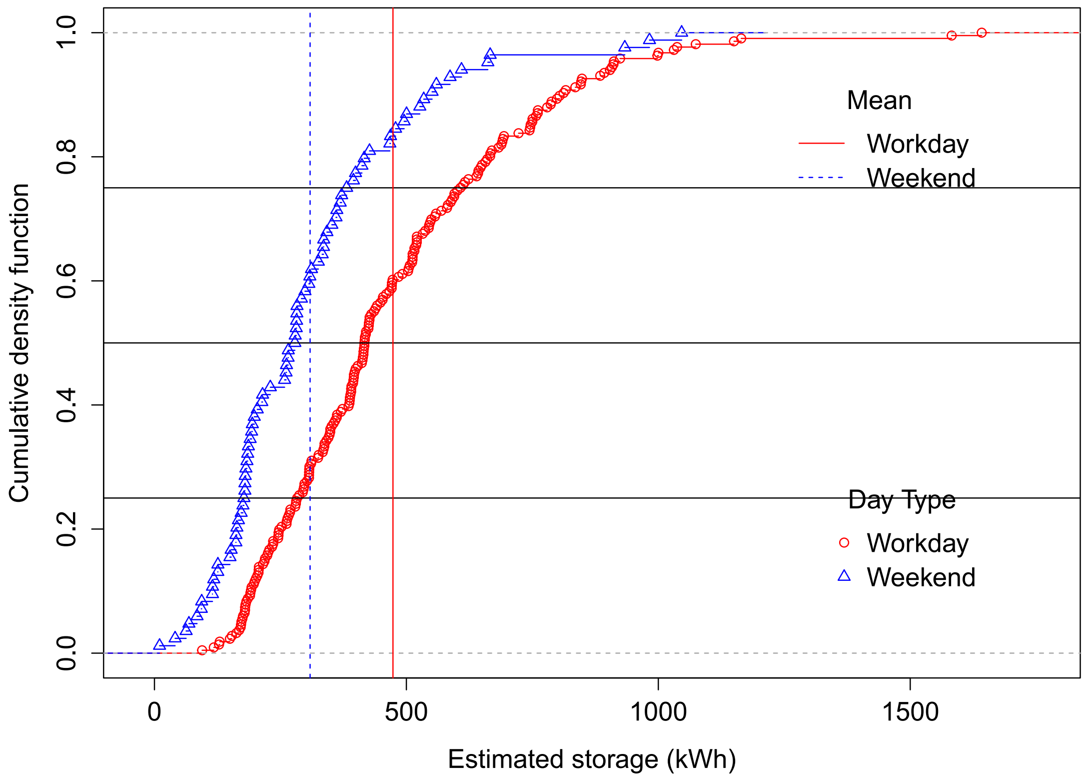


Figure 42.: Distribution of the estimated storage from daily estimations

of 10% of the individual daily consumption. Approximately 92% of the forecast error in X_n can be covered by the capacity of 10% the consumption, while the capacity of 14% the consumption covers entirely the incurred error. Workdays are $\approx$ 83% covered by c_e sizing 10% of daily consumption. With 15% of the consumption, 97% of the set can be covered, while certain days required a storage of 18% to entirely absorb the error. Such results raise even further the importance of the intraday relevance on availability of a storage capacity, that will be further investigated in section 4.4.4.

4.4.4 *Impact of Intraday Storage Availability*

Many factors can impact the accuracy of a forecasting algorithm, whose consequence will propagate to the storage sizing [95]. Although SARIMA resulted with an average MAPE of 8.2%, the high propagation of the forecast errors had a significant impact on the resulting estimations. As such, the extracted c_e from the estimation method applied should be observed through its efficiency, or storage re-usage. The re-usage average of every daily estimation resulted similar for workdays $\langle \frac{w_{tot}}{c_e} \rangle_{X_w} \approx 169\%$ to that of the nonworking days $\langle \frac{w_{tot}}{c_e} \rangle_{X_n} \approx 152\%$. Although a low re-usage rate was identified for most of the days, some did result in a higher re-usage rate, e.g. 250%.

Intraday Capacity Estimation

Since reusing of storage capacity was identified as low, the origin for the storage requirement due to absorption of the forecast errors gains importance. Identifying how the error propagates within a day, will help to better understand the role of the variable storage shape over time. Some instances in Figure 40 resulted in greater forecast errors, but it is not clear how they mostly propagate to the resulting storage estimations c_e. Hence, the same estimation method from section 4.4.3 is hereby used on shorter (or intraday) time frames, in order to assess their individual impact to the overall storage estimated.

Based on the slope variation in Figure 39, it was decided to estimate the requirements c_e^i for six intra-day intervals (4-hour each). The impact of each interval ($i \in [1, \ldots, 6]$) is calculated as $\frac{c_e^i}{c_e}$, where c_e^i and c_e occur during the same day. As such, their impact percentage is evaluated for all days (for both X_w and X_n sets) and depicted in Figure 43. As it can be seen, certain intraday intervals have much higher impact than others, e.g. as high as threefold impact. For evaluation of the workday dataset, the estimated capacities c_e are mostly inflated by midday intervals, in particular from 08:00–12:00, 12:00–16:00, and 16:00–20:00 intervals. The results depicted in Figure 43 assist towards understanding the storage distribution relevance on intraday basis. Therefore, the continuous dynamic adjustment of the storage availability [95] should be considered even for the intraday intervals. As such, one can conclude that requirements for storage sizing can be further reduced if the total capacity of storage is properly distributed on intraday basis.

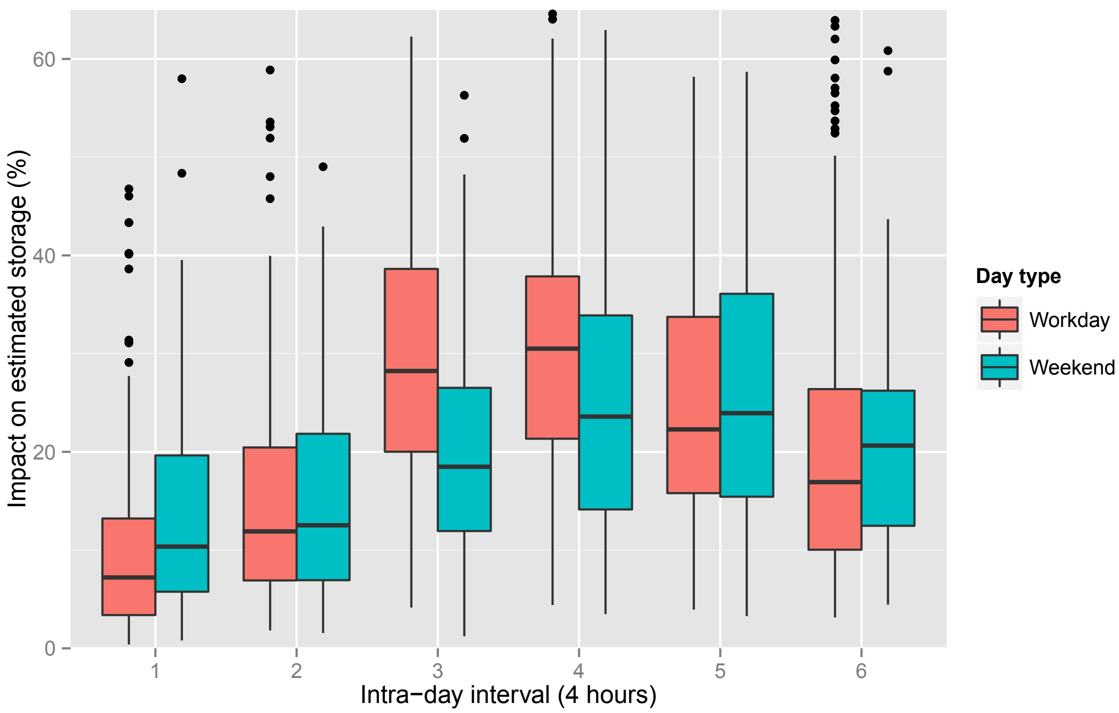


Figure 43.: Impact of intraday intervals to the storage requirements

Proposing Storage Shapes

To further investigate this, the capacities available within the intraday intervals vary over time, respecting the overall shape of the variable storage. As such, the X_w set is used to validate the hypothesis, where integrals of shapes presented in Figure 44 equal to the each other. If the capacity shape functions are indicated as $c_a[n]$ and $c_b[n]$, dependency is described as:

$$\sum_{n=1}^{l} c_a[n] \equiv \sum_{n=1}^{l} c_b[n] \ \forall a, b, \tag{3}$$

for all the intervals of a day, where a and b are the shape identifiers. The shapes for the variable storages hereby are selected from X_w (for demonstration purposes) as: (i) constant, (ii) identified peak, (iii) all intervals from Figure 43, and (iv) finally the actual error measured in Figure 40. It should be noticed that a final capacity over time $c[n]$ is calculated based on a capacity shape $c_a[n]$ and a selected capacity c (in kWh). For example, the constant shape will always have c value, while the peak function will either have $0\% * c$ or $200\% * c$ available for an interval n. Therefore, one can directly compare efficiency of different shapes (to address errors of a forecast algorithm) only by varying the overall storage capacity c.

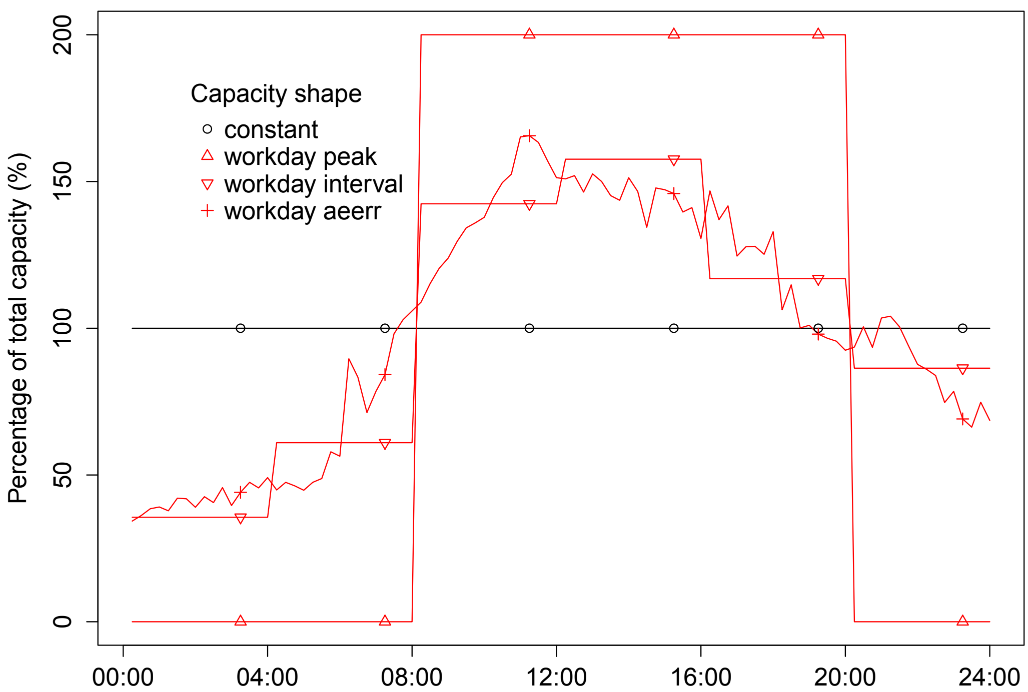


Figure 44.: Shapes for a variable storage capacity for intraday intervals based on experiments for X_w

Property of a Variable Storage

To measure the efficiency of the each storage shape proposed in Figure 44, the methodology detailed in [69] is adopted for charging and discharging behaviour. Hence charging/discharging efficiency of a specific storage technology is not considered and $w[n]$ is absorbed if the storage can absorb it (based on its state of charge). A variable storage introduces an increased complexity of unit management, in particular towards estimating the connection and disconnection State of Charge (SoC) of an individual asset [113], which is not addressed in this work. SOC cannot be treated individually (based on an asset), and if $n_1 < n_2$, the overall SOC of the storage is expressed as

$$SOC[n_2] = \frac{SOC[n_1]c[n_1] + q(c[n_2] - c[n_1])}{c[n_2]}, \tag{4}$$

where state at n_2 is inherited by its previous condition (at n_1) and forecast error is added. As the SOC per unit is not available, the variable q is introduced for the overall SOC over time. For the cases in this work, variable q is considered as

$$q = \begin{cases} 50\%, & c[n_2] \geq c[n_1] \\ SOC[n_1], & \text{otherwise.} \end{cases}$$

Management of connected units can be based on numerous factors e.g. selecting a EV to be charged [30]. This is similar to the power plant management, where "dispatch" refers to the timing turning on and off power plants to match grid's needs. Although considered interesting, evaluation on individual SOC of variable storage units is left for future work.

Evaluation of Storage Shapes

Hereby, every shape from Figure 44 is evaluated individually per day, while c is fixed (for any day in X_w) based on the total yearly load and their resulting MAPE is averaged to understand the overall impact. In the following experiment c is chosen based on the percentage of the average daily consumption, introduced in section 4.4.1. Hence, 2% of storage capacity is calculated as $c = (2.7\text{GWh}/365) * 0.02 \approx 145\text{kWh}$. By considering the X_w set and other values of c, the verification of the assumption for the proposed shapes is illustrated in Figure 45.

Although the peak shape had a fast convergence rate, it converges towards a $MAPE > 0\%$. This was somehow expected, as many intervals in the shape had $c_i[n] = 0\%$ of the total capacity c. The constant storage resulted in high drop for low capacities, while almost linear drop is noticed for $c \geq 4\%$. The interval based capacities had a slightly better performance, however the shape of capacity from Figure 40 was highly efficient. The impact of having the capacity distributed as absolute energy error (or "aeerr" in the figure) may pose as a reliable indicator of where the focus should be on the effort to improve efficiency of storage. It is critical to note the indications of Figure 45, where even the slightest variation of the capacity distribution provoked significant results to

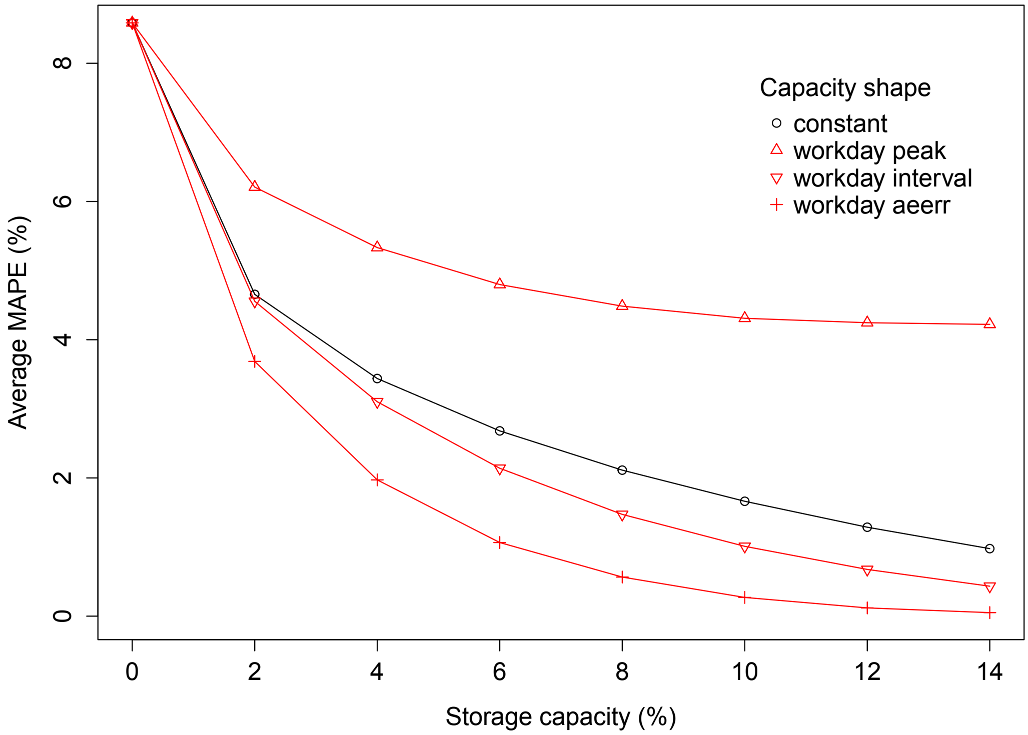


Figure 45.: Improvement rate of the absorbed forecast errors for the selected storage shapes

the overall performance. As an example, figure shows that MAPE for $c = 6\%$ of the "aeerr" case approximates MAPE of $c = 14\%$ for the constant case. In this example the difference of 8% results in significant capacity size i.e. of $c = (2.7\text{GWh}/365) * 0.08 \approx 580\text{kWh}$. Therefore, the evaluation of the variable shapes from Figure 44 resulted in a significant difference on efficiency of a capacity available (e.g. if capacity of connected EVs is considered), showing that slightest availability over identified intervals can bring the critical reductions in storage sizing.

4.5 VARIABLE ENERGY STORAGE

Although traditional battery storage systems proved to be efficient, they are considered to be expensive solutions [114], thus some investigated opportunities in reducing the costs of such systems [65]. In this section, a commercial stakeholder will be analysed with respect to achieve sufficient forecast accuracy, in order to be able to participate to the new business opportunities of Smart Grids. The primary aim is to capitalize on internally available assets, in particular the storage capabilities of stakeholder's EVs [32]. They gain even more attention as they can absorb forecasting errors not only by charging/discharging, thus loosing energy on storage efficiency, but load shifting by rescheduling. To do so, identified significance of storage shaping in section 4.4.4 will be defined by real-world assets i.e. an EV fleet will be used to achieve the accuracy.

Since vehicles are expected to exist anyhow within company assets, their involvement should lead to a significant cost reduction. This reduction will economically enable stakeholders to profit from their predictability within the Smart Grid era [25]. The empirical investigation here is based on real-world data, for stakeholder obtained from section 4.4.1 and presence of EVs at charging stations will be shown in section 4.5.1. It will show how the presence of its employees provoke higher forecasting errors, while section 4.5.2 will show that capacity of present EVs at can assist in mitigating such errors. As such, results in section 4.5.3 will show that an EV fleet can be used for absorbing forecast errors, even entirely eliminate a need for a static solution. Nevertheless, in intervals of low fleet presence, availability of static storage unit may be more beneficial than increasing the fleet size. All steps of this approach will be provided with definitions as well as empirical data to make it possible to follow it through.

4.5.1 *Presence of Storage Units*

With the electrification of transportation networks [45], we are witnessing an increase in the penetration of highly mobile electricity storage units e.g. electric vehicles [32]. However, the transportation vehicles of today are highly underutilized as they are idle 96% of their time [30]. If same usage rates are applied to future EVs, this imply that the majority them can be connected to the grid and be available to power systems. From a company perspective, employee cars which are in the garages of the employer buildings can compose a Variable Energy Storage (VES). In section 4.4.4 one could see relevance of an intraday capacity available and here a method to describe presence of mobile storage units is proposed.

Unit Presence Definition

Every mobile storage unit is able to connect to the electricity grid at some point in time and this connection time frame is called the grid session. Each grid session s is instantiated by connecting a unit to the grid at time t_c and is terminated by its disconnection at time t_d. Sessions of each individual unit can only occur sequentially, where for one session the storage unit is considered to be present for any time t as $t_c \leq t < t_d$. The step function [115] is used to model a single grid session of a storage unit. It is an elementary function denoted by $u(t)$, which holds one for positive side and zero for negative. A single grid session s is represented by two step functions as

$$p^s(t) = u(t - t_c^s) - u(t - t_d^s). \tag{5}$$

As such, the function returns one only if a unit is present on the grid, otherwise zero is returned. Numerous such sessions are actually the components for composition of the unit presence function $p(t)$. This function will return the total count of units present at time t. It is mathematically represented as

$$p(t) = \sum_{s \in S} p^s(t), \tag{6}$$

where S is the set of sessions from all mobile units considered. As such, the function returns $\mathbb{N}^0$, where zero indicates that none of the units is present.

Statistical Presence

In contrast to the unit presence function, one may be interested in understanding how many units are expected to be connected at a selected point in time. Furthermore, without knowing the number of present units of an entire fleet, the presence rate cannot be calculated. Thus, a function $v(t) \in \mathbb{N}^0$ indicates the count of individual units in ownership over time. The presence function is represented by

$$f(t) = \frac{p(t)}{\min(v(t), q(t))}, \tag{7}$$

having limitation of the charging points at the premise noted as $q(t)$. As such, the function is used for a statistical assessment of fleet's behaviour allowing scaling of their presence. If the set X contains time points of interest (e.g. 00:00 of all working days in year 2012), and Δ indicates the season length, then statical presence for all points in X is calculated as

$$F(t)_X = \begin{cases} \frac{1}{|X|} \sum\limits_{\tau \in X} f(\tau - t), & 0 \leq t < \Delta \\ 0, & \text{otherwise}, \end{cases} \tag{8}$$

where $|X|$ is cardinality and the return value is $\mathbb{R}_0^+$. Once calculated, the statistical model can be used for any fleet size to estimate an expected presence at a point in time, from its statistical model from Equation 8, as

$$\bar{f}(t) = \sum_{\forall i} \sum_{\tau \in X_i} F(t - \tau)_{X_i}, \tag{9}$$

where $X_i \cap X_j = \varnothing$ where all $i \neq j$. It is important to note that the model is not prone to errors introduced by an inappropriate selection of the points in each X_i. In fact, better selection of these points (e.g. only working days, without holidays) will result to a more accurate statistical model of the fleet's presence.

Presence of a Real-World EV Fleet

After quantitatively identifying forecast errors of the stakeholder in section 4.4.1, it would be interesting to see if fleet's presence can assist at the times of the highest errors. In this section, the presence curves are produced from 1044 grid sessions $s \in S$ of a real-world EV fleet. The data is collected from 5 January 2012 to 10 August 2013 (585 days), where 18 working days were marked as holidays (thus nonworking days). The fleet was continuously composed of five A Class E-Cell vehicles ($v(t) = 5$) which were the production result between Mercedes Benz and Tesla. These compact cars are pure EVs in ownership of the same stakeholder presented in section 4.4.1. As cars were not directly assigned to employees, other cases may expect slightly different mobility patterns [116].

Since many different variations in the session duration were noted (15 minutes $\leq t_d^s - t_c^s \leq$ 4 days), the first observation is done through their duration. The distribution function of the complete set S is depicted in Figure 46. A peak of short sessions can be immediately noticed, but these sessions have no significant impact on the unit presence curve $p(t)$. Further investigation let to depicting sessions initiated on Friday, which have much greater impact on the fleet's presence. As Figure 46 indicates, the mean duration from the complete set S averaged around 10 hours, while for Fridays (depicted from 179 sessions) resulted in more the 13 hours. As it can be observed, small peaks around 72 hours have significant impact as vehicles are present over the entire weekend.

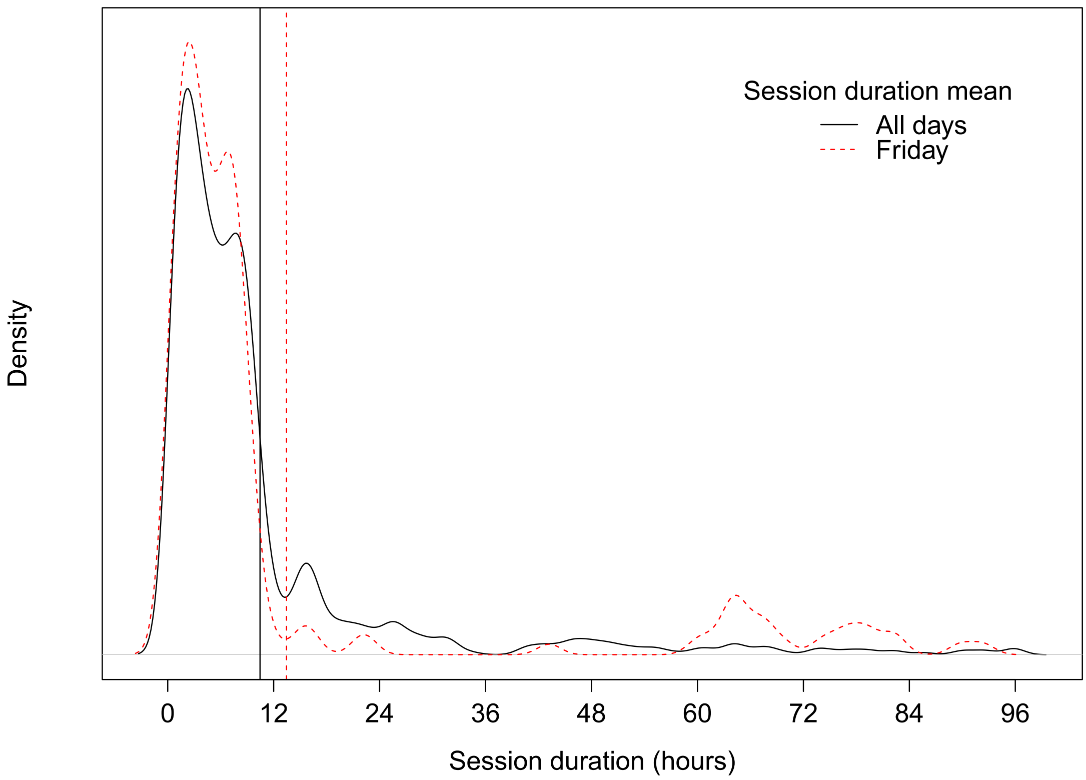


Figure 46.: Distribution function of the session duration for the complete set S and set of session initiated on Friday

A second observation is made for the duration of a grid session $t_d - t_c$ over its connection time t_c. Such investigation will help understand intraday behaviour of the units, having $S' \subseteq S$ where duration of all sessions is limited to 1 day. In Figure 47 one can see the movement of EVs for all sessions $s \in S'$, where most vehicles are connected within the stakeholder's working time. The trend of availability to the end of working time can be noticed from the drop of hours on the grid if moving along the t_c axis. All of the connections above this drop are considered as storage units being available over midnight, what appears to be more often for nonworking days.

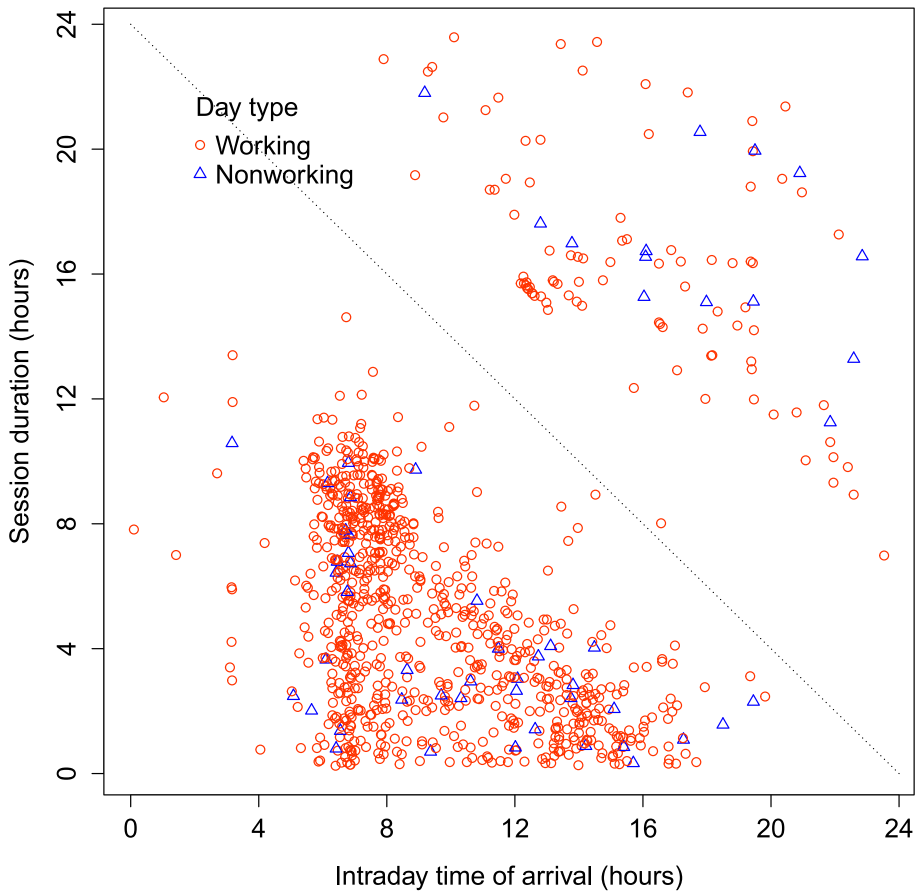


Figure 47.: Duration of the grid connection session in respect to time of their intraday initialization

Points in set X can be set to many different variations, however for experiments with a commercial stakeholder the weekly points should be used due to the significant difference over weekdays. For a better understanding of fleet's behaviour, the defined sets S and S' can be both observed through their statistical presence function $F(t)_X$. As previously depicted in Figure 46, such limitations are expected to be significant in the overall presence of the fleet. Figure 48 depicts the resulting statistical presence for both complete set S and the reduced one S' for Δ of 1 week. A significant improvement of the presence for the complete set is observed. Interestingly all workdays of the week lookalike (in average 18.9% for S and 11.3% for S'), while significant drop is noticeable over weekend days (in average 9.1% for S and 1.8% for S'). Such a small difference allows the distinction of X to working days X_w and nonworking days X_n (including holidays). Later in experiments the statistical presence will be used to scale the fleet size and therefore evaluate impact of its shape.

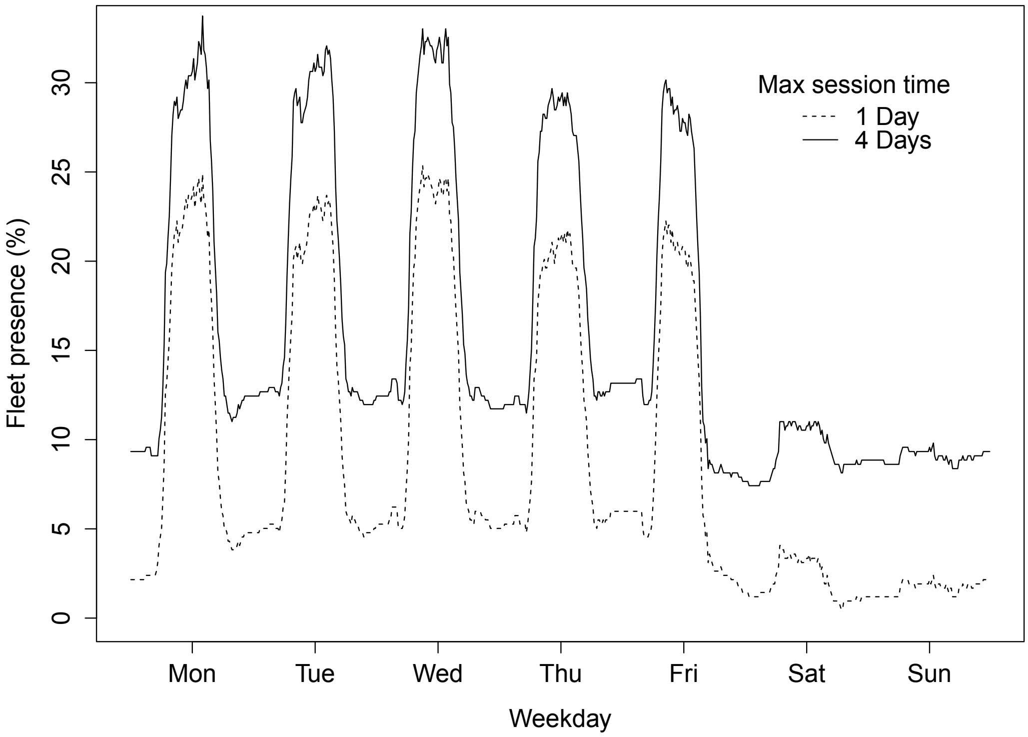


Figure 48.: A real-world EV fleet statistical presence at weekly Δ for both complete S and reduced set S'

4.5.2 *Variable Storage Capacity*

The definition of the presence curves in the previous section is further used to address stakeholder's forecasting errors. Once presence curves are computed, they are used to calculate the resulting capacity composed of storage units. If capacity of a single unit is denoted as c, simple multiplication as

$$p_c(t) = cv(t)f(t) = cp(t), \tag{10}$$

will give the total capacity available over time. From the experiment described in section 4.5.1, we have in Figure 49 the capacity availability from the fleet that is calculated for $c = 36$ kWh. The statistical capacity present over X_w and X_n resulted with an average capacity of 36.9 kWh and 16.6 kWh, respectively for the complete session set S. As such, $p_c(t)$ can be used in the assessment simulation for improvement through an existent EV fleet. For further scaling of its fleet, statistical presence curves $F(t)_X$ can be used only for one classification of storage characteristics c. Since many fleets are expected to have units of different c, the equations need to be further expanded.

Units of Different Capacities

Although the calculation of the presence curves can be done through capacity, it is not applicable to fleets with units of different capacities. For example, if only two vehicles of capacity c and $10c$ are available, the presence of smaller unit

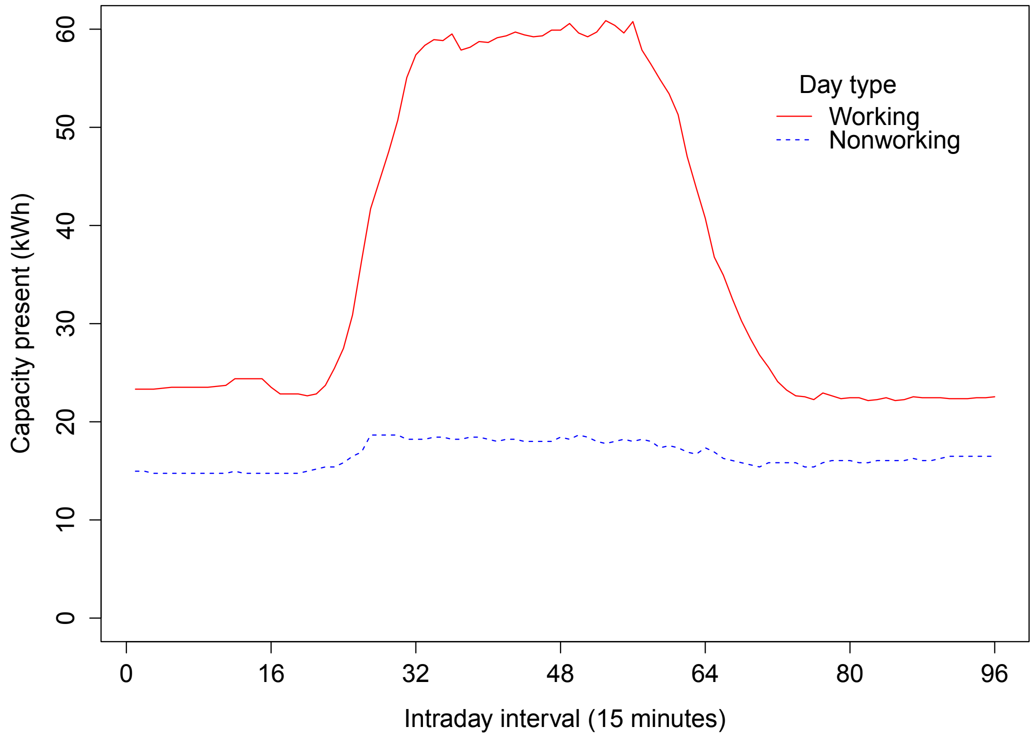


Figure 49.: Storage capacity present from the fleet's statistical presence

may jeopardize the estimation of the actual capacity available. With that in mind, every grid session s is expanded with the classification j of invariable capacity c_j. All sessions s_j are therefore populating the set of $S_j \subseteq S$. The classified statistical presence is expanded from Equation 8 as $F_j(t)_X$, where only $s \in S_j$ are considered. Although c_j is considered to be invariable, the total count of classified mobile units is to be scaled in simulations and is represented as $v_j \in \mathbb{N}^0$. Scaling v of each classification will contribute to total capacity present and is mathematically represented as

$$\bar{p}_c(t) = \sum_{\forall j} c_j v_j \bar{f}_j(t). \tag{11}$$

The total capacity available is expected to grow by an increasing number of units within the stakeholder's fleet, so one can assess their impact on the achievement of greater forecast accuracy. As such, individual variation of v_j can be used for assessment of the capability to address forecast errors by simulations with scenarios utilizing different classes of vehicles.

Presence of Static Storage

Presence of EV fleets is expected to differ between stakeholders, especially due to their diversity. For any case, especially for fleets that suffer from low presence, a continuously present storage can be critical to reach required forecast accuracy [69]. This is somehow similar to what was shown in the "peak" case in Figure 44. If the fleet from Figure 49 is used to absorb the forecast errors of the stakeholder (as inFigure 40), the improvement rate will lookalike the "peak" case in Figure 45.

Particularly for example of Figure 48, even a significant fleet scaling might still result to insufficient capacity to cover the forecast errors in the intervals of low presence. Hence the model should adopt a static storage capacity in parallel with the variable one to fill in such gaps. The total capacity present can be noted as

$$\bar{P}_c(t) = C + \bar{p}_c(t), \tag{12}$$

where C is the constant capacity present at stakeholders premises. Scaling C and unit count v_j is therefore used for assessment of impact in absorbing the forecast error. It is important to notice that a continuously present storage can also be represented as another classification j (in duration of the entire set), so $F_j(t)_X = 1 \ \forall X, t$. Instead, for clarity of formulas it is decided to observe C individually.

4.5.3 *Assessment on Actual Storage Requirement*

According to the evaluation results from section 4.4.4, the statistical presence of EVs in section 4.5.1 appears to have good fit to absorb the forecast errors of the stakeholder from section 4.4.1. In this section, the methods from section 4.5.2 are applied to evaluate the stakeholder in respect to different fleet sizes. Since the energy data of the stakeholder is a discrete-time signal of $T = 15$ minutes, the presence curves are sampled at the same frequency. This is important as a forecast error $w[n]$ is quantitatively absorbed by an estimated capacity $\bar{P}_c[n]$ for the same interval n. Using these discreet signals, the assessment of EV fleet as a storage via $\bar{P}_c$ is done through variations of the mobile unit in ownership v_j (or their capacity c_j for any classification j) and the static storage capacity C.

Individual Capacity Scaling

It is expected that static and dynamic part, i.e. its EV fleet, of storage will significantly differ in their relevance. As the forecast algorithm in 4.4.1 resulted to greater errors within the working hours of the stakeholder, Figure 45 indicates that capacity available from its fleet will play an important role in the error absorption. Same as what was defined by Equation 3, the efficiency of the dynamic storage shape P_c cannot be directly compared to C on the efficiency to absorb the errors. As such, the dynamic storage shape available is to be averaged and scaled by a constant over all working X_w and nonworking X_n days in X. The average capacity presence resulted respectively in 20.5% and 9.2% for the overall shape efficiency, having an average (on weekly basis) of 17%. For their direct comparison, as described in [43], the following equation needs to be considered

$$\bar{P}'_c(t) = m \int_a^b \bar{P}_c(t) \, \mathrm{d}t \equiv \int_a^b C \, \mathrm{d}t = C(b - a), \tag{13}$$

where m is the scaling factor used to align the areas of the dynamic and static shapes. For the follow-up experiments $m = 100\%/17\% = 5.88$ is used to calculate the efficient dynamic storage. This overall shape efficiency is used to depict the approach comparison in Figure 50, by individually scaling of v_j and C in $\bar{P}_c$.

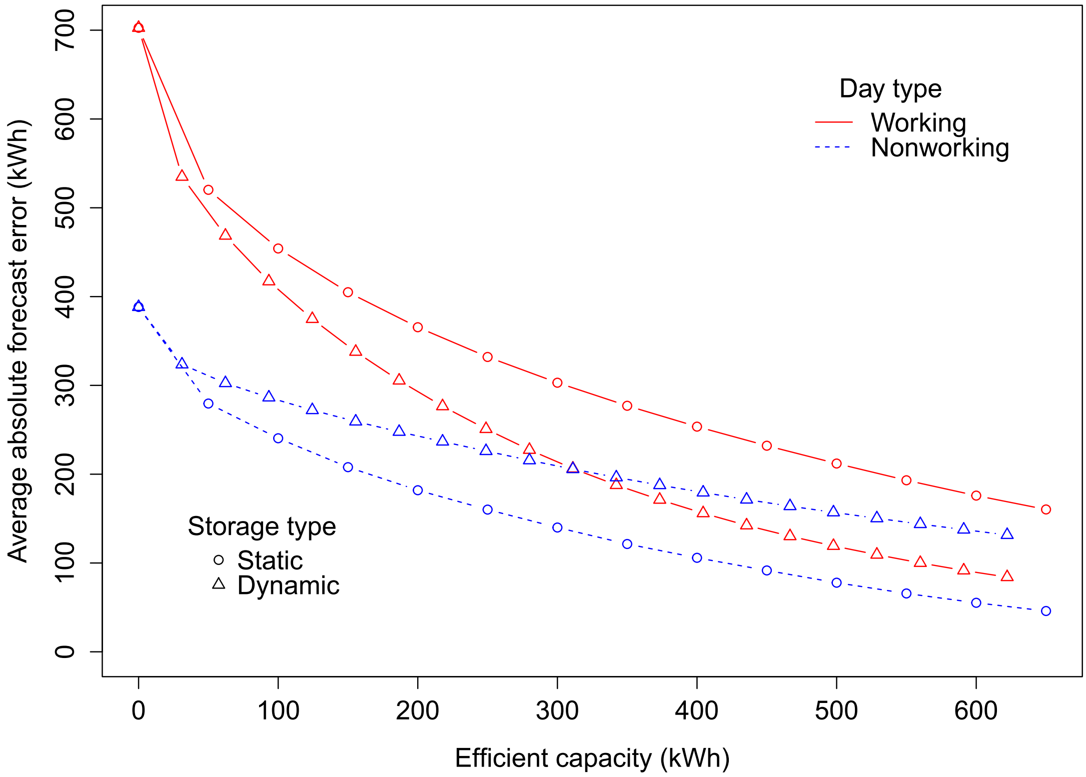
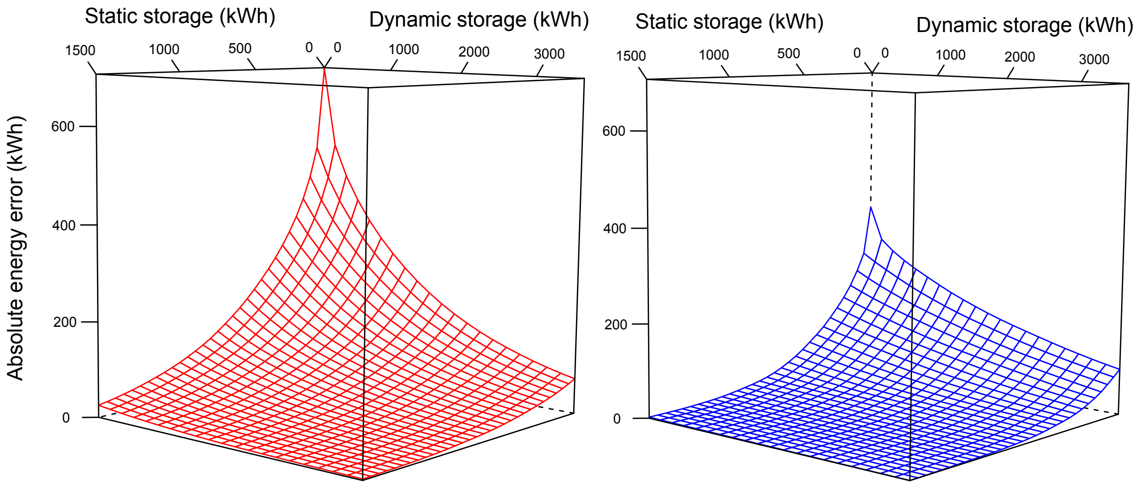


Figure 50.: Individual impact from static and dynamic storage approach in absorbing the forecast errors

As demonstrated by results in Figure 45, the fleet successfully overlaps with the working hours in X_w and over-performs in comparison to the static storage. However, the error reduction on nonworking days is significantly higher with the static storage approach. Obviously, it might be expected that static storage availability is required for the intervals of low presence [69]. The Equation 13 is not to be omitted, as average efficiency of the dynamic capacity resulted only in 17% for the complete set S. However, a constant storage unit brings additional cost [114], while the EV fleet might result in only slightly higher investment (as employee vehicles will be present anyway). Combing these two approaches will not only reduce costs, but their individual advantages are expected to complement each other.

Interdependent Storage Relationship

The individual scaling depicted in Figure 50 already emphasize weaknesses and strengths of the two approaches. Their combination is expected to fill in the performance gaps of the other approach, as the results strongly indicate that two approaches are advantageous either for X_w or X_n. Applying Equation 12, the storage scaling is performed for both C and a_j together, thus affecting the total capacity available $P'c$. Figure 51 depicts how their scaling reduce the absolute forecast error for working X_w and nonworking X_n days. The values are selected so that one can observe the advantage of combing the two approaches, having $0 \leq C \leq 1500$ and $a_j = 100$ (or 72% of employees). Aligned axis values of Figures 51a and 51b are important for a quantitative observation in forecast error reduction from their daily consumption (of 8.2 MWh and 5 MWh respectively). The capacity selected from both approaches is insufficient for X_w, while the static storage was embolden for X_n. As expected, both figures show high convergence towards the average forecast error of zero once approaches are combined.



(a) For X_w with the presence of S_w sessions (b) For X_n with the presence of S_n sessions

Figure 51.: The forecast error reduction on daily average with the interdependent storage approaches

It is important to point out Equation 13, where C of the constantly present storage takes 100% of capacity availability while dynamic capacity efficiency is measured to 20.5% for X_w and 9.2% for X_n. Understanding the benefits of the presented approaches to address the uncertainty of forecast can help obtaining a most economical settings to achieve an adequacy of the resource [77]. Omitting the potential of stakeholder's EV fleet would increase the requirement of a constantly present storage, which is expected to rise the overall system costs [92].

Achieving Forecast Accuracy Levels

Envisioning a longer-term role of capability of a variable storage as backup on forecast errors, and opportunity (dependent on the actual SOC) to accommodate an intermittent energy resource, can take us to the rethinking of the roles in the energy systems of today. A question is where to draw a line for the stakeholder in between the two approaches presented, in particular from an economical aspect [92]. Perhaps their performance dependency needs to be evaluated as a resource adequacy [77], as retailers today deal with MAPE greater than 2% for most sophisticated algorithms [14]. This is not to be omitted when a stakeholders is aiming for the accuracy of its energy retailer. In dependence to their goals, stakeholders can achieve sufficient accuracy for daily, weekday or even intraday requirements [90]. Hereby only continuous accuracy is observed, where the forecast error levels will be always below a certain MAPE limit. Figure 52 depicts few examples resulted from the result combination of Figure 51.

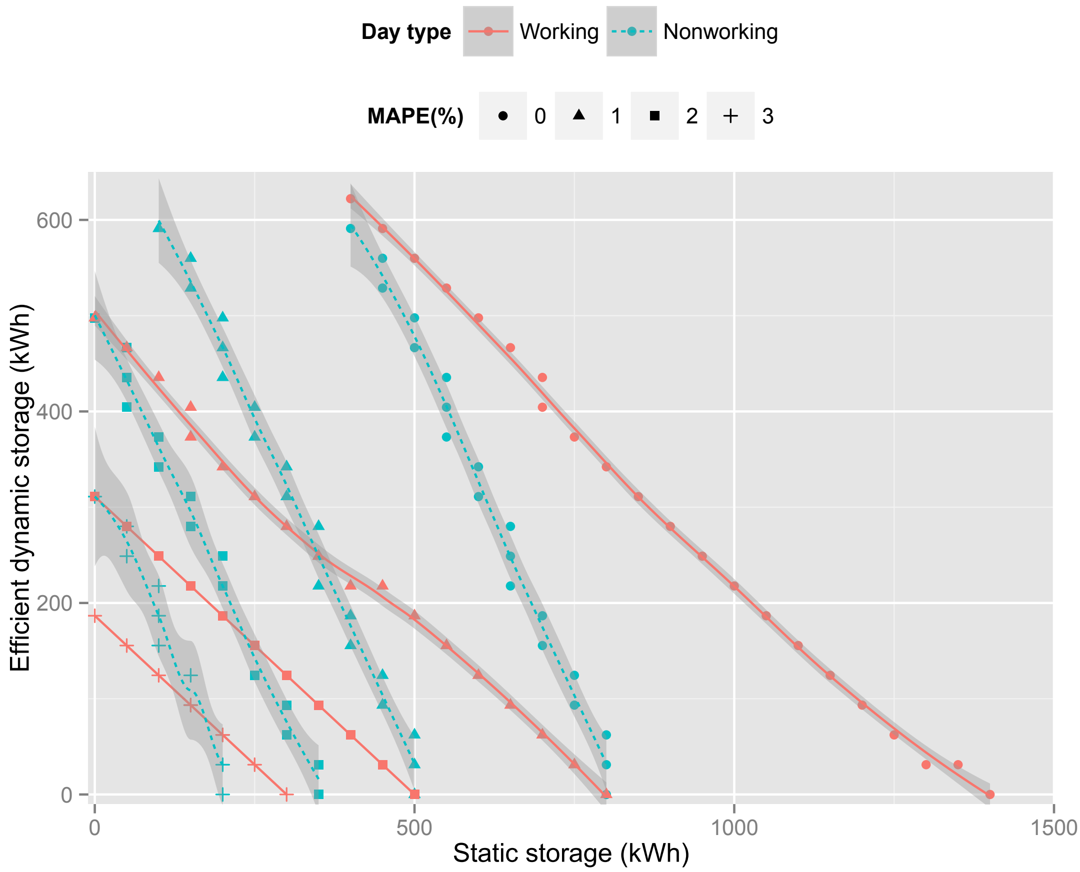


Figure 52.: Example of forecast accuracy levels of the commercial stakeholder

The limits are minimal for each depicted accuracy level of MAPE, thus on the right side of the limits a lower MAPE is expected (for both X_w and X_n). Furthermore, losses due to the storage efficiency ξ can be omitted if the observation is made through the entire consumption of a stakeholder as $\frac{(1-\xi)}{2}|w[n]| + y[n]$. For example, if forecasted load is $\hat{y}[n] = 150$ kWh and the actual energy consumed is $y[n] = 160$ kWh (having MAPE of 6%), with an $\xi = 90\%$ the actual energy consumed deviates only 0.3% (or 160.5 kWh) from the stakeholder's original consumption. With this in mind, the storage efficiency has almost no impact to costs of a stakeholder; hence stakeholders can apply the proposed methodology to evaluate their dependency on static and dynamic storage in their target to become an economically sustainable resource.

Quality of Presence Curves

If statistical presence is applied to a dynamic storage, as done for simulations of Figure 51, the quality of presence curves need to be considered. As an example, one can have high presence over one day while none next day. Their statistical presence curve (from section 4.5.1) will still result as their average. With that in mind, Figure 53 depicts the quality of statistical presence for the entire experimental fleet ($v(t) = 5$) for all the sessions in S. As it can be seen, the negative side of Figure 53 indicates the overestimated intervals as $f(t) < \bar{f}(t)$, for which the forecast errors are considered to be absorbed. The positive side indicates more presence than expected (thus $f(t) > \bar{f}(t)$), which provokes no problem in error absorption. Although this fluctuation exists for the experiential fleet used in this work, other fleets may result with a much more accurate presence curves. It should be mentioned that number of EVs within a fleet play a significant role in approximation of these curves, as $\lim_{v(t)\to\infty} f(t) = \bar{f}(t)$ for any t is expected. Nevertheless, in case of a low or a high number of EVs one should consider availability of an amortization if quality of needed intervals (to address the errors) is not sufficient.

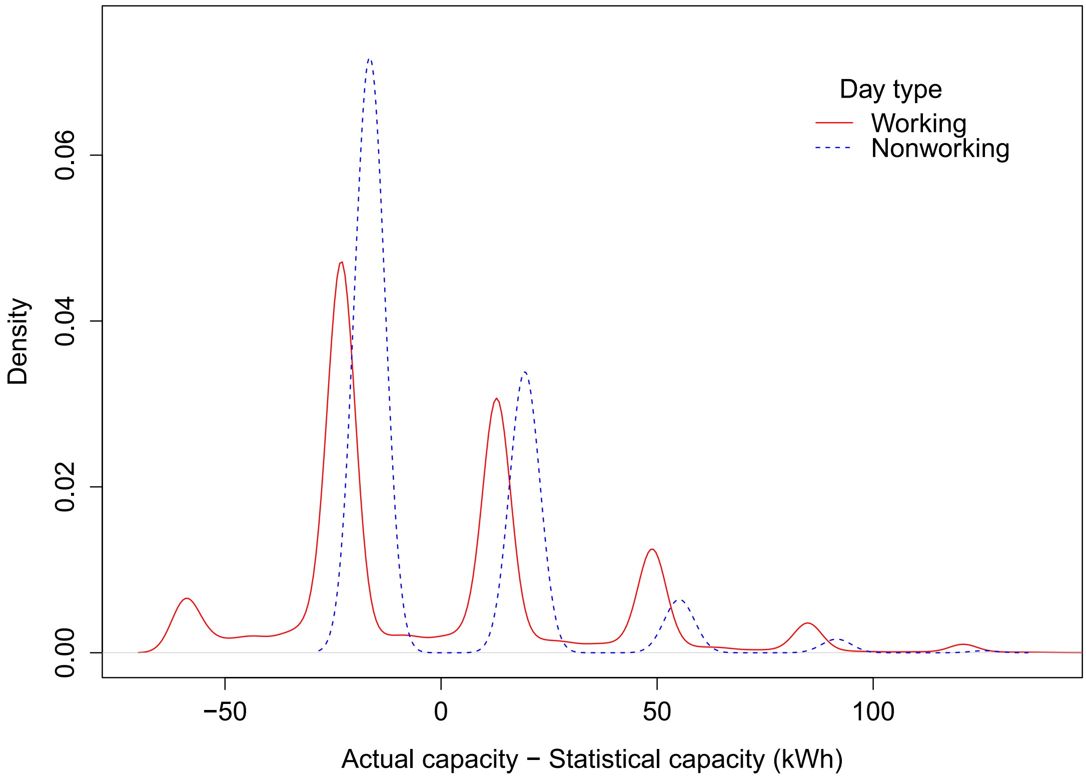


Figure 53.: Measured capacity deviations from the statistical presence

4.6 CONCLUSION

The transition towards an information-driven smart grid will empower its stakeholders to be active in the electricity supply-chain [28]. A high degree of their forecast accuracy will allow them to benefit by slipping into new roles envisioned by the Smart Grid [18]. It was shown how grouping offers clear benefits for stakeholders to reduce their "potential" costs on local energy markets, such as reducing demand imbalances. In a consumer group of 50 participants, the percentage of purchased energy that could be attributed to penalties was reduced from 21.58% to 3.90%. This case showed that certain accuracy can already be achieved on smaller scales of aggregation and not necessarily from aggregating tens of GW, which in [14] resulted in MAPE of approximately 2%. Thus, detailed experiments were conducted to reveal the improvement rate of the prediction accuracy by smaller scales of aggregation. In section 4.2 it was shown that forecasting accuracy converges rapidly to an overall accuracy of the set (for the particular algorithm), concluding that this is not a random effect. The results show good accuracy even for small groups, e.g. 200 households, as well as how individual prediction accuracy impacts the overall cluster accuracy.

Although a significant forecast error reduction can be achieved by clustering prosumers [68], based on the well known RES cases [92], it was decided to use static storage availability of different sizing to further enhance the forecast. The experiments carried out resulted in a significant improvement on accuracy, even with capacities that match only few percent of cluster's average daily consumption. As forecast error absorption propagates to the overall storage sizing [95], this chapter presented how a proper intraday distribution of the storage can significantly impact the accuracy [43]. Although this work in particularly focus on batteries, these solutions may result as expensive and economically irrelevant [70]. To reduce potential costs, in section 4.5 is shown that EVs are a good fit to improve self-forecast of a commercial stakeholder. Furthermore, the concept of VES is empirically evaluated with an EV fleet whose presence was measured at 34% for hours resulting in greater forecast errors (which are more relevant [43]). Nevertheless the VES concept is not explicitly linked to a storage technology. In future any asset that can somehow appear to "store" energy (directly or indirectly) can contribute to a VES. In section 6.3 more assets of this kind are proposed.

Results of this chapter are achieved by (simple and robust) off-the-shelf forecasting algorithms, and customization and inclusion of additional parameters [91] can be used for further improve. Nevertheless, equipping a pVPP with a variable energy storage unit was identified as of high importance, and their significance for Self-Forecasting EneRgy load Stakeholder (SFERS) need to closer assess even with the off-the-shelf algorithms. Once a forecast accuracy level can be achieved, one may realize self-forecasting infrastructures in the future [94]. Such autonomous systems will be presented in chapter 5, and will be shown how they lead to a greater energy optimization of resources and the system in overall, while stakeholders may expect many energy related benefits [69].

5 Self-Forecasting Stakeholders

The energy industry is undergoing many changes [117, 20, 52]. Beside rising costs and pressure to reduce the carbon footprint, they are straining an already vulnerable infrastructure that is struggling to keep pace with increasing demand [116]. Great expectations are therefore put upon Smart Grids and integration of more Renewable Energy Sources (RES), in order to bring the promise of a more effective grid infrastructure [18] and low carbon resources. This however brings also the challenges mainly introduced by the complexity of Distributed Generation (DG), intermittent production and finally the unpredictability of loads. Therefore the Smart Grid efforts try to bring the stakeholders closer to each other and empower them to consume energy services more efficiently [118]. Wide adoption of Information and Communication Technologies (ICT) by the traditionally passive consumers also connected them digitally [72], and therefore attractive to the energy industry. Their communication goes beyond the sampling of a smart meter, to an area where accuracy and higher resolution of samples can be delivered (if needed [73]) in a timely manner. Equipped as such and beyond e.g. with capability of bidirectional communication, an investigation on how the traditionally passive consumers can contribute to needs of other stakeholders is made.

According to the Smart Grid vision [8], improved energy management may stem from the near real-time bidirectional communication between, and within, stakeholders. Research on Smart Grids heavily invests in this direction [50], with the majority of ongoing trials relying upon participation of individual stakeholders, or even small groups [90]. For one to capitalize on the opportunities, some also presented in section 3.2.2, a high forecasting accuracy is needed. Still, for individual stakeholders (or even groups) an accurate forecast is hard to achieve [40], so methods from chapter 4 need to be applied. This calls for highly predictable loads, and/or full utilization of the assets available in order to achieve an artificially predictable load. Otherwise, if a load cannot be accurately predicted, one would not be able to verify the load changes and therefore could not measure the contribution [12]. This will, of course, limit their capability to participate and benefit from many Smart Grid opportunities [56, 27, 119].

To find the equilibrium towards achieving the previously mentioned objectives of energy industry, while in parallel lowering the overall stakeholder costs, the systems of the future will need to have a higher degree of flexibility and carefully consume all of its resources [94]. The traditionally passive stakeholders can internally execute their forecast, or so called self-forecast, in pursue of new sources of energy related revenue [49, 25]. Since the predictability of those stakeholders is achieved internally through their assets, externally they still appear unpredictable. If the self-forecast would be reported upwards, many other stakeholders would be able to optimize their processes (as described in Appendix A). Not only the accurate loads would help the overall system reliability [15], but predictable stakeholders can also offer flexibility of their loads. Although many Demand Side Management (DSM) mechanisms have been in place for some years [49], offering flexibility is not a simple task. Still, flexibility of some assets is easier to express. In particular, the focus here will be on batteries of Electric Vehicles (EVs), whose flexibility can be expressed directly from their State of Charge (SoC). To benefit from these empowers (as shown in this chapter) a smart energy system that enables the realisation of the deterministic behaviour. The latter is introduced, to empower next-generation electricity networks with effective collaboration among the stakeholders.

An overview on how flow of information in power networks changed with adoption of ICT is made in section 5.1.1, while section 5.1.3 describes how Internet of Things (IoT) can be applied to move beyond sampling – to event driven information acquisition. For this new capability, a closer look is given in section 5.2, where the flexibility of prosumer infrastructures is investigated for generating new revenues. The use case of section 5.2.3 uses the flexibility of Public Lighting System (PLS) to prove the potential. With that in mind, in section 5.3 system architecture is proposed to enable the active contribution of the traditionally passive stakeholders. Few scenarios for enabling facility management are proposed in section 5.3.2, and one strategy is assessed for the proposed system in section 5.4. The focus there is to investigate the Key Performance Indicators (KPIs) of such systems and how (potentially) available assets, i.e. EVs in this case, can impact its performance of self-forecasting. Finally, the discussion is made in section 5.5 and the chapter is concluded in section 5.6.

5.1 SENSING IN POWER NETWORKS

Due to the increased fine-grained information acquisition as well as the high quality and frequency of it [120], we are moving towards real-time view of the whole network. Such view get importance as ageing infrastructures will require significant investments in order to cover renovation and reinforcement required in near future [117]. The capabilities will empower traditional approaches for estimating the network state to analyse it down to the device level [17]. While Supervisory Control And Data Acquisition (SCADA) systems e.g. in UK generally extend only down to $33kV$ networks, on $11kV$ such systems are seldom available [9]. Therefore pseudo-measurements are introduced for estimating the system state, which may introduce uncertainties due to possible operating and environmental conditions.

As an example, being able to monitor voltage stability is a key part in preventing voltage collapses [11]. From the Quality of Service (QoS) perspective, one of the most important constraints on the distribution system design is the voltage level at the end-user (residential customer) point. To avoid voltage deviations, due the uncertainty of the network state estimation, remote points can be added to a significant number of points on the $11kV$ distribution networks. However, there is a significant cost associated with the acquisition of such real-time measurements. Hence, careful choice of location from which measurements can be acquired is needed [121]. If such measurements (or their approximation) may be accommodated using the existing smart grid infrastructure and technologies (e.g. the currently deployed smart meters), such expenses may be significantly reduced or avoided. The device level accessibility offered by Advanced Metering Infrastructure (AMI) may address these concerns, even though they are expected to operate with performance significantly lower than SCADA systems [73], with large number of sensing points distributed all over the network [53].

For the NOBEL project (`www.ict-nobel.eu`), the smart metering platform was developed [21] that among other functionalities, specializes in near real-time acquisition of metering information from the grid as it is reported directly by the smart meters [74]. A plethora of additional information apart from consumption meter readings are reported, including voltage, frequency, current, active power etc. Based on this information, as well as more advanced capabilities for direct interaction with the meters e.g. in order to increase the frequency of data reading [73], one can get a much finer view on the grid infrastructure that can complement the view of dedicated equipment.

5.1.1 *Sensing by Large Number of Smart Meters*

Although within the NOBEL trial every smart meter reading received contains many electrical measurements, this section will demonstrate AMI capabilities on some voltage deviations, while the complete assessment can be found in [53]. In particular the voltage drops are a well know issue [121], and by using the data collected from a large number of distributed voltage sensors one is able to identify areas of the grid with repeatable deviations. As the trial occurs within the European Union, the nominal voltage of 230V $\pm 6\%$ at 50Hz [122] is used as the point of reference. In practice, this means that received voltage deviations at consumer's premise may deviate in the range of 216.2V – 243.8V. With application of the European standard EN 50160 [123] where voltage characteristics at the customer's supply terminals under normal operation conditions are specified, the range of variation of the 10 minute RMS of the supply voltage is $\pm 10\%$ for 95% of a week. In other words, for more than 8 hours a week no boundaries are applied to the received voltage. By analysing the real smart meter data of the trial, many meters were identified to exceed the regulatory boundaries of 230V $\pm 6\%$. As shown in Figure 54, many measurements were detected crossing higher voltage boundary ($\geq 243.8V$). The data covers again a period of more than a month, and the set contains 1355 measurements in total. Out of these 1355, 617 measurements were identified going over the higher voltage limit ($\geq 243.8V$), which results in 45.54% of the measurements outside the allowed deviation zone.

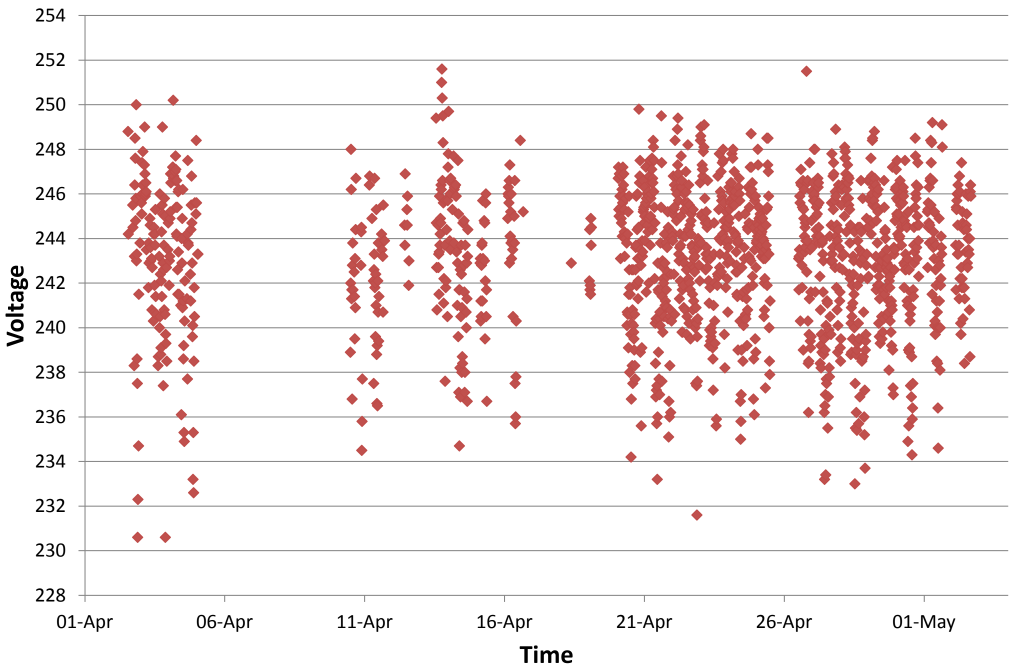


Figure 54.: Example of high resolution of a single device that violates the higher voltage boundary

As one can see, having a large number of such distributed sensing points within the entire grid can provide accurate information of its operations. Although today none of the commercially available residential smart meters can replace high-precision dedicated equipment e.g. network analysers, on large scale one may expect highly reliable information at a fraction of the cost over an existing infrastructure (already deployed for smart metering). One drawback of such approach may be the communication delay, which may be too high for reacting to critical grid events. Hence one should not consider this approach for critical grid events; still in [34] a good quality of information and timely delivery was successful so that new insights in infrastructure operations can be obtained.

5.1.2 *Delay in Monitoring a Small City*

In late 2011 the first trial of the NOBEL project ran a platform to monitor and manage approximately 5000 smart meters that connect and communicate meter readings and additional information. All acquired data, after some sanity checks, is stored in the "cloud" and made available via the services of the platform. The developed services [21] are results of complex queries and assessment done transparently on the data and customized for business purposes. These 5000 distributed sensing devices streamed their data into the system and a fine-grained analytics of acquired data was timely and highly detailed visible to the grid stakeholders [34]. Furthermore, a fraction of these customers had access to their energy prediction to perform more advanced tasks, such as energy trading i.e. buying and selling energy in a marketplace using the brokering services shown in section 3.2.2.

Timely delivery for such system required high performance, thus technologies and methods from [38] have been adopted. Results of the trial from Figure 55 show the cumulative distribution function for measurement data delivery. The delay was measured from the moment the data was generated by the smart meter, up to the final stage that it was collected by the energy service platform. This basically shows how long one can expect to wait for a particular percentage of measurements available to the metering platform. For instance, the figure suggests that one can expect to have around 90% of the measurements within 240 seconds for this experimental environment. The realistic communication assessment – without any effort to optimize it on the Distribution System Operator (DSO) side – is especially interesting for the consumers of the data e.g. the state estimation algorithms [9] may have reliable calculation already on lower percentages of the collected measurements.

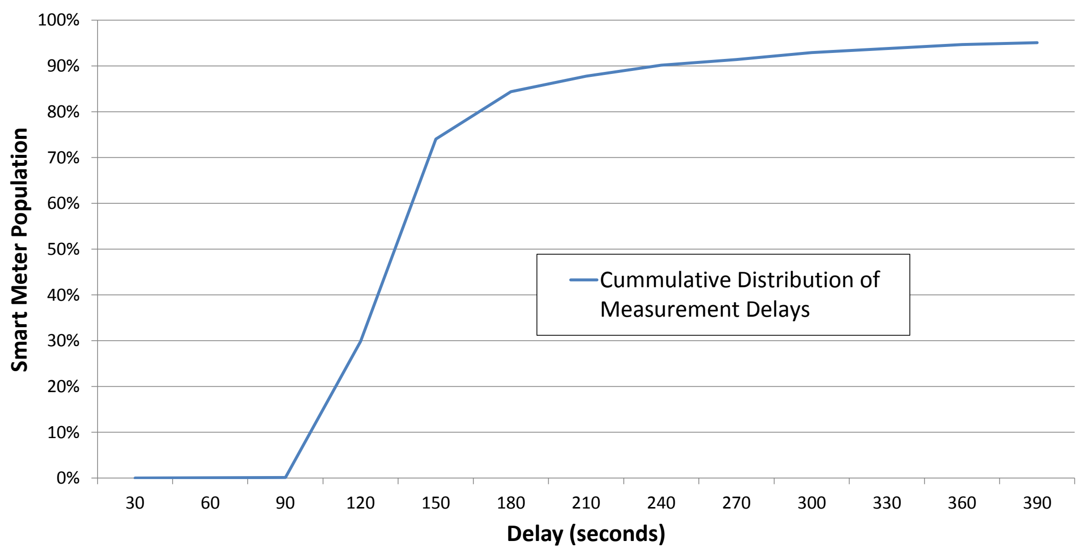


Figure 55.: Cumulative distribution function for measurement delays

5.1.3 *Event-Driven Sensing*

To deliver a high QoS, grid operators of today mostly rely on multi-sensor systems named Phasor Measurement Unit (PMU) that are placed on points of interest within the grid. Some have enough intelligence to automatically respond to QoS events within the network. Eventually, these could provide the emerging smart grid with the desired features of self-healing for detected anomalies. To do so, one would have to create a network of PMUs and Phasor Data Concentrator (PDC) to collect the information and transmit it to a SCADA system at control centres. As an example the FNET project [124] utilizes a network of approximately 80 high-precision Frequency Disturbance Recorders to collect syncrophasor data from the U.S. power grid [125]. However, these power quality systems are costly and are operated by the DSOs.

With new capabilities in the smart grid, as well as rapid advancement of modern Internet technologies at application and protocol level, complementary solutions become possible. The usage of smart meters for voltage monitoring has been investigated [126] and shown to be feasible approach, although the power line communication technology used puts severe constraints on the bandwidth. However, if such measurements may be provided along with the smart metering data, then analytics may provide a good insight on the infrastructure [53]. This section investigates the potential combination of the modern smart meter sensory devices and application logic over an event-driven Internet infrastructure using wireless low power technologies. The focus is explicitly on a promising set of emerging technologies, i.e. IPv6 and more specifically 6LoWPAN (defined in RFC4944), REST and a publish/subscribe model to further improve channel efficiency [120], to assess the feasibility and potential benefits of such solutions.

Architecture

The metering infrastructure is composed of smart meters in a mesh network and, optionally, concentrators. A smart meter measures the amount of energy consumed or produced by a customer and submits periodic readings of the amount produced or consumed. It may also be able to measure and report other important measurements such as power, frequency, voltage and power factor. Furthermore, it can issue events, such as for a change in state, e.g. on or off, and when a threshold is violated, such as a customer drawing more power than his contract allows.

The multi-layered architecture from section 3.2.1 is here used and Figure 56 depicts higher details on lower layers of it. As a remark, several layers exist i.e. the device layer, the middleware and the enterprise services and applications. Embedded devices (in this case smart meters, concentrators etc.) are composed from a hardware as well as a software part that enables their low level programmability. Top layer depicts the enterprise services of Integration and Energy Management system (IEM) and applications that can form mash-ups e.g. NOEM from section 3.3.3. Between the two, there is a middleware layer partially at infrastructure level and partially at device level, in this case the Data Capturing and Processing (DCP). To tackle heterogeneity, the architecture is also able to deal with smart meters that communicate with proprietary protocols through concentrators that connect to the smart meter network transparently for the IEM. Each smart meter hosts a developed communication and hardware adaptation platform (IPC) that allows end-to-end connectivity in the entire network.

Although meters can directly report their measurements to IEM, it makes sense to have in-network DCP to abstract network details and improve performance of

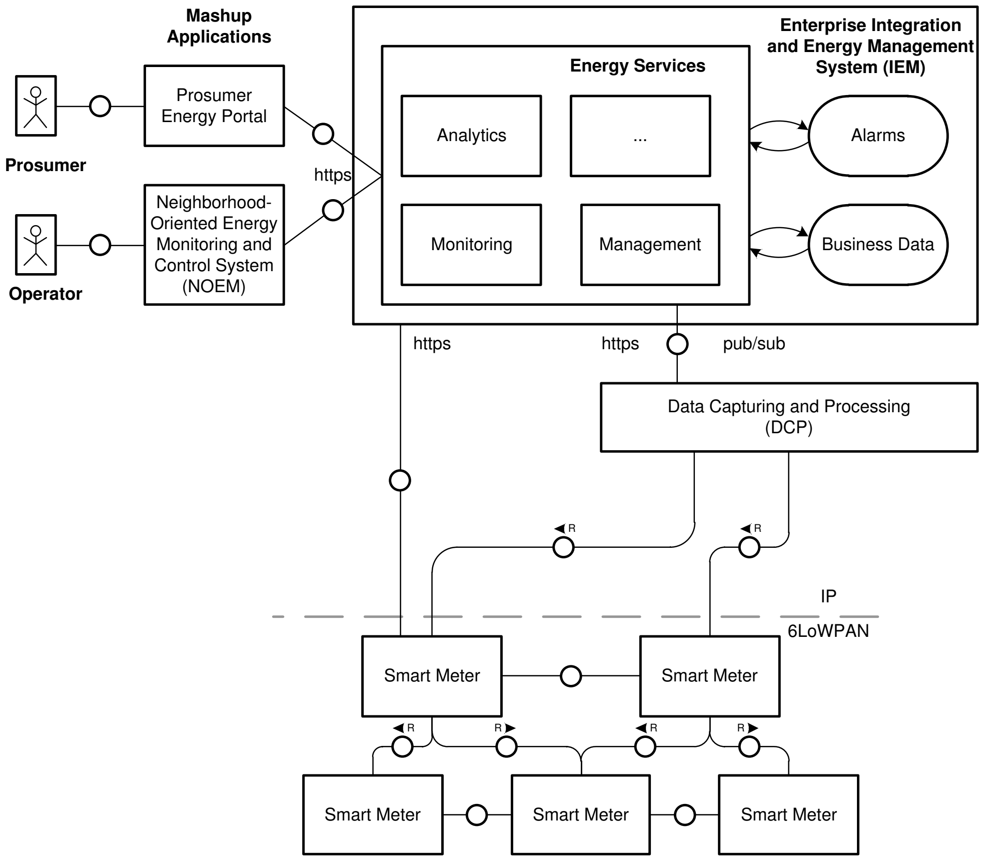


Figure 56.: Architecture overview with focus to the embedded devices

IEM [74]. DCP captures information from both IP-connected meters and meters connected via proprietary protocols with a concentrator. The middleware follows the publish/subscribe paradigm that enables efficient gathering and in-network processing of information and fosters the implementation of loosely-coupled fully distributed systems. The business services express their information needs, e.g. getting meter readings with a certain frequency for a selection of smart meters. The middleware is responsible for adaptively using the network to efficiently fulfill them.

Experimental Results

The main goals of the carried out experiments in this section focus on the aspect of evaluating a subset of communication aspects of the 6LoWPAN mesh smart meter network and its interaction with the energy services, i.e. the delay between issuing a subscription, e.g. due to an alarm raised in the grid, and the reception of data from the smart meters. By doing so it is possible to evaluate if such actions for monitoring can be done in an acceptable time-frame for the operators to capture certain grid events. The meter triggers an event whenever a particular measure violates pre-configured thresholds. The event is processed by the enterprise system, and a decision is made on whether the frequency should be increased or not. This approach safe-guard the network from being flooded in case of a wider problem, given that bandwidth in the current systems is limited.

The nodes send the meter readings with the dynamically adjustable sampling frequency, which was specified by the subscriptions. The data is pushed to the IEM via its REST web services dedicated for acquiring the metering data [21]. Since all readings (including their timestamp) are stored together with the reception time by the IEM monitoring service, and this information is used to determine delays to make or change subscriptions, as well as ratios of message loss.

As for on-demand power quality monitoring of the grid a high frequency and quick response time are necessary, the DCP is configured to start publishing metering data (at frequency requested) as soon as a subscription arrives. These activities are evaluated using an Internet connected IPv6 distributed testbed i.e. the IEM services are hosted in Walldorf, Germany, while the mini-testbed and the DCP root instance are running in Stockholm, Sweden.

An iteration of the data collecting experiment used to validate the design consists of the following process (which in the experiment is repeated 80 times):

- *Step 1:* The IEM issues a $60sec$ interval subscription for all meters. The time until all meters are reporting data is measured.

- *Step 2:* After $5min$, one of the nodes is chosen, in a round-robin fashion, as the target to increase its sampling rate to $10sec$. Again the time until the chosen node is reporting data with the selected (higher) frequency is measured.

- *Step 3:* All subscriptions are removed.

Figure 57 shows a histogram with the distribution of completion times for making a new subscription to all meters, and having received new data from every node. The histogram shows that in the majority of the tests all the meters responded within a $30sec$ period, with a tail of cases with higher delays. The higher delays, especially the little peaks at $70 - 90sec$, stem from cases where the subscription request for one meter was lost in the radio network. Since DCP uses periodic beacon messages both for discovery and information about the state of the nodes including the currently known subscription, it takes some time for DCP to become aware of the message loss and reissue the subscription. This time is configurable and set to $60sec$ for this test. Therefore, the DCP subscriber might need up to $60sec$ (or even more if the beacon messages are lost) to realize a request was lost, the re-send is subsequently delayed. This is a deliberate trade-off; technically a smaller refresh interval can be chosen, but at the cost of increased administrative messaging overhead.

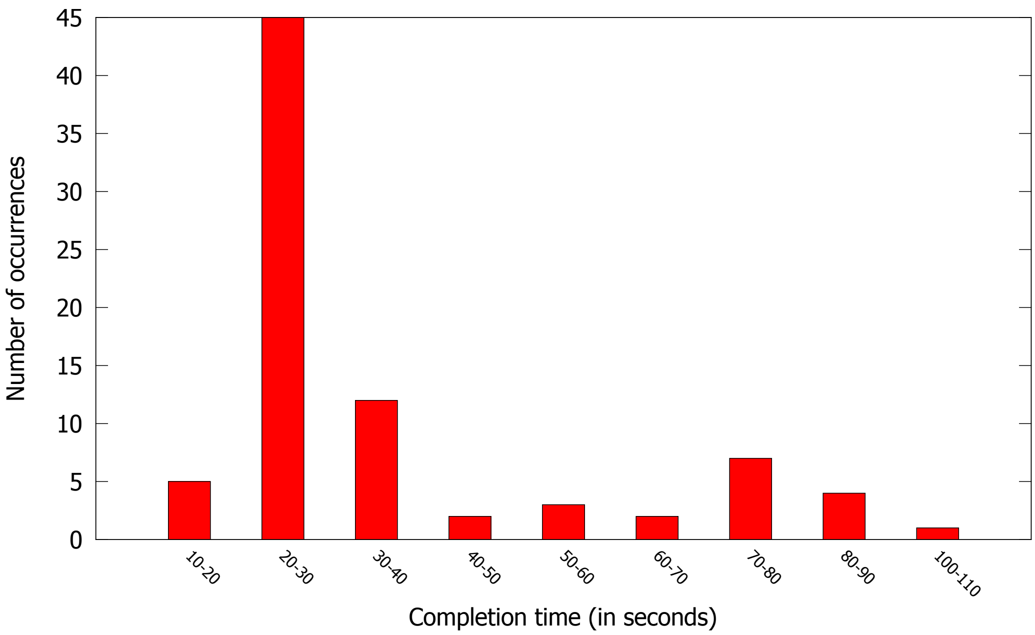


Figure 57.: Histogram of the time it takes for all meters to report data after a new subscription

5.2 FLEXIBLE LOADS

In this emerging context, infrastructure owners of, for instance, industrial facilities, buildings, wind parks, electric car fleets, offices, arenas, schools, convention centres, shopping complexes, hospitals, hotels, public lighting etc. look for new business opportunities [28] depending on the capabilities of the infrastructure they operate. Today most of them try to minimize their costs by, for instance, adjusting their energy consumption when its possible [119]. However, the emergence of the Smart Grid may provide new capabilities for increased revenues for stakeholders. By making their energy footprint flexibility available to grid managers, stakeholders can charge for their energy behaviour adjustments [67]. A typical example is the electric car fleet manager, who traditionally would try to minimize costs by charging the cars when the electricity prices are low. However now the trend is towards a multi-constraint goal, where the customer-needed QoS has to be guaranteed, e.g. an EV charged sufficiently to accomplish user's next goal, but also take into consideration the broader context i.e. the management of a variable energy storage facility [43] that can store and feed-in energy back to the grid depending on specific KPIs e.g. on cost-benefit, performance, green energy usage, etc. In the same train of thought other infrastructures, such as the PLS, which although is much more constrained in comparison to other facilities, it may still be used as energy balancing party by adjusting its behaviour by, for instance, adjusting illumination according to their technical and regulatory capabilities.

This section sheds some light in the new opportunities for active participation of prosumer infrastructures. In order to take advantage of a flexible energy footprint, potential scenarios for monetizing the available flexibility are investigated. As an exemplary case, section 5.2.3 expresses such a stakeholder flexibility over the use-case of PLS, as is highly predictable prosumer. Later, in section 5.3 and section 5.4.2, the focus will be on enabling unpredictable prosumers to be flexible, as municipalities and public authorities strive towards cost reduction and identification of new revenue sources.

5.2.1 *Energy Behaviour Flexibility*

Every prosumer on an electricity grid is introducing a certain load. Independently of the load's nature (consumption or production), this may have a time-dependent flexibility associated with it, which depends on the nature of the underlying task producing or consuming energy. Shifting loads is a fundamental aspect in the global Smart Grid vision, and a typical example often given is that of being able to turn devices ON or OFF for specific times. However, there are many more possibilities in modern intelligent devices and systems [127] apart from a binary state, which are spread between the two extremes (ON and OFF) and in principle can be depicted with a variable load profile over time. Being able to correlate the load profile with the tasks executed, and the life-cycle of the device, may enable flexible energy management [128] depending on external criteria such as performance, energy efficiency, costs etc. Any process that can be split to timeslots with distinctive loads that can be adjusted, on the time or magnitude, is a good flexibility candidate as its execution time may be extended with lower overall load or shifted load for specific timeslots.

Some infrastructures may be highly unpredictable, e.g. a wind or solar parks, while being measurable [12] and "adequate as resource" [77] is considered the key point for stakeholder to provide flexibility. Other infrastructures, such as PLS, are highly predictable due to their behaviour patterns. The PLS consumption is easy to predict as its load is usually constant (within a zone) for many hours with negligible deviations. However, from the overall consumption there is a lower limit depending on the regulatory framework (e.g. at least 70% illumination from 20:00 to 06:00). The difference between the lower limit and the maximum load that can be imposed to the grid may be flexibly adjusted. This flexibility is now becoming a potential business enabler [28] and may be used for balancing the grid while in parallel offering benefits to its owner, such as additional revenue, or contributions towards the community's sustainability goals.

One key application area falls under general efforts in aggregating and disaggregating flexibility objects offered by stakeholders [129]. In that context, every prosumer in the Smart Grid city may offer certain flexibility, which comprises of loads that can be adjusted, as well as the corresponding cost that the requester will have to bear for negotiation. These flexibility offers (of individuals) are formatted as in [27] and aggregated to the same format, finally forming a single flexibility offer. Figure 58 depicts the aggregation of two discrete loads and their respective costs to a single load and cost curve. As flexibility offered may only partially fit to the need of a requester, several negotiation steps may be inevitable. Once a prosumer offers his flexibility, the requester may accept the offer, or propose to accepts only parts of it. At the end of the negotiation, the final negotiated load, as well as the corresponding price curve, are agreed. The complexity of managing very large numbers of processes and costs (a typical task of an aggregator [130]), as well as considering the specific conditions of each stakeholder, can be a daunting task [129] which is not in the context of this investigation.

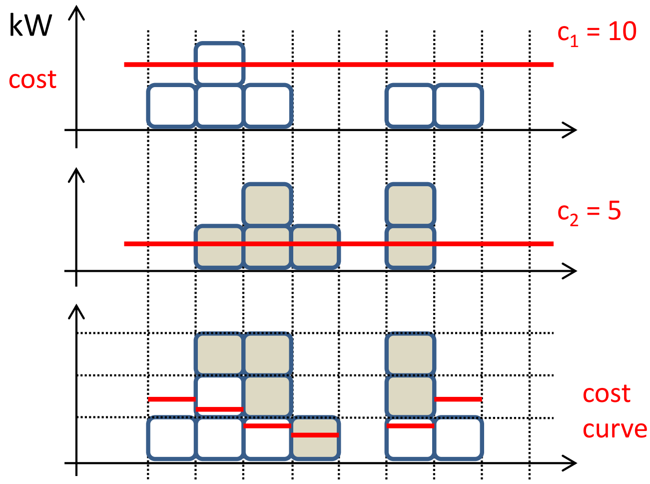


Figure 58.: Example of the aggregation step for two flexible process

5.2.2 *Flexibility-Driven Scenarios*

Being able to disaggregate, assess and adjust energy behaviour at process or device level, may yield significant benefits in the Smart Grid era. Such flexible prosumers can participate in various DSM scenarios [28], at a level that either was not possible before or was done only at small-scale proprietary systems and uniformly controlled infrastructures. Focus here is on three example scenarios to show how the energy flexible infrastructures may be utilized within a Smart Grid city. While some parts of these scenarios may be partially realizable today, the most sophisticated version of them assumes the existence of diverse energy services [21] available to the stakeholders. As depicted in Figure 59, the focus is on three key scenarios i.e.:

- *Scenario 1* – Bilateral negotiation of flexibility
- *Scenario 2* – Market-traded flexibility
- *Scenario 3* – Energy flexibility outsourcing

These scenarios are indicative on the new capabilities and interactions that are possible over a service-based Smart Grid infrastructure and its stakeholders [33].

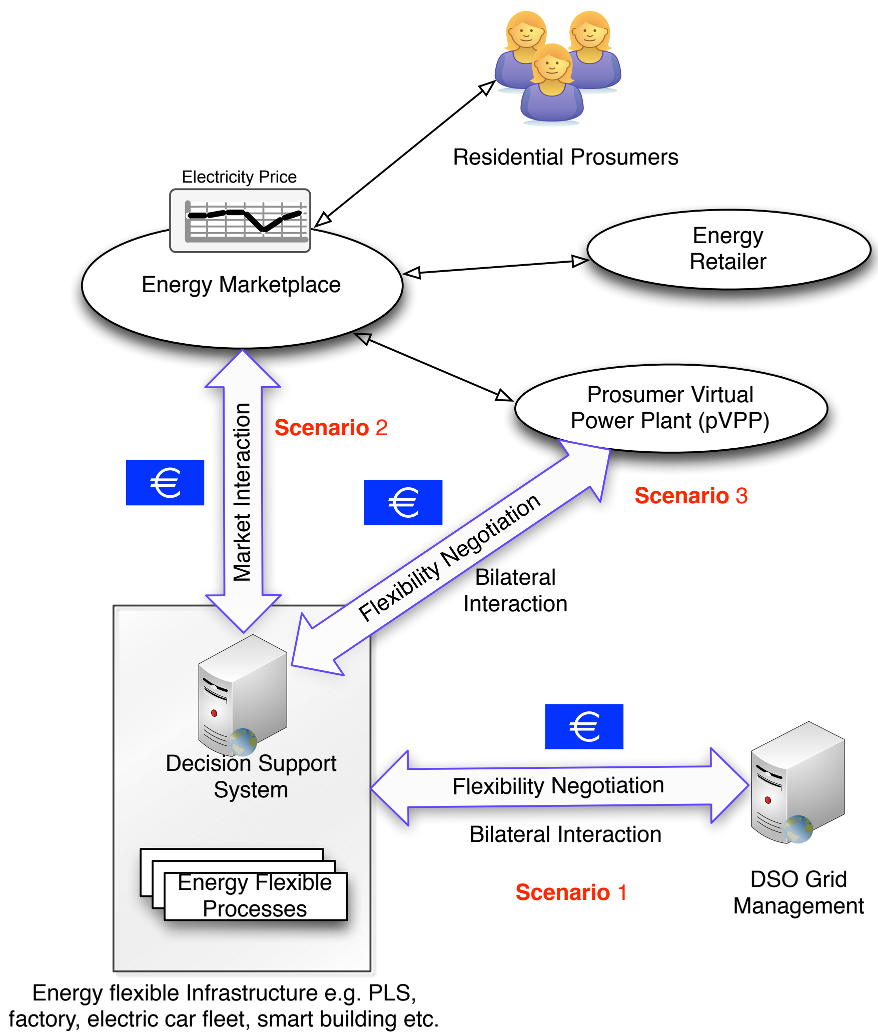


Figure 59.: Scenario overview for interacting with flexible energy systems

All of these (and many more), are not exclusive and can coexist depending on the business models, available means and goals of the respective participating stakeholders. While one may recognize partially current practices in industrial energy management, the aim here is to address it from the viewpoint of flexible energy infrastructures. The latter usually are considered to be larger prosumers e.g. Photovoltaic (PV) parks, wind farms, smart buildings, public lighting systems, public facilities. However, in the Smart Grid era, such infrastructures may be composed from a (very) large number of prosumers (e.g. residential users), who as standalone do not have any real impact, however as a group, they may significantly impact the grid and its operations as they can act as a prosumer Virtual Power Plant (pVPP) [61]. How these pVPPs are created e.g. based on social, economic, geographic or other criteria is beyond the scope of this research; however these should not be neglected as they may empower third party service providers that act on their behalf (as for instance depicted in *scenario* 3).

Scenario 1: Bilateral Negotiation

The first scenario (depicted as *scenario* 1 in Figure 59) aims to bring together flexible prosumers and those who can benefit from an adjustment of the energy load in the network. Typically the main stakeholder is the DSO who aims at keeping the network in balance and may use large flexible prosumers (usually industrial facilities) as balancing partners. However, in more advanced scenarios this role could also be assumed by others, e.g. an energy retailer that has over-provisioned energy within his network and seeks to reduce energy consumption of a large player in order to guarantee uninterrupted supply to residential users, or not significantly deviate from his contract with the DSO (which may be costly due to penalties).

For this scenario to be realized, real-time energy monitoring, management and assessment services need to be in place. With the IEM services (from section 3.2.2) in place for the NOBEL project, such a flexibility scenario was realized, while an example of the information exchanged is presented in [27]. Additionally, micro-contracting and legally binding should be possible. As this approach assumes bilateral interactions, any stakeholder seeking a comparative analysis with similar contracts offered by other stakeholders would have to contact them directly. The absence of standardized workflows and interaction protocols may hinder him and lead to an integration nightmare. Furthermore it is questionable to what extend open behaviours may be realized as each stakeholder will have to develop his own system, and also heavyweight stakeholders may impose their offers. Although such approaches can be implemented today, one has to consider several aspects in order to create open systems and standardized interactions that may be able to accommodate new business models in the future.

Scenario 2: Trading Flexibility

One key visions of smart cities is that of energy prosumers to be able to optimize usage of local energy resources [28] by trading them on-line e.g. in local smart city wide marketplaces [23]. This vision is the core of *scenario* 2 from Figure 59, where a user knows his energy behaviour (potentially assisted by advanced prediction services), and buys or sells energy he needs on a local market. If the prosumer knows and can shift his energy signature by, for instance, deferring or cancelling processes (or parts of them), he could benefit as he can offer this flexibility as a tradable good in the market. Although a single prosumer may not have significant impact if sufficient forecast accuracy cannot be achieved, [25] suggest that groups of them can transact on such markets as a prosumer virtual power plant [28].

Energy flexibility can be traded, i.e. the prosumer may offer the option to consume less or consume more depending on the benefits, such as additional revenue that he can get. Such a market based negotiation is possible; however, it entails the agreement on future behaviour among the participants. Although this constitutes a longer term approach, it has significant benefits as it enables the applicability of economic models and strategies towards shaping future energy behaviour on the prosumer side. Sophisticated approaches may be realized, while economic products similar to what we are accustomed from the stock exchange may be created [23]. Since these will be well-known platforms that will handle such transactions, one can expect them to evolve rapidly and integrate functionalities (e.g. compliance, payments, micro-contracting) that may be made available to its participants.

Scenario 3: Flexibility Outsourcing

Another interesting way to approach energy flexible prosumers is *scenario* 3 as depicted in Figure 59, which complements both *scenario* 1 and *scenario* 2. As discussed, pVPPs may arise in the Smart Grids, and may act as a larger prosumer. The overall behaviour of the pVPP may be disaggregated to the specific users (or groups of users) constituting it. Based on the flexibility knowledge for each of these users, the pVPP will be able to adjust its overall behaviour and offer this flexibility (the continuously changing sum of the flexibility of its members), in a local energy market. Third party service providers will be needed to manage such pVPPs and provide the basic services needed e.g. for users to join/leave, informational services, energy monitoring, energy management, prediction, billing etc. These service providers will act on behalf of their members and ensure benefits on their behalf.

Comparing directly *scenario* 3 with the other two scenarios, one can see clearly that here is the case of outsourcing the energy behaviour, while maintaining some per customer preferences. Many surveys [36] bring up the issue of energy management automation at residential prosumers, as many users although enthusiastic at the beginning, fail to be actively engaged for longer periods of time and clearly wish for automated systems that will consider both their needs (e.g. comfort preferences), but in parallel will be able to autonomously consider external information (e.g. price signals) and manage their energy signature accordingly. This scenario accommodates exactly that, i.e. the outsourcing of energy management to a third party (leader of pVPP) who act on their behalf. Of course incentives will need to be considered in conjunction with new business models in order to attract users to join a specific service provider.

5.2.3 *Case Study: The Public Lighting System*

For demand side management approaches to work, some prosumers must be able to adjust their energy behaviour [94]. This implies that each prosumer has (i) knowledge of his own processes as well as the energy prosumed associated with them, (ii) the capability to do timely monitoring on his infrastructure and (iii) the capability to apply energy management related decisions to it [49]. The PLS may have a maximum energy consumption level as well as a minimum level (depending on regulation or dynamic conditions such as traffic, weather etc.). The difference between these two defines the stakeholder's "flexibility" in adjusting the energy footprint of the system. Such flexibility cannot only be considered in order to lower costs but also to increase revenue in other settings. So the public lighting system could act as an energy balancing partner in various settings e.g. turn-on consumption in case of significant energy availability e.g. from wind parks or adjust its behaviour also in correlation with energy prices benefit from it [47].

Providing some insights on the role of public infrastructures such as the PLS is a timely issue, as in municipalities cost-effective approaches to provide a public service but reduce the costs are sought [131, 132]. However, existing approaches typically target the reduction of usage (in order to lower the cost) e.g. in several cities in United Kingdom, public lighting system parts are simply turned off in the after midnight hours or significantly dimming the lights (as reported by newspapers e.g. in Table 2). Over England and Wales over half a million street lights are switched-off in order to save money. This approach has created in many cases a public outcry as the fear for impact on civilian safety is debated. Apart from safety [133], full street lighting goes beyond practical issues (e.g. road safety, crime etc.) and also addresses social aspects.

Table 2.: Public Lighting System turn-off to reduce costs in UK. Source: Daily Mail Newspaper, 09 July 2011

City	# Lights	Cost Decision Taken
Buckinghamshire	1600	switched off after midnight
Cornwall	30000	dimmed
Durham	12000	dimmed
Essex	91000	switched off after midnight
Gloucestershire	15000	dimmed or switched off after midnight
Leicestershire	51000	dimmed or switched off
Norfolk	27000	switched off 00:00–05:30
North Yorkshire	30000	to be switched off after midnight
Nottinghamshire	90000	to be dimmed or switched off
Suffolk	40000	dimmed or switched off

Apart from centrally controlled decisions to turn on/off the lights based on time, some others have experimented with user-driven management. For instance in some cities in Germany (e.g. Lemgo) citizens may turn on the lights across a street by sending Short Message Service (SMS) via their mobile phones (each street light has a 6-digit code that is sent to a centrally administered number). Other approaches try to reduce consumption by combining factors such as pedestrian flow with safety guidelines [134]. Although such solutions prove the concept, having not fully automated systems is not to be expected to be adopted by consumers [36]. A more pragmatic approach is that of dimming the lights, which attempts to provide a trade-off between cost and usage. Today, with the prevalence of Light-Emitting Diodes (LEDs) used in public lighting systems, this makes increasingly sense, not only because of the overall energy savings (which could be in the range of 40% [132]), but also the additional capabilities they provide in flexibly managing the system. Control, by simply turning on/off specific LEDs within a street light hence dimming it, can be easily applied, and can be done instantly due the very fast reaction of LEDs on the power-on/power-off signals. Other, more advanced solutions, involve usage of street sensors and adjust overall lighting based on requirements for the lighting conditions i.e. weather, traffic, etc. and even the human visual perception [135].

These approaches target again locally autonomous systems for reducing energy consumption according to the current conditions. If *scenario* 1 from Figure 59 is applied, several steps need to be taken in the bilateral communication between the DSO and the Public Lighting System. This will lead them to an agreement on the future behaviour and benefits for both of them. In the specific case, where bidirectional communication between the DSO and the PLS exists, the PLS offers its flexibility, while DSO proposes the reduction of the energy signature of the PLS. Independent of who actually initiates the communication, the PLS firstly assesses its own energy requirements, in order to understand the available flexibility that can be negotiated. Subsequently it requests from the DSO potential flexibility curve as well as a price curve describing the cost range for each adjustment. The DSO makes a potential offer on the adjustment and prices willing to pay (he may coordinate with multiple other stakeholders), and then the PLS does a cost benefit analysis to assess his situation. Finally for the cases where a positive cost benefit analysis is achieved, the PLS negotiates with the DSO the behaviour adjustment, which at the end is sealed with a micro contract for the expected behaviour. This interaction in between two parties is depicted in Figure 60.

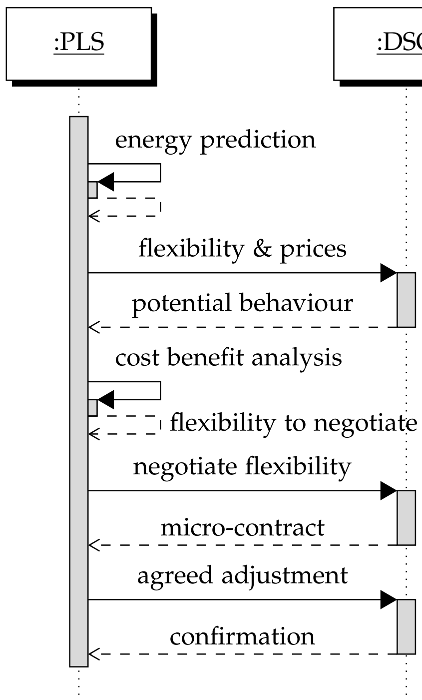


Figure 60.: Energy flexibility negotiation according to *scenario* 1

In *scenario* 2, the overall goal of trading the flexibility available on the prosumer side (in this case the PLS) in order to create new revenues is investigated. In the NOBEL project, prosumers make use of a local energy marketplace where energy can be bought and sold. A service platform, globally available to all prosumers, is offering energy services for real time monitoring, management, billing etc. as well as a marketplace [23] has already been evaluated [38]. Here several interactions are possible, and one such is depicted in Figure 61. The PLS system may subscribe to informational events coming from the market itself and delivered via the energy platform services. Such information includes current energy prices, historical information, available buy/sell offers etc. Together with information obtained from the PLS, e.g. flexibility assessment, a cost benefit analysis can be made and then a trading strategy is defined. Once the decision is taken in the PLS side, it can configure an agent (as offered by the platform [25]) who takes over and tries to satisfy the behaviour defined by the PLS. This could be for instance a way to procure energy at the lowest possible price or sell the flexibility of the PLS at the highest possible one. The PLS manager can receive the notifications and performance of the agent in his monitoring screen [34], while the automated PLS management system adjusts the behaviour of the PLS to correspond to the results of the auctions on the marketplace.

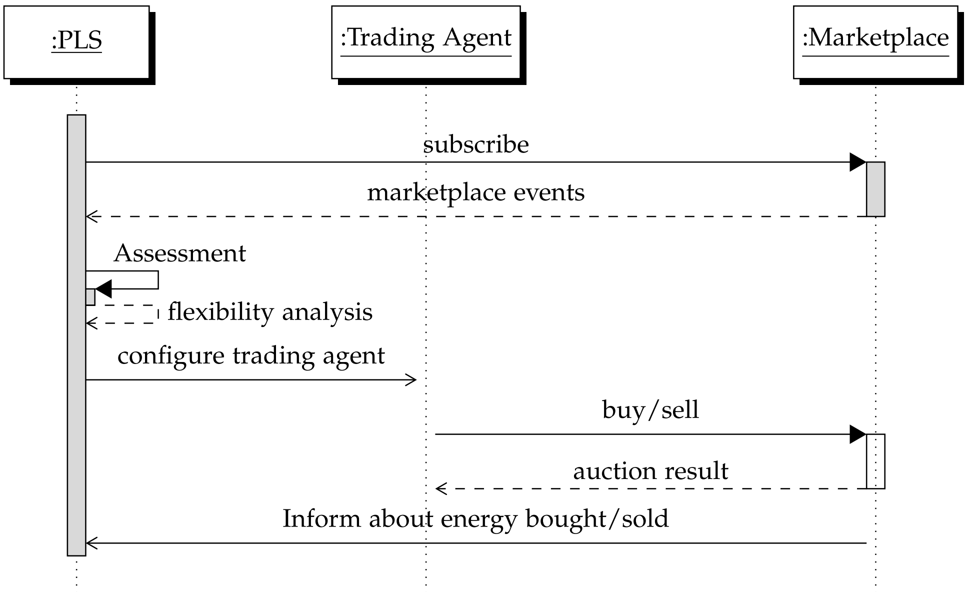


Figure 61.: Energy flexibility negotiation according to *scenario* 2

Figure 62 depicts a result of such scenario interactions. As shown, although the PLS was willing to adjust its behaviour at a much lower price, the final agreed price was generally higher, which yields out some additional financial benefits for the PLS. For some slots where no consensus was achieved, as there was a significant difference on the conditions the transacting partners had set. For the latter, no flexibility has been traded (although available from the PLS side), as it is not a financially viable solution for the PLS, hence no adjustments on the energy signature of the PLS will be done.

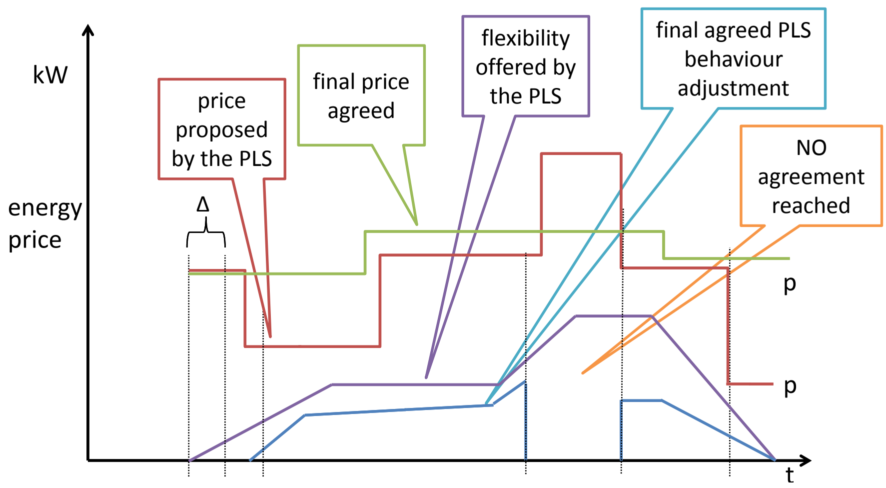


Figure 62.: Public Lighting System negotiation for intervals of duration Δ

5.2.4 *Summary*

While several other scenarios are possible with a varying degree of complexity, it is important to understand the huge potential brought by *scenario* 2. Not only sophisticated strategies may be defined but also market/economic models and strategies can be utilized similar to what is done in stock exchange. Hence we move towards a highly dynamic system that may readjust itself according to the interactions of its stakeholders and is business driven. Additional levels of interactions may be introduced between the stakeholders with the pros and cons that they bring with them e.g. as shown in *scenario* 3 where prosumers enable aggregators to act on their behalf. All of the example scenarios mentioned are complementary and can co-exist.

Simplicity of the process made by stakeholders as PLS can be easily adopted to offer flexibility. Furthermore, high predictability of such stakeholders allows them to be measurable and therefore one can verify their execution [12]. Although other stakeholders may like also to offer flexibility, they may suffer from extreme management complexity of their internal processes to offer it. On top of this they may be unpredictable, what adds additional requirements to the grid operations, as energy supplied and demanded need to be in balance. This, however, can be overachieved if the forecast uncertainty is addressed with stakeholder aggregation (as in section 4.2) and their empowerment with storage solutions, as section 4.3 suggests. In such a setup, the flexibility offered by stakeholders can be based on the current SoC of their solution, reducing the management complexity of their internal processes.

5.3 ENABLING DETERMINISTIC ENERGY BEHAVIOUR OF STAKEHOLDERS

A typical industrial building can be seen as an ecosystem [136], its internal (e.g. building infrastructure) as well as the new extended components (e.g.

electric vehicles, static storage etc.) can cooperate to improve energy management [49]. This in turn can enable new forms of business interaction with other stakeholders that are either currently impossible, or incur high integration costs. Of particular interest is a facility's ability to keep-up with previously planned [90], or forecast [109], levels of energy consumption and/or production, and its flexibly in adjusting to new situations while trying to minimize costs, or increase revenue for its owners [92]. By being able to perform a reliable forecast, such a facility could generate revenue through effective participation in, for instance, local energy markets [25], or demand response programs [67].

Forecasting the electricity consumption and/or production (either internally or externally) behaviour of a stakeholder will of course lead to errors [26]. However, future on-premise capabilities, such as on-site energy generation or EV fleets [32], will provide such stakeholders with new business and management opportunities [137]. As such, these load imbalances may not need to be propagated to external stakeholders as it is done today. The challenge is on how to leverage the facility's capabilities [138] and external interactions in order to bring benefit to the stakeholder in ownership. More specifically, how the existing and new assets that are under the control of the facility management can be empowered with Smart Grid technologies and services, and be effectively used to address any energy shortage or excess caused by the on-site prediction errors [95].

To address this problem, this section proposes a system that takes advantage of existing (including temporal) assets and Smart Grid services, and enables facility management to actively adjust its energy consumption/production behaviour as seen by external stakeholders, while adhering to its internal goals and strategies. The proposed system considers a stakeholder with variable storage and energy trading capabilities, which may be the norm in the years to come. Also several management strategies are described to demonstrate capabilities that can be realized with such a system. Although individual aspects may exist in ongoing research work, the proposed system combines several of them together i.e. forecasting, storage and trading. Hereby clear applications in facility management (i.e. industrial buildings) will be tackled, but one can apply the same observation to any other stakeholder, or a cluster of them [69].

5.3.1 *System Architecture*

The proposed system is modular and designed to empower the collaboration of the independently operating sub-systems, as well as the homogenization of their functionalities in a mash-up end-user application. As depicted in FMC notation (`www.fmc-modeling.org`) in Figure 63, one can distinguish the interactions of the end-user via the cockpit, the core system components involved in the back-end i.e. Energy Load Forecast (ELF), Variable Energy Storage (VES) and Energy Trading (ET), as well as the reliance on external parties such as energy market or an external energy stakeholder.

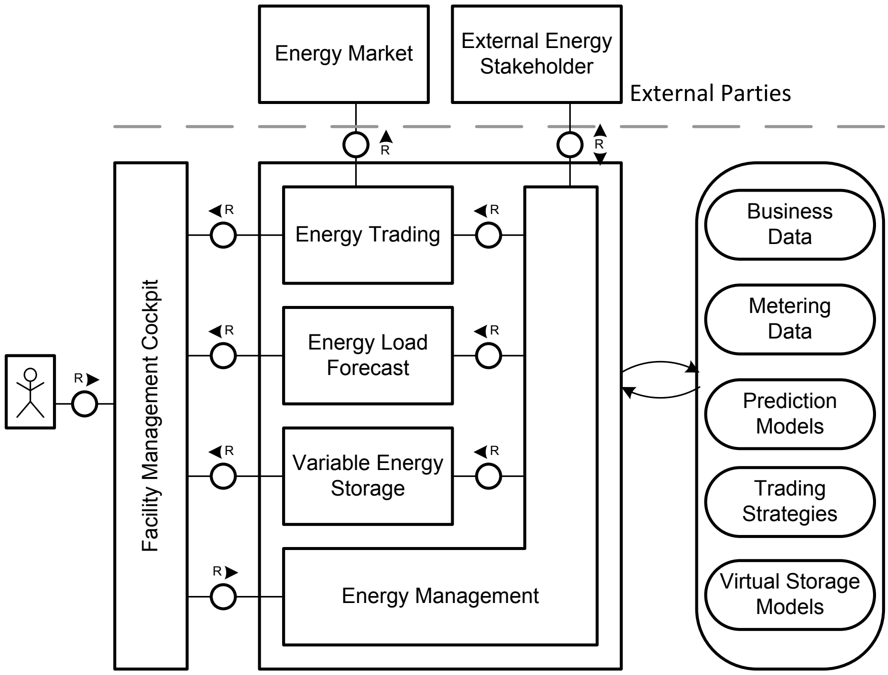


Figure 63.: System architecture overview

The *Energy Load Forecast (ELF)* is the sub-system responsible for forecasting the energy signature based on historical smart metering data (residing in the metering system) as well as real-time data acquired by the infrastructure [74]. Its results form the basis for the decision making process of how to handle the excess or shortage of energy predicted.

The *Variable Energy Storage (VES)* consists of managing the available "storage" on premise [69]. The latter may include static as well as dynamic energy storage (such as a fleet of EVs) within its overall constellation [43]. The VES is also envisioned to have the capability of managing processes that could store or re-feed energy, or even to reschedule charging/discharging of storage units and virtually absorb a desired load.

The *Energy Trading (ET)* is able to trade energy on smart city marketplaces, that is, intelligently buying or selling energy depending on the needs of the overall system [27].

The *Energy Management (EM)* is a coordinating entity which enables the collaboration among the different sub-systems, in this case ELF, VES and ET, while in parallel taking the decisions on the actions to be enforced. Based on the enterprise goals and strategies set by the facility manager, it may dynamically decide between the portions of energy that can be "stored" in the VES or traded by ET in an electricity market. In fact, as trading on energy markets of today is done in blocks of units, forecast accuracy gets further challenged and therefore tightens the relationship in between VES and ET.

The *Cockpit* is the User Interface (UI) that the end-user, i.e. the facility manager dealing with the energy related aspects, interacts with. The cockpit is envisioned as a mash-up application depicting key aspects of the status of the underlying infrastructure, including enterprise related key performance indicators. It can depict in real-time all information related to the utilization of the storage, the energy forecasting as well as the achieved energy accuracy, the energy traded and related costs, the currently available and followed energy management strategies etc. The cockpit is considered to be easily realised as a web application hosted in the cloud, easily accessible via the browser e.g. of a mobile device or laptop.

Finally, the system is envisioned to be able to communicate with external parties and services such as an *energy market* and *external energy stakeholders* in order benefit from its active participation on power networks. This also implies the role of being part of a larger ecosystem and the capability of being easily integrated in its business processes; for instance the goals pursued by the facility management could be adjusted to reflect dynamically changing enterprise needs.

Energy Load Forecast (ELF)

Forecasting is a well-known component of every energy management system. Imbalances provoked due energy load forecast errors may result in a shortage or excess of energy that must be accommodated, e.g. in form of charging or discharging a battery. In order to perform the forecast, the ELF requires the availability of the actual energy load $y[n]$ of a stakeholder (an interval n of size T) in the past, i.e. its smart metering data, and potentially other information such as weather data, asset specific behaviour or participation in processes, etc. Once current energy load is measured $y[n_0]$, the self-forecast can be reported with a minimum offset of Δ, thus always reporting a single interval load $\dot{y}[n_0 + \Delta]$. The reported value $\dot{y}$ is always suffering the error of the offset as can be observed in Figure 64.

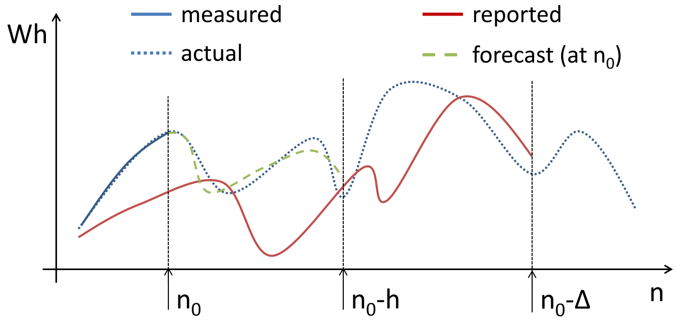


Figure 64.: Forecasting on different horizons and intervals to improve the forecast accuracy

Additionally, ELF utilizes advanced forecasting algorithms that continuously provide accurate predictions $\hat{y}[n]$ on smaller horizons (as also indicated in Figure 64) Its accuracy depends on multiple conditions such as the applied forecast algorithm, the required horizon etc. In this work, ELF provides a forecast of the system for any horizon h in the future, so the continuous load forecast is done for an interval $\widehat{y}[n-h]$. It is expected that many components of the system will require different horizons, as $h \approx \Delta$ might not be of interest. Figure 64 depicts the accuracy of $\widehat{y}[n]$ as being higher than reported load $\dot{y}[n]$ (since $h < \Delta$). The ELF configuration is expected to be done internally based on historic accuracy of the achieved performance.

Variable Energy Storage (VES)

The described in section 4.5, this component combines both static and dynamic storage into one (virtual) unit of capacity. If performance degradation is not considered, a static storage has a constant capacity. In contrast, dynamic storage is composed of multiple (potentially mobile) units that are at some point in time connected to the grid [43]. While static storage can charge or discharge in dependence to its actual SoC, these dynamic units are the actual energy-flexible components, when of course connected to the grid. This flexibility is gained by controlling the amount of energy that they charge or discharge as well as rescheduling such activities over an interval n of length T. As it will not always be possible to compensate the exact energy needed, e.g. due to technical restrictions, on every reschedule request, the error that should be absorbed $\dot{s}[n]$, is not necessary always to be fully addressed, but reduced to what is actually stored $s[n]$. This gap can be however improved with properly managing the variable storage as a whole, since it is combined from its dynamic and its static part (which does not have the same temporal restrictions). A potential usage of the VES might be to use its dynamic part to compensate the closest value possible, while the static part can correct the uncompensated part of the error by charging or discharging the amount of energy needed. However, the exact usage may depend on various other technical or financial constraints, and is out of the scope of this work.

A stakeholder owned EV fleet (for which it is assumed that the facility management has full control over) is a good example of the dynamic part of a variable storage, while it is limited by scheduling and vehicle restrictions. For the rescheduling step, different priorities will need to be satisfied in order to ensure that these EVs are always within the fleet requirements. As such, any EV fleet can be used to calculate the maximal shiftable load to positive $\Delta s^{+}[n]$ and negative $\Delta s^{-}[n]$. For this calculation it has to be considered, that EVs can only vary their charging between the maximal and minimal power, or interrupt the charging completely. Within these limits, the fleet can react on energy shortage or surplus at stakeholder's premise, e.g. by interacting with an energy market or even compensating forecast errors by rescheduling or shifting loads for different intervals n. Therefore in case of an energy demand change, discharging of EVs would be a secondary option, while rescheduling has precedence. This is mainly possible due the great capacities EV batteries (so they can complete multiple trips), thus a VES can artificially deliver a higher round-trip efficiency as no losses are made due to the charging/discharging.

Energy Trading (ET))

Local energy markets may emerge as a scalable methodology for controlling the levels of consumption and production on the grid [23, 99], in particularly as a response to the increasing deployment of distributed energy resources e.g. PV panels, wind farms, μCHP generators, etc. Within the architecture proposed in Figure 63, a local energy market is considered as an opportunity for a stakeholder not only to maintain its predictability, but to also, in some cases, better utilize and capitalize on its storage facilities. With that in mind, the ET system component interfaces with the local market to buy/sell energy by applying different trading strategies, such as [101, 56].

The stakeholder calculates, on an interval basis, the energy trading target $\dot{\tau}[n]$ based on its internal strategies and goals. For instance, the trading targets could be based on the forecasting errors provided by the ELF. A limit price, $\tau_p[n]$,for either buying or selling is optionally set with each target to indicate the maximum (minimum) buying (selling) price for an interval n. If the pricing information for a particular interval is undefined, the ET will trade aggressively on the market to ensure that the targets are met, so $\tau[n]$ presents the net quantity traded by the ET with the interval. Otherwise, each target can only be met within the bounds of its pricing constraints.

Current targets can be updated as more accurate information from the system components is available, such as forecasts of $\widehat{y}$ to better assess the error of $\dot{y} - y$. In such cases, the ET updates its market position to meet the new targets. For instance, if the target is set to $\dot{\tau}[n] = 50$ Wh, of which current trading is $\tau[n] = 20$ Wh, when a new target of -30 Wh is received, the ET should then sell $\tau[n] = -50$ Wh to meet the new target. The performance of the ET can be tracked by requesting the total traded quantities $\tau[n]$. Furthermore, for purposes of a cockpit (thus assistance to an operator), the ET provides interfaces to access the overall market information, as prices $p[n]$ and trading volumes per time interval.

Energy Management (EM)

Although the system is expected to act autonomously, as illustrated in Figure 63, a facility manager can interact with it via a cockpit. An example of such a cockpit and information it offers is depicted in Figure 65. The facility manager can consume the (real-time) information depicted and by calibrating or setting the overall goals can exercise high-level control over the infrastructure. Such goals could be the optimization of the infrastructure reaction to the energy surplus or shortage reported by ELF towards economic objectives such as minimization of cost. Other, corporate social responsibility related objectives could exists, e.g. maximization of usage of green electricity or even simpler ones such as making sure that the EVs of the employees are fully-charged by the end of their workday.

The transformation of user goals (calibrated via the cockpit) to strategies are processed by the EM, which takes into consideration all other constraints of the system and takes the overall decision on the appropriate strategies to be followed.

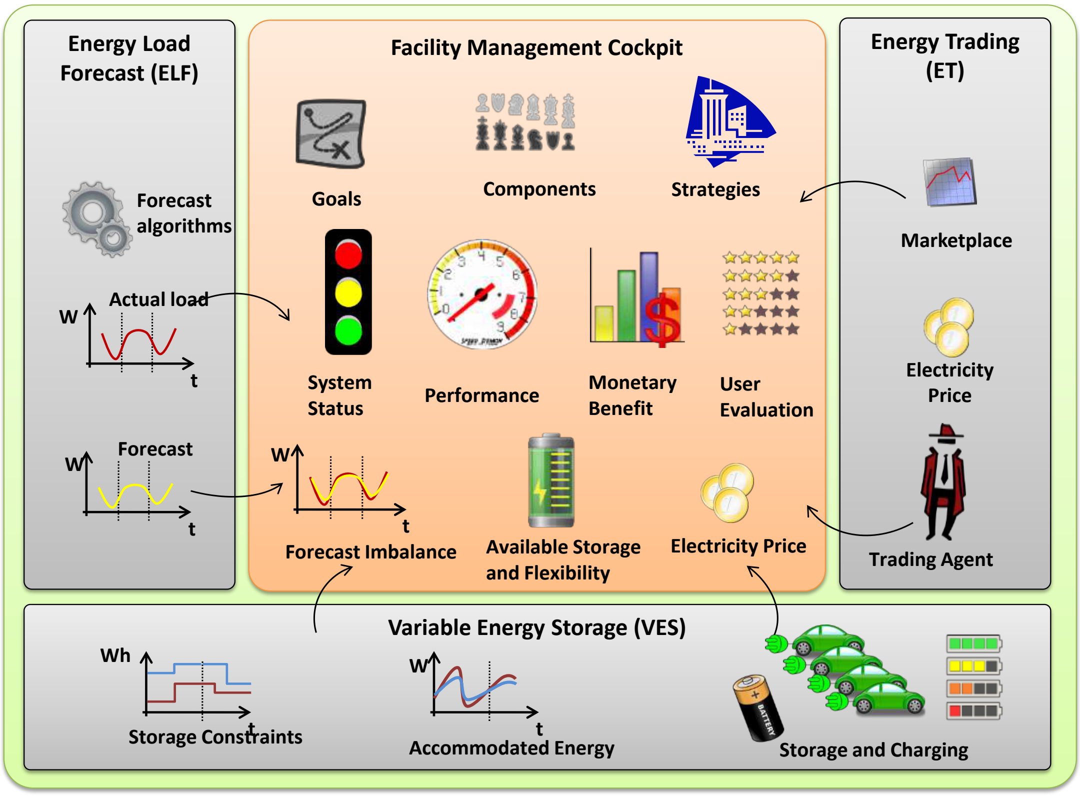


Figure 65.: The envisioned system cockpit as a mash-up application

EM acts also as a communication broker among the different parts of the system as it holds the system-wide knowledge that is not available to the individual parts i.e. the ELF, ET, VES, enabling the latter scaling or extension of the system with other components or variations of the existing ones.

EM acts as the coordinator and decision engine, which communicates with ELF, ET and VES, and provides them with the operational context info. As an example, in a scenario where the EM is informed about the energy surplus available due to a forecast error, it may decide to redirect part of it towards charging the EVs while another part may be redirected to the ET (by charging schedule adjustment) in order to be traded to the market (because the price is high or can not be covered wholly by the VES). It is important to note that trading (as of today) is done in energy blocks, thus consumer and producer sometimes cannot ideally match their market orders [23], while energy price of a producer is appealing. In such cases, as Appendix A describes, VES operation my not even be required to address $\dot{y} - y$, but ET could sell the energy in favour of stakeholder's revenue. In section 5.3.2, some envisioned strategies are presented, while deterministic behaviour of a stakeholder is kept as system's primary objective.

5.3.2 *Energy Management Strategies*

The system proposed, whose main components are illustrated in Figure 63, is flexible enough to accommodate several envisioned scenarios, depending on the goals set by the user, the available at time capabilities, and actions to be enforced. The scenarios this section focus upon, are in no way exhaustive, but serve to provide some understanding of the potential strategies that could be followed by the facility management. The aim is to showcase the system's flexibility, which is a key part of realizing agile enterprises in the future.

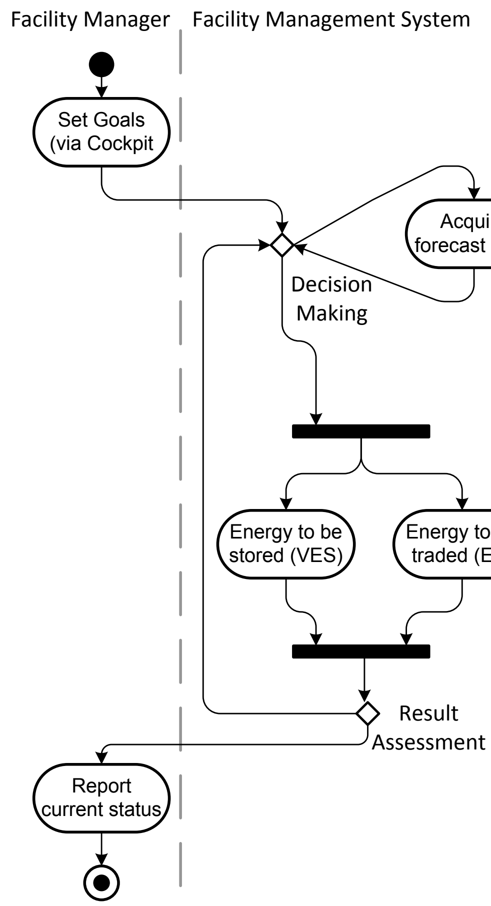


Figure 66.: General view of the activity involving the architecture components

A general view of the workflow is depicted in Figure 66. The user input is acquired, which together with the forecast error and the underlying status and constraints of the sub-systems, are used to reach a decision for either trading or storing of energy (or both). Some key strategies will be discussed in more detail here, while others can be found in [94].

Generally, each envisioned strategy may not involve all parts of the system, as this depends on the actual constraints imposed at the time of the decision making. This also signals that an organization does not have to wait until all of the architecture parts are deployed and become operational to start realising (a limited set of) energy management strategies. As an example, the ELF and the VES could be realized today, as will be presented in chapter 5, while the ET could be realized some years later when energy markets are available at smart city level and it makes economic sense for the facility managers to participate in them. Hence, the system architecture accommodates the "migration" i.e. incremental evolution of the infrastructure towards the fully-fledged Smart Grid vision.

Storing Energy not Traded by ET

The decision making process depicted in Figure 66 may consider a strategy that is described as follows: *after the estimation of the energy error within an interval by the ELF, try to trade the difference via the ET and differ any non-traded energy to the VES for storage*. The workflow of such strategy is illustrated in Figure 67. ET accommodated $\tau[n]$ for the interval in question (potentially even at different prices p), and VES is contacted in order to absorb the remaining $\dot{s}[n]$.

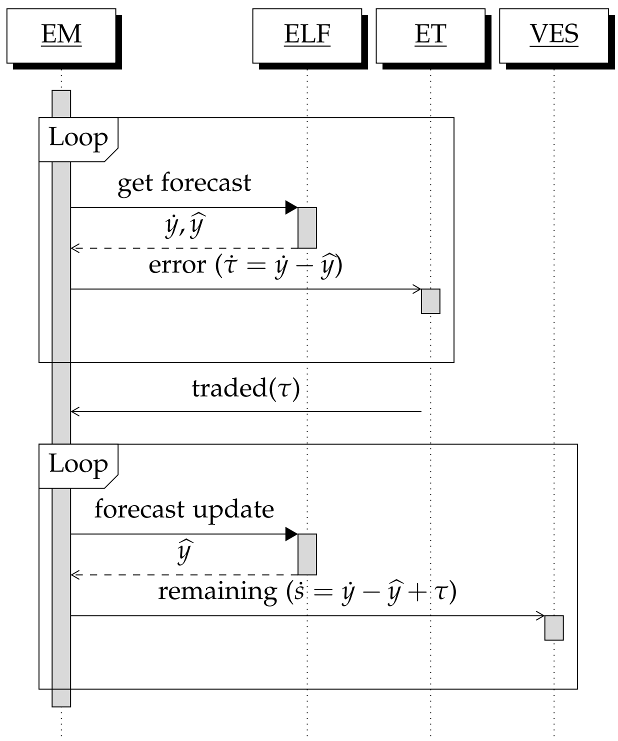


Figure 67.: Strategy of Storing Energy not Traded by ET

The actual outcome of the trading done by ET depends on the real-market conditions (law of demand and supply) and hence strategy adaptation might be needed over time e.g. acting alone or as part of a larger group [25]. As the ET might not be able to fully trade the energy needed to balance the forecast error $\dot{\tau}[n] = \dot{y}[n] - \widehat{y}[n]$, a part of it remained non-traded. The traded quantity $\tau[n]$ is then communicated back to the EM, which instructs VES to accommodate the remaining $\dot{s}[n] = \dot{y}[n] - y[n] + \tau[n]$ energy. This process leads to a new state where the error is minimized as a "best effort" procedure is followed by

ET (interaction with external stakeholders) and VES (internal stakeholder) to minimize its impact.

Depending on the business motivation, this strategy may be followed when the Return of Investment (ROI) by selling the energy on the marketplace is high. This may be a result of high prices on the energy market, inability or no need of storing the energy internally, etc. The actual decision-making process will be dynamic and the exact fine-tuning is not considered here.

Trading Energy not Accommodated by VES

In compliance with the decision making process depicted in Figure 66, here is the focus on a strategy that can be described as follows: *after calculation of the energy due to the incurring error by ELF, try to accommodate the excess or shortage of energy via the VES and for the remaining part not accommodated by the VES, use the ET*. In this strategy, the ET acts as a mitigating agent for any part of the error that could not be absorbed by the VES, which is shown in Figure 68. In detail, the EM acquires the forecasting error ($\dot{s} = \dot{y} - \hat{y}$) from the ELF and informs the VES, which attempts to accommodate the imbalances introduced by the errors, and informs the EM of any amount that could not be accommodated due to its internal constraints (s). These amounts are then given to the ET to be mitigated on the market ($\dot{\tau} = \dot{y} - y + s$).

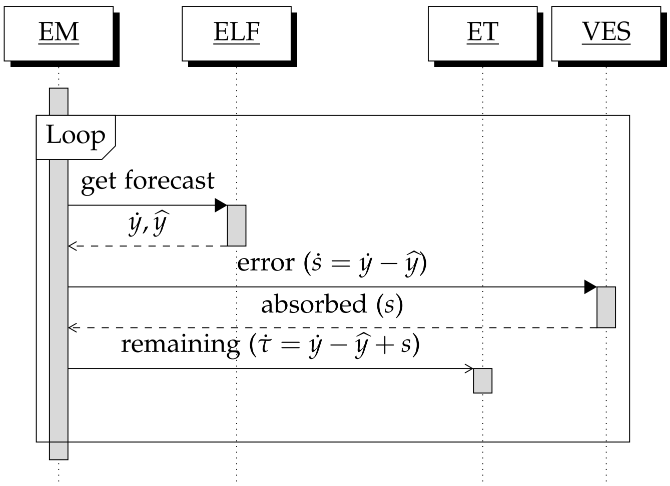


Figure 68.: Strategy of Trading Energy not Accommodated by VES

This strategy is expected to be used when the enterprise has the capability to store energy extensively for its own use. For instance a significant number of EVs at the disposal of facility management means that the VES can rely on storing energy there and acquiring it back again when needed. Even if the energy is not needed during the day for tackling imbalances, the EVs are charged and the energy can be used for the enterprise's processes in the future; an action that enhances better planning of energy-relevant actions. If a local marketplace is available, the ET tries to trade the energy difference in order to meet the reported load.

Deterministic Behaviour for Flexibility

Independent of the strategies followed by the facility management, deterministic behaviour is a prerequisite to be measurable on the grid [12]. In compliance to the decision making process, as depicted in Figure 66, once stakeholder's load is reported the decisions can be made on changing the load on demand, or so called flexibility. As an example, consider a strategy that is described as follows: *continuously report load on Δ whose forecast errors are absorbed by VES and offer load flexibility to them third parties based on the state of VES*.

This example goes beyond the traditional process of trying to cover the energy imbalances and but tries maximize revenue through the available assets on stakeholder's premise. The VES might reach high or low SOC, while trying to cover the occurring imbalances, and based on its SOC it could be transformed into economic benefit for the company. The VES may have its own models for estimating reliability of the component, and hence can act on a flexibility request of the external energy stakeholders. The system is flexible enough to accommodate such actions, as managing complexity of VES is less complex than to manage business processes of a company. However, to avoid conflicts or side-effects, additional analysis on the resource utilization is needed which is not part of the investigation presented here.

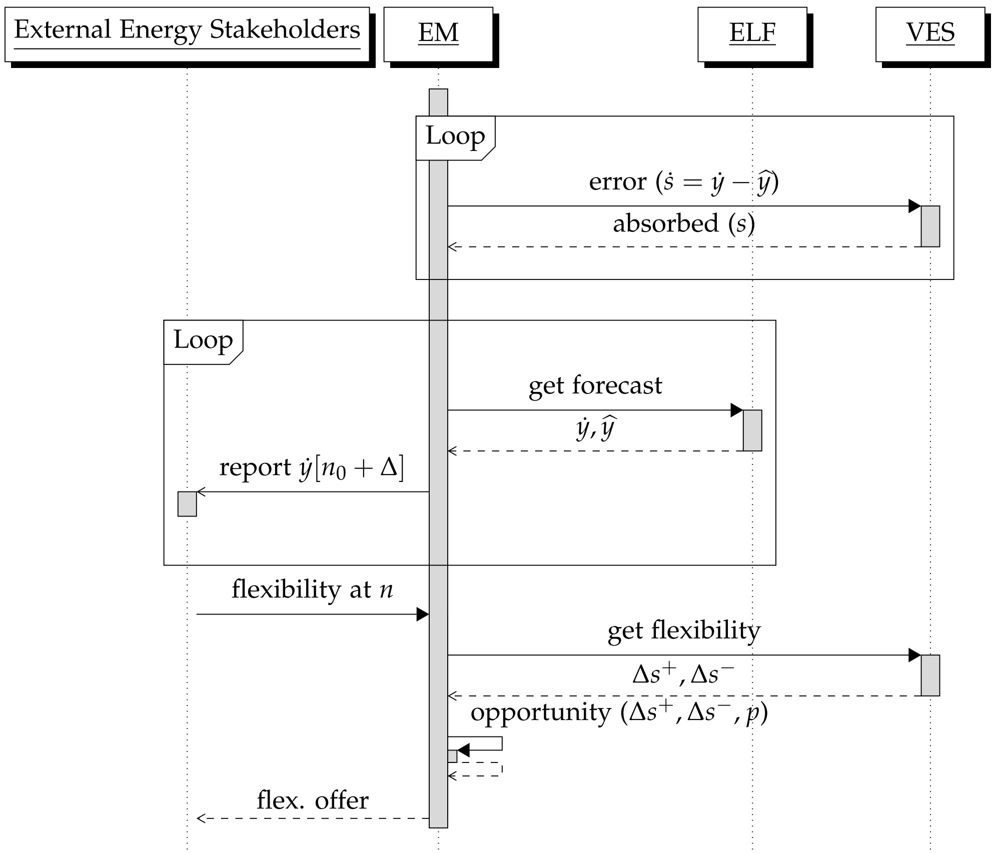


Figure 69.: Strategy of Deterministic Behaviour for Flexibility

Since the VES is trying to compensate the error produced by the ELF, a certain SOC will be achieved. Based on the actual flexibility levels with consideration of SOC, the VES can offer a certain capacity within an interval n for charging/discharging in order to increase the enterprise's revenue. Furthermore, instead of only offering flexibility based on SOC left from the error compensation, the VES may calculate the maximal and minimal shiftable energy $\Delta s^{+}[n]$ and $\Delta s^{-}[n]$. This potential flexibility can be then offered in benefit for a price p. In section 5.4 the deterministic behaviour of a real world case will be assessed, such that a stakeholder can benefit from this strategy.

5.4 Self-Forecasting Energy Load Stakeholder

Forecasting is the key part of an efficient management, and if it is reliable, improvements in planning of energy relevant processes can be realised [49]. As an example, static storage solutions are already widely adopted to balance the effects of unpredictability [64]. Others meet their planning by, if technically possible, control (e.g. starting/stopping/rescheduling) their heavyweight energy processes. The latter requires deep knowledge of processes, assets and full-understanding of interdependencies, which is a highly complex endeavour; in addition fine-grained control should also be available. Hence, the first approach which is largely agnostic to these and does not have such extended requirements i.e. making the storage solutions attractive for the strategy presented in Figure 69.

The nature of storage is however changing and potential of substitutions is already evaluated in section 4.5. Penetration of EVs and the potential coordinated usage of their storage capacity [70], poses them as attractive alternatives to traditional static storage. They become even more attractive if one considers the high cost of the static storage solution, and in particular the Battery Energy Storage System (BESS) [65]. Therefore, the non-utilized storage capacities owned and present at stakeholder's premise, such as the storage capabilities of EVs, can be considered as "wasted" resources. As organizations strive towards fully utilizing their resources, in order to achieve higher efficiency and return of investment, new solutions need to be realized.

A solution that will take into consideration stakeholder's assets and couple with state of the art analytics and forecasting, to be able to contribute to efficiency of power networks, is needed [15]. To meet those requirements, the system from section 5.3.1 is proposed for enabling facility management to achieve deterministic energy behaviour; however that system includes several futuristic aspects such as the capability of trading on local energy marketplaces [94]. Here the focus is on the subset of that proposal that can be realised via VES and this section evaluates the key performance indicators of such system. A stakeholder adopting this specific solution is hereby referred to as a Self-Forecasting EneRgy load Stakeholder (SFERS). It is envisioned that they fully utilize the capabilities of Smart Grid and modern IT systems including smart metering, EV integration, energy load reporting, real-time forecasting and management etc. Experimental results show that such system has a promising potential for real world cases.

5.4.1 *The SFERS system*

The system is expected to be utilized by a stakeholder (e.g. managing a commercial building), or a cluster of stakeholders (e.g. active in a residential neighbourhood [69]), where its components can access to metering data, business data, and energy management agreements. It constitutes an extension of some parts of the more general architecture proposed for holistic energy management [94], also presented in section 5.3.1. Here, the focus is on evaluating the proposed architecture as a running system.

System decisions are made by the EM component, sophisticated forecast is done by ELF and management of both dynamic and static storage units is done by the VES unit. Figure 70 depicts the detailed view on architecture extension of one introduced in Figure 63. Although the system runs autonomously, the operator can interact with it via a facility management cockpit where strategy selections are made. Additionally, strategy from Figure 69 suggests integration of external services by a SFERS in order to enhance its capability on the larger vision and interworking [37], while also it can provide input to other services e.g. load reporting (instead of measuring it [139]) for DSM/DR verification.

Operational Context

To clarify the operational aspects of the system, the context and issues it addresses are here described in greater details. The forecast horizon h is the future number of intervals for which a demand forecast is generated. If an energy load $y[n]$ is forecasted from time series and is executed at interval n_0, the return forecast series $\widehat{y}[n]$ are for all $n \in [n_0 + 1, n_0 + h]$. Greater horizons are expected to result in higher errors, which however converge [14]. Even though a forecast can be observed by MAPE, the mean observation from the intervals in a horizon hides the actual error of the different intervals. As an example, the forecast at n_0 is expected to result to much higher absolute percentage error at $n_0 + h$ than the one at $n_0 + 1$, but MAPE in overall will hide it. Nevertheless, the internal system components as VES are expected to benefit from the continuous update of $\widehat{y}$, as scheduling algorithms may use it to better address the errors of ELF.

Since the SFERS system will report the energy load in the same fashion as a smart meter with an offset, a new parameter is introduced. The offset is observed as time, and is linked to the metering resolution, thus it is observed via $\Delta \in \mathbb{N}_1$ i.e. at end of an interval n_0 the load forecast for $n_0 + \Delta$ will be reported as $\dot{y}$. As an example, at resolution $T = 15$ minutes, the five hours offset will have $\Delta = 5hours/15min = 20$. In this way all the intervals reported will suffer from the error an offset introduces. The smaller the offset, the greater is the accuracy that can be achieved. Furthermore, having greater errors at intervals

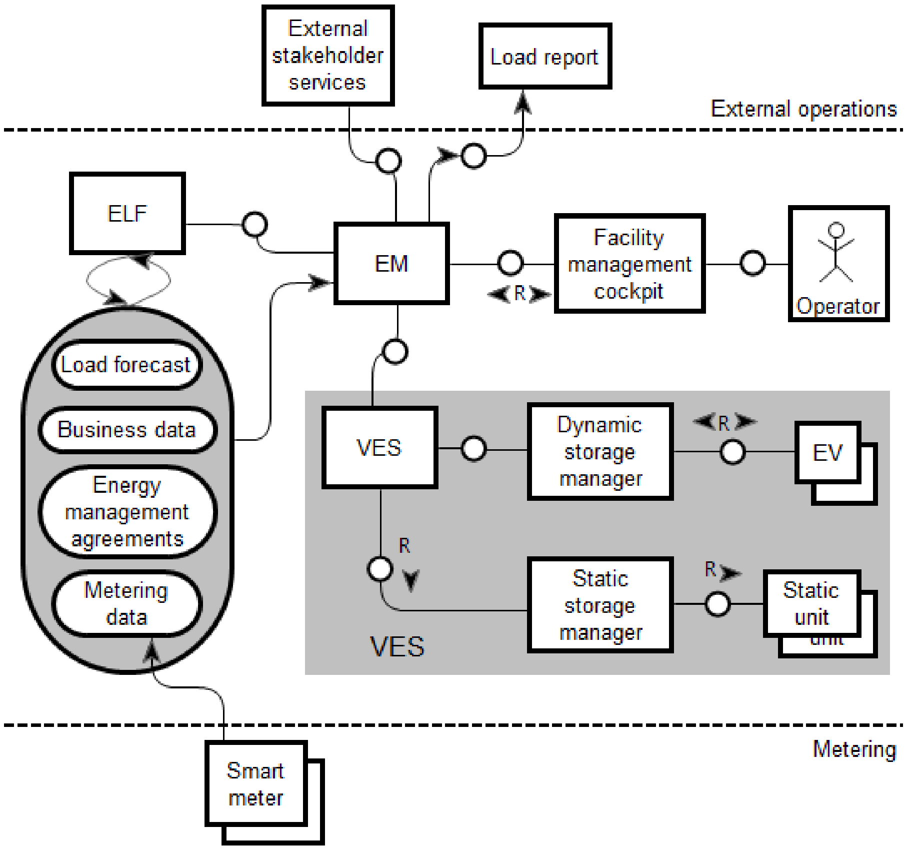


Figure 70.: SFERS system view

will also affect the capacity required by the VES to absorb them [43]. Hence the forecast will be $\dot{y}_\Delta$, such that $|\dot{y}_{\Delta_1}[n] - y[n]| < |\dot{y}_{\Delta_2}[n] - y[n]|$ is expected (but not necessarily resulting) for $\Delta_1 < \Delta_2$. Figure 71 demonstrates the effect of Δ on intraday intervals on a real-world example.

VES Controller

The management of VES is inevitable for a live system [43], in particular when a dynamic unit of VES is disconnected and connected with its individual SOC. The evaluation shown in section 5.4.2 addresses this issue by simulating individual units. For the actual management of VES, a controller for charging/discharging connected storage units and storage load adjustment (required to keep the SFERS system reliable) are required.

Since the strategy from Figure 69 makes SFERS highly dependent on VES, charging schedules of EVs have a significant impact. In this section the individual SOC of a unit is considered through entire period of the evaluation, where error from reported load $\dot{y} - y$ is changing the state of VES; thus an algorithm has the goal to keep it reliable. In order to achieve such goal, algorithm 1 is utilized, where maximum charge (positive) and discharge (negative) are the steps of change for a unit. Although better solutions may exist, algorithm 1 distributes the SOC equally such that theoretical assessment for SFERS can be

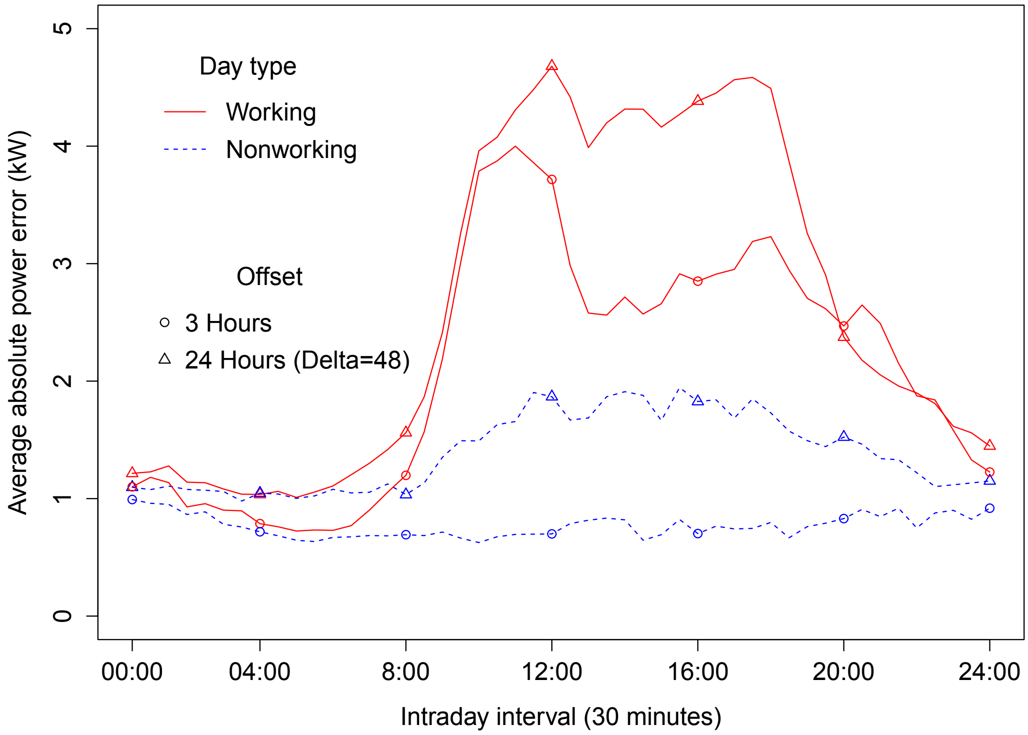


Figure 71.: Impact of different Δ on forecast accuracy over intraday intervals

made. Further investigations on improving the proposed algorithm, such as by driver requirements, are expected in future work (as noted in section 5.5).

As VES absorbs the forecast errors of $\dot{y}$, its SOC is affected. Furthermore, every individual storage unit which is part of the VES, has its targeted SOC, independently if it is static or dynamic. Hence, to keep the system in balance through the entire period of the evaluation (e.g. one year), every report $\dot{y}$ need to be adjusted on requirements of the VES. In this work, the SOC-based adjustment of $\dot{y}$ is made for any offset value Δ. The load requested by VES is based on the current SOC of storage units that are available at $n_0 + \Delta$. The $n_0 + \Delta$ interval takes an equal fraction over Δ for the forecast adjustment, and in this case in order to set the SOC to 50%. The adjustment controller is mathematically described as:

$$\dot{y}[n] = \widehat{y}_\Delta[n] - \frac{1}{\Delta}\left(\bar{c}[n] \cdot \overline{SOC}_n[n-\Delta] - \frac{\bar{c}[n]}{2}\right), \tag{14}$$

where $\bar{c}[n]$ is the available capacity at n, e.g. coming from connected cars, and $\overline{SOC}_n[n-\Delta]$ is the SOC at $n-\Delta$ of all units available at interval n. This controller resulted in good performance, however more sophisticated controlling methods could be applied [140] and need to be considered in future work.

Runtime Simulation

To assess the potential of the proposed system, all parts of it are simulated using the real data. In SFERS system, calculating $\widehat{y}$ or $\dot{y}$, the forecasting at n_0 will depend only on the actual load $y[n]$ for $n \leq n_0$. From the signal description,

Data: Connected storage units
Result: Remaining energy imbalance
while *imbalance* > 0 **do**
  get storage unit with min(SOC);
  if *stored energy* $=$ *unit capacity* **then**
    exit;
  else
    charge min(imbalance, maximum charge);
    update energy imbalance;
  end
end
while *imbalance* < 0 **do**
  get storage unit with max(SOC);
  if *stored energy* $= 0$ **then**
    exit;
  else
    discharge max(imbalance, maximum discharge);
    update energy imbalance;
  end
end

Algorithm 1: A SOC-based control algorithm for VES

an interval forecast $\dot{y}[n_0 + \Delta]$ needs to be reported once the sample at $n_0 - 1$ is available. As such, a forecasting algorithm would be executed 48 times to produce $\dot{y}$ for a day in intervals of 30 minutes. As a result, excessive times may be needed and computation requirement of sophisticated forecasting algorithms [107] may heavily impact the system performance. Hence, we decided to execute forecast on preselected offsets Δ and use them in the simulation environment. The reported energy is calculated as indicated in Equation 14, which is considered as production unit. The consumer of the simulator is the measured energy consumed by the stakeholder in evaluation. The imbalance produced is to be absorbed by VES (from the following section) or will result as an overall system imbalance – used for depicting the graphs in section 5.4.2.

For the variable storage, the actual disconnection/reconnection of uniquely identifiable units is critical [43]. A dynamic storage unit is only available within VES once it is on-premise and it holds its own SOC. As one can imagine, different scales of company EV fleets may exist and hereby they will be reproduced from real world data. Same as for results in section 4.5, the data from five "Mercedes-Benz A-Class E-Cell" (used as pool vehicles with battery capacity of 36 kWh) is used here, as well as measuring the number of traditional (individually dedicated) vehicles in the garage of the stakeholder. Both individual employee and pool vehicles are considered, where pool EVs were not directly assigned to individual employees and different mobility pattern may be expected [116]. The storage shapes used for the evaluation are presented in Figure 72.

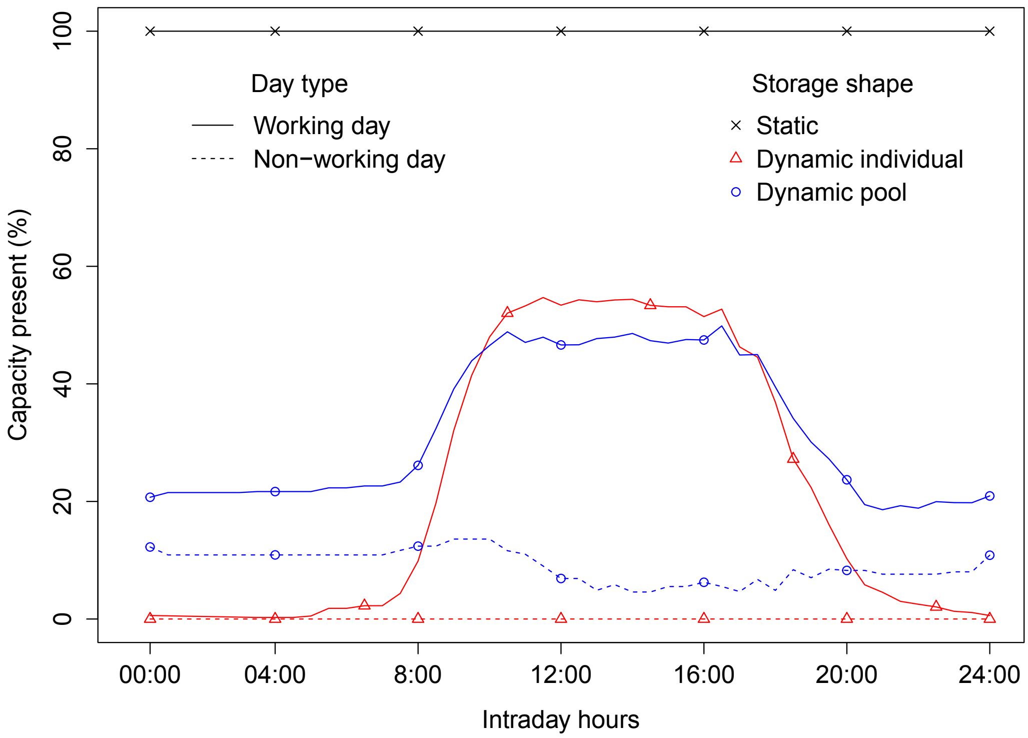


Figure 72.: Intraday presence of static and dynamic storage (from pool and individual vehicles)

As it can be seen, the static storage units are not disconnected, while the dynamic part is composed based on data from working and non-working days. Illustrated presence curves are built from few yearly schedules of 100 different units (≈ 800 charging sessions on-premise). In a continuous evaluation, these individual EVs will assist us towards producing more realistic experimental results. However, has to be pointed out that several abstractions are undertaken that may be considered as limitations. For instance, round-trip efficiency of the batteries is equal to 1. It is also assumed that all batteries allow discharging down to a SOC of 0%.

5.4.2 *System Evaluation*

The key indicators of SFERS have impact on its performance; hence the results they yield need be assessed. The experiments will use a commercial building (with offices occupied by approx. 100 employees) and assess it over the entire year 2011. Its consumption in 2011 was 234.4 MWh with an average daily power consumption of 29 kW for working and 20 kW for non-working days. This building is mainly used on working days, which are responsible for 80% of its yearly consumption. Although working hours (08:00–17:00) cover less than 26% time of a year, they are responsible for 37.8% of energy used and 50.8% of all the forecasting errors. The system is simulated on 30 minute resolution

($T = 30$) over time frame of an entire year, where different offset parameters Δ and configurations of VES units are evaluated.

Assessing Metric Impact

Using the methodology proposed in section 5.4.1, the offset effect on system efficiency is evaluated. Few standard forecasting algorithms were utilized to measure how an offset affects the Mean Absolute Percentage Error (MAPE) of the stakeholder in evaluation. Tests were made with Holt-Winters (HW) and Seasonal AutoRegressive Integrated Moving Average (SARIMA) models for weekly season, while SARIMA was also evaluated with the extra daily seasoning (also used by others [14]). In order to enable a direct comparison with evaluations of others, the offsets selected are $\Delta \in [3, 6, 12, 24, 48]$. Finally, the measured energy load from previous 4 weeks was used to train the forecast model of each interval. The experimental results acquired, for both horizon and offset forecasting, can be seen in Figure 73.

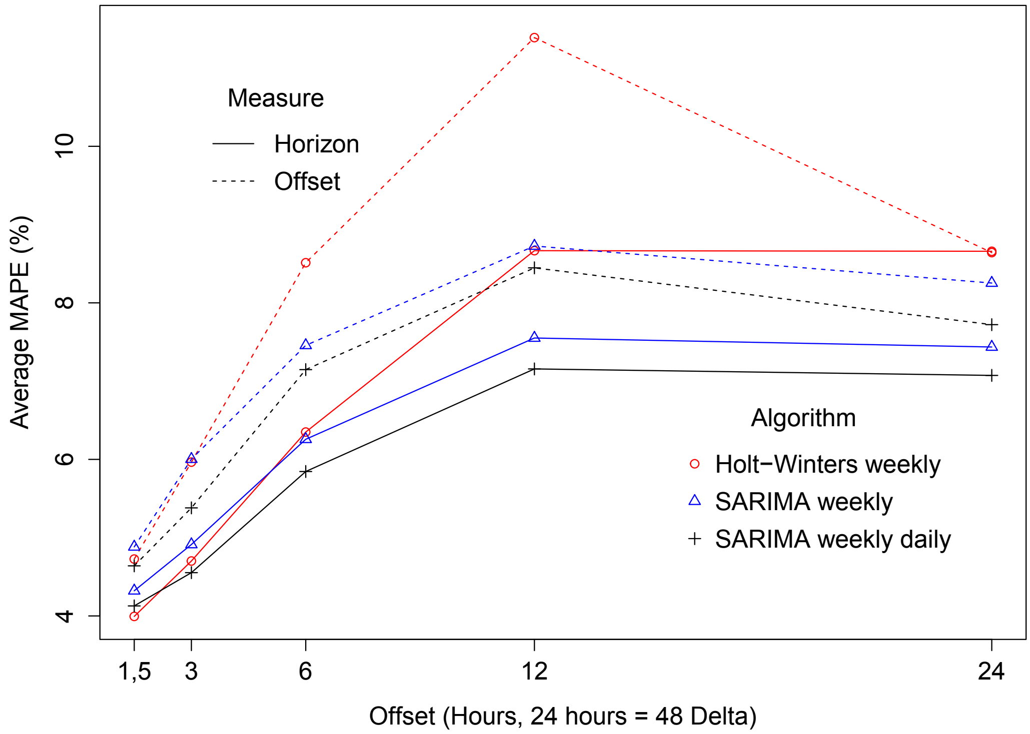


Figure 73.: Impact on stakeholder's forecast accuracy by horizon and offset forecasting

It is noted that the average MAPE of the SARIMA models resulted to lower Δ. For HW one can clearly notice significant growth of MAPE for $\Delta = 24$, or 12 hours offset, while for $\Delta = 48$ the results significantly improve. Although it is not easy to identify the reasons behind the performance degradation, one may hypothesize that it is due to the daily seasonality of the data (while only weekly seasoning is considered). Since the SARIMA resulted in better performance, SARIMA weekly was selected for the subsequent experiments with VES.

For comparison with work of others [14], Figure 73 also depicts the forecast accuracy over horizons for all Δ. In other words, the MAPE resulting of $\widehat{y}[n]$ for all intervals n forecasted from n_0, thus for $\Delta = 24$ the MAPE for horizon would be the mean of the set $\{\widehat{y}[n_0+1], \widehat{y}[n_0+2], \ldots, \widehat{y}[n_0+24]\}$. As such, when the mean value is observed, the forecasted intervals closer to n_0 will improve the overall accuracy. As we can see, function $\dot{y}[n]$ on average resulted to higher errors than from $\widehat{y}[n]$, what was also mentioned in operation part of section 5.4.1.

Absorbing Errors with a Static Storage

To improve the forecast accuracy of SFERS, experiments were conducted during which the total capacity c of VES (that is owned by a SFERS) will be increased. As depicted in Figure 72, the shape of static storage solutions is constant, thus capacity growth is linear. Figure 74 presents the forecast accuracy achieved for all offsets already evaluated in Figure 73. It is important to notice that 1% of the horizontal axis represents the capacity of $\frac{234.4\text{MWh}}{365} \cdot 1\% = 6.42$ kWh for this stakeholder.

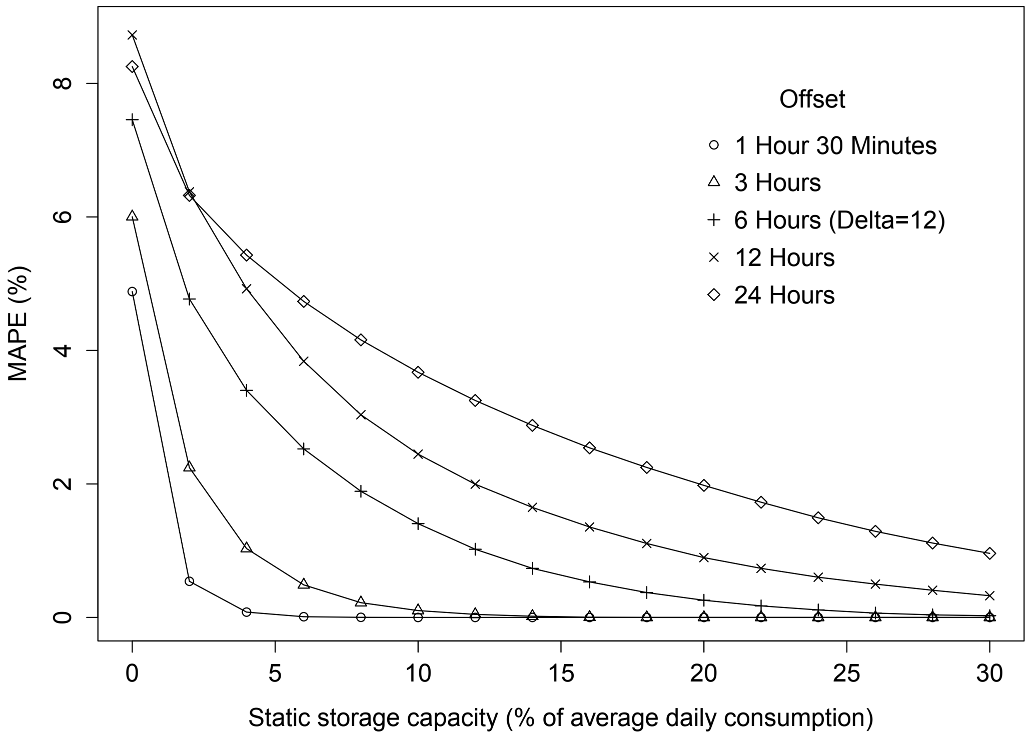


Figure 74.: Absorbing forecast errors with a static storage

The significant difference in accuracy progress on different Δ can be noticed. Interestingly, Figure 73 suggests that $\Delta = 24$ has worse MAPE than $\Delta = 48$ for the selected algorithm, while greater improve rate can be noticed. If observed through numbers, for $\Delta = 48$ at $c = 20\%$ an error of $\approx 2\%$ was measured, while for $\Delta = 24$ the same accuracy was already achieved at $c \approx 12\%$. Since MAPE for $\dot{y}$ approximates for both Δ, the VES controller was identified to be of critical importance. Of course, the controller at $\Delta = 24$ has only half the delay of $\Delta = 48$, but the capacity measured for MAPE $\approx 2\%$ is almost half as well. For all the other offsets, the VES charge adjustment (from Equation 14) brought better performance, such that SFERS in real world implementations can reach a sufficient accuracy with an extremely low c within VES.

Absorbing Errors with Dynamic Storage Units

Assessment done for static storage in the previous experiment is done here for a dynamic storage composed from both, pool and individual vehicles. In Figure 72 the average presence of pool vehicles corresponds to only 24% of the static one, while the individual one is way lower. Still, this capacity is generally considered to be available "for free" and should not be omitted. As analysed in [43], the availability of the dynamic capacity of both vehicle types is correlated with the source of imbalances and hence good-enough to address the stakeholder's forecast errors depicted in Figure 71. The results of Figure 75 show evaluation for both individual and pool vehicles using batteries of 36 kWh (or 5.6% of average daily consumption). Same as in [43], one can immediately notice how low presence fleets tend to MAPE $> 0\%$. It is important to note that horizontal axis represents the total capacity of the fleet and not only the present part nor its average.

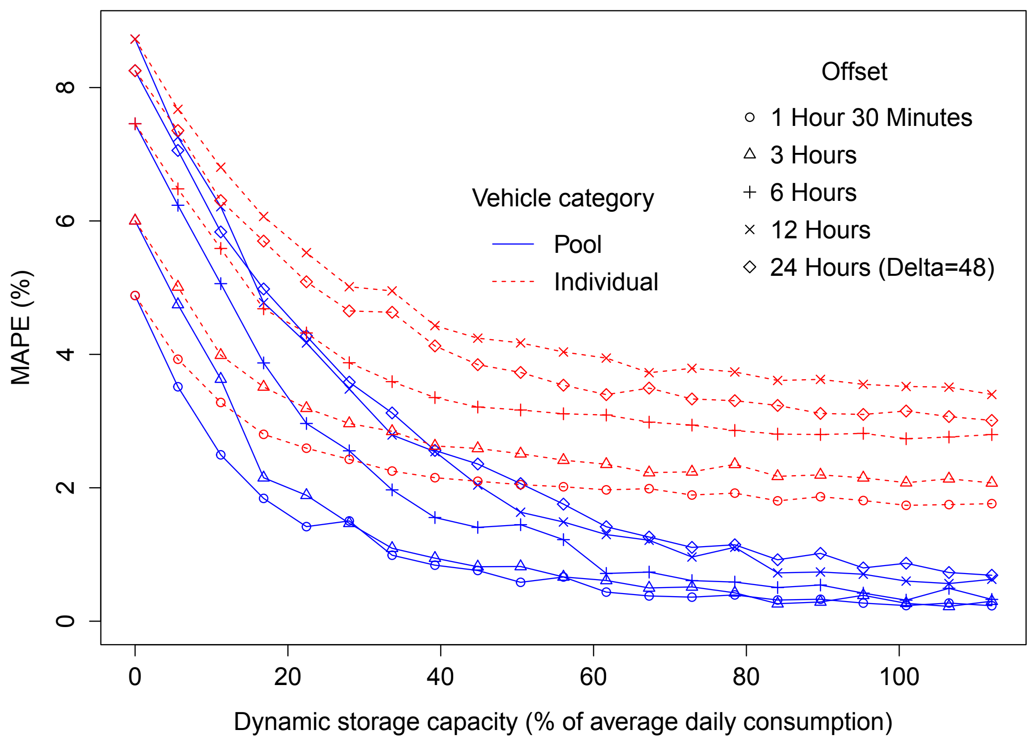


Figure 75.: Absorbing forecast errors with a dynamic storage

Understanding the relevance of results in Figure 75 to the consumption pattern of the stakeholder is important. The office has a relative fixed number of employees which corresponds to a stable number of EV cars. At some point the number of EVs may result to dynamic storage capacity that can go beyond the 100% of stakeholder's daily consumption. As shown in Figure 75, depending of the vehicle category, one can rapidly achieve the desired accuracy levels. As an example, in [69] the capacity estimated was around 12% (approx. 181 kWh), which corresponds to the capacity of only 5 EVs for the 183 households in evaluation. That work used the daily Seasonal Naïve algorithm, thus the offset is already $\Delta = 48$.

Enabling a Real World Stakeholder

As the stakeholder in evaluation is the building where author was located, it was decided to evaluate the real world case of the offices. The location has 100 employees and average presence of company vehicles on-premise was measured at 27 for peak hours on working days. According to the presence curves shown in Figure 72, the total fleet size equals to 46 vehicles, which is the reference point for evaluation in this section. These vehicles, however, suffer from zero presence for non-working hours and non-working days (74% of the time). In [43] similar cases converged to $MAPE > 0\%$ and such accuracy may not be acceptable for the SFERS system. With that in mind, the overall VES will contain a certain number of dedicated individual EVs (within the entire fleet), that will be complemented with a static storage solution. Figure 76 shows how different compositions of VES with individual EVs and different sizing of the static storage, have resulted to enhanced system reliability.

One needs to note that Figure 76 indicates the convergence of system without static storage to $MAPE > 0\%$. However, this accuracy is significantly higher than those of Figure 74. Positioned as such, one can immediately notice that only small fraction of the static solution is required. As an example, at $\Delta = 48$ with 20% of EVs in the fleet, accuracy of 1% is already achieved at 10% (64.2 kWh), while static solution on its own achieves it around 28%. This significant difference already justifies the relevance of considering the company EVs on-premise, rather than using costly static solutions [65].

An additional experiment, where individual EVs are replaced by pool EVs (such that the total number of vehicles stays the same), has been realised. In Figure 77, the assessment is depicted and the obvious impact of non-working time presence of pool vehicles can be noticed for all the evaluated cases. One should immediately notice that there is a really small initial difference if 50% and 100% of the total fleet are EVs of individuals. This is due the already fast convergence depicted on Figure 75, where there are enough cars to address all the error produced within working hours. Now, depending on the case, if pool vehicles replace the individual ones, significant impact can be already seen.

What is important here to note is that there is no management system behind the pool vehicles. If properly managed, their presence on-premise at non-working hours can be more robust by not releasing all the vehicles to employees. In this particular case, as Figure 74 suggests, if only one vehicle of 36kWh is present at

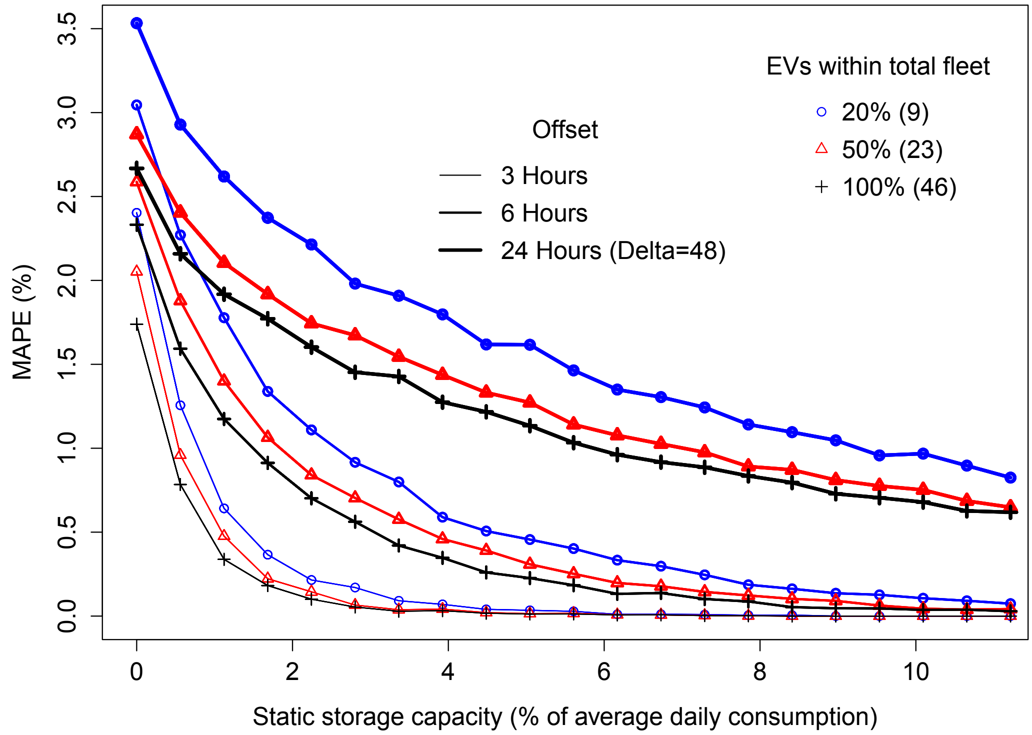


Figure 76.: Addressing low presence of individual EVs by adding static storage capacity to VES

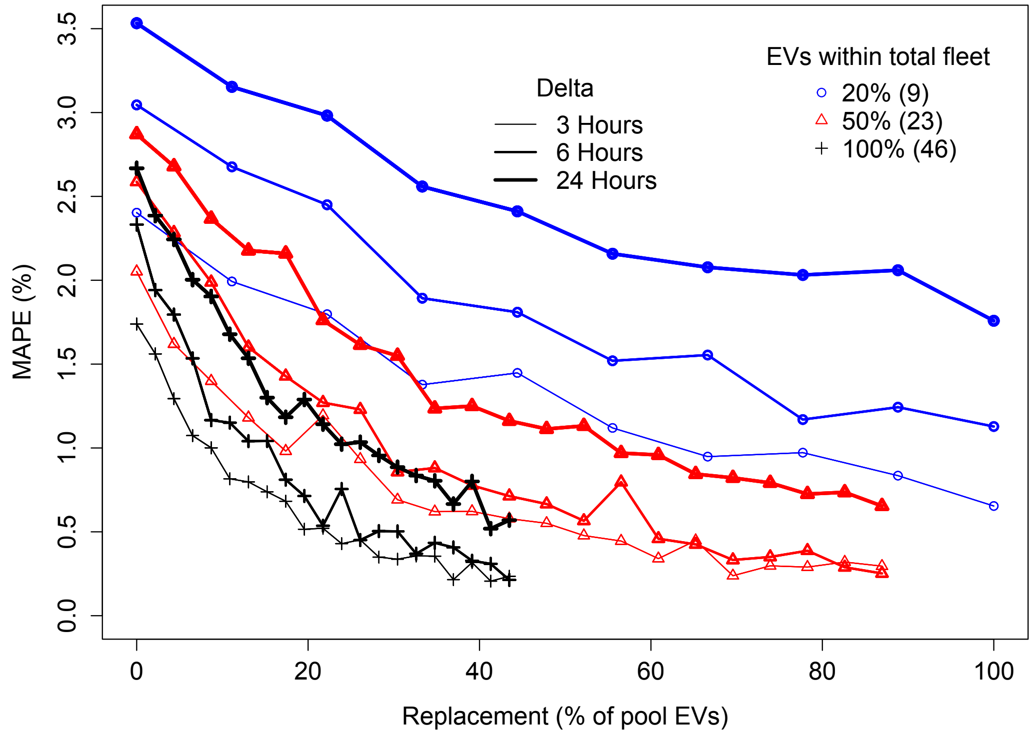


Figure 77.: Addressing low presence of individual EVs by their replacement with pool EVs

the location, the accuracy of 1.5% can be already achieved for the case of 20% EVs in the fleet (and not $\approx$ 3.2% from Figure 75). If two pool vehicles are properly managed, this case goes to $\approx$ 1%, and so on. Therefore management is to be considered as important as usage of the static storage and, as one can imagine, already via software for booking the pool vehicles.

5.5 DISCUSSION

As assessed in section 5.1, distributing a large number of smart meters and their eventing capabilities can take Smart Grids to next step in QoS. Not only that they can report measurements (on device level), but their sampling frequency can be dynamically adjusted on-demand of an operator [73]. The timely delivery of these measurements, however, raise a question if stakeholder's flexibility can be expressed for these smart meter owners. Furthermore, Figure 55 shows how collection of measurements from approximately 5000 units can be achieved in matter of few minutes, thus one may even expect flexibility being offered in timely manner by a cluster of stakeholders, as in section 4.3.3. With such infrastructure in place, one may further exploit its capability and enable traditionally passive stakeholders to be actively involved in power networks, what is the main focus of chapter 5.

Main contribution of the thesis is proposed in section 5.3, however considerations on its main components i.e. ELF, VES, ET need to be adequately addressed. Forecasting done by the ELF, cannot only be based solely on historical data, but needs to include real-time information. To this end, the Internet of Things coupled with the Cloud [141, 33] and the vast resources for analytics will help. Additionally, more specific knowledge of the processes involved, their scheduling at enterprise level, as well as their potential interdependencies may lead to better forecasting and planning. An appropriate combination of intelligent algorithms with (real-time) fine-grained data may enable the better adjustment of the infrastructure behaviour prediction.

Another key part of the system, the VES, demonstrates that the temporal storage availability e.g. coming from an EV fleet, can be used to acquire additional benefits for the enterprise. Although charging/discharging of EV batteries or even rescheduling (in order to achieve accuracy) may sound promising, at the moment few, if any, companies have adopted the EVs in their fleets. This can be mitigated through the addition of static storage, or another buffer-like component in order to make SFERS possible. However, EVs would tremendously improve the potential of the overall system and section 5.4.2 gives some indication that company-controlled fleets are the right target group for such concepts.

If a company utilizes an EV fleet as a storage solution, the system has to make sure that individual and global constraints are met e.g. that each car will be charged for its next trip. To ensure the latter, mitigation actions need to be planned e.g. adding more cars than the minimum needed. In this way each unit can provide a certain percentage of the battery capacity to the variable storage and still can guarantee that the EV is ready whenever the user needs it. As an example, if a desired SOC for SFERS featuring a static storage solution would be at 50%, the clustered available storage from EVs would have to be also at that level. However this does not necessarily mean 50% SOC for the individual EVs as this might conflict with the owner's goals which are e.g. to be at least 80% in order to cover his travel plans. Such constraints are not considered in this assessment, and are left as future work.

A limitation of this work is that no actual technical aspects dealing with the EV charging are considered. Today, charging or discharging sessions might not be as flexible as assumed in this work, and EV constraints may enforce specific behaviours e.g. once connected charge at least 20% of the capacity per session etc. Additionally, often charging/discharging may have a significant impact on the EV battery charging cycles and degradation might occur [142] which may result in financial costs. These aspects are explicitly left out from this evaluation as we did not want to link the results to a (currently available) technology, but rather to evaluate the concept from a more theoretical/general point of view. However, in the future, for commercial implementations, one needs to investigate what technologies may be considered and their impact.

Further investigation for components to enable deterministic behaviour of a stakeholder made the ET component being considered in section 5.3.1. This is in particularly important since the local energy markets are hot topic in Smart Grid research [23, 143, 59]. The interaction with other local stakeholders can not only aid SFERS in reducing its forecast error, but also create additional opportunities though energy and storage capacity trading [94]. With this in mind, the ET component is considered to be an important element of SFERS, but not the key one. As no real-world deployment of such markets currently exist, no evaluation was made in this thesis, but operational assumptions were made; however in a real-world assessment the underlying trading behaviour must be anchored in a clear understanding of the market's rules and protocols. Additionally, in order for the ET to meet the a wide range of strategies, such as the ones described in section 5.3.2, it must be able to adequately handle dynamically changing trading goals in conjunction with market-forecasts and enterprise's needs.

Generally, the author considers that there is an added value if deterministic systems are operational and would assist towards informed and automated decision-making processes in domain of power networks. The realization of a stakeholder becoming a SFERS however, will need to be assessed and fine-tuned in real-world trials once the required Smart Grid services are in place.

5.6 Conclusion

Resource reliability and active consumer contribution gained value due the dynamics and complexity introduced by RES and Distributed Energy Resources (DER). Therefore in the next generation of power grids, a greater awareness of stakeholders, especially since communication in between them is possible, needs to be is improved [46]. Although the level of isolation of consumers in the traditional power networks confined them to passive behaviour, with the introduction of Smart Grids, new intelligent systems can be designed for stakeholders [94]. Positioned as such, new roles will emerge in the Smart Grid era, given that the traditionally passive stakeholders are able to be active on the grid by accurately assessing and adjusting their own energy behaviour [94] to the needs of other stakeholders. With this goal in mind, the intelligent deterministic system SFERS is introduced in this chapter, in order to contribute to the realisation of the Smart Grid vision. Furthermore, the ability to capitalize on business opportunities is vital for the success of modern enterprises too. To that extent, fully utilizing all the capabilities offered by assets in ownership of a stakeholder is pivotal.

It was investigated what information, accuracy, resolution and capability of such a smart energy system are needed. Flexibility scenarios are proposed and investigated for the traditionally passive stakeholders. It has been shown how flexibility-driven scenarios can be realised with various degrees of interaction, e.g. bilateral interaction among interested stakeholders or even flexibility trading on envisioned energy marketplaces [24]. Requirements triggered by the need for deterministic behaviour of stakeholders are gathered, and architecture to achieve determinism is proposed. Thereby the system presented is built upon the orchestration (by the EM) of three key independent components i.e. ELF, VES, ET. Strategies to become SFERS are proposed and the one from section 5.3.2 is evaluated on a real-world case. It was demonstrated how a stakeholder's EVs in a fleet can collectively compose the VES that is seen as a promising alternative to the traditional static storage energy solutions available. With evaluation, the KPIs and potential of such systems are identified by simulating the running system.

Simulation results have shown that KPIs are the offset of reporting the energy load, as well as SoC adjustment of a VES. To achieve the same accuracy, the VES load adjustment required 2% and 20% of capacity, for an offset of 3 and 24 hours respectively. Even though the initial forecast accuracy for the 24 hour offset had MAPE of 8%, the accuracy of a retailer could be achieved with storage of energy capacity between 5 and 15% of his daily energy consumption. If the fleet of pool EVs was used (instead of BESS), the achieved accuracy of a retailer is reached already at 40% (7 vehicles for 100 employees) of daily consumption. If only traditional vehicles were used as EVs for 20% (9) of the current fleet size (46), the accuracy already approached the one of a retailer. If enhanced with a BESS, for only 2% of daily consumption a significant performance improvement is achieved. Smaller forecast offsets resulted with a significantly greater efficiency.

Not only the deterministic behaviour of SFERS will bring energy related revenue [92], but also will help to better operate and plan the usage of existing infrastructures, and empower decision making processes for many stakeholders involved in grid operations. However, several considerations are also raised as the detailed aspects of the system need to be further investigated both technically and financially. It is also clear that one-size-fits-all solution might not be available and customization needs to be done depending on the real-world case constraints, e.g. predictability of a stakeholder is in direct relation to its storage sizing [43]. As shown, various combinations for the desired forecast accuracy can be realised, but every stakeholder should be individually evaluated depending on his assets and their usage, before the actual deployment of the SFERS system.

6 Conclusion and Outlook

According to the Smart Grid vision [8], efficiency improvement may stem from the near real-time bidirectional communication between stakeholders. Many research and development projects [50] adopt the capabilities offered by Smart Grids to achieved better grid management, integration of smart-houses [33] and smart-buildings, accommodation of intermittent energy resources including Electric Vehicles (EVs), demand-response schemes [26], local energy markets for business interactions [23], etc.

This dissertation focus on actively involving the traditionally passive stakeholders to contribute challenges of electricity grids [9, 10, 11, 12], that are expected to grow with penetration of Distributed Energy Resources (DER) and Renewable Energy Sources (RES). To achieve this, the Smart Grids are used as foundation to obtain the deterministic and flexible energy loads of stakeholders. This concept is hereby called Self-Forecasting EneRgy load Stakeholders (SFERSs), and rises the research challenges of (1) enabling an efficient communication in between stakeholders, (2) reaching sufficient forecast accuracy of individuals or small scale of aggregations, and (3) building a system to enable active involvement of traditionally passive stakeholders. This work answers the challenges and the author expects for SFERS to eventually lead to better utilization of resources, improved management and energy efficiency, as well as benefits from new energy related revenues (as discussed in Appendix A) in Smart Grids.

This chapter will revise how research challenges of the thesis were addressed in section 6.1. A short summary of overall contribution is made in section 6.2 and future work in section 6.3 calls for many other research questions to be answered.

6.1 Addressing the Challenges

To address the Challenge 1 – Active Stakeholders – data initially has to be collected by a metering platform. In section 3.1 the proposed metering platform, as part of an Advanced Metering Infrastructure (AMI), is evaluated from the perspective of its individual components. The results show that bulk transfer of meter readings can significantly improve the receive rate of the platform. Although for the experimental setting in this work converged towards approximately 120 readings per message, with only 60 the performance starts to converge significantly (96% of receive rate of the 120 case). These performance improvements also can be claimed on the TCP payload efficiency of the messages transferred as bulk – if meter readings are transmitted individually, the actual metering data covered only 9% of the payload.

Once data can be collected, the added-value services of the Integration and Energy Management system (IEM) were proposed in section 3.2. Their design and implementation was proven in the real world trial of the NOBEL project [37]. The services were consumed by customer related solutions and dedicated operator solutions as presented in section 3.3. Less than 5% of response times were above 1 second for the Neighbourhood Oriented Energy Management (NOEM). Great part of these 5% delays came directly from monitoring services, since smart meter active energy readings of a stakeholder were fetched from the large DB table, measured to take 51.64% of the entire DB size as shown in section 3.3.2. Furthermore, NOEM tended to monitor smart meters in groups, which further affected the performance. In section 3.4 it was shown how to reduce (up to 60 times) the time taken in (group) aggregation of meter readings, or even beyond using more tuned DBMS. Finally, the quality of collected data, such as problems with missing readings, and near real-time performance, were identified as the main driver for operating successfully many services of the IEM platform.

The relevance of the Challenge 2 – Achieving Forecast Accuracy – was shown in the case of envisioned neighbourhood trading solutions for Smart Grids. The evaluation from section 4.1 demonstrated that groups can perform significantly better in trading together than by its individuals. This was mainly due the improved forecast accuracy, from an average individual Mean Absolute Percentage Error (MAPE) of 50% to 10.6% as a group. Resource sharing also played an important role in trading improvement. Most importantly, the increased forecasting accuracy contribute significantly to the group reduction of 68% in unnecessary buys and of 100% in demand imbalance, if compared to the individual trading cases of the participants.

Grouping was thus proposed on smaller scales in section 4.2. It is identified that a convergence point can be reached quickly e.g. around 200 households. It was shown that predictable individuals will contribute to the overall predictability of a cluster. As an example, MAPE of a 24 hour forecast for 160 households was measured at 4.09% and 5.93%, for clusters composed of individual of greater and less predictable behaviour respectively. In section 4.3 the usage of Battery Energy Storage System (BESS) solutions was identified important for the further improve in forecast accuracy – converging faster already for small storage capacities. Furthermore, results from section 4.4 suggests that forecast errors can be addressed more efficient if capacity is properly distributed, thus raising the relevance of the Variable Energy Storage (VES) concept introduced. In numbers, the forecast accuracy of the constant capacity shape equalling 8% of daily consumption (580 kWh) can be achieved with capacity shaped as the average daily forecast error with capacity of only 4% of daily consumption (290 kWh). Later in section 4.5, EVs are evaluated and significant capacity potential was noted from fleet's presence on-premise that was measured to peak above 34% (of the total fleet size on average) for hours where forecast errors are greater.

With increased forecast accuracy and communication infrastructure in place, Challenge 3 – Deterministic Behaviour – was investigated. In section 5.1 one can see how accurate and frequent sampling can be, even to the point of embedding intelligence to raise Quality of Service (QoS) (by adjusting the sampling frequency of a smart meter). This infrastructure is evaluated in a trial of the NOBEL project for sampling resolution of 15 minutes for 5000 meters over more than 6 months. The same infrastructure can be used for load reporting of the self-forecasting stakeholders. Therefore with IEM in place, a stakeholder that can achieve deterministic behaviour (such that is also measurable [12]) can benefit from many Smart Grid opportunities, e.g. from flexibility scenarios [27] described in section 5.2.

It is proposed to achieve the deterministic behaviour of the traditionally passive stakeholders by executing a self-forecast whose errors are absorbed locally by assets. Architecture is proposed in section 5.3 that will allow stakeholders to be deterministic by reporting (or smart metering with an offset) their loads to third parties, or even trade energy on their own [23]. With the architecture in place, a determinism strategy with VES is proposed in section 5.3.2 and evaluated in section 5.4. Simulation results showed that Key Performance Indicators (KPIs) are the offset of reporting the energy load, as well as State of Charge (SoC) adjustment of a VES. To reach bottom accuracy of a retailer (2% of MAPE [40]), the VES load adjustment required 2% and 20% of capacity, for an offset of 3 and 24 hours respectively. If the fleet of pool EVs was used (instead of BESS), the achieved accuracy of a retailer already at 40% (7 vehicles for 100 employees) of daily consumption. If only traditional vehicles were used as EVs for 20% (9) of the current fleet size (46), the accuracy already approached the one of a retailer. Further enhancement with BESS of capacity of only 2% of daily consumption a significant performance improve is met. Still, section 5.4.2 demonstrates in a real world case, many stakeholders are expected to be able to become SFERS (even for reporting on the 24 hour offset).

6.2 Summary and Applications

As the emerging Smart Grid increases integration of (highly intermittent) RESs, that are also DERs, the scheduling complexity and overall production unpredictability will continue to rise [15]. This calls for more deterministic behaviours within power grids and even active contribution of the traditionally passive stakeholder. Already in section 2.1 many mechanisms that deal with these challenges were presented. Some stakeholders already apply these concepts, but not all stakeholders can join such programs as predictability is one of the prerequisites. As this thesis showed, further penetration of flexible assets in Smart Grids, such as EVs [45], will increase the opportunities in combining methodologies from section 2.1 and allow many stakeholders to meet the prerequisites individually or as a group.

This work shows that traditionally passive consumers can "artificially" be deterministic [94] and flexible [49] in their energy behaviour. It was shown that (1) an efficient communication in between stakeholders can be achieved, (2) that sufficient forecast accuracy can be achieved on individuals or small scale of aggregations, and (3) system to enable active involvement of traditionally passive stakeholders can be designed with the widely adopted ICT of today. A comprehensive view, from practical and scientific angle, on answers of the research questions showed that SFERS can be achieved. The evaluations result positive on real world cases, by simulations conducted with real world data, indicating the possibilities of how determinism can be achieved for different classes of stakeholders. In the final evaluation, this thesis focuses on a commercial stakeholder (thus VES is to be composed of company-owned EVs), however future applications of SFERS may have huge potential if also composed of community-owned assets e.g. private EVs in a neighbourhood.

The findings of this thesis will further contribute to many development projects to further expand the spectrum of stakeholder opportunities already brought by Smart Grids. By enabling traditionally passive consumers to actively contribute needs of other stakeholders, and therefore their collaboration [28], the overall reduction in costs and greater efficiency can be achieved. Further efforts of the SFERS concept towards better grid management, integration of smart-houses [33] and smart-buildings, accommodation of intermittent energy resources including EV, demand-response schemes [26] and local energy markets for business interactions [23] are envisioned too. Knowledge of revenue opportunities gained while conducting this work is documented in Appendix A, but still other opportunities may exist that the author is not aware of. However, the already identified business relevance has to be investigated to understand the potential of a stakeholder to become SFERS and finally deploy the system for cost related benefits.

6.3 FUTURE WORK

Applying storage to improve resource reliability was done before [95]. Researchers even worked on controlling algorithms for different benefits of storage availability, e.g. price related [47]. The SFERS system, on the other hand, uses storage to keep its determinism and, therefore, system's reliability, while offering flexibility based on its SoC. Although configuration of stakeholders is expected to vary, one may expect that VES capacity will play a great role in maintaining determinism. In order to reduce storage size algorithms that focus on Mean Percentage Error (MPE), and not only on MAPE, should be considered. As the resulting evaluation on re-usage efficiency from section 4.4.4 pointed out, keeping the MPE around zero would increase efficiency of the capacity units. Current re-usage efficiency is measured to 169% for working and 152% for non-working days.

The introduced concept of VES calls for future work from many aspects. Controlling of BESS is possible for practical applications [64], still managing complexity of VES is higher due its dynamic part. The great part of managing focus is on selection of storage units to absorb forecast errors. An example of management would be to give priority to the first leaving units (of VES) to achieve their next goal, but many other strategies can be envisioned and need to be investigated in future. Another type of management is the SoC adjustment by Equation 14, whose for a dynamic storage unit this would primarily focus on meeting the usage goals e.g. driver requirements of an EV. Besides these, other strategies proposed in section 5.3.2 need to be investigated. Finally, the VES capacity forecasting poses a great challenge. Such algorithms need to be in place for management of a VES, both for knowing the potential to absorb the forecast errors, as well as for load adjustments.

Other questions, such as the effect of technological barriers arise for the batteries composing a VES , e.g. maximum depth of discharge. Therefore mixtures of technologies to fulfil requirements need to be investigated. Technologies that are not impacted by number of cycles [70], or even charging depth per cycle, need to be investigated e.g. compressed air. Additionally in the future one should consider rescheduling techniques to avoiding the technological barriers. In fact, this will "artificially" deliver a higher round trip efficiency for VES overall – as there is no actual charge nor discharge. Finally, not only well known storage technologies can help forecast accuracy, but the nature of the storage can be achieved from any flexible unit, e.g. supermarket freezers, data centres or interior/exterior lighting, and should be considered in future.

The general architecture from Figure 63 already considers trading functionality for SFERS. This has a great potential for self-forecasting stakeholders, as presented in Appendix A, and should be investigated in the future. Furthermore, since SFERS is envisioned to be autonomous, software trading agents [101] need to considered too. This is particularly important from the perspective of the decision made by the Energy Management (EM) component in deciding whether forecast errors are absorbed by stakeholder's VES or traded by Energy Trading (ET).

Bibliography

[1] ITU. Itu internet report 2005: The internet of things. Technical report, International Telecommunication Union, 2005.

[2] acatech. Cyber-Physical Systems: Driving force for innovation in mobility, health, energy and production. Technical report, acatech – National Academy of Science and Engineering, December 2011.

[3] Stamatis Karnouskos and Orestis Terzidis. Towards an information infrastructure for the future internet of energy. In *Kommunikation in Verteilten Systemen (KiVS 2007) Conference*. VDE Verlag, 26 February - 02 March 2007.

[4] Fiona Fui-Hoon Nah, Keng Siau, and Hong Sheng. The value of mobile applications: A utility company study. *Commun. ACM*, 48(2):85–90, February 2005.

[5] Luis M Vaquero, Luis Rodero-Merino, Juan Caceres, and Maik Lindner. A break in the clouds: towards a cloud definition. *ACM SIGCOMM Computer Communication Review*, 39(1):50–55, 2008.

[6] Dawud Gordon, Jan-Hendrik Hanne, Martin Berchtold, Ali Asghar Shirehjini, and Michael Beigl. Towards collaborative group activity recognition using mobile devices. *Mob. Netw. Appl.*, 18(3):326–340, June 2013.

[7] Hong Sheng, Fiona Fui-Hoon Nah, and Keng Siau. Strategic implications of mobile technology: A case study using value-focused thinking. *The Journal of Strategic Information Systems*, 14(3):269 – 290, 2005. The Future is UNWIRED: Organizational and Strategic Perspectives.

[8] Xinghuo Yu, C. Cecati, T. Dillon, and M.G. Simões. The new frontier of smart grids. *Industrial Electronics Magazine, IEEE*, 5(3):49–63, September 2011.

[9] A. Shafiu, N. Jenkins, and G. Strbac. Measurement location for state estimation of distribution networks with generation. *Generation, Transmission and Distribution, IEE Proceedings-*, 152(2):240 – 246, march 2005.

[10] C.N. Rasmussen. Improving wind power quality with energy storage. In *Sustainable Alternative Energy (SAE), 2009 IEEE PES/IAS Conference on*, pages 1–7, September 2009.

[11] Jinquan Zhao, Wenying Huang, Zhaoxiong Fang, Feng Chen, Kewen Li, and Yong Deng. On-line voltage stability monitoring and control (vsmc) system in fujian power grid. In *Power Engineering Society General Meeting, 2007. IEEE*, pages 1 –6, june 2007.

[12] Miriam L. Goldberg and G. Kennedy Agnew. Measurement and verification for demand response. Technical report, DNV KEMA Energy and Sustainability, Feb 2013.

[13] Eugene A. Feinberg and Dora Genethliou. Load forecasting. In J.H. Chow, F.F Wu, and J.J Momoh, editors, *Applied Mathematics for Restructured Electric Power Systems: Optimization, Control, and Computational Intelligence*, pages 269–285. Springer, 2005.

[14] James W. Taylor. Triple seasonal methods for short-term electricity demand forecasting. *European Journal of Operational Research*, 204(1):139 – 152, 2010.

[15] Nick Frydas, Konrad Borkowski, Goran Strbac, Jakob Helbrink, Niklas Damsgaard, Len Borjerson, Peter Styles, and David Holding. Impact assessment on european electricity balancing market. Technical report, Mott MacDonald, 2013.

[16] SmartGrids European Technology Platform. Smartgrids: Strategic deployment document for europe's electricity networks of the future, April 2010.

[17] D. Backer. Power quality and asset management the other "two-thirds" of ami value. In *Rural Electric Power Conference, 2007 IEEE*, pages C6 –C6–8, may 2007.

[18] European Commission. SmartGrids SRA 2035 – Strategic Research Agenda: Update of the SmartGrids SRA 2007 for the needs by the year 2035. Technical report, European Technology Platform SmartGrids, European Commission, March 2012.

[19] John Santaferraro. Offloading analytics. *Business Intelligence Journal*, 17(4):43–48, 2012.

[20] Federation of German Industries (BDI). Internet of Energy: ICT for energy markets of the future. BDI publication No. 439, February 2010.

[21] Stamatis Karnouskos, Per Goncalves da Silva, and Dejan Ilić. Energy services for the smart grid city. In *6th IEEE International Conference on Digital Ecosystems Technologies (DEST)*, pages 1–6, 2012.

[22] Frede Hvelplund. Renewable energy and the need for local energy markets. *Energy*, 31(13):2293–2302, 2006.

[23] Dejan Ilić, Stamatis Karnouskos, Per Goncalves da Silva, and Martin Griesemer. An energy market for trading electricity in smart grid neighbourhoods. In *Digital Ecosystems Technologies (DEST), 2012 6th IEEE International Conference on*, pages 1–6, 2012.

[24] Caroline Julian and Julian Dobson. Re-energising Our Communities: Transforming the Energy Market Through Local Energy Production. A ResPublica Green Paper, February 2012.

[25] Per Goncalves da Silva, Dejan Ilić, and Stamatis Karnouskos. The impact of smart grid prosumer grouping on forecasting accuracy and its benefits for local electricity market trading. *IEEE Transactions on Smart Grid*, 5(1):402–410, 2014.

[26] J.L. Mathieu, P.N. Price, S. Kiliccote, and M.A. Piette. Quantifying changes in building electricity use, with application to demand response. *Smart Grid, IEEE Transactions on*, 2(3):507–518, September 2011.

[27] Stamatis Karnouskos, Dejan Ilić, and Per Goncalves Da Silva. Using flexible energy infrastructures for demand response in a smart grid city. In *The third IEEE PES Innovative Smart Grid Technologies (ISGT) Europe, Berlin, Germany*, pages 1–7, 2012.

[28] Stamatis Karnouskos. Demand side management via prosumer interactions in a smart city energy marketplace. In *IEEE International Conference on Innovative Smart Grid Technologies (ISGT 2011), Manchester, UK*, December 5–7 2011.

[29] S.V. Papaefthymiou, E.G. Karamanou, S.A. Papathanassiou, and M.P. Papadopoulos. A wind-hydro-pumped storage station leading to high res penetration in the autonomous island system of ikaria. *Sustainable Energy, IEEE Transactions on*, 1(3):163–172, Oct 2010.

[30] Willett Kempton and Jasna Tomić. Vehicle-to-grid power implementation: From stabilizing the grid to supporting large-scale renewable energy. *Journal of Power Sources*, 144(1):280–294, 2005.

[31] Kejun Qian, Chengke Zhou, M. Allan, and Yue Yuan. Modeling of load demand due to ev battery charging in distribution systems. *IEEE Transactions on Power Systems*, 26(2):802–810, May 2011.

[32] Jasna Tomić and Willett Kempton. Using fleets of electric-drive vehicles for grid support. *Journal of Power Sources*, 168(2):459–468, 2007.

[33] Stamatis Karnouskos. Smart houses in the smart grid and the search for value-added services in the cloud of things era. In *IEEE International Conference on Industrial Technology (ICIT 2013), Cape Town, South Africa*, 25–27 February 2013.

[34] Stamatis Karnouskos, Per Goncalves da Silva, and Dejan Ilić. Developing a web application for monitoring and management of smart grid neighborhoods. In *IEEE 11^{th} International Conference on Industrial Informatics (INDIN), Bochum, Germany*, pages 408–413, July 29–31 2013.

[35] Y H. Katz, David A. Culler, Seth S, Sara Alspaugh, Yanpei Chen, Stephen Dawson-haggerty, Prabal Dutta, Mike He, Xiaofan Jiang, Laura Keys, Andrew Krioukov, Ken Lutz, Jorge Ortiz, Prashanth Mohan, Evan Reutzel,

Jay Taneja, Jeff Hsu, and Sushant Shankar. An information-centric energy infrastructure: The berkeley view. Sustainable Computing: Informatics and Systems, 2011.

[36] Per Goncalves da Silva, Stamatis Karnouskos, and Dejan Ilić. A survey towards understanding residential prosumers in smart grid neighbourhoods. In *Innovative Smart Grid Technologies (ISGT Europe), 2012 3rd IEEE PES International Conference and Exhibition on*, pages 1–8, 2012.

[37] SEVENTH FRAMEWORK PROGRAMME: Information and Communications Technologies. Neighbourhood Oriented Brokerage ELectricity and monitoring system (NOBEL). www.ict-nobel.eu, 2010. FP7- 247926 - ICT-2009-4.

[38] Stamatis Karnouskos, Dejan Ilić, and Per Goncalves Da Silva. Assessment of an enterprise energy service platform in a smart grid city pilot. In *IEEE 11^{th} International Conference on Industrial Informatics (INDIN), Bochum, Germany*, July 29–31 2013.

[39] José Ramón Cancelo, Antoni Espasa, and Rosmarie Grafe. Forecasting the electricity load from one day to one week ahead for the spanish system operator. *International Journal of Forecasting*, 24(4):588 – 602, 2008. <ce:title>Energy Forecasting</ce:title>.

[40] R. Sevlian and R. Rajagopal. Value of aggregation in smart grids. In *Smart Grid Communications (SmartGridComm), 2013 IEEE International Conference on*, pages 714–719, 2013.

[41] M.K. Celik and W.-H.E. Liu. An incremental measurement placement algorithm for state estimation. *Power Systems, IEEE Transactions on*, 10(3):1698 –1703, aug 1995.

[42] Stamatis Karnouskos. Communityware smartgrid. In *21st International Conference and Exhibition on Electricity Distribution (CIRED 2011), Frankfurt, Germany*, 6-9 June 2011.

[43] Dejan Ilić and Stamatis Karnouskos. Addressing energy forecast errors: An empirical investigation of the capacity distribution impact in a variable storage. *Energy Systems, Springer*, pages 1–14, 2014.

[44] Ad van Wijk and Leendert Verhoef. *Our Car As Power Plant*. IOS Press BV, 2014.

[45] Wencong Su, H. Eichi, Wente Zeng, and Mo-Yuen Chow. A survey on the electrification of transportation in a smart grid environment. *Industrial Informatics, IEEE Transactions on*, 8(1):1–10, 2012.

[46] Srinivasan Keshav and Catherine Rosenberg. How internet concepts and technologies can help green and smarten the electrical grid. In *Green Networking*, pages 35–40, 2010.

[47] Reza Arghandeh and Robert Broadwater. Distributed energy storage control for optimal adoption of solar energy in residential networks. In *20th International Conference on Nuclear Engineering*, 2013.

[48] C.W. Potter, A. Archambault, and K. Westrick. Building a smarter smart grid through better renewable energy information. In *Power Systems Conference and Exposition, 2009. PSCE '09. IEEE/PES*, pages 1–5, March 2009.

[49] 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, 2011.

[50] Vincenzo Giordano, Alexis Meletiou, Catalin Felix Covrig, Anna Mengolini, Mircea Ardelean, Gianluca Fulli, Manuel Sánchez Jiménez, and Constantina Filiou. Smart Grid projects in Europe: Lessons learned and current developments 2012 update. Joint Research Center of the European Commission, JRC79219, 2013.

[51] Vic Shao. The next energy savings: Power efficiency. White Paper, 2014. Green Charge Networks.

[52] A. Nourai. "Distributed Bulk Storage!": The future of batteries in grid applications. In *Power and Energy Society General Meeting, 2011 IEEE*, pages 1–1, 2011.

[53] Dejan Ilić, Stamatis Karnouskos, and Per Goncalves da Silva. Sensing in power distribution networks via large numbers of smart meters. In *Innovative Smart Grid Technologies (ISGT Europe), 2012 3rd IEEE PES International Conference and Exhibition on*, pages 1–6, 2012.

[54] Vasughi Sundramoorthy, Qi Liu, Grahame Cooper, Nigel Linge, and Joshua Cooper. DEHEMS: a user-driven approach for domestic energy monitoring. In *Internet of Things 2010 Conference, Tokyo, Japan*, 29 Nov - 01 Dec 2010.

[55] SAP AG. Transforming Power Delivery in Generation, Transmission, and Distribution, January 2014. Mobile strategies light the way for 21st century successes.

[56] Perukrishnen Vytelingum, Sarvapali D. Ramchurn, Thomas Voice, Alex Rogers, and Nicholas R. Jennings. Trading agents for the smart electricity grid. In *9th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2010), Toronto, Canada*, pages 897–904, 10-14 May 2010.

[57] comverge. Intelligent energy management. White Paper, 2011.

[58] Salman Kahrobaee, Rasheed A. Rajabzadeh, Leen-Kiat Soh, and Sohrab Asgarpoor. Multiagent study of smart grid customers with neighborhood electricity trading. *Electric Power Systems Research*, 111(0):123 – 132, 2014.

[59] Yong Ding, Martin Alexander Neumann, Per Goncalves da Silva, Michael Beigl, and Lin Zhang. A control loop approach for integrating the future decentralized power markets and grids. In *IEEE International Conference on Smart Grid Communications (SmartGridComm)*, 2013.

[60] M.H. Albadi and E.F. El-Saadany. A summary of demand response in electricity markets. *Electric Power Systems Research*, 78(11):1989 – 1996, 2008.

[61] D. Pudjianto, C. Ramsay, and G. Strbac. Virtual power plant and system integration of distributed energy resources. *Renewable Power Generation, IET*, 1(1):10–16, March 2007.

[62] A. Nourai, R. Sastry, and T. Walker. A vision & strategy for deployment of energy storage in electric utilities. In *Power and Energy Society General Meeting, 2010 IEEE*, pages 1–4, 2010.

[63] E. Matallanas, M. Castillo-Cagigal, E. Caamaño-Martín, D. Masa-Bote, A. Gutiérrez, and F. Monasterio. Analysis of the self-consumption possibilities in small grid-connected photovoltaic systems in spain. In *Proceedings of the 26th European Photovoltaic Solar Energy Conference*, pages 4619–4624, Munich, Germany, 2011. WIP-Renewable Energies.

[64] T.K.A. Brekken, A. Yokochi, A. von Jouanne, Z.Z. Yen, H.M. Hapke, and D.A. Halamay. Optimal energy storage sizing and control for wind power applications. *IEEE Transactions on Sustainable Energy*, 2(1):69–77, January 2011.

[65] V. Alimisis and N.D. Hatziargyriou. Evaluation of a hybrid power plant comprising used EV-batteries to complement wind power. *Sustainable Energy, IEEE Transactions on*, 4(2):286–293, 2013.

[66] Charul Vyas and Bob Gohn. Smart grid consumer survey preferences and attitudes about smart meters, home energy management, demand response programs, and smart appliances. Technical report, PIKE Research, 2012.

[67] US Department. Benefits of demand response in electricity markets and recommendations for achieving them. Technical report, US Department of Energy, 2006.

[68] Dejan Ilić, Per Goncalves da Silva, Stamatis Karnouskos, and Malte Jacobi. Impact assessment of smart meter grouping on the accuracy of forecasting algorithms. In *Proceedings of the 28th Annual ACM Symposium on Applied Computing (SAC)*, SAC '13, pages 673–679, New York, NY, USA, 2013. ACM.

[69] Dejan Ilić, Stamatis Karnouskos, and Per Goncalves da Silva. Improving load forecast in prosumer clusters by varying energy storage size. In *IEEE Grenoble PowerTech 2013, Grenoble, France*, 16 – 20 June 2013.

[70] K.C. Divya and Jacob Østergaard. Battery energy storage technology for power systems – an overview. *Electric Power Systems Research*, 79(4):511 – 520, 2009.

[71] Martin LaMonica. Cisco: Smart grid will eclipse size of internet. *Interview, May*, 2009.

[72] V Mirchandani. The next billion SAP users will be smart meters. *Interview online, Jul*, 2009.

[73] Joel Hoglund, Dejan Ilić, Stamatis Karnouskos, Robert Sauter, and Per Goncalves da Silva. Using a 6lowpan smart meter mesh network for event-driven monitoring of power quality. In *Smart Grid Communications (SmartGridComm), 2012 IEEE Third International Conference on*, pages 448–453, 2012.

[74] Stamatis Karnouskos, Per Goncalves da Silva, and Dejan Ilić. Assessment of high-performance smart metering for the web service enabled smart grid era. In *Proceedings of the 2nd ACM/SPEC International Conference on Performance engineering*, ICPE '11, pages 133–144, New York, NY, USA, 2011. ACM.

[75] Vishal Sikka, Franz Färber, Wolfgang Lehner, Sang Kyun Cha, Thomas Peh, and Christof Bornhövd. Efficient transaction processing in SAP HANA database: the end of a column store myth. In *Proceedings of the 2012 ACM SIGMOD International Conference on Management of Data*, SIGMOD '12, pages 731–742, New York, NY, USA, 2012. ACM.

[76] Roy Thomas Fielding. *Architectural Styles and the Design of Network-based Software Architectures*. PhD thesis, University of California, Irvine, Irvine, California, 2000.

[77] *PJM Capacity Market*, 2013.

[78] Stamatis Karnouskos. Agent-based mediated control in smart grids. In *IEEE International Conference on Innovative Smart Grid Technologies (ISGT 2011), Manchester, UK*, dec 2011.

[79] Dejan Ilić, Stamatis Karnouskos, and Martin Wilhelm. A comparative analysis of smart metering data aggregation performance. In *IEEE 11^{th} International Conference on Industrial Informatics (INDIN), Bochum, Germany*, pages 434–439, July 29–31 2013.

[80] Jian Yin, Anand Kulkarni, Sumit Purohit, Ian Gorton, and Bora Akyol. Scalable real time data management for smart grid. In *Proceedings of the Middleware 2011 Industry Track Workshop*, Middleware '11, pages 1:1–1:6, New York, NY, USA, 2011. ACM.

[81] Stamatis Karnouskos, Armando Walter Colombo, Jose L. Martinez Lastra, and Corina Popescu. Towards the energy efficient future factory. In *Proc. 7th IEEE International Conference on Industrial Informatics INDIN 2009, Cardiff, UK*, pages 367–371, June 23–26, 2009.

[82] Hasso Plattner. A common database approach for OLTP and OLAP using an in-memory column database. In *Proceedings of the 2009 ACM SIGMOD International Conference on Management of data*, SIGMOD '09, pages 1–2, New York, NY, USA, 2009. ACM.

[83] Daniel J. Abadi, Samuel R. Madden, and Nabil Hachem. Column-stores vs. row-stores: how different are they really? In *Proceedings of the 2008 ACM SIGMOD international conference on Management of data*, SIGMOD '08, pages 967–980, New York, NY, USA, 2008. ACM.

[84] Peter A. Boncz, Martin L. Kersten, and Stefan Manegold. Breaking the memory wall in MonetDB. *Commun. ACM*, 51(12):77–85, December 2008.

[85] Franz Färber, Sang Kyun Cha, Jürgen Primsch, Christof Bornhövd, Stefan Sigg, and Wolfgang Lehner. SAP HANA database: data management for modern business applications. *SIGMOD Rec.*, 40(4):45–51, January 2012.

[86] Guido Moritz, Dirk Timmermann, Regina Stoll, and Frank Golatowski. Encoding and compression for the devices profile for web services. In *The Fifth Workshop on Service Oriented Architectures in Converging Networked Environments (SOCNE 2010), Perth, Australia*, pages 514–519. IEEE Computer Society, 20-23 April 2010.

[87] Nicholas C. Zakas. The evolution of web development for mobile devices. *Queue*, 11(2):30:30–30:39, February 2013.

[88] Matjaz B. Juric, Ivan Rozman, Bostjan Brumen, Matjaz Colnaric, and Marjan Hericko. Comparison of performance of web services, ws-security, rmi, and rmi-ssl. *J. Syst. Softw.*, 79:689–700, May 2006.

[89] Stamatis Karnouskos. Cyber-Physical Systems in the SmartGrid. In *IEEE* 9^{th} *International Conference on Industrial Informatics (INDIN), Lisbon, Portugal*, July 26–29 2011.

[90] Marita Blank and Sebastian Lehnhoff. Assessing reliability of distributed units with respect to the provision of ancillary services. In *IEEE* 11^{th} *International Conference on Industrial Informatics (INDIN), Bochum, Germany*, July 29–31 2013.

[91] Saima Aman, Yogesh Simmhan, and Viktor K. Prasanna. Improving energy use forecast for campus micro-grids using indirect indicators. In *Proceedings of the 2011 IEEE 11th International Conference on Data Mining Workshops*, ICDMW '11, pages 389–397, Washington, DC, USA, 2011. IEEE Computer Society.

[92] Magnus Korpaas, Arne T. Holen, and Ragne Hildrum. Operation and sizing of energy storage for wind power plants in a market system. *International Journal of Electrical Power & Energy Systems*, 25(8):599–606, 2003.

[93] IFMA Foundation. *Technology for Facility Managers: The Impact of Cutting-Edge Technology on Facility Management*. John Wiley & Sons, Somerset, NJ, USA, September 2012.

[94] Dejan Ilić, Stamatis Karnouskos, Per Goncalves Da Silva, and Sarah Detzler. A system for enabling facility management to achieve deterministic energy behaviour in the smart grid era. In *3^{rd} International Conference on Smart Grids and Green IT Systems (SmartGreens), Barcelona, Spain*, 3–4 April 2014.

[95] P. Pinson, G. Papaefthymiou, B. Klockl, and J. Verboomen. Dynamic sizing of energy storage for hedging wind power forecast uncertainty. In *Power Energy Society General Meeting, 2009. PES '09. IEEE*, pages 1–8, 2009.

[96] J. K. Kok, M. J. J. Scheepers, and I. G. Kamphuis. Intelligence in electricity networks for embedding renewables and distributed generation. In Rudy R. Negenborn, Zofia Lukszo, and Hans Hellendoorn, editors, *Intelligent Infrastructures*, volume 42 of *Intelligent Systems, Control and Automation: Science and Engineering*, pages 179–209. Springer Netherlands, 2010.

[97] N.D. Hatziargyriou, A. Dimeas, A.G. Tsikalakis, J.A.P. Lopes, G. Karniotakis, and J. Oyarzabal. Management of microgrids in market environment. In *Future Power Systems, 2005 International Conference on*, 2005.

[98] D. Menniti, Anna Pinnarelli, and N. Sorrentino. Operation of decentralized electricity markets in microgrids. In *Electricity Distribution - Part 1, 2009. CIRED 2009. 20^{th} International Conference and Exhibition on*, pages 1–4, 2009.

[99] Per Goncalves da Silva, Stamatis Karnouskos, and Dejan Ilić. Evaluation of the scalability of an energy market for smart grid neighbourhoods. In *IEEE 11^{th} International Conference on Industrial Informatics (INDIN), Bochum, Germany*, pages 380–385, July 29–31 2013.

[100] EPIA. Global market outlook for photovoltaics until 2016. Technical report, European Photovoltaic Industry Association, 2012.

[101] D. Cliff and J. Bruten. Less than human: Simple adaptive trading agents for CDA markets. In S. Holly, editor, *Computation in Economics, Finance and Engineering: Economic Systems*, pages 117–122, Oxford, 2000. Pergamon.

[102] Lacir J. Soares and Marcelo C. Medeiros. Modeling and forecasting short-term electricity load: A comparison of methods with an application to brazilian data. *International Journal of Forecasting*, 24:630–644, 2008.

[103] M.Y. Cho, J. C Hwang, and C-S Chen. Customer short term load forecasting by using arima transfer function model. In *Energy Management and Power Delivery, 1995. Proceedings of EMPD '95., 1995 International Conference on*, volume 1, pages 317–322vol.1, 1995.

[104] Rob J. Hyndman and George Athanasopoulos. Forecasting: principles and practice, 2012.

[105] James W. Taylor, Lilian M. de Menezes, and Patrick E. McSharry. A comparison of univariate methods for forecasting electricity demand up to a day ahead. *International Journal of Forecasting*, 22(1):1–16, January 2006.

[106] Rob J. Hyndman and Anne B. Koehler. Another look at measures of forecast accuracy. *International Journal of Forecasting*, 22(4):679 – 688, 2006.

[107] James W. Taylor. Short-Term Electricity Demand Forecasting Using Double Seasonal Exponential Smoothing. *Journal of the Operational Research Society*, 54(8):799–805, 2003.

[108] Georgios Chalkiadakis, Valentin Robu, Ramachandra Kota, Alex Rogers, and Nicholas R. Jennings. Cooperatives of distributed energy resources for efficient virtual power plants. In *The 10^{th} International Conference on Autonomous Agents and Multiagent Systems - Volume 2*, AAMAS '11, pages 787–794, Richland, SC, 2011. International Foundation for Autonomous Agents and Multiagent Systems.

[109] B. M J Vonk, P.H. Nguyen, M.O.W. Grond, J.G. Slootweg, and W.L. Kling. Improving short-term load forecasting for a local energy storage system. In *Universities Power Engineering Conference (UPEC), 2012 47th International*, pages 1–6, 2012.

[110] Fabio Genoese and Massimo Genoese. Assessing the value of storage in a future energy system with a high share of renewable electricity generation. *Energy Systems*, 12667:1–26, 2013.

[111] T. M. Peng, N.F. Hubele, and G.G. Karady. Advancement in the application of neural networks for short-term load forecasting. *Power Systems, IEEE Transactions on*, 7(1):250–257, 1992.

[112] Suling Zhu, Jianzhou Wang, Weigang Zhao, and Jujie Wang. A seasonal hybrid procedure for electricity demand forecasting in china. *Applied Energy*, 88(11):3807–3815, 2011.

[113] P. Sanchez-Martin, G. Sanchez, and G. Morales-Espana. Direct load control decision model for aggregated EV charging points. *Power Systems, IEEE Transactions on*, 27(3):1577–1584, 2012.

[114] P. Poonpun and W.T. Jewell. Analysis of the cost per kilowatt hour to store electricity. *Energy Conversion, IEEE Transactions on*, 23(2):529–534, 2008.

[115] Alan V Oppenheim, Alan S Willsky, and Syed Hamid Nawab. *Signals and systems*, volume 2. Prentice-Hall Englewood Cliffs, NJ, 1983.

[116] Mustafa Dogru, Matthew Andrews, John Hobby, Yue Jin, and Gabriel Tucci. Modeling and optimization for electric vehicle charging infrastructure. In *24th annual POM conference, Denver, Colorado, USA*, 2013.

[117] J. Lassila, T. Kaipia, and J. Lohjala. New investment strategies in the modern electricity distribution business - reliability in the long-term planning. In *Power Engineering Society General Meeting, 2007. IEEE*, pages 1 –8, june 2007.

[118] I. Lampropoulos, G.M.A. Vanalme, and W.L. Kling. A methodology for modeling the behavior of electricity prosumers within the smart grid. In *Innovative Smart Grid Technologies Conference Europe (ISGT Europe), 2010 IEEE PES*, pages 1 –8, oct. 2010.

[119] Alexandra-Gwyn Paetz, Birger Becker, Wolf Fichtner, and Hartmut Schmeck. Shifting electricity demand with smart home technologies – an experimental study on user acceptance. In *30th USAEE / IAEE North American Conference Online Proceedings*, 2011.

[120] Stamatis Karnouskos, Per Goncalves Da Silva, and Dejan Ilić. Performance evaluation of web service enabled smart metering platform. In *1st International ICST Conference on E-Energy, 14-15 October 2010 Athens Greece*. Springer, October 2010.

[121] E. Lakervi and E. J. Holmes. *Electricity Distribution Network Design*. Institution of Engineering and Technology, 1996.

[122] Cenelec standard voltages (en 60038:2011), 2011.

[123] Voltage characteristics of electricity supplied by public electricity networks (en 50160:2010), 2010.

[124] R.M. Gardner and Yilu Liu. FNET: A quickly deployable and economic system to monitor the electric grid, may 2007.

[125] Reynaldo Francisco Nuqui. *State Estimation and Voltage Security Monitoring Using Synchronized Phasor Measurements*. PhD thesis, Virginia Tech, July 2001.

[126] Andreas Abart, Daniel Burnier, Benoit Bletterie, Matthias Stifter, Helfried Brunner, Andreas Lugmaier, and Alexander Schenk. Power snapshot analysis: A new method for analyzing low voltage grids using a smart metering system. In *21st International Conference and Exhibition on Electricity Distribution (CIRED 2011), Frankfurt, Germany*, 2011.

[127] Stamatis Karnouskos. The cooperative internet of things enabled smart grid. In *Proceedings of the 14th IEEE International Symposium on Consumer Electronics, Braunschweig, Germany*, 07-10 June 2010.

[128] Alessandro Cannata, Stamatis Karnouskos, and Marco Taisch. Energy efficiency driven process analysis and optimization in discrete manufacturing. In *35th Annual Conference of the IEEE Industrial Electronics Society (IECON 2009), Porto, Portugal*, 3-5 November 2009.

[129] Laurynas Šikšnys, Mohamed E. Khalefa, and Torben Bach Pedersen. Aggregating and disaggregating flexibility objects. In *24th International Conference on Scientific and Statistical Database Management (SSDBM), Chania, Crete, Greece*, pages 379–396, 25-27 June 2012.

[130] Oleg Gulich. Technological and business challenges of smart grids – aggregator's role in current electricity market. Master's thesis, Lappeenranta University of Technology, Faculty of Technology, Electrical Engineering, Electricity Market and Power Systems, 2010.

[131] A. Ceclan, D.D. Micu, E. Simion, and R. Donca. Public lighting systems an energy saving technique and product. In *Clean Electrical Power, 2007. ICCEP '07. International Conference on*, pages 677 –681, may 2007.

[132] Michael Mutmansky, Todd Givler, Jessica Garcia, and Nancy Clanton. Advanced Street Lighting Technologies Assessment Project – City of San Diego. Technical report, Clanton and Associates, Inc., January 2010.

[133] Mark S. Rea, John D. Bullough, Charles R. Fay, Jennifer A. Brons, John Van Derlofske, and Eric T. Donnell. Review of the safety benefits and other effects of roadway lighting. Technical report, National Cooperative Highway Research Program Transportation Research Board of The National Academies, 2009.

[134] Reinhard Müllner and Andreas Riener. An energy efficient pedestrian aware smart street lighting system. *Int. J. Pervasive Computing and Communications*, 7(2):147–161, 2011.

[135] C.R.B.S. Rodrigues, P.S. Almeida, G.M. Soares, J.M. Jorge, D.P. Pinto, and H.A.C. Braga. An experimental comparison between different technologies arising for public lighting: Led luminaires replacing high pressure sodium lamps. In *Industrial Electronics (ISIE), 2011 IEEE International Symposium on*, pages 141 –146, june 2011.

[136] Stefano Carosio, Matti Hannus, Christian Mastrodonato, Elisabetta Delponte, Andrea Cavallaro, Francesco Cricchio, Stamatis Karnouskos, Jorge Pereira-Carlos, Charles Bastos Rodriguez, Oskar Nilsson, Isabel Pinto Seppä, Tomasz Sasin, Juliusz Zach, Heidy van Beurden, and Tashweka Anderson. *ICT Roadmap for Energy-Efficient Buildings – Research and Actions*. EU Project ICT4E2B Consortium, January 2013.

[137] H. Kanchev, Di Lu, F. Colas, V. Lazarov, and B. Francois. Energy management and operational planning of a microgrid with a PV-based active generator for smart grid applications. *Industrial Electronics, IEEE Transactions on*, 58(10):4583–4592, 2011.

[138] S. Teleke, M.E. Baran, A.Q. Huang, S. Bhattacharya, and L. Anderson. Control strategies for battery energy storage for wind farm dispatching. *Energy Conversion, IEEE Transactions on*, 24(3):725–732, 2009.

[139] Travis Walter, Phillip N. Price, and Michael D. Sohn. Uncertainty estimation improves energy measurement and verification procedures. *Applied Energy*, 130(0):230 – 236, 2014.

[140] S. Teleke, M.E. Baran, S. Bhattacharya, and A.Q. Huang. Rule-based control of battery energy storage for dispatching intermittent renewable sources. *Sustainable Energy, IEEE Transactions on*, 1(3):117–124, Oct 2010.

[141] Edward Curry, James O'Donnell, Edward Corry, Souleiman Hasan, Marcus Keane, and Seán O'Riain. Linking building data in the cloud: Integrating cross-domain building data using linked data. *Advanced Engineering Informatics*, 27(2):206 – 219, 2013.

[142] Scott B. Peterson. *Plug-in hybrid electric vehicles: battery degradation, grid support, emissions, and battery size tradeoffs*. PhD thesis, Carnegie Mellon University, May 2012.

[143] Steffen Lamparter, Silvio Becher, and Jan-Gregor Fischer. An agent-based market platform for smart grids. In *Proceedings of the 9th International Conference on Autonomous Agents and Multiagent Systems: Industry track*, AAMAS '10, pages 1689–1696, Richland, SC, 2010. International Foundation for Autonomous Agents and Multiagent Systems.

[144] Eric Hirst. Real-time balancing operations and markets: Key to competitive wholesale electricity markets. Technical report, Edison Electric Institute, 2001.

[145] Energy and Climate Change Committee. Energy prices, profits and poverty: Fifth report of session 2013–14. Technical report, House of Commons, 2013. Volume I: Report, together with formal minutes, oral and written evidence.

[146] EPEX Spot. Epex spot, 2010.

[147] T. Vinnal, L. Kutt, and H. Kalda. Analysis of power consumption and losses in relation to supply voltage levels. In *Power Quality and Supply Reliability Conference, 2008. PQ 2008*, pages 23 –28, August 2008.

[148] Bei Gou and A. Abur. An improved measurement placement algorithm for network observability. *Power Systems, IEEE Transactions on*, 16(4):819 –824, nov 2001.

[149] W. H. Kersting. *Distribution System Modeling and Analysis*. CRC Press, 2nd edition, 2007.

[150] A Tolón-Becerra, X Lastra-Bravo, and F Bienvenido-Bárcena. Proposal for territorial distribution of the eu 2020 political renewable energy goal. *Renewable Energy*, 36(8):2067–2077, 2011.

[151] J. Stragier, L. Hauttekeete, and L. De Marez. Introducing smart grids in residential contexts: Consumers' perception of smart household appliances. In *Innovative Technologies for an Efficient and Reliable Electricity Supply (CITRES), 2010 IEEE Conference on*, pages 135 –142, sept. 2010.

[152] E. Buchmann, S. Kessler, P. Jochem, and K. Bohm. The costs of privacy in local energy markets. In *Business Informatics (CBI), 2013 IEEE 15th Conference on*, pages 198–207, July 2013.

[153] Jens Grossklags and Alessandro Acquisti. When 25 cents is too much: An experiment on willingness-to-sell and willingness-to-protect personal information. In *6th Annual Workshop on the Economics of Information Security, WEIS 2007*, 2007.

[154] Michael Valocchi and John Juliano. Knowledge is power – driving smarter energy usage through consumer education. Technical report, IBM Institute for Business Value, 2012.

[155] Jurek Pyrko and Sarah Darby. Conditions of energy efficient behaviour – a comparative study between sweden and the uk. *Energy Efficiency*, 4:393–408, 2011. 10.1007/s12053-010-9099-x.

[156] José Javier Garcia, Juan José Cardenas, Regina Enrich, Dejan Ilić, Stamatis Karnouskos, and Robert Sauter. Smart City Energy Management via Monitoring of Key Performance Indicators. In *"Challenges Of Implementing Active Distribution System Management" CIRED Workshop 2014, Rome, Italy*, 2014.

List of Acronyms

6LOWPAN IPv6 over Low power Wireless Personal Area Networks

AMI Advanced Metering Infrastructure

AS Application Server

BESS Battery Energy Storage System

BRP Balance Responsible Party

CO_2 Carbon dioxide

DB DataBase

DBMS DataBase Management System

DCP Data Capturing and Processing

DER Distributed Energy Resources

DG Distributed Generation

DR Demand Response

DSM Demand Side Management

DSO Distribution System Operator

EJB Enterprise Java Beans

EM Energy Management

ELF Energy Load Forecast

ET Energy Trading

EV Electric Vehicle

FMC Fundamental Modelling Concepts

HTML HyperText Markup Language

HTTP Hypertext Transfer Protocol

HW Holt-Winters

ICT Information and Communication Technologies

IEM Integration and Energy Management system

IOT Internet of Things

IP Internet Protocol

KPI Key Performance Indicator

LED Light-Emitting Diode

MAPE Mean Absolute Percentage Error

MDS Metering Data System

MPE Mean Percentage Error

NOBEL Neighbourhood Oriented Brokerage ELectricity and monitoring system

NOEM Neighbourhood Oriented Energy Management

OLAP On-Line Analytical Processing

OLTP On-Line Transaction Processing

QOS Quality of Service

PC Personal Computer

PDC Phasor Data Concentrator

PLS Public Lighting System

PMU Phasor Measurement Unit

PV Photovoltaic

PVPP prosumer Virtual Power Plant

RES Renewable Energy Sources

REST Representational State Transfer

ROI Return of Investment

SAAS Software-as-a-Service

SARIMA Seasonal AutoRegressive Integrated Moving Average

SCADA Supervisory Control And Data Acquisition

SFERS Self-Forecasting EneRgy load Stakeholder

SLOC Source Lines Of Code

SMS Short Message Service

SN Seasonal Naïve

SOAP Simple Object Access Protocol

SOC State of Charge

SQL Structured Query Language

TCP Transmission Control Protocol

TSO Transmission System Operator

UI User Interface

V2G Vehicle-to-Grid

VES Variable Energy Storage

XML eXtensible Mark-up Language

ZIP Zero-Intelligence Plus

A Business Relevance

Highly accurate forecast plays a pivotal role to any strategic or business decisions the Smart Grid stakeholders will take. As this work suggest that SFERS can be achieved (even on lower scales of aggregation), one can access new business opportunities [28]. Some envisioned by the author and analysed with many experts in the field are discussed in the following sections.

A.1 Load flexibility

Traditional energy consumers are isolated from operations of electricity grids, while they continuously affect them. In fact, consumers pay the unpredictable part of their loads through their energy bills. In fact, power networks are heavily supported by Balance Responsible Partys (BRPs), to the extent that balancing costs are included in the costs of the network usage [15]. These responsive operations keep the electricity grid in balance [144] by addressing the stochastic loads that can occur. Instead of paying for this service, consumers could address the unpredictable part by flexibility of their loads. However, even if a stakeholder would like to act flexible on the grid, and is equipped with a smart meter, its flexibility cannot be verified for stakeholders that have have no "predictable" behaviour [26].

This work showed that "predictable" behaviour can be artificially achieved, by determinism of SFERS, thus one can measure (even individuals) on flexibility they executed [12]. As an example, since load of SFERS is reported in advance the DR effectiveness does not need to be approximated [26], but rather directly measured. As such, their flexible part can be used to reduce the requirement for energy balancing, or indirectly the energy costs. The SFERS flexibility is as such used on soft basis, but also can contribute to capacity reduction and peak shaving activities. Hence it is in the benefit also of the infrastructure managers e.g., DSO to have larger clusters of prosumers (e.g., pVPPs) that would have the necessary footprint (due to the high number their members) to assist in critical situations by adjusting their load, and potentially appear as spinning reserves.

A.2 ENERGY RETAIL

Energy retailers provide a valuable service for their customers by forward hedging much of their wholesale energy purchases, smoothing the impact of wholesale price volatility for customers and reducing price shocks. However, in some cases, other supplier operating costs and profit margins are estimated around 15–19% [145]. Due to the non significant percentage, many researchers call for potential interconnection of producers and consumers over trading platforms of Smart Grids [25].

As presented in this work, forecast accuracy is hard to achieve via forecasting algorithms, thus the SFERS concept is presented. As work manly focus on the current accuracy of retailers today, or 2–3% [40], those SFERS that can achieve it may be able to join such programs and trade energy on their own. An additional question goes to accuracies of SFERS that go beyond what retailers face today. As an example, the balancing costs a forecast error bears can be lower than what retailers pay today. Nevertheless, if accuracy of a retailer is reached they might be able to trade energy on their own. One example would be to trade on day-ahead markets, such as EPEX Spot [146], where uniform pricing model is adopted so everyone pays and gets the market clearing price (thus even retailers). Furthermore, based on current flexibility of SFERS, one may benefit from energy prices on intraday markets, e.g. sell energy they bought on the day ahead market if price is convenient. In some cases, according to experts, is more beneficial absorbing the errors of forecast by trading, via the ET component described in section 5.3.2, than storing it. However, to enable fully automatized brokerage agents, that act on the behalf of SFERS in energy markets [25], limitation of trading units (e.g. 0.1MW on [146]) need to be removed.

A.3 INTEGRATION OF RES

Energy bills are rising and are likely to continue to rise in the future [145]. The wholesale price of fuel has been the largest contributing factor, driven by rising global gas prices. Several other factors are also contributing to price rises, including climate change policies, therefore introduction of RES is significantly penetrating world wide. However, RES are unpredictable, and if we keep on penetrating them, a fully accurate forecast of consumption may not play the key role. Instead, a desired behaviour would be to meet the current equilibrium of consumption and production.

The SFERS concept is expected to further support RES with their flexibility features e.g. addressing near-real time surplus in production or consumption. Even today, balancing of intermittent RES is supported by energy storage, such as BESS. Its application is well known in wind farm prediction scenarios [92] to meet the expected (or reported) behaviour. As SFERS are based on their asset flexibility, they might be preferred to be used to address the intermittent behaviour, than deployment of new BESS systems [144]. One can even go beyond and consider renting available VES capacity, as proposed by one of the scenarios in [94]. Nevertheless, the traditional tariff model offered by retailers can be challenged by volatile energy production brought by integration of RES. With that in mind, high flexibility in load behaviour (of traditionally passive consumers) might be desired in future e.g. energy is consumed when wind is blowing.

A.4 Power Network Operations

Delivering an uninterrupted supply and high QoS is among the goals of every Distribution System Operator (DSO). Deviations from acceptable quality levels [122] can cause blackouts, or damage equipment leading to financial impact for stakeholders. As voltage may vary significantly in distribution networks, this has an impact on the energy efficiency side [147]. Over-voltage can result in a reduction of equipment lifetime and increased energy consumption without any performance improvements. Transients, i.e. large and brief voltage increases, can destroy electronics and degrade equipment parts. With that in mind, DSO must understand the network state, which is currently estimated, thus measurement placement techniques are used to get an accurate state estimation.

Some of the techniques involve starting with already available measurements and try to reduce system's non-observability by adding pseudo-measurements [148]. Applying such techniques can become inaccurate if applied to distribution systems [9]. In [9] few reasons were identified, but it was noted that in some cases observability cannot be overcome by the addition of few pseudo-measurements. These challenges are noted even without adoption of DER and RES, so one can expect growth of their significance in future. This is why determinism and flexibility of a SFERS is to be rewarded due to reliability of its reported load. Detecting and defining facilities to become SFERS will ensure the overall system stability, thus reducing their needs for estimation [149]. In fact, the earlier the demand forecast/report of SFERS is given to a DSO, the better. This way, grid problems, such as critical power-line congestion, can be detected early (preventive maintenance), and corrective actions can be planned and realised on-time, ideally before bigger problems occur.

A.5 Sustainability

A great part of stakeholders in electricity grids are challenged with sustainability goals [150]. However, sustainability goals not necessarily have to constrain their current businesses. As previously mentioned, adopting SFERS is expected to lower the energy costs, therefore one can benefit from reduced operational costs. As an example, a company can reduce price of their products and therefore be even more competitive on their markets. Further reduction in costs is due to adoption of existing assets to compose a VES for SFERS. As EVs are used in the evaluation case, stakeholder's costs are further reduced due to lower cost per kilometre of an EV(than if oil is used for traditional vehicles). Adoption of EVs is expected to further support the climate change policies of today (that many companies try to reach) and help them reduce their CO_2 emissions by using environmental friendly energy mixes [150].

B Stakeholder Acceptance

In the smart grid there are several ideas of what could be offered in terms of technology, but strong cases for their business benefits still need to be proven in practice. For SFERS not only the business benefits need to be proven, but the actual stakeholder acceptance behind the system presented in section 5.3. One way to gain some insight into the thoughts and dispositions of the consumers is through focused surveys, which constitute a practical way of gauging stakeholder expectations and inclinations and are routinely performed to this end. In the process of designing and realizing futuristic concepts that allow the prosumers to interact via smart grid services [21], a survey was conducted in order to evaluate and understand the interest, impact, and willingness of prosumers. In section B.1, the result and analysis of the survey is presented, and the insights are summarized and discussed in section B.2. As will be shown, the results can be fully adopted in understanding the existence of SFERS.

B.1 Survey Analysis

The analysis presented in this section is according to the methodology from [36], and aims at highlighting the key aspects that were brought to surface. For analysis purposes, the questions are divided into 4 categories: willingness to change, automated control, value-added services and privacy. Due the relevance for the thesis, following sections present only partial results of [36].

B.1.1 *Willingness to Change*

One of the main pillars upon which the smart grid promise is built assumes that the prosumers are willing to adjust their behaviour based on new timely information they have access to. Although this is a multifaceted problem, it is important to understand if the prosumers want to adjust their behaviour, under what conditions, and in what way. As smart grids envision highly distributed generation, the increased participation of the demand side to stabilize the grid is a highly relevant area of research [151], as it will greatly impact the way end-users interact with the grid. Understanding the willingness of end-users to transition into this new paradigm of thinking and acting in the smart grid, whether it be responding to price signals, actively trading energy resources, or simply paying a little more to consume more "green" energy, is of paramount importance.

As the nature of the generation and distribution of electricity changes, end-users will have to take a more active role in managing their usage to manage costs and diminish their impact on the environment [94]. Part of the survey questions were pertaining the willingness of the end-consumers to change and adapt their consumption behaviour, to engage with each other to reduce costs, and to provide usage information to their retailer in order to reduce costs. As depicted in Figure 78, depending on the information they acquire, the overwhelming majority of people are willing to modify their own behaviour.

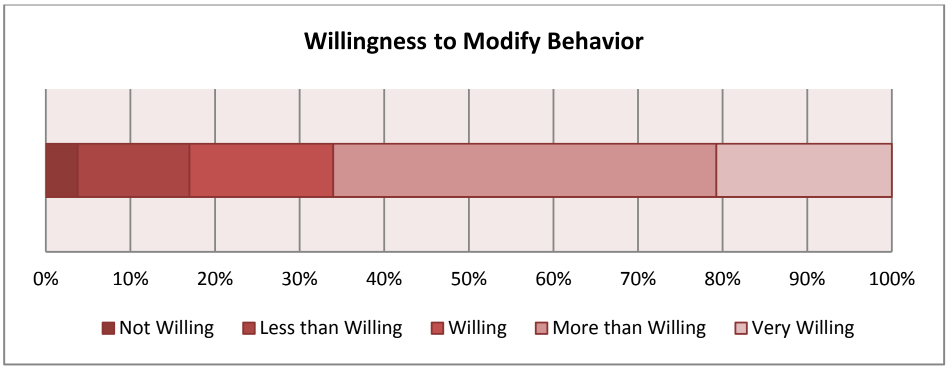


Figure 78.: The willingness of participants to modify their consumption behaviour based on external signals such as price

These are some of the key aspects of the smart grid, where people are expected to adjust their behaviour in order to assist reducing energy at peak times, as well as maximize the use of intermittent renewable energy, such as wind of solar photovoltaic. Additionally, the majority of participants would be willing to pay slightly more to reduce environmental impact by using green energy, as Figure 79 depicts. Therefore, in principle the prosumer has an interest in modifying their behaviour; however, to what extent, and by what means, needs to be further investigated.

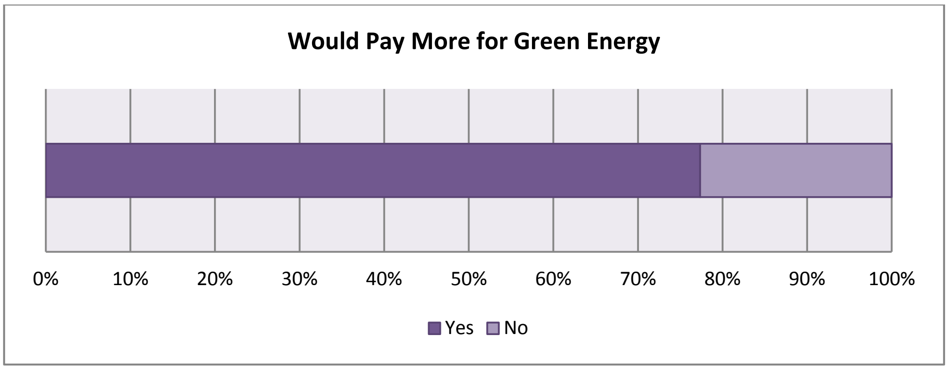


Figure 79.: The percentage of participants that would pay more for green energy

An interesting aspect in the envisioned smart grid is based on the willingness of "prosumers" to share resources (for example, unused ones) or trade them on an electricity market [25]. The major goal here is the understanding of prosumers' energy behaviour both as individuals as well as part of groups (defined by social, economic, geographic, etc., criteria). The aforementioned objective may be greatly assisted by having better prediction and real-time analytics on the provided and vast smart-grid information. As shown in Figure 80, there is overwhelming support for sharing unused resources, especially if some monetary benefit can be obtained. Additionally, about 2/3 of the prosumers seem positive towards participating in shared-interest groups. This is especially interesting in the cases where service providers may act on behalf of a larger group of users (such as prosumer Virtual Power Plant [61]), and perform actions such as bidding into energy markets [23] or actively managing their participants' energy devices according to bilateral service contracts.

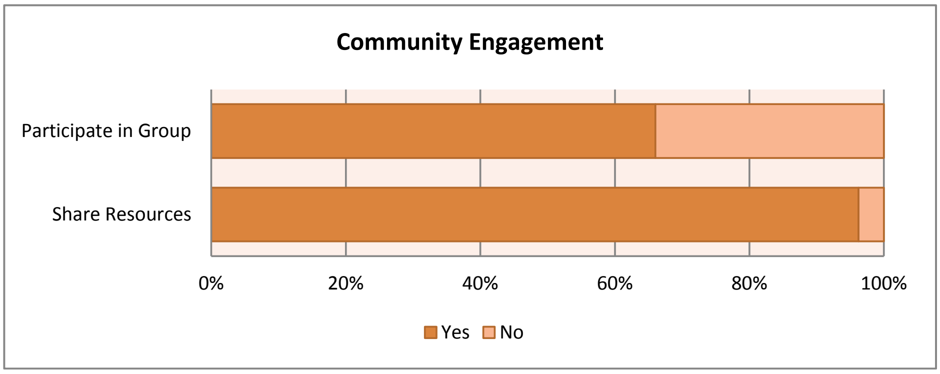


Figure 80.: The percentage of participants that would like to engage with their community to form groups and share resources

As the smart grid is expected to be information-centric [35], one has to look at the broader picture and not only the technical information that may be acquired by the infrastructure. The increasing trend towards bilateral communication between retailers and their customers means new interaction patterns can emerge, and new approaches in handling dynamic changing situations as required in Demand Side Management and Demand Response can emerge. For instance, customers may reduce their energy costs by providing extra information about themselves, which in turn might help their retailers better assess situations and reduce costs incurred for example by forecasting errors.

The survey results as depicted in Figure 81 reveal that the majority of participants are willing to provide information about their energy-usage expectations to third parties. However, only about half of them are willing to classify in detail their behaviour pattern, for example being on vacation. This seems to suggest that new tools need to be offered to prosumers that allow them to model and understand their energy usage patterns so that they may convey their usage expectations to retailers without revealing detailed, privacy-infringing aspects [152].

Hence, the right balance between privacy and rich user-information provision that the smart grid promises is based upon needs to be striven towards to, and supported by, the necessary tools.

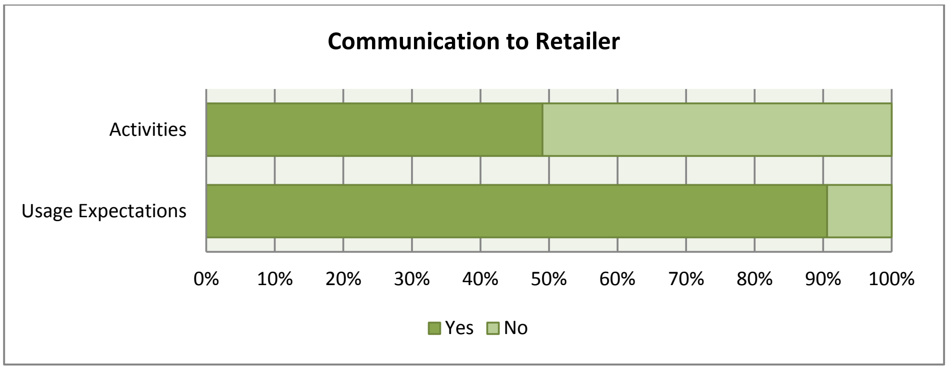


Figure 81.: The percentage of participants that would communicate their activities and their usage expectations to their retailer

B.1.2 *Automated Control*

Although information-rich real-time monitoring of energy aspects is a key promise of smart grid, in order to be effective this needs to be strongly coupled with real-time control and management of the infrastructure. This will make possible large-scale energy-management approaches, as now situations can be monitored and reacted upon in much more sophisticated ways [34]. There are several promising scenarios here,for instance independent service providers would be able to remotely control household devices to curb usage in peaks times. This idea may not be new, as it is already implemented in commercial and industrial sectors, but applying it at large-scale residential areas and infrastructure that could not be monitored and controlled in real-time is new ground. EnerNOC (`www.enernoc.com`) is a good example of a company offering DR in the commercial and industrial sphere. It bids the energy flexibility of their customers in the energy market; whereby in some cases, its customers can generate more revenue by shutting down machinery to curb energy usage, than by continuing production.

Figure 82 depicts that the survey participants are willing to allow automatic management of devices as far as this does not affect any loss of comfort. This open the door for optimisation approaches between usage-patterns and device operation (which may lead to increased energy efficiency), effectively moving away from "one-size-fits-all" design and operational assumptions of appliances towards user-specific adaptations. However, the findings point out that people are more willing to allow their own devices to automate their energy consumption (based on external signals, such as price), than to allow external parties to manage their behaviour. This puts forward a clear message that the user wants to be in control of his own infrastructure but would happily engage to automatic control approaches that do not negatively impact the accommodated lifestyle. Such results are of significant interest if EVs of customers are used for composing a variable storage solution, who's potential was shown in section 4.5.

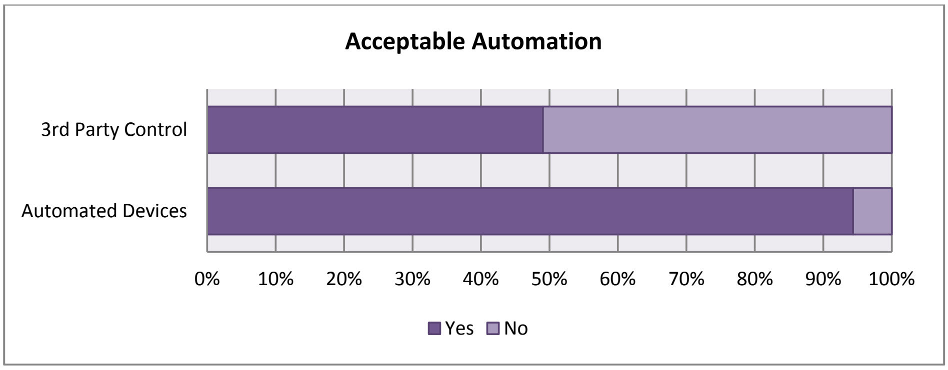


Figure 82.: The percentage of participants that would like automates devices and would accept 3rd party management of devices

Interestingly, in a follow-up question "If you could trade any excess photovoltaic production in a small market, would you be willing to allow another party to manage that task for you in the same way a managed fund might manage your investments?", 81% of participants said yes. This seems to indicate a disparity in the willingness to allow third parties control between consumption and production devices. It also suggests that neighbourhood level energy aggregators may be a viable business model for managing local energy requirements in the future. However, this reaction might also be result of inexperience with energy-producing devices and their tight integration with in-house consumption, something that has been fortified with the existing feed-in tariffs in several countries that led to users considering the energy-generation sources as a third-party infrastructure that is just co-located to their premises and hence fail to make the connection between the energy produced by such systems and their own consumption.

B.1.3 *Value-Added Services*

As well as providing end-users with an in-depth view of their energy consumption, fine grained metering data together with artificial intelligence and data-mining algorithms can provide end-users with novel added-value services [38]. Such services are expected to play a pivotal role in retailer offerings, as they might serve as key differentiators between competing stakeholders. Examples of these services could be: enabling end-users to compare their consumption with that of similar households in the region, allowing the retailer to provide their customers with suggestions on how to improve their behaviour, as well as bill shock services (which notify the customer early enough that s/he is on track for a larger than usual bill), or vacation services, which allow the customer to be informed of any unexpected energy usage in the house during a period of absence, such as when travelling. Although innovative creative thinking might come up with new ideas, in order for them to materialize one would have to heavily rely on monitoring, assessment, and management of the infrastructure, its stakeholders, and the information it holds as indicated multiple times in this section.

As can be seen in Figure 83, there is a high level of interest in value-added services such as recommendation and comparison services. In order to catalyse this process, it would be important to outfit consumers with tools that give them access to their consumption data, as well as the ability to manage it, which implies sharing it via user-controlled policy access [66]. With such enabling approaches, innovative on-line services could be created that leverage this data to create value for the customer and the service provider, much in the same way several providers operate today, for instance Facebook and Google in the social media domain.

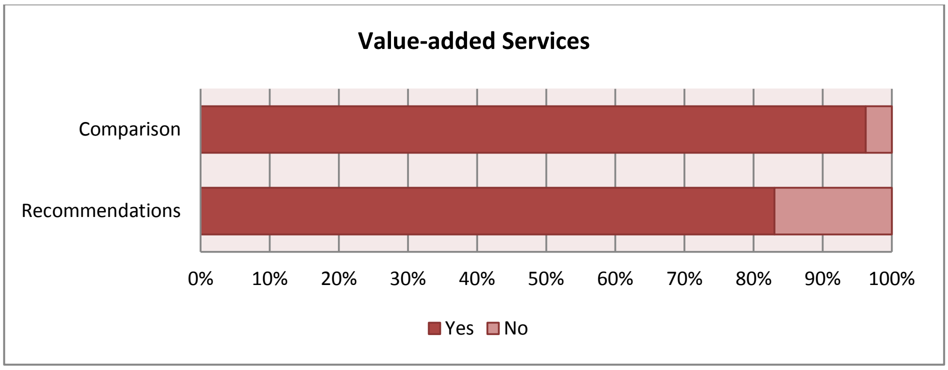


Figure 83.: The percentage of participants that would like value added services such as comparison and recommendation

B.1.4 *Privacy*

Privacy is a key area in the emerging smart grid that needs to be properly addressed in order not to pose as a roadblock. Experience so far both on telecommunications and Internet services has shown that value can be created for the users who may be willingly (or simply unaware of the compromises they get to) sacrifice part of their privacy in order to enjoy such services. Similarly, here the privacy concerns versus the services offered will be a battlefield, and approaches that offer a user-controllable balance between functionality and (private) information provided are sought.

As depicted in Figure 84, the finding is that users may share information and partly trade their privacy if this is done in a controllable visible way, such as sharing data with the energy provider. However, over 90% said that this should be done under privacy preserving measures (e.g. anonymization, etc.). This is in line also with the interest in sharing information on social networking sites, for which most of the users do not see the benefit of simply sharing their energy consumption at the moment, probably due to absence of real value-added applications in these. However, this lack of interest dropped to about 50% if additional benefits were given, such as better pricing or access to additional value-added services. Concluding, the finding is that while privacy is paramount, it is still negotiable; however, it is still unclear how much privacy would the participant be willing to sacrifice, and for what level of benefits [153].

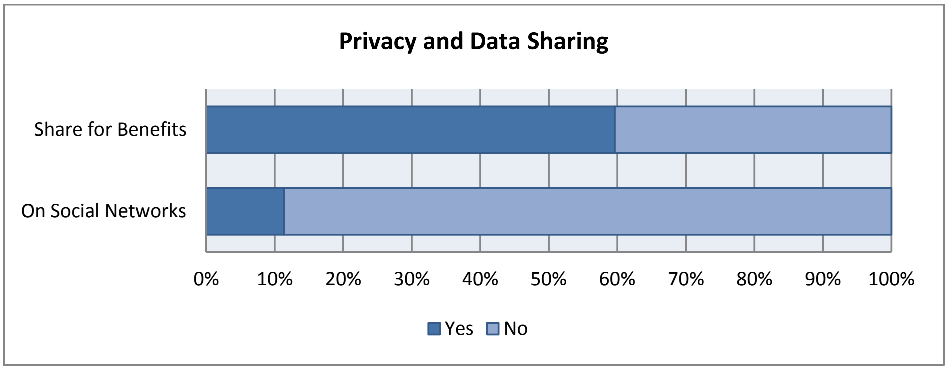


Figure 84.: The percentage of participants that would like share their usage information on social networking sites or for additional benefits.

B.2 Towards Prosumer Energy Services

The analysis in section B.1 has provided some key messages for the stakeholders actively involved in realising the smart grid. The need to go beyond the fundamentals, that is, smart metering and couple the smart grid with an advanced energy service infrastructure, is eminent. This should not be a standalone one for the sake of the smart grid, but amalgamated with the existing Internet applications and services so can further evolve by taking into account energy information, while the traditional grid processes may also benefit from prosumer interactions at other levels. The latter holds especially true for the three directions dealing with (i) monitoring, (ii) assessment/analytics, and (iii) control, where significant work still needs to be invested.

In a more detailed fashion some of the *findings* in the survey point out towards the following:

I. there is a need for better and more fine-grained access to data acquired by monitoring, even down to the device layer, as already presented in chapter 3

II. although there is a need to preserve privacy, there is also the necessity of sharing information and trading part of it in order to enjoy value-added services

III. users are willing to share their energy resources with the local community, in an effort to reduce their own energy costs, e.g. by local energy trading [25]

IV. users would allow third parties to manage and trade their energy resources (solar photovoltaic panels, etc.), by forming pVPP [61]

V. think favourably of the idea of smart and self-managed devices, but are unfavourable to third-party direct control of their consumption devices

These findings are in line with the findings from other surveys and reports. For instance in [66], fewer participants seemed interested in obtaining more usage information (in this case through an energy information display), and also in participating in demand response programs. However this interest is growing [28] and, as these findings indicate, goes beyond simple cost interest towards the community. In [154], the participants did not seem to have a satisfactory understanding of the electricity grid delivery, something that may be depicted also in the results. Smart appliance usage and participation in energy efficiency actions are in-line with the findings of others [66], also noting similar concerns about privacy and the appliance controlling. It is clear that multi-disciplinary research that goes beyond technology is needed, towards economics and behavioural science [154]. The final success of course is also bound to the specific conditions on user acceptance [119] in each country or region that can stipulate the uptake of the smart grid benefits [155]. Significant effort will need to be invested in modelling behaviour of prosumers [118] in order to be able to correlate it with key performance energy indicators and business scenarios of smart grids, as the main contribution of this thesis presented in chapter 5.

An interesting issue is how one should approach these findings, especially from the view of developing new applications and functionalities for the emerging smart grid. The traditional approach in the energy domain is to create monolithic applications, since usually the whole value chain, that is, the data acquisition, analysis, and partially control, were in the hands of the same stakeholder. However, with the liberalization of the energy market as well as the vision of the smart grid, there are now multiple stakeholders competing in multiple layers. Therefore, integration and interaction based on the traditional models would be not only anachronistic but impossible in the future. The quest then is towards finding commonalities, such as at the functional level, that may be realized by open platforms and services and may provide various views on the acquired data and enable further composition of them to more sophisticated ones [27]. Hence, section 3.2 addresses this eminent need for the so called common energy services that can be used as a basis for future development.

As electricity gets more expensive and technologies improve, the amount of internal generation, at the household level, is likely to rise. This will create new challenges for distribution-grid managers, as the power flow will originate from several points in the distribution grid. This is a big shift from the traditional model where power flowed in one direction. The good news is, at least, that the participants in this survey are willing to share their resources for a cost benefit (*findings* 3,4). Providing a convincing case to the users, especially tackling the aspects of intelligent device control (self o external) and usefulness of having it as part of a broader DR action, is a key area that needs to be addressed [151]. This also indicates that new business models [130] and services are required to enable this type of behaviour [21], so in chapter 5 the concrete capitalization proposals are made.

B.3 CONCLUSION

Coupled with the deployment of AMI and the increasing penetration of RES, new services and tools will be created to ease the new level of engagement customers will have with the system [34]. In order to target such efforts adequately and in the right direction, previous sections has presented the results of a survey directed at electricity end-users. The goal was to understand what types of information and services they would like to have access to, where they would like to access it, and how important privacy was for them. Additionally, the survey tried gauge how willing people would be to engage with their community [28] and join their energy resources [25]. This is in particularly important for SFERS, as some in scenarios collaboration of smaller stakeholders to achieve required forecast accuracy is envisioned (e.g. what section 4.3 experiments with).

The major insights analysed in section B.1 and some selected ones outlined in section B.2 show that customers want a better level of understanding of their behaviour. It was shown that the participants are willing to engage with their community and share their production surplus, with an aim to help the community or reduce their overall electricity costs. Furthermore, while the results have re-emphasized the need for strong privacy practices, they have indicated that privacy is negotiable, and that more effort is needed to understand exactly to what extent and in exchange for what. Additionally, methodologies to enable secure fine-grained sharing of data need to be investigated to accelerate innovation in the service space. The results were successfully applied to the realization of NOBEL energy services [21] and are considered for the main contribution of this thesis, in particular for pVPP scenarios due the importance of aggregation shown in section 4.2.2.

C About the Author

CURRICULUM VITAE

Surname and name	Mr. **ILIĆ Dejan**
Birth date and place	31.08.1983. Serbia, Belgrade
Residence	Karlsruhe (Germany)
Contact	E: dejan.ilic@ymail.com

Objectives:

Since 2010, I worked on EU research projects with the responsibility of pioneering solutions for Smart Grids in close collaboration with (international) partners from industry and academia. My current interests are towards **empowering active contribution** of the customers that are passive and isolated in businesses of today. I would like to leverage my knowledge with **next-generation technologies** in order to prototype the concepts that will enable active involvement of isolated millions.

Education:

Ph.D. in Computer Science	Karlsruhe, Germany
Karlsruhe Institute of Technology (KIT), Informatics (cum laude)	9. 2010 – 7. 2014

At **SAP SE** (in Karlsruhe) on numerous Smart Grid projects

Self-Forecasting Energy Load Stakeholders for Smart Grids
Prof. Dr. BEIGL Michael, Prof. Dr. TERZIDIS Orestis, M.Sc. KARNOUSKOS Stamatis (SAP)

M.Sc. in Computer Engineering	Trieste, Italy
University of Trieste, Faculty of Engineering (107/110)	10. 2006 – 2. 2010

At **SAP Research** (in Karlsruhe) on the Secure SCM project

Secure Business Computation by using Garbled Circuits in a Web Environment
Prof. Dr. BARTOLI Alberto, Dr. SCHRÖPFER Axel (SAP)

B.Sc. in New Information Technologies	Belgrade, Serbia
Advanced School of Electrical Engineering (7.81/10)	10. 2003 – 9. 2006

Management Program for Sectored Service Facilities
Prof. M.Sc. CIROVIĆ Zoran, Prof. Dr. OBRADOVIĆ Slobodan

Mindset and skills:

Team player that will take over responsibility for dedicated project tasks. Strong analytical, communication and decision making skills. Innovative thinker, open-minded and autonomous problem solving attitude. Ability to quickly understand challenges and to generate and promote new and creative ideas. Willing to learn new technologies and acquire proficient development skills.

Languages (reading, writing, speaking):

English:	advanced,	advanced,	advanced
Italian:	advanced,	intermediate,	advanced
German:	beginner,	beginner,	beginner
Spanish:	intermediate,	beginner,	intermediate
Serbian/Croatian:	native,	native,	native

Technical knowledge:

- Practical knowledge in the complete lifecycle of ICT solutions
- Programming languages: **Java**, C#, C++, JavaScript
- Other (modeling) languages: **SQL** (Transact SQL), UML, BPMN
- Strong knowledge of **R language** for statistical computing and data analysis
- Working experience with **forecasting algorithms** as Holt-Winters, ARIMA
- Good knowledge of in-memory computing and distributed systems
- Deep understanding of **SAP HANA DB** and **HANA Cloud**
- Working experience of web technologies like **REST**, HTML5, XML, **AJAX**
- Profound knowledge of **DBMS**, such as MySQL, and ORM, such as **Hibernate**
- Good knowledge of protocols as **TCP**, IP, UDP, HTTP(S), SOAP
- Working experience with **MS Dynamics ERP** solutions

Work experience:

Product & Innovation, SAP SE — Karlsruhe, Germany

Researcher, delivering state-of-the-art for Smart Grid projects — 2.2010 – 8.2014

- **SmartHouse/SmartGrid** – design and development of Enterprise Metering Platform
- Neighbourhood Oriented Brokerage ELectricity and monitoring system (**NOBEL**) – pioneering ideas for the complete IT solution (and evaluate in a real-world trial)
- SMARTgrid KeY nEighborhood indicator cockpit (**SmartKYE**) – business and technological M2M development with SAP technologies (HANA DB and HANA Cloud)
- All projects required: decision making, continuous **collaboration with partners**, writing **documentation**, knowledge dissemination, **supervision** of students

Research, SAP Research — Karlsruhe, Germany

Master thesis, Secure SCM project, Security&Trust research — 6.2009 – 11.2009

- Learning **cryptographic techniques** as Oblivious Transfer, SHA1, Garbled circuits
- Dynamic application development (Java, JSP, Servlet and JavaScript) with Tomcat
- Team problem solving based on specifications and confidential tasks

IT Consulting and Development, Adacta Italia s.r.l. — Trieste, Italy

Internship, MS Dynamics business ERP solutions — 2.2009 – 4.2009

- Experiencing **MS Dynamics ERP** business solutions – Navision 5.0 and Axapta 2009
- Hand on experience with component installation, configuration and customization

Edu. Institution, Advanced School of El. Engineering — Belgrade, Serbia

Bachelor thesis, complete management application project — 6.2006 – 9.2006

- C/S management application for service facilities as restaurants and bars
- Design and realization of DB using MS SQL Server
- Web Service development for **Bluetooth communication with Smart Devices**
- Smart Device and Touch Screen application development

Sales and Servicing, Elektro Elit d.o.o. — Belgrade, Serbia

Work experience, complete IT solution development — 2004

- Development with MS Access 2003 and MS SQL Server 2000
- Managing team members, decision making

Internet point, multiple premises — Belgrade, Serbia

Work experience, infrastructure planning and deployment — 2001 – 2003

Internet point, DEHU's Place d.o.o. — Belgrade, Serbia

Proprietary, management and business experience — 2000 – 2002

- Managing experience, business and employment organization
- Choosing appropriate workstations, servers and network equipment

List of Publications

2010

Stamatis Karnouskos, Per Goncalves Da Silva, and Dejan Ilić. Performance evaluation of web service enabled smart metering platform. In *1st International ICST Conference on E-Energy, 14-15 October 2010 Athens Greece*. Springer, October 2010

2011

Stamatis Karnouskos, Per Goncalves da Silva, and Dejan Ilić. Assessment of high-performance smart metering for the web service enabled smart grid era. In *Proceedings of the 2nd ACM/SPEC International Conference on Performance engineering*, ICPE '11, pages 133–144, New York, NY, USA, 2011. ACM

2012

Per Goncalves da Silva, Stamatis Karnouskos, and Dejan Ilić. A survey towards understanding residential prosumers in smart grid neighbourhoods. In *Innovative Smart Grid Technologies (ISGT Europe), 2012 3rd IEEE PES International Conference and Exhibition on*, pages 1–8, 2012

Joel Hoglund, Dejan Ilić, Stamatis Karnouskos, Robert Sauter, and Per Goncalves da Silva. Using a 6lowpan smart meter mesh network for event-driven monitoring of power quality. In *Smart Grid Communications (SmartGridComm), 2012 IEEE Third International Conference on*, pages 448–453, 2012

Dejan Ilić, Stamatis Karnouskos, and Per Goncalves da Silva. Sensing in power distribution networks via large numbers of smart meters. In *Innovative Smart Grid Technologies (ISGT Europe), 2012 3rd IEEE PES International Conference and Exhibition on*, pages 1–6, 2012

Dejan Ilić, Stamatis Karnouskos, Per Goncalves da Silva, and Martin Griesemer. An energy market for trading electricity in smart grid neighbourhoods. In *Digital Ecosystems Technologies (DEST), 2012 6th IEEE International Conference on*, pages 1–6, 2012

Stamatis Karnouskos, Per Goncalves da Silva, and Dejan Ilić. Energy services for the smart grid city. In *6th IEEE International Conference on Digital Ecosystems Technologies (DEST)*, pages 1–6, 2012

Stamatis Karnouskos, Dejan Ilić, and Per Goncalves Da Silva. Using flexible energy infrastructures for demand response in a smart grid city. In *The third IEEE PES Innovative Smart Grid Technologies (ISGT) Europe, Berlin, Germany*, pages 1–7, 2012

2013

Per Goncalves da Silva, Stamatis Karnouskos, and Dejan Ilić. Evaluation of the scalability of an energy market for smart grid neighbourhoods. In *IEEE 11th International Conference on Industrial Informatics (INDIN), Bochum, Germany*, pages 380–385, July 29–31 2013

Dejan Ilić, Per Goncalves da Silva, Stamatis Karnouskos, and Malte Jacobi. Impact assessment of smart meter grouping on the accuracy of forecasting algorithms. In *Proceedings of the 28th Annual ACM Symposium on Applied Computing (SAC)*, SAC '13, pages 673–679, New York, NY, USA, 2013. ACM

Dejan Ilić, Stamatis Karnouskos, and Per Goncalves da Silva. Improving load forecast in prosumer clusters by varying energy storage size. In *IEEE Grenoble PowerTech 2013, Grenoble, France*, 16 – 20 June 2013

Dejan Ilić, Stamatis Karnouskos, and Martin Wilhelm. A comparative analysis of smart metering data aggregation performance. In *IEEE 11th International Conference on Industrial Informatics (INDIN), Bochum, Germany*, pages 434–439, July 29–31 2013

Stamatis Karnouskos, Per Goncalves da Silva, and Dejan Ilić. Developing a web application for monitoring and management of smart grid neighborhoods. In *IEEE 11th International Conference on Industrial Informatics (INDIN), Bochum, Germany*, pages 408–413, July 29–31 2013

Stamatis Karnouskos, Dejan Ilić, and Per Goncalves Da Silva. Assessment of an enterprise energy service platform in a smart grid city pilot. In *IEEE 11th International Conference on Industrial Informatics (INDIN), Bochum, Germany*, July 29–31 2013

2014

Per Goncalves da Silva, Dejan Ilić, and Stamatis Karnouskos. The impact of smart grid prosumer grouping on forecasting accuracy and its benefits for local electricity market trading. *IEEE Transactions on Smart Grid*, 5(1):402–410, 2014

Dejan Ilić and Stamatis Karnouskos. Addressing energy forecast errors: An empirical investigation of the capacity distribution impact in a variable storage. *Energy Systems, Springer*, pages 1–14, 2014

Dejan Ilić, Stamatis Karnouskos, Per Goncalves Da Silva, and Sarah Detzler. A system for enabling facility management to achieve deterministic energy behaviour in the smart grid era. In *3rd International Conference on Smart Grids and Green IT Systems (SmartGreens), Barcelona, Spain*, 3–4 April 2014

José Javier Garcia, Juan José Cardenas, Regina Enrich, Dejan Ilić, Stamatis Karnouskos, and Robert Sauter. Smart City Energy Management via Monitoring of Key Performance Indicators. In *"Challenges Of Implementing Active Distribution System Management" CIRED Workshop 2014, Rome, Italy*, 2014