Predictive Data Analytics for Energy Demand Flexibility

Ph.D. Dissertation

Bijay Neupane

Dissertation submitted August, 2017

A thesis submitted to the Faculty of IT and Design at Aalborg University (AAU) and the Faculty of Computer Science at Technische Universität Dresden (TUD), in partial fulfilment of the requirements within the scope of the IT4BI-DC programme for the joint Ph.D. degree in computer science. The thesis is not submitted to any other organization at the same time.

Abstract

The depleting fossil fuel and environmental concerns have created a revolutionary movement towards the installation and utilization of Renewable Energy Sources (RES) such as wind and solar energy. The RES entails challenges, both in regards to the physical integration into a grid system and regarding management of the expected demand. The flexibility in energy demand can facilitate the alignment of the supply and demand to achieve a dynamic Demand Response (DR). The flexibility is often not explicitly available or provided by a user and has to be analyzed and extracted automatically from historical consumption data. The predictive analytics of consumption data can reveal interesting patterns and periodicities that facilitate the effective extraction and representation of flexibility. The device-level analysis captures the atomic flexibilities in energy demand and provides the largest possible solution space to generate demand/supply schedules.

The presence of stochasticity and noise in the device-level consumption data and the unavailability of contextual information makes the analytics task challenging. Hence, it is essential to design predictive analytical techniques that work at an atomic data granularity and perform various analyses on the effectiveness of the proposed techniques. The Ph.D. study is sponsored by the TotalFlex Project (http://www.totalflex.dk/) and is part of the IT4BI-DC program with Aalborg University and TU Dresden as Home and Host University, respectively. The main objective of the TotalFlex project is to develop a cost-effective, market-based system that utilizes total flexibility in energy demand, and provide financial and environmental benefits to all involved parties. The flexibilities from various devices are modeled using a unified format called a flex-offer, which facilitates, e.g., aggregation and trading in the energy market. In this regards, this Ph.D. study focuses on the predictive analytics of the historical device operation behavior of consumers for an efficient and effective extraction of flexibilities in their energy demands.

First, the thesis performs a comprehensive survey of state-of-the-art work in the literature. It presents a critical review and analysis of various previously proposed approaches, algorithms, and methods in the field of user behavior analysis, forecasting, and flexibility analysis. Then, the thesis details the flexibility and flex-offer concepts and formally discusses the terminologies used throughout the thesis.

Second, the thesis contributes to a comprehensive analysis of energy consumption behavior at the device-level. The key motive of the analysis is to extract device operation patterns of users, the correlation between devices operations, and influence of external factors in device-level demands. A novel cost/benefit trade-off analysis of device flexibility is performed to categorize devices into various segments according to their flexibility potential. Moreover, device-specific data preprocessing steps are proposed to clean device-level raw data into a format suitable for flexibility analysis.

Third, the thesis presents various prediction models that are specifically tuned for device-level energy demand prediction. Further, it contributes to the feature engineering aspect of generating additional features from a demand consumption timeseries that effectively capture device operation preferences and patterns. The demand predictions utilize the carefully crafted features and other contextual information to improve the performance of the prediction models. Further, various demand prediction models are evaluated to determine the model, forecast horizon, and data granularity best suited for the device-level flexibility analysis. Furthermore, the effect of the forecast accuracy on flexibility-based DR is evaluated to identify an error level a market can absorb maintaining profitability.

Fourth, the thesis proposes a generalized process for automated generation and evaluation of flex-offers from the three types of household devices, namely Wet-devices, Electric Vehicles (EV), and Heat Pumps. The proposed process automatically predicts and estimates times and values of device-specific events representing flexibility in its operations. The predicted events are combined to generate flex-offers for the device future operations. Moreover, the actual flexibility potential of household devices is quantified for various contextual conditions and degree days.

Fifth, the thesis presents user-comfort oriented prescriptive techniques to prescribe flex-offers schedules. The proposed scheduler considers the trade-off between both social and financial aspects during scheduling of flex-offers, i.e., maximizing the financial benefits in a market and at the same time minimizing the loss of user comfort. Moreover, it also provides a distance-aware error measure that quantifies the actual performance of forecast models designed for flex-offers generation and scheduling.

Sixth, the thesis contributes to the comprehensive analysis of the financial viability of device-level flexibility for dynamic balancing of demand and supply. The thesis quantifies the financial benefits of flexibility and investigates the device type specific market that maximizes the potential of flexibility, both regarding DR and financial incentives. Henceforth, a financial analysis of each proposed technique, namely forecast model, flex-offer generation model, and flex-offer scheduling is performed. The key motive is to evaluate the usability of the proposed models in the device-level flexibility based DR scheme and their potential in generating a positive financial incentive to markets and customers.

Seven, the thesis presents a benchmark platform for device-level demand prediction. The platform provides the research community with a centralized

repository of device-level datasets, forecast models, and functionalities that facilitate comparisons, evaluations, and validation of device-level forecast models. The results of the thesis can contribute to the energy market in materializing the vision of utilizing consumption and production flexibility to obtain dynamic energy balance. The developed demand forecast and flex-offer generation models also contribute to the energy data analytics and data mining fields. The quantification of flexibility further contributes by demonstrating the feasibility and financial benefits of flexibility-based DR. The developed experimental platform provide researchers and practitioners with the resources required for device-level demand analytics and prediction.

Dansk Resumé (Summary in Danish)

Brugen af fossile brændsler har skabt en revolutionær bevægelse inden for vedvarende energikilder som for eksempel vind og sol. Brugen af vedvarende energi skaber dog også nye udfordringer, da de fysisk skal integreres i den eksisterende infrastruktur og dernæst skal produktionen styres så den matcher det nuværende behov. Da behovet for energi er fleksibelt er det muligt at opnå ligevægt mellem produktion og behov, såkaldt dynamisk Demand Response (DR). Graden af fleksibilitet er dog ikke information som direkte er til rådighed eller leveret fra forbrugere, og det er i stedet nødvendigt automatisk at udtrække denne viden ved at analysere eksisterende data om energiforbrug. Gennem forudsigende analyse af det eksisterende energiforbrug kan man udtrække interessante mønstre, som kan bruges til at kortlægge graden af fleksibilitet i energinettet. Ved at indsamle strømforbruget for enkelt enheder bliver strømforbruget kortlagt på det lavest mulige niveau, hvilket giver et større løsningsrum af mulige planer for produktion i forhold til behov.

Analysering af forbrugsdata på enhedsniveau besværliggøres dog på grund af støj i det indsamlede data, samt manglende information om den aktuelle kontekst da dataene blev indsamlet. Det er derfor nødvendigt at designe nye metoder for forudsigende analyse specifikt til data indsamlet på enhedsniveau, samt at analysere ydeevnen for de foreslåede metoder. Dette Ph.D. studie er sponsoreret af TotalFlex Projektet (www.totalflex.dk), og er en del af IT4BI-DC programmet med Aalborg Universitet og TU Dresden som det respektive Hjemme og Værtsuniversitet. Det primære mål for TotalFlex projektet er at udvikle et omkostningseffektivt system baseret på det eksisterende marked, der ved brug af viden om det fleksible energibehov opnår både økonomiske og miljømæssige fordele for alle involverede parter. Det fleksible energibehov er for alle enheder er modelleret ensartet i form af et flex-offer. Brugen af et ensartet format gør det muligt at foretage aggregering og handel på tværs af energimarkedet. Med dette Ph.D. studie fokuserer vi på at analysere eksisterende forbruges data på enhedsniveau gennem brugen af forudsigende analyse, for effektivt at klargøre fleksibiliteten i deres energibehov.

For det første, indeholder denne afhandling en omfattende analyse af state-

of-the-art fra den eksisterende litteratur. Den præsenterer en kritisk analyse
og vurdering af de foreslåede algoritmer, metoder inden for brugeradfærd anal-
yse, forudsigelse og fleksibilitetsanalyse. Dernæst bliver koncepterne flex-offer,
fleksibilitet og den øvrige terminologi, som efterfølgende vil blive anvendt i
afhandlingen, formelt beskrevet.

For det andet, bidrager afhandlingen til en gennemgående analyse af hvor-
dan energi forbruges på enhedsniveau. Motivation for denne analyse er at
etablere brugsmønstre for enheder, korrelationer mellem brugsmønstre for flere
enheder og effekten af eksterne faktorer for energibehovet på enhedsniveau. En
cost/benefit analyse af enheders fleksibilitet bliver udført for at kategorisere en-
heder i forskellige segmenter ud fra deres potentiale for fleksibilitet. Derudover
præsenteres nye enhedsspecifikke præprocesserings metoder til at rengøre og
transformere data på enhedsniveau til et format brugbart for fleksibilitetsanal-
yse.

For det tredje, præsenterer afhandlingen forskellige modeller optimeret til
forudsigelse af energibehovet på enhedsniveau. Derudover bidrager den til fea-
ture engineering som en del af processen med at generere features fra tidsserier
over forbrugsbehov, som repræsenterer præferencer for bestemte enheder og
mønstre. Forudsigelsen af behovet gøres ved brug af disse features og anden
kontekstuel information til at forbedre præcisionen af forudsigelsesmodellen.
Dernæst evalueres flere forudsigelsesmodeller for at bestemme hvilken kombi-
nation af model, prognose horisont og data granularitet er bedst egnet til flek-
sibilitetsanalyse på enhedsniveau. Derudover bliver effekten af forudsigelses-
modellens på fleksibilitets baserede DR evalueret, for at identificerer hvilken
grad af fejl et marked kan absorbere før det ikke længere skaber profit.

For det fjerde, foreslår afhandlingen en generel proces for automatisk gener-
ering og evaluering af flex-offers fra tre forskellige gængse enheder, våd-enheder
, elektriske køretøjer og varmepumper. Den foreslåede proces forudsiger og es-
timerer automatisk tiden og værdien for enhedsspecifikke begivenheder. Disse
begivenheder repræsenterer hver enhed muligheder for fleksibilitet. De forud-
sagte begivenheder bliver kombineret for at generere flex-offers til brug for en-
hederne i fremtiden. Derudover bliver det konkrete potentiale for fleksibilitet
af gængse enheder kvantificeret i forhold til forskellige kontekster og dage.

For det femte, præsenterer afhandlingen teknikker til at ordinere flex-offers
med fokus på brugerkomfort. De foreslåede teknikker tager udgangspunkt i
balancen mellem de sociale og finansielle aspekter ved planlægning af flex-
offers. Konkret gøres dette ved at maksimere markedets finansielle fordele og
samtidig minimerer tabet af komfort for brugerne. Derudover præsenteres også
et afstands bekendt mål for fejl som kan kvantificere den faktiske ydeevne af
forudsigelsesmodeller designet for generering og planlægning af flex-offers.

For det sjette, bidrager afhandlingen til den omfattende analyse af de fi-
nansielle fordele ved dynamisk balancering af produktion og behov ved brug af
fleksibilitet på enhedsniveau. Afhandlingen kvantificerer de finansielle fordele
ved fleksibilitet og undersøger hvilket enheds specifikt marked maksimerer
potentiallet for fleksibilitet, med fokus på både DR og de finansielle incita-

menter. Herefter bliver en finansiel analyse af hver af de præsenterede teknikker forudsigelse model, generering af model for flex-offer og planlægning af flex-offer, udført. Den primære motivation er at udfører en brugbarhedsanalyse af de foreslåede modeller til DR på enhedsniveau, og deres potentiale for at skabe positive finansielle incitamenter for både markeder og forbrugere.

For det syvende, præsenterer afhandlingen en platform til at benchmarke forudsigelse af energibehovet på enhedsniveau. Platformen giver forskningsmiljøet et centralt lager af datasets på enhedsniveau samt forudsigelsesmodeller. Derudover indeholder platformen også funktionalitet til at facilitere sammenligning, evaluering og validering af forudsigelsesmodeller på enhedsniveau. Resultatet af denne afhandling bidrager til energimarkedet ved at materialisere visionen om at bruge fleksibiliteten i energiproduktion og energiforbrug til dynamisk at opnå en balance mellem produktion og forbrug af energi. De udviklede modeller til at forudsige energibehov og generation flex-offers bidrager også til både analyse af energidata og data mining. Kvantificeringen af fleksibilitet bidrager også ved at demonstrere muligheden for, og de finansielle fordele ved, DR baseret på fleksibilitet. Den udviklede eksperimentelle platform giver forskere og praktikere de nødvendige resurser til at analyserer og forudsige energibehovet på enhedsniveau.

Thesis Details

Thesis Title:	Predictive Data Analytics for Energy Demand Flexibility
Ph.D. Student:	Bijay Neupane
Supervisors:	Prof. Torben Bach Pedersen, Aalborg University
	Prof. Dr.-Ing. Wolfgang Lehner, Technische Universität Dresden

The main body of this thesis consist of the following accepted and submitted papers.

[1] Bijay Neupane, Torben Bach Pedersen, and Bo Thiesson, "Evaluating the Value of Flexibility in Energy Regulation Markets". *In Proceedings of the 2015 ACM Sixth International Conference on Future Energy Systems (e-Energy '15), Bangalore, India, pages 131-140, 2015.* (Input for Chapter 4)

[2] Bijay Neupane, Torben Bach Pedersen, and Bo Thiesson, "Towards flexibility detection in device-level energy consumption" *In Proceedings of the Second ECML/PKDD Workshop, DARE'14, pages 1-16, 2014.* (Input for Chapter 3 and 5)

[3] Bijay Neupane, Laurynas Šikšnys, and Torben Bach Pedersen, "Device-level Demand Forecasting for Flexibility Markets," *Submitted for journal publication*, Submitted on August, 2017.(Input for Chapter 3 and 6)

[4] Bijay Neupane, Laurynas Šikšnys, and Torben Bach Pedersen, "Generation and Evaluation of Flex-Offers from Flexible Devices," *In Proceedings of the ACM Eighth International Conference on Future Energy Systems (e-Energy '17), Hong Kong, pages 131-140, 2017*(Input for Chapter 3 and 7)

[5] Davide Frazzetto, Bijay Neupane, Thomas Dyhre Nielsen, and Torben Bach Pedersen, "User-comfort Oriented Prediction and Scheduling of Flex-Offers from Flexible Units," *Submitted for a Conference publication*, Submitted on April, 2017.(Input for Chapter 8)

[6] Bijay Neupane, Laurynas Šikšnys, and Torben Bach Pedersen, "DeMand: A Tool for Evaluating and Comparing Device-Level Demand and Supply Forecast Models". *In proceeding of the Workshops of the EDBT/ICDT 2016 Joint Conference, EnDM 2016, Bordeaux, France*, pages 1-6, CEUR-WS.org, ceur-ws.org/Vol-1558/paper17.pdf. (Input for Chapter 9)

This thesis has been submitted for assessment in partial fulfilment of the joint Ph.D. degree. The thesis is based on the submitted or published scientific papers which are listed above. Parts of the papers are used directly or indirectly in the extended summary of the thesis. As part of the assessment, co-author statements have been made available to the assessment committee and are also available at the Technical Doctoral School of IT and Design at Aalborg University and the Faculty of Computer Science at Technische Universität Dresden. The thesis is not in its present form acceptable for open publication but only in limited and closed circulation as copyright may not be ensured.

Acknowledgements

This thesis would not have been possible without the help and support of many people, whom I would like to acknowledge in this section.

I am especially grateful to my academic supervisor professor Torben Bach Pedersen for this research opportunity and for his continuous support and guidance throughout the entire Ph.D. study. I have a high regard for his inspiration, intellectual ability, and competence at work. Lots of gratitude is due for his intricate comments and recommendations on the content of this thesis. During my Ph.D. study, I visited Database Technology Group at Technische Universität Dresden for six months. I am again grateful to Prof. Dr.-Ing. Wolfgang Lehner for his creative and productive research collaboration which was of immense help to structure and improve this document.

I would also like to express my sincere gratitude and appreciation to my Ph.D. co-advisors Associate Professor Bo Thiesson, Post. Doc. Laurynas Šikšnys, and Dr.-Ing. Martin Hahmann for their guidance and support in learning different technologies during my Ph.D. research. I am very grateful for their valuable and constructive comments and reviews on the articles. I would like to say that this work would not have been possible without their advice and support.

I would like to thank the entire colleagues and staffs of the Database, Programming and Web Technologies group at Aalborg University for helpful atmosphere and productive collaboration. Special thanks to Søren Kejser Jensen and Kim Ahlstrøm Meyn Mathiassen for helping me in writing the Danish summary of the thesis.

I would like to express my sincere gratitude to my family and friends for constant encouragement and playing a positive role in completing this thesis. Finally, I would like to acknowledge the financial support of the Faculty of IT and Design, Aalborg University and the TotalFlex project, which was funded by the ForskEL program of Energinet.dk.

Contents

Chapter 1

Introduction

1.1 Background and Motivation

Concerns about the environment, energy security, and the depletion of fossil fuels have motivated many countries to increase the share of renewable energy sources in their total energy production and reduce the number of sources that depend on fossil fuels. This ambition has resulted into various challenges raised by European and international policies and programs. The EU member states' aim is to produce 20% of their total energy demand from renewable energy sources (RES) by 2020 and attain a more ambitious target of 27% of their total energy by 2030. Moreover, various countries have set their own RES targets, as summarized in Table 1.1.

Table 1.1: 2020 RES target and current status of various European countries (2015) (Source: Eurostat [1])

Country	Status2015	Target2020
EU	**16.7**	**20**
Belgium	8.0	13
Bulgaria	18.2	16
Czech Republic	15.1	13
Denmark	30.8	30
Germany	14.6	18
Estonia	28.6	25
Ireland	9.2	16
Greece	15.4	18
Spain	16.2	20
France	15.2	23
Croatia	29.0	20
Italy	17.5	17
Cyprus	9.4	13
Latvia	37.6	40
Lithuania	25.8	23
Luxembourg	5.0	11
Hungary	14.5	13
Malta	5.0	10
Netherlands	5.8	14
Austria	33.0	34
Poland	11.8	15
Portugal	28.0	31
Romania	24.8	24
Slovenia	22.9	25
Slovakia	12.9	14
Finland	39.3	38
Sweden	53.9	49

Achieved RES Target (%), 2014
0.0   100.0
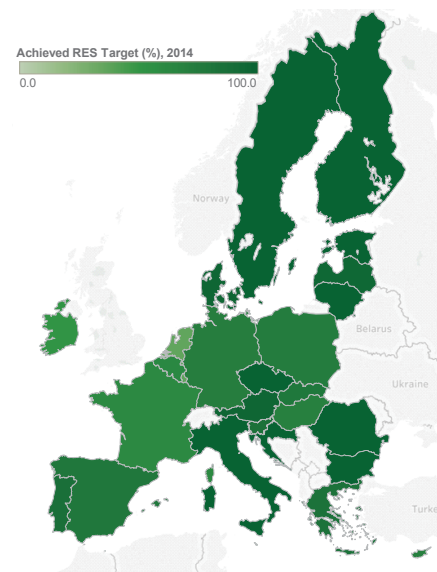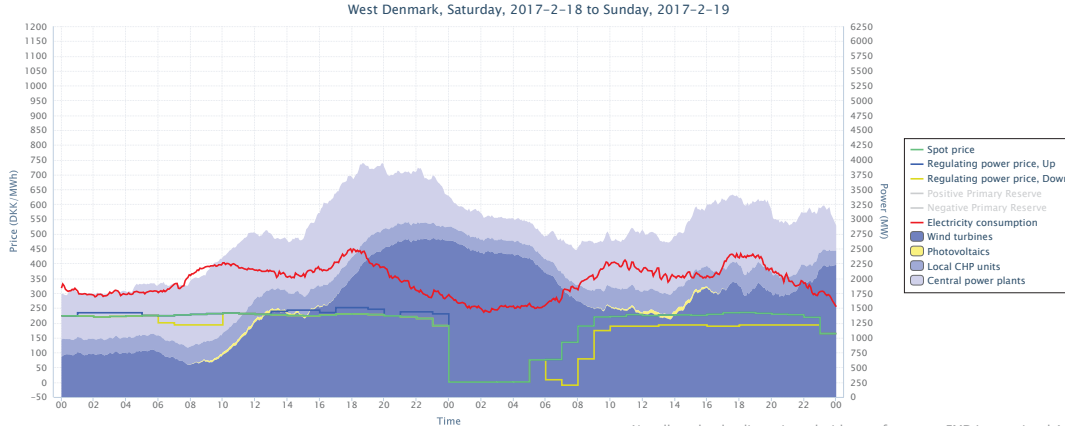
Figure 1.1: RES supply, electricity demand, and market prices for DK1 price area
Denmark. (Source: www.emd.dk/el/)

So as to achieve these targets, a major transformation in the grid system
has to be achieved. Firstly, a higher percentage of RES have to be installed
and integrated into the grid system. Secondly, the electrification of most of the
energy-drawing load units such as electric vehicles (EV), heat pumps (HP),
household appliances, etc., are required. However, the integration of a higher
percentage of RES into the grid system is challenging due to the changes in
market dynamics that come with it. Furthermore, it exposes the market players
to a financial risk due to uncertainties in supply, grid congestion, and changes in
market dynamics (prices and balances). Similarly, the electrification of energy
consumption units (mentioned above) increase uncertainty in demand, which,
in combination with intermittent RES, creates greater Demand Response (DR)
challenges. Figure 1.1 depicts an actual scenario of RES integration for the DK1
price area in Denmark. The RES supply for some of the hours is higher than
total demand, and the excess energy is either sold at a lower price or curtailed to
balance the market. However, for some hours, the supply from RES is only 30%
of the total energy demand, and the rest of the demand is fulfilled by utilizing
expensive fossil fuel-based energy sources. Hence, the low dispatch capacity
of RES makes it costly to support generations at peak periods. Further, the
introduction of distributed energy resources, e.g., grid connection of household
PV, has introduced additional stress on the existing grid system, which was
originally designed for the one-way flow of electricity.

The development of an intelligent grid system and the exploitation of avail-
able flexibility in demand from load units are key to confronting the challenges
of integrating larger percentages of RES and electrification of load units. In
this regard, various technologies and solutions have been proposed in different
areas, such as Demand Response schemes [2, 3], Energy Storage systems [4, 5],
Energy Management Systems [6–10], etc. Specifically, DR schemes propose
various incentives to motivate energy consumers (prosumers) to actively par-
ticipate in the demand response so that their energy demand (or supply) can
be varied to obtain a dynamic energy balance in the market. The PhD work

was being carried out as a part of one of the DR initiatives called TotalFlex [8], that provides a mechanism to extract and utilize the flexibility in demand units as an efficient solution using the concept of flex-offers proposed in the EU FP7 project MIRABEL [11]. The TotalFlex project addresses the challenges of balancing the energy consumption and demand by utilizing the novel concept of the flex-offer, which utilizes the flexibility in the operation of household and industrial load units to create the flexible consumption offers. The aim of the Ph.D. work is to exploit the flexibility potential of the household electrical loads for a dynamic energy planning and demand response and contribute to confronting the challenges of intermittent RES supply.

The flex-offer captures the flexible part of an energy demand for a household load, which can be exploited by market players such as TSO, DSO, and Aggregator to balance demand and supply better or to delay the costly grid upgrade. Consumers also directly benefit from this market approach by being offered better electricity prices and reduction in their total bills. The customers also contribute to the society by adopting or providing flexibilities in their consumption and helping to solve the load management problem for RES integration.

However, load usage patterns are stochastic and depend on user behavior and contextual information such as working habits, availability, etc. Furthermore, special events such as gathering and holidays might drastically change load usage, which influences the performance of energy demand and flexibility prediction models. The flexibility based demand response relies on the accuracy of the predicted flexibility as represented by a flex-offer. Hence, a higher error in the forecasted flexibility will increase market imbalances instead of solving the problem of integrating higher RES. The performance of a forecast model and confidence in the predicted flexibility can be drastically improved simply by moving the flexibility analysis to a lower data granularity, i.e., household or transformer level. However, a market player prefers to capture the best available flexibility which is only provided at the device-level (atomic flexibility) that gives the market more control over the load and manage the market balance. The device-level load data along with the contextual information are rarely available, and if available, they are very noisy. Hence, the extraction of useful patterns that provide the flexibility in load usage is a challenging task.

Different loads have different operating behavior in terms of operating duration, energy consumption, operating cycles, etc. Thus, a flexibility extraction and prediction model should be general enough to capture flexibility from a variety of loads. Further, a general format to represent the flexibility of various load types is required, such that trading and aggregation (if required) are simplified. Demand flexibility can be utilized by a market player for various purposes from day-ahead planning to intra-day balancing and auxiliary services. Thus, market players are always interested in prediction of flexibility for various horizons.

The introduction of the flex-offer concept and the requirement of designing a mechanism for user behavior analysis and prediction of future energy demand
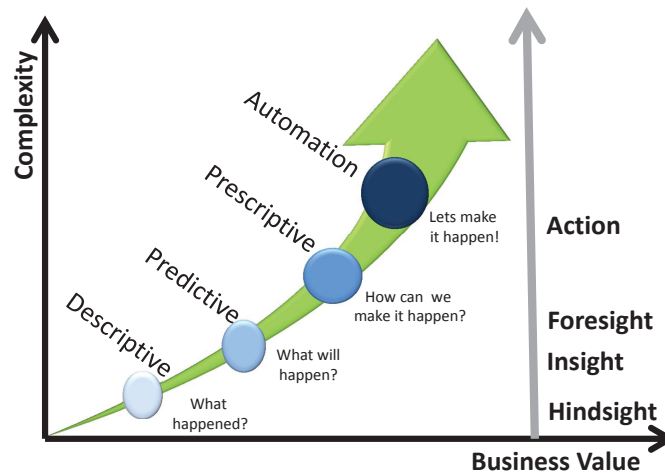
Figure 1.2: The steps of business analytics
(http://www.gartner.com/newsroom/id/2881218)

to extract flexibility in energy demand have led to his Ph.D. work. This Ph.D. work focuses on predictive energy data analytics with a focus on load operation pattern analysis and demands flexibility prediction. More specifically, it deals with the issue regarding accurate and precise demand forecast, flexibility detection, and flexibility forecast at the device-level, and uses the forecast for automated generation of flex-offer. Further, it aims to validate the concept of flexibility-based demand response by quantifying its financial viability.

1.2   Energy Data Analytics

Data analytics is the qualitative and quantitative process of transforming data into insights to enhance productivity and business gains. It is a part of the decision support system that helps to add to the human cognitive capacity for decision making by providing valuable insight into a business's performance, its existing behavior, and its future path. Data analytics involves discovering, interpreting, and representing meaningful patterns in data, and it covers the journey from data collection to decision making and ends with action. Depending on the business's requirements, complexity, and values and the intelligence it provides, data analytics can be divided into different segments (see Figure 1.2).

Descriptive analytics is the first stage of data analytics and focuses on the collection, aggregation, and consolidation of data. It is the most commonly used analytics and lays the foundation for more advanced analytics. It aggregates raw data to yield useful information that provides a clear and accurate picture of the past and current performance of the business. Descriptive analytics provides valuable insight into the past and helps to answer "What has happened or is happening now and why?".

Predictive analytics focuses on utilizing past information to predict probable future outcomes. It utilizes a variety of data mining, machine learning, and statistical methods to detect patterns and relationships in data and provides actionable insights that help businesses with effective planning and decision making. The predictive analytics estimates the likelihood of the future event and helps to answer: "What could happen in the future?"

Prescriptive analytics is the part of the analytics that focuses on finding the best course of actions for the future or current situations. It acts on the outcomes of the descriptive and prescriptive analytics and explores a number of possible actions to suggest a solution that could lead a business into a position of taking advantage or minimizing the risks of the future opportunities. It also provides the impact of each possible future action to ensure a business takes the best course of decision. It not only provides answers to what will happen but also answers why it will happen.

Data analytics for energy domain is a hot topic and fundamental in implementing flexibility-based demand response to achieve the RES target. With the increase in the number of energy resources (distributed generation), the connectivity of the grid becomes complex, and the structure will generate an enormous amount of heterogeneous data every second, such as device-level energy consumption data, household production data, grid capacity state data, etc. The right set of data analytics will help to transform these data into insights that help to maintain the grid stability. Furthermore, the predictive analytics will provide both short-term and long-term predictions of energy demands at various points of the grid that will enable the local optimization of the resources for balancing demand and supply dynamically. The general steps of predictive analytics are depicted in Figure 1.3. Moreover, the prescriptive analytics will give support in taking the best action for the future scenario by optimizing the market objective.

Within the energy data analytics, the Ph.D. work will focus on a comprehensive analysis of the energy consumption behavior of household devices so as to provide insights into the available flexibilities in their operations. Furthermore, it will perform predictive analytics of device-level consumption and develop solutions for demand and flexibility forecasts, leading to the generation of flex-offers for the flexible part of the energy demands. The provided solutions will help a market to obtain a clear picture of the available flexibilities within their portfolio and utilize the predicted flex-offers for various purposes, such as to plan their supply better, to obtain local balances, to correct any deviation in the spot market, etc.
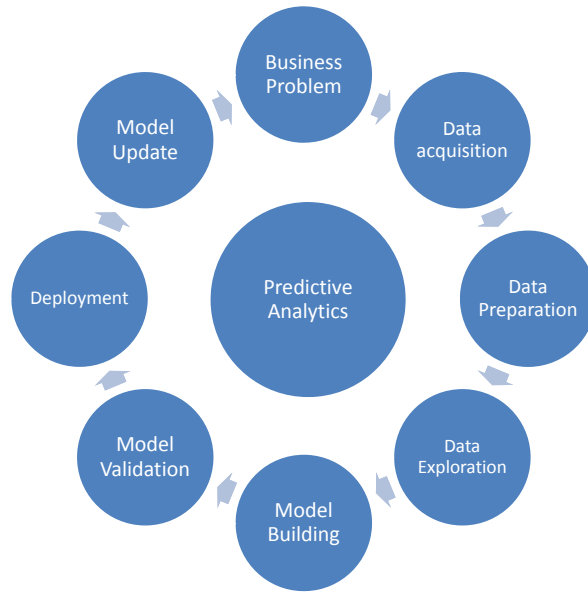
Figure 1.3: Predictive analytic process cycle
(http://blogs.sas.com/content/subconsciousmusings/2013/01/11/)

1.3   Demand Flexibility

Demand flexibility refers to the possibility of preponing or postponing a portion
(or even whole) of electrical energy demands from consumption and production
of various Distributed Energy Resources (DERs), satisfying user imposed and
other constraints. The TotalFlex project proposed a general way of capturing
and representing flexibilities in DERs using the concept of flex-offer proposed
in MIRABEL. Flex-offer offers a robust and generic way to describe flexibility
in electricity consumption and production of various DERs. An advantage of
flex-offers is that they can explicitly specify available flexibilities in a gener-
alized way and later they can be aggregated and disaggregated across various
dimensions, e.g., different DERs. A single flex-offer typically includes:

- Energy profile, having a number of discrete slices, specifies electricity
  consumption and production options over a device's active period of op-
  eration.

- Time flexibility interval specifies a time duration in which device's oper-
  ation can be preponed or postponed.

When using a flex-offer, no specific knowledge about the underlying DERs
is needed, whether the demands come from heat pumps, EVs, wet-devices, etc.
An example of a basic flex-offer is shown in Figure 1.4.

Figure 1.4 shows a simple example of a device-level flex-offer generated
from the extracted flexibility of a dishwasher. The flex-offer in the figure states
that the dishwasher could be activated anytime between 9 PM - 4 AM and
it operates for 2-time units. The figure also shows the energy profile for the
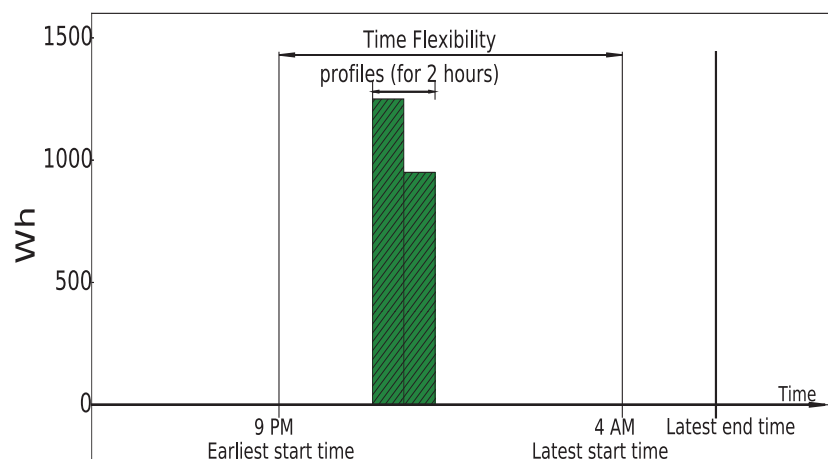
Figure 1.4: A sample flex-offer for dishwasher.

dishwasher, which represents the flexible demand for each unit time of the dishwasher's operation. It further has a constraint that the energy profile of the dishwasher cannot be changed, i.e., once activated it should be operated continuously for 2-time units.

Within demand flexibility, this Ph.D. work focuses on the flexibility extraction and flex-offer generation from the household consumption units, specifically dishwasher, washer dryer, Electric Vehicles, and Heat Pumps. The household appliances and other devices are becoming intelligent, and already today they can be set to run later, or response to energy price signals, e.g., washer dryer can be scheduled to operate later or operate freezer at cost optimized mode. Hence, the key assumptions of the Ph.D. work are the operation of some of the devices can be automatically controlled and users are willing to provide flexibility as their contribution to the demand management for some financial benefits. The control can be either performed by the existing market players such as BRP, DSO or may be delegated to a new entity such as an Aggregator. Further, the key working hypothesis of this Ph.D. work is, there exists detectable and predictable flexibility in usage patterns. The Ph.D. thesis models the flexibilities from various device types into a unified format of flexible consumption offers, i.e., flex-offers.

recre

1.4 Device-level Behavior Analysis

Device usage behavior analysis is the task of understanding when, how long, and how frequently devices are used in a household. A behavior analysis should provide a clear insight into the patterns and preference of users over a device usage. A better and sound understanding of the device usage behavior is fundamental to evaluate the possibility of device (user) participation in the dynamic DR schemes. A better understanding of the usage behavior enables the market
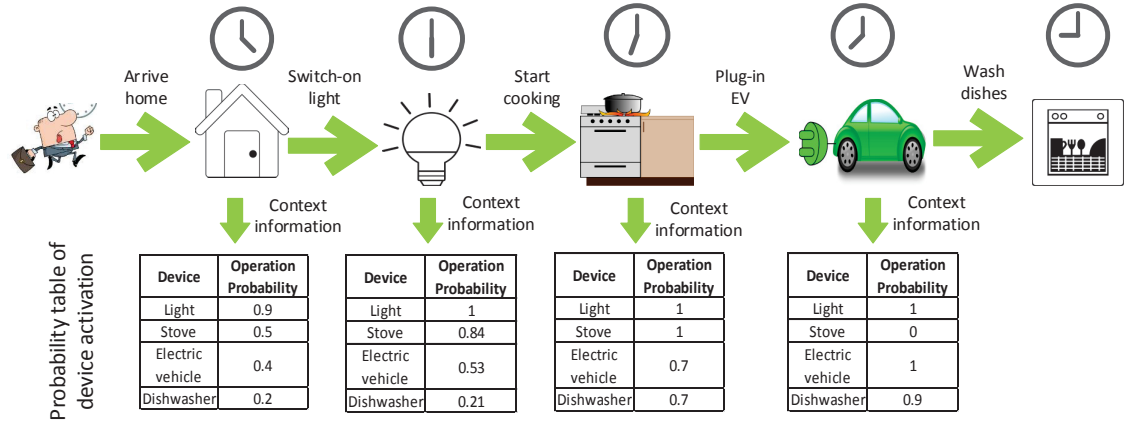
Figure 1.5: Influence of user activities on device activation probabilities.

players to exploit the information for more efficient portfolio planning, which aligns with the RES integration goals. Further, it helps in deciding what devices can be considered flexible enough to participate in DR and what devices are not, e.g., washing machine can be flexible but not television. Henceforth, to get the clear understanding of flexibility potential of a device, we should have a clear understanding how and when a user operates the device.

However, an increase in types of electrical appliances in households (EV, heat pump, and dishwasher), flexible working environment, the purpose of use, etc. makes the behavior analysis a complex task. Similarly, the energy consumption behavior of a household varies according to the contextual property of the house and the people living there, e.g., the size of the house, age, the number of occupants, their occupation, etc. Further, it is essential to understand the interaction and inter-dependencies between the various devices with perspective to everyday activities, e.g., cooking, cleaning, heating, traveling (EV). This information is crucial in providing foresight on the probable device usage based on the current user activity (see Figure 1.5). Insight into the energy consumption behavior and saving potential is a useful tool for market players to motivate users to contribute their flexibility.

The device-level data are often very noisy and obscure, creating difficulty in tracking the actual operating behavior of devices (when switched on and switched off). Further, as discussed before, a device usage behavior is interlinked with the various contextual information of a household, and this information is usually not available, increasing uncertainty over the analysis. This Ph.D. thesis provides solutions on the device-level data preprocessing. Further, the Ph.D. thesis utilizes the descriptive analytics on device-level data to extract meaningful flexibilities from devices usage behavior.

1.5 Device-level Demand Prediction

Prediction of future demand, supply, RES production, grid conditions at various nodes of the grid system help to improve the observability and effective planning for anticipated disturbance in the system. Specifically, an optimal scheduling of supply and demand in flexibility-based DR immensely relies on the prediction of the future device usages and the associated energy profiles. Moreover, this information is also fundamental to estimate available flexibilities in energy demands and generate (potential) flex-offers for the flexible part of future device operations.

The potential of flex-offers to be used to solve the local problems at the various hierarchy of existing grid system requires forecast models capable of handling data of various granularities. Previous research has shown that the energy-demand data at an aggregated level exhibits more regular patterns than the disaggregated data at the device-level. Most of the previous solutions that deal with a data characteristic of aggregated household or deregulated market level cannot be easily reused to formulate models for device-level data with highly fluctuating and uncertain characteristics. Similarly, as discussed before, depending on the requirement, flex-offers can be used in various markets such as the day-ahead market for planning, intraday and balancing market to correct deviations, etc. and market needs forecast models able to predict at various horizons. The capacity of flex-offers to capture flexibilities from a wide variety of devices (both producing and consuming energy) demands a model able to cope with the heterogeneous operating properties and evolving users' behavior. As mentioned in [12] most of the existing techniques are designed for static analysis of time series data and fails to maintain the efficiency with the evolving time series in distributed system architectures.

The complex energy market model combined with hierarchical grid architecture reflects the need of efficient and robust forecast models. Specifically, demand forecast along with automated detection of flexibility for a wide variety of flexible objects requires more dynamic, efficient, maintainable, and highly accurate models for seamless adaptation of flexibility based DR by the existing energy market. The main challenge of device-level forecasting is incorporated highly stochastic user consumption behaviors in a prediction model. Within the device-level demand forecasting, this Ph.D. thesis focuses on designing and evaluation of generalized device-level demand forecast models and also designing models for predicting all parameters required for generation of flex-offers. In this regards, the thesis evaluates the usability and effectiveness of existing forecast models for device level forecasting. Further, the thesis extends the flex-offer generation technique to the device-level, utilizing the input from the device-level demand forecast and flexibility forecast. In addition, this thesis will evaluate forecast horizon and data granularity best suited for the flexibility market regarding utility it brings to the market players.
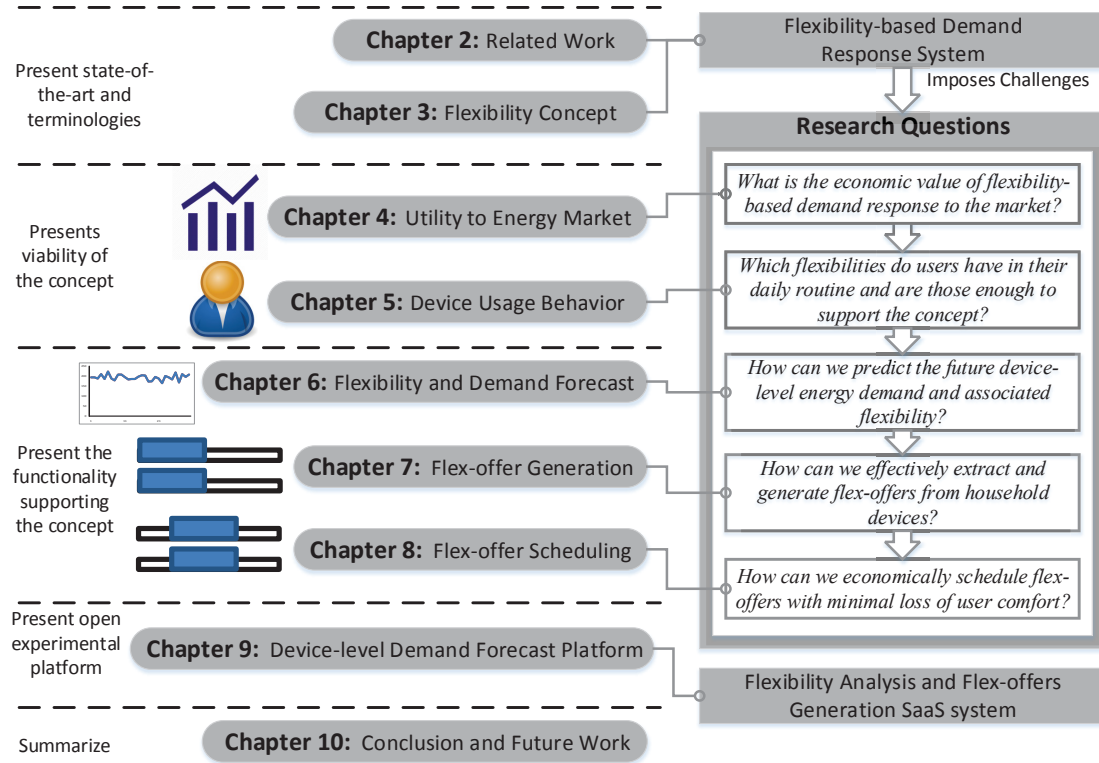
Figure 1.6: The structure of the thesis.

1.6   Thesis Overview

The remaining part of the thesis follows the structure shown in the Figure 1.6. The chapters of the thesis focus on exploring solutions specific to research questions imposed by flexibility based demand response (i,e., prescriptive analytics of flexibility based market) and have different goals, summarized below.

Literature Review Chapter 2 provides an overview of the existing works in the literature focusing on components of demand flexibility concept and provide directions towards solutions to the research questions shown in Figure 1.6. Specifically, the chapter presents review and analysis of various previously proposed approaches, algorithms, and methods in the field of user behavior analysis, forecasting, and flexibility analysis.

Flexibility Concept Chapter 3 details the flexibility and flex-offer concepts proposed in the MIRABEL and TotalFlex projects. The flexibility-based DR concept proposed by the projects leads to the research questions (shown in Figure 1.6) that have to be addressed for its effective implementation. In this regards, the chapter presents a strategy of mapping device operation sequences from various household devices into a generalized flex-offer format. Further, it introduces datasets used in the experiments and proposes a *cost-benefit* based

categorization of flexibility for devices in the datasets. Moreover, the chapter also discusses the spot and regulation markets and details the use of demand flexibility by these markets to confront RES challenges.

Utility to Energy Market Chapter 4 unravels the first research question regarding the economic value of flexibility based demand response to a market. The chapter performs an analysis on the impact of introducing demand flexibility in the energy market. Specifically, it investigates the changes in the market dynamics (energy prices and balances) brought by changing or shifting a certain amount of energy demand. Further, the chapter analyses the best flexibility type for a given market objective. The impacts are quantified based on the financial benefit or loss incurred to a market by adopting the demand flexibility.

Device Usage Analysis Chapter 5 investigates the second research question of assessing the types and sizes of flexibilities available in users daily routine. In this regard, the chapter presents a descriptive analytics on the device-level energy consumption, utilizing real-world data sets. Specifically, it explores users' device usage preference and the possibility of extracting flexibility from their usage patterns. Further, it discusses different device specific data pre-processing steps required for device-level flexibility analysis and quantifies the available flexibility in devices operations.

Flexibility and Demand Forecasting Chapter 6 resolves the third research question regarding the feasibility of predicting device-level energy demands and associated flexibilities. In this regards, the chapter assesses the accuracy and usability of the state-of-the-arts forecast models for device-level energy demand and flexibility forecasting. Henceforth, the chapter investigates the data granularity and forecast model best suited for the proposed flexibility based DR. The flexibility based DR depends on the accuracy of proposed forecast model, where an error in predicted demand may lead to higher market imbalances generating a substantial loss to the market. In this regard, the chapter analyses the financial benefits of flexibility in relation to the intrinsic quality of the demand forecast models.

Flex-offer Generation Chapter 7 explores solutions to the fourth research challenge of extracting and generating flex-offer from flexible household devices. In this regard, the chapter presents a generalized flex-offer generation and evaluation process (FOGEP). The FOGEP maps a device operation sequences to flex-offer attributes and presents device-specific steps involved in the flex-offer generation. It further discusses the input and output in each step and also present forecast models and context information required for the step. The chapter demonstrates the use of FOGEP for three different household device types, namely wet-devices, heat pumps, and electric vehicles.

Flex-offer Scheduling Chapter 8 investigates the solutions for the fifth research challenge of efficient scheduling of flex-offers with minimal loss of user comfort. In this regards, the chapter presents a novel user-comfort oriented prescription technique for scheduling of flex-offers. The scheduling technique performs a trade-off between the loss of user comfort and financial gain and prescribes a schedule that optimizes the combination of these two factors. The chapter also proposes a Flexibility-Aware Error (FAE) measure to quantify the performance of forecast models designed for a flexibility market.

Device-level Demand Forecast Platform Chapter 9 packages some of the functionalities and systems proposed during the exploration of solutions for the research questions (see Figure 1.6) into a device-level forecasting platform. In this regards, the chapter presents the DeMand system, an open online tool for device-level data analysis, design and evaluation of prediction models, and flex-offer generation.

Conclusion Chapter 10 summarizes the contributions, concludes the thesis, and presents future work.

The papers included in this thesis are listed in the following. Chapter 3 is based on Paper 2, 3, and 4, Chapter 4 is based on Paper 1, Chapter 5 is based on Paper 2, Chapter 6 is based on Paper 3, Chapter 7 is based on Paper 4, Chapter 8 is based on Paper 5, and Chapter 9 is based on Paper 6. As mentioned earlier, paper 3 and 5 are under review.

[1] Bijay Neupane, Torben Bach Pedersen, and Bo Thiesson, "Evaluating the Value of Flexibility in Energy Regulation Markets". *In Proceedings of the 2015 ACM Sixth International Conference on Future Energy Systems (e-Energy '15), Bangalore, India, pages 131-140, 2015*(DOI: `http://doi.acm.org/10.1145/2768510.2768540`)

[2] Bijay Neupane, Torben Bach Pedersen, and Bo Thiesson, "Towards flexibility detection in device-level energy consumption" *In Proceedings of the Second ECML/PKDD Workshop, DARE'14, pages 1-16, 2014.* (DOI: `http://dx.doi.org/10.1007/978-3-319-13290-7_1`)

[3] Bijay Neupane, Laurynas Šikšnys, and Torben Bach Pedersen, "Device-level Demand Forecasting for Flexibility Markets," *Submitted for journal publication.*

[4] Bijay Neupane, Laurynas Šikšnys, and Torben Bach Pedersen, "Generation and Evaluation of Flex-Offers from Flexible Devices," *In Proceedings of the ACM Eighth International Conference on Future Energy Systems (e-Energy '17), Hong Kong, pages 131-140, 2017*(DOI: `http://doi.acm.org/10.1145/3077839.3077850`)

[5] Davide Frazzetto, Bijay Neupane, Thomas Dyhre Nielsen, and Torben Bach Pedersen, "User-comfort Oriented Prediction and Scheduling of Flex-Offers from Flexible Units," *Submitted for a Conference publication.*

[6] Bijay Neupane, Laurynas Šikšnys, and Torben Bach Pedersen, "DeMand: A Tool for Evaluating and Comparing Device-Level Demand and Supply Forecast Models". *In proceeding of the Workshops of the EDBT/ICDT 2016 Joint Conference, EnDM 2016, Bordeaux, France*, pages 1-6, CEUR-WS.org, (`http://ceur-ws.org/Vol-1558/paper17.pdf`)

Chapter 2

Related Work

2.1 Literature Review

Demand Response: The integration of a higher percentage of renewable energy sources into the power grid creates huge challenges, both in regards to the physical integration and in regard to the management of the expected demand. Numerous smart grid projects are aiming at the efficient utilization of intermittent RES production to mitigate the effects of imbalances due to the integration of Renewable Energy Sources (RES), [13–16]. Further, various optimization techniques for bidding strategies considering the imbalance penalty has been studied in [17] [18] [19]. An optimal bidding strategy recognizing the imbalance penalty and allowed imbalance band has been discussed in [18]. Further, two different approaches for tuning the optimal bidding strategy to account the fluctuations of the generated power has been discussed in [17]. Concurrently, there are many types of Demand Response (DR) programs that have been adopted by energy markets [20], e.g., price-based DR [21–23], demand reduction bidding [24, 25], load shift strategy [26–28], etc. Further, techniques for integrating household devices into demand side management for leveling of fluctuating RES production has been explored in [29, 30]. In particular, the TotalFlex [8] project proposes a DR technique to actively control electricity consumption, including individual household devices, to confront the challenges of intermittent RES supply.

The mandatory requirement for Distribution System Operators (DSO) to install smart meters in all households and the introduction of smart devices has enabled an avenue of utilizing devices flexibilities for dynamic DR. Market players can utilize flexibility to compensate their deviation to a previous commitment, delaying the huge investment, or any other purpose. At the same time, consumers can participate in the flexibility-based DR contributing flexibilities in their device usage in response to financial or other incentives. Dynamic pricing schemes such as time-of-use, peak-time, consumption level pricing, etc. set energy prices such that users are motivated to shift the energy

consumption to off-peak period or align with RES production in response to financial benefit [21, 31–33]. However, users are insensitive towards price signals that limits the applicability of such approaches for a large deployment [34]. Moreover, these approaches require frequent user involvement and thus might suffer response fatigue (reduce in a number of participants) in the long run [35].

As an alternative to avoid frequest user involvement, direct centralized control of flexible devices has been extensively proposed for an automated management of flexibilities. Demand management with centralized control relies on a central unit responsible for direct control of energy consumption of devices. The unit controls devices at an individual or an aggregated level [16, 27, 36] to reduce user cost [37], to level peak loads [21, 31, 38], or to generate financial benefits to the distribution utilities [34], etc. For the individual device-level, the central unit collects the base and flexible energy consumption requests of devices and optimally schedules them within the flexibility bounds [39, 40]. The proposed approaches rely on the flexibility information to be provided by the consumer along with the demand request and thus still possesses challenges of response fatigue. At an aggregated level, the devices are grouped according to type or are clustered based on their energy consumption behavior [16, 29]. The central unit estimates the base and flexible energy demand from the group of devices and generates the load reduction or shift plan accordingly. This has challenges of satisfying individual user constraints such as temperature comfort range, minimum charge level for EVs, and maximum shiftable time for wet-devices [41]. Further, the flexibility highly depends on the individual user behavior and device technical details. Thus, the analysis at an aggregated level cannot accurately capture the flexibility, i.e., loses a major portion of the flexibility.

Device-level Consumption Data: The dynamic extraction and utilization of atomic flexibilities for DR require the access to the historical consumption data for devices. However, the device-level data are rarely available and even if available are very noisy to perform any analysis. There have been various efforts on disaggregating household energy demand into demands from individual load [42–46]. However, the disaggregation is a challenging task due to various reasons such as multiple operating states of devices, variation in energy demand for similar devices, similar consumption signal from different devices, etc. Furthermore, an initiative for providing a publicly available disaggregated energy dataset has been taken in [47, 48] together with a proposed hardware architecture for collecting household energy data.

Flexibility Analysis: Demand reduction and shifting potential and opportunity from the household devices have already been demonstrated through various studies, specifically for wet-devices [16, 32, 49], heat pumps (HP) [50–53], and electric vehicles (EV) [27, 38, 39, 54, 55]. Further, positive attitude of users to adapt to the changes by providing flexibility in their consumption behavior has been established in [16, 56]. However, users should be aware of their incentive for behavioral change, and it should be large enough to drive them towards offering flexibility [35]. Most of the work in the literature discusses the

potential and benefits of device-level load-shifting. However, the main question of how device flexibility in the next days can be automatically predicted from users' past consumption behavior is still unanswered. Various methods for extracting flexibilities from electricity time series has been discussed in [57], but the analysis utilizes the household level data. Further, [52,58] present methods for estimating the demand flexibility from heat pumps. However, it does not explicitly discuss the modeling flexibility in a way it can be effectively traded in the energy market. Further, the analysis only focuses on one flexible device type and misses the analysis on device type such as wet-devices, electric vehicles, etc.

Device-level Forecasting: Most of the flexibility-based DR projects (discussed above) depend on accurate forecast of both non-shiftable and shiftable energy demands. However, the stochasticity associated with device-level demand makes forecasting a difficult task, and the situation worsens in the absence of context information. Further, a market highly depends on the forecast accuracy for bidding. There have been improvements in the forecasting techniques for RES [59,60], but the forecast error remains too high. The higher forecast error causes a higher imbalance in the market, consequently increasing the loss. An experiment for estimating the cost associated with wind power prediction error [61] shows that, depending on the forecast horizon and granularity, the prediction error can reduce the total revenue by up to 10%.

In the past few decades, forecasting methods focusing on various research domains have been proposed. Forecast methods proposed in the literature include quantitative time series methods (e.g. Moving Average, Exponential Smoothing, ARIMA) and Machine Learning models (Artificial Neural Network, SVM, Fuzzy Logic). Various linear and non-linear regression methods have been proposed in the literature [62], as well as autoregression [63], time-varying (time-varying auto-regression [64]), etc. has been proposed to deal with forecasting in various domains. Recently, we have seen a surge in research and development of various forecasting techniques focusing on the energy domain. Existing forecasting techniques in the energy domain are based on both classical time series models and Machine Learning models [65,66] implementing various methods, e.g., ANN, FUZZY Logic, SVM, expert system, statistical model, etc. over a variety of different energy dataset and time horizon. Most of the work related to forecasting energy demand and supply are tailor-made and tuned to specific problems, regional data, or time-horizon with the result that a slight variation according to these specific modeling assumptions will degrade performance. Previous research [67,68] has further proposed incorporating other external information is to obtain robustness and generalizability of the model, but still, these assumptions fail under the highly flexible and dynamic market for energy supply and demand management as proposed by TotalFlex.

A signature-based pattern matching technique to predict the device future consumption is proposed in [69]. The proposed system only predicts the power consumption for the next couple of hours for currently operating devices. Sim-

ilarly, [70] narrows the scope to predicting the deactivation times of currently operating devices. These approaches conflict with the forecast horizon requirement of flexibility based DR, which requires beforehand flexibility information for efficient scheduling of supply and demand. A probabilistic model for predicting the device status, operation time, and duration is proposed in [71]. The analysis was performed with simulated data and does not address the dynamic and stochastic behavior associated with real-world device operation data.

Forecast Evalutaion: The unpredictability of the user behavior creates challenges in achieving a higher accuracy for device-level demand forecasting. Further, the traditional error measures such as RMSE, MAE, Precision, etc., which perform a point-wise comparison of the actual and predicted values do not quantify the actual performance of the forecasting models in relation to the flexibility market. For example, 1-hour difference between predicted and actual activation of a device is an error with a traditional metrics. However, since devices are in an idle state for most of the time between consecutive activations, 1-hour difference might not be an error considering demand flexibility as long as the operation of the device finishes before next activation. A distance-based device-level forecast method has been proposed in [71]. However, the work is evaluated on simulated devices that do not capture the stochasticity associated with real device usage, and no evaluation of the effects of incorrect predictions on the market and users has been performed.

Utility Evaluation: The implementation and acceptance of the flexibility-based DR depends on the financial benefits it generates to its stakeholders. Although there have been some works on quantification of the benefit of load shifting [34, 72], the analyses are based on markets with a less integration of RES compared to Denmark where RES fulfills more than 40% of the electricity demand. Further, most of the previous work focuses on the quantification of the reduction in the customer's energy bill [72–75]. However, in a grid system with higher percentages of RES, it is rather players like BRPs and DSOs that generate substantial savings by avoiding the regulation market and network congestion, respectively.

Summary: Despite the immense possibility of economically extracting flexibility, to obtain dynamic demand and supply balance, and users' willingness to provide flexibility in their energy demand, there is a number of important challenges that have been either only partially addressed or still remain unexplored.

(a) The economic assessment of the benefits of demand flexibility in various energy markets and for various stakeholders needs to be further explored.

(b) Descriptive analytics of device-level data with a focus on flexibility market also need to be further explored.

(c) A general representation of the flexibility is needed, which has been discussed in [76] but not generalized to all device types [77–79].

Forecasting Platform (Chapter 9)

Utility Modeling (Chapter 4)

Optimization(Chapter 8)

Flex-offer Generation (Chapter 7)

Forecasting (Chapter 6)

User Behaviour Analysis (Chapter 5)
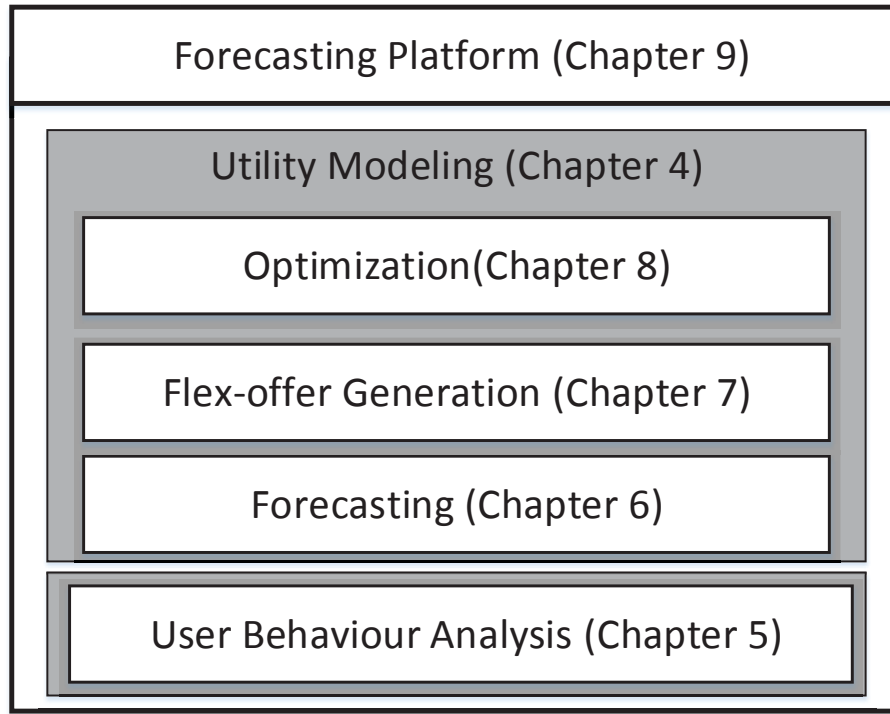
Figure 2.1: Thesis contributions

(d) An assessment of effectiveness and usability of the standard forecast models for device-level energy demand forecasting at various data granularity and horizon need to be further explored. Further, a metric to evaluate the performance of forecast model that consider the demand flexibility is still missing.

(e) A simple and robust process to capture flexibility from all household devices with minimal user intervention is still missing [80].

(f) A comprehensive open platform for benchmarking, comparing, and reevaluating device-level forecast models is still missing.

2.2   Thesis Contribution

To fill the missing gaps for effective implementation of device-level flexibility-based demand response market, this thesis makes the following contributions:

- This thesis models the flexibilities from various device types in a unified format represented as flex-offers that facilitate, e.g., aggregation and trading (Section 3.1).

- This thesis proposes a number of structural models to capture the relationship between the market power prices and regulation volumes. Further, it formulates the effect on the regulating power prices that are

caused by the fluctuation in the market balance as a consequence of shifting flexible demands (Section4.1).

- This thesis quantifies the utility of flexibility-based demand response to energy market for various flexibility types and market objectives. Further, it decomposes the actual utility into the direct utility and the utility due to the side effects of shifting flexible demand (Section 4.3).

- This thesis proposes data pre-processing techniques specific to the device level data analytics (Section 5.2).

- This thesis presents state-of-the-art analytics of device level energy consumption that analyzes the patterns and periodicity associated with individual devices with a focus towards characterizing and detecting flexibility to be used in the flex-offer framework (Section 5.3).

- This thesis quantifies the flexibility potential of devices in various dimensions and experimental settings (Section 5.3).

- This thesis proposes a number of features that reliably capture usage patterns and address the requirements of a device-level demand forecast (Section 6.1).

- This thesis formulates a set of equations for quantifying the financial benefits of flexibility in relation to the intrinsic quality of a demand forecast model (Section 6.2).

- This thesis proposes a generalized flex-offer generation and evaluation process to capture flexibility from a variety of household devices (Section 7.1).

- This thesis presents an economic assessment of flexibility in the spot and regulation markets and proposes the best market to trade flexibility (Section 7.3).

- This thesis proposes a novel flexibility-aware error (FAE) measure to evaluate the performance of forecasting models (Section 8.2).

- This thesis proposes a novel user-oriented prescription technique for scheduling of predicted flexible demand that considers both the social and financial aspects of demand shifting (Section 8.3).

- This thesis presents an open benchmark platform for device-level demand and flexibility forecasting (Section 9.1).

Chapter 3

Flexibility Concept and Terminology

In this chapter, we detail the flexibility and flex-offer concepts proposed in MIRABEL and was further enhanced by the TotalFlex project. The concepts are discussed in relation to the operating sequence of a smart device. The presented concept and flex-offer representation will be used throughout the thesis as a basis for predictive analytics on the flexibility potential of household energy demand. In this chapter, we first present various actions representing a general operating sequence of a smart device and model flex-offer as an attribute representing a combination of these actions. Then, it presents an overview of all the dataset used in this thesis along with the cost-benefit categorization of flexibility potential of household devices. Finally, it discusses the Nordic electricity market that will be used for demonstrating and validating market potential of the flex-offer concept.

3.1 Smart Devices

In this section, we consider different types of smart devices, whose electricity consumption can be externally controlled (e.g., dishwashers, HPs, EVs). We describe the sequence of activities involved in an operation of a smart device. Finally, we present the *flex-offer* (FO) concept that captures the flexibility incorporated in the various activities.

3.1.1 General Operation of Smart Devices

A smart device is a (IoT enabled) device that can be controlled by an external controller such as smart meter, etc. Figure 3.1 depicts the general sequence of actions performed during an operation of a smart device. An operation of a smart device starts with a user performing the *Switch-on* action that signals
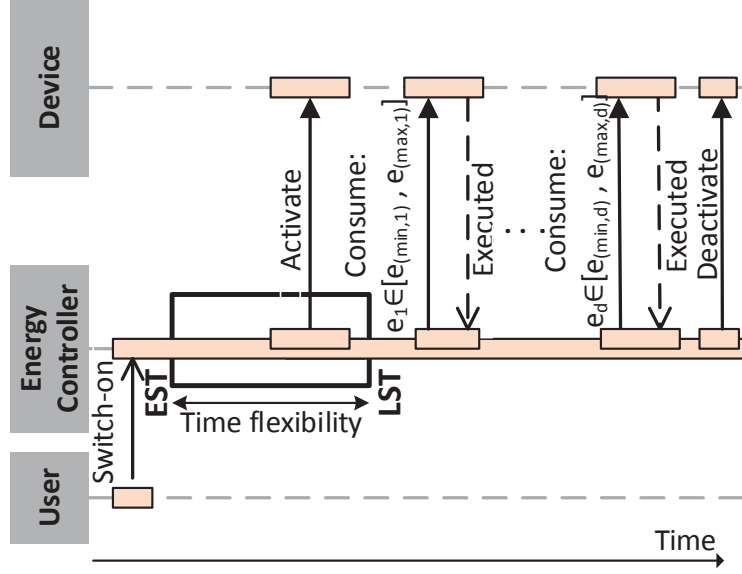
Figure 3.1: Sequence diagram for smart devices.

an Energy Controller (EC) to utilize the device to perform a certain task(-s). An EC is a device-aware external unit that decides on the execution time and the amount of energy to be consumed for a task. Henceforth, to operate the device, the EC triggers the *Activate* action within a time range bounded by the Earliest Start Time (EST) and Latest Start time (LST). The EST and LST are the earliest and the latest timestamp for the EC to trigger the device activation. In this thesis, the time is discretized into equal-sized units. Immediately after the activation of the device, the EC sends the first *Consume* action to the device.

The *Consume* action signals the device to perform a task consuming $e_t \in \mathbb{R}$ amount of energy at the time t. Further, the value of e_t, selected by the EC, lies within a range $[e_{(min,t)}, e_{(max,t)}]$ defined by a minimum $e_{(min,t)}$ and a maximum $e_{(max,t)}$ energy bounds, i.e., $e_{(min,t)} \leq e_t \leq e_{(max,t)}$. After the execution of the task, the device sends an *Executed* acknowledgment to the EC, reporting the actually consumed amount of energy and other device/task status details. The EC consecutively executes the *Consume* action for every 1 hour interval until the completion of all task(-s). Hence, $\langle e_1, \ldots, e_d \rangle$ represents an energy profile for $d \in \mathbb{N}$ consecutive time intervals, denoted as the *operation duration*. For example, the device could be an electric vehicle (EV) that consumes electrical energy e_t for each time unit of charging, i.e., each Consume action. Then, $\langle e_1, \ldots, e_d \rangle$ represents the energy profile for the duration of d time units of charging. Finally, at the end of the operation, the EC sends the *Deactivate* signal to the device.

3.1.2 Flexibility and Flex-offers

The term flexibility denotes a flexible energy demand and is represented in two dimensions. The first dimension is the *time flexibility*, which represent the possibility of preponing or postponing a portion of a demand for energy. The second dimension *amount flexibility*, is the range between maximum and minimum energy demand at a particular point in time. For example, if we consider a demand for a dishwasher, *time flexibility* represents the possibility of shifting the activation time to better match an anticipated surplus production from RES. Similarly, the *amount flexibility* represents the volume of power demand from, say, electric heating that can be scaled up or down according to the market requirement.
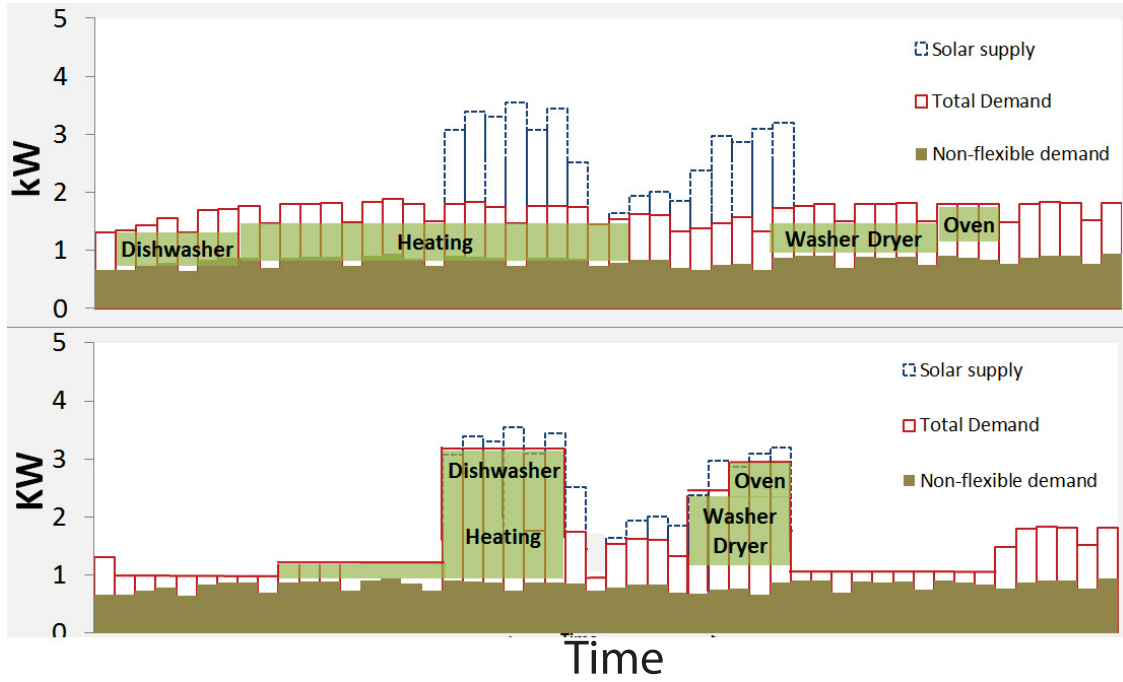
Figure 3.2: Energy demand and supply, before and after demand flexibility management (using flex-offer). The non-flexible part of the demand includes lightning, TV, etc.

Hence, in relation to Figure 3.1, we define the *flexibility* as the potential to amend the energy profile and the time the *Activate* action occurs, and we represent this flexibility as a so-called *flex-offer* (FO) [11].

Definition 1. *A flex-offer f is a tuple $f = ([t_{es}, t_{ls}], p)$, where $[t_{es}(EST), t_{ls}(LST)]$ is the time interval during which to trigger the Activate action and p is the energy profile. p is a sequence of slices $\langle s_1, \ldots, s_d \rangle$, where a slice s_t is a continuous range $[e_{min}, e_{max}]$ defined by the minimum e_{min} and maximum e_{max} energy bounds, and d is the number of slices in p.*

Figure 3.3 shows an example of an FO f representing the energy demand for
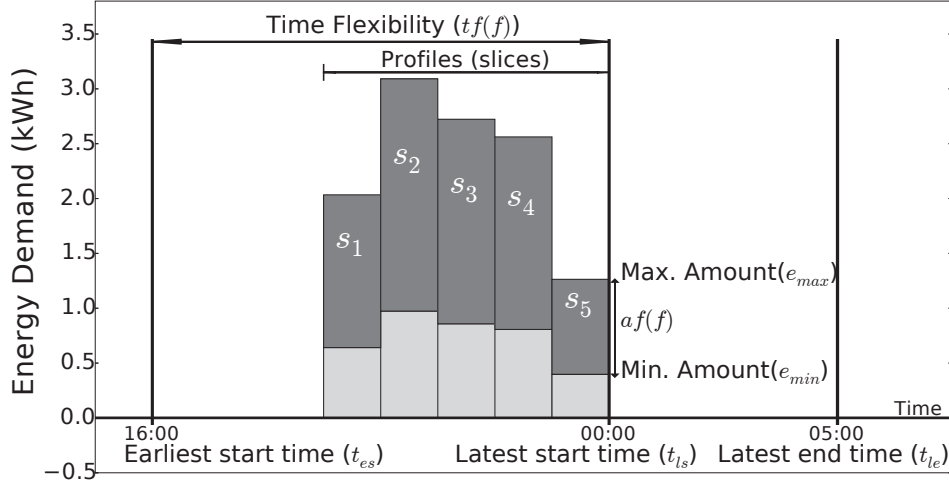
Figure 3.3: A sample FO from EV.

a single charging event of an EV. The FO in the figure states that the EV could be charged starting anytime between 4 PM and 12 AM. The FO has an energy profile with five consecutive slices: $s_1 = [0.6, 2], s_2 = [1, 3], s_3 = [0.9, 2.7], s_4 = [0.8, 2.6]$, and $s_5 = [0.4, 1.2]$, where the first and the second elements represent the height of the light-shaded (e_{min}) and dark-shaded (e_{max}) bars in the figure, respectively. The time flexibility $tf(f)$ of the FO f is the difference between LST and EST, i.e., $tf(f) = t_{ls} - t_{es} = 8$ hours. Further, t_{le} represents the Latest End Time (LET) of the last slice and is calculated as $t_{le} = t_{ls} + d$. Similarly, the amount flexibility $af(f)$ is given by the sum of the difference between amount bound of all slices, i.e., $af(f) = \sum_{i=1}^{d} e_{(max,i)} - e_{(min,i)} = 5.8$ kWh.

Similar to Figure 3.3, an FO can represent the flexibility of *all* the aforementioned device types, by applying the device-specific flexibility constraints. For example, e_{min} and e_{max} will be equal for a washer-dryer as it operates at a fixed power level for a given setting. Depending on the device type the time flexibility can be of three types:

- *Forward Time Flexibility*: The flexible energy demand can only be shifted forward in time, to a time later than the initial planned start time for the demand.

- *Backward Time Flexibility*: The flexible energy demand can only be shifted backward in time, to a time before the initial planned start time for the demand.

- *Bi-directional Time Flexibility*: The flexible energy demand can be shifted in both directions of time.

The flex-offers extracted from individual devices represents *micro flex-offers*. The computational complexity of optimally utilizing the flexibility of each micro flex-offer derived from millions of individual devices is too high to make

direct scheduling feasible. Thus, multiple micro flex-offers are aggregated into fewer larger flex-offers, known as *macro flex-offers*. The quantity of energy flexibility for each timestamp in a macro flex-offer will depend on the aggregation technique used, the profile of the individual micro flex-offers, and the market requirement. Various techniques to aggregate micro flex-offers to macro flex-offers have been described in [81]. However, we will apply a simpler view on aggregation (where required) in this thesis in order not to take the focus away from the detection and generation of atomic flex-offers and investigating its impact on the energy market. The simplified aggregation technique assumes that the energy profile of a flex-offer spanning multiple time units can be broken into multiple independent offers, one for each time unit. In this case, the aggregation simplifies as a simple grouping of the flex-offers (with the same time flexibility) at each time unit, resulting in the energy profile of a macro flex-offer being simply the sum of the values for the underlying micro flex-offers. For example, a macro flex-offer with an amount flexibility of 100 MWh could be aggregated from $50K$ micro flex-offers each with 2 KWh of amount flexibility. Although this simplified view may seem unrealistic at first, we emphasize that the thesis aims to investigate the case of a huge amount of micro flex-offers, were far from all flex-offers need to be considered to sustain the market balance. The above assumptions are therefore reasonable for the selected smaller part of the flex-offers that are activated during the balancing.

3.2   Device-level Data

This section presents detailed device-level energy consumption datasets used in this thesis along with generalized device-level data collection architecture. Further, we categorize the available devices into different categories according to their flexibility potential. Finally, we present the device type specific attributes that represent activities involved in an operation of a smart device.

3.2.1   Data

1) Zensehome [82]: Zensehome is a smart house system manufactured by the Zense Technology, a partner of the Totalflex consortium. Zensehome data is a closed device-level dataset containing the average power readings in watts for individual devices. The dataset is available for 13 different households and buildings; each containing 3 to 100 devices that include devices like dishwasher, washer dryer, lights, freezers, etc. The dataset is logged at a frequency of once every 15 minutes and is collected from January 2014 to October 2015. Table 3.1 shows characteristics of the collected data across some of the individual houses.

2) INTrEPID [48]: This is an open data that contains energy consumption profiles for household devices recorded at various frequencies. The dataset contains energy consumption profiles for 61 households in Denmark and Italy,
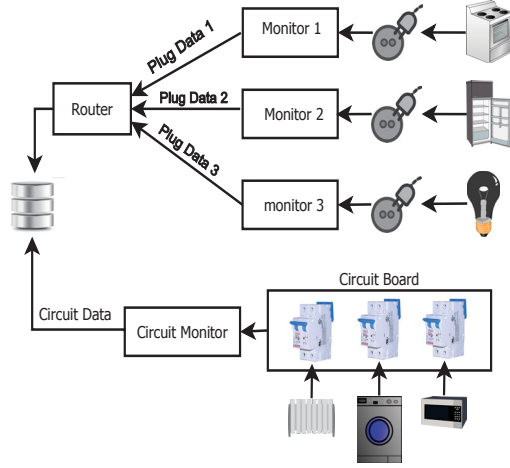
Figure 3.4: The REDD hardware architecture for data collection (adapted from REDD [47]).
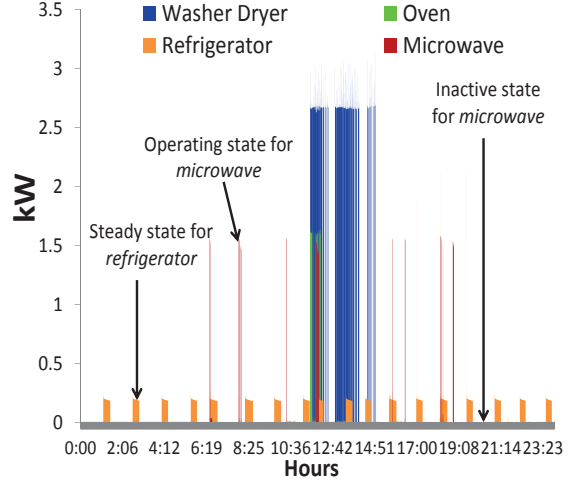
Figure 3.5: Power demand for selected devices over the course of a day (April 30, 2011; *house 1*).

Table 3.1: Data details for each house.

Device Type/ House Number	4351	12680	41222	50404
Lightning	25	18	15	78
Refrigerator	3	1	1	4
Washer Dryer	0	2	2	2
Heating	1	5	2	8

each containing time series profiles from 10 - 15 devices. However, in this thesis, we only use the energy profiles for wet-devices from 30 different households, each containing at least one of the devices washers, dryer, or dishwasher.

3) Heat Pumps: This is a closed data of power demand for HPs from 50 households in Denmark collected at a 5-min resolution. The HP dataset includes the ambient and room temperatures, and are annotated with various context values such as family size ranging from 1-5 adults per house, house area ranging from 80-700 m^2, etc. Table 3.2 depicts the parameter values for HPs and contextual information for some of the houses used in the experiments. The thermostat parameters of a particular HP and house are determined by fitting a model to the historical dataset, discussed in Section 7.2.

4) EVnetNL [83]: This is a closed EVs charging data of privately owned cars, charged at public charging stations. The dataset contains the charging profiles for 30K Smart Card (30K EV assuming each EV represented by one smart card) collected over a duration of 1 year (Jan-Dec 2015). The dataset consists of 11M charging log points collected at a resolution of 15 minutes.

5) REDD [47] : This is an openly available data consisting of energy

Table 3.2: LTI model parameters values and contextual information for selected houses.

House No.	$R[°C/kW]$	$C[Wh/°C]$	η	Adult	Children	Area(m^2)	Model
1	2.22	60.43	3.58	2	0	199	DHP-L 8
2	2.3	81.13	4.36	-	-	215	Fighter 1245
3	3.4	106.12	4.6	3	1	180	Fighter 1245-8
4	9.01	13.66	2.2	2	2	146	DHP-OPTI PRO
5	5.41	184.54	2.06	-	1	201	Queen VV9 DC

consumption profiles for six different houses, each containing profiles for 16 to 24 individual devices, and is collected in April to June, 2011. Data is available for a varying number of days for the different houses; ranging from 15 to 35 days; see the third column of Table 3.3. Some days only have partial records, and across all houses, the total number of days with a record for at least one hour of the day is 130 days. The REDD dataset is collected at the main level, circuit level, and plug level, where the plug level data is used to log devices in cases, where multiple devices are grouped to a single circuit. The plug level data were collected using a wireless plug monitoring device and circuit level data were collected using a emonitor connected to a circuit breaker panel, as illustrated in Figure 3.4. The available dataset was recorded at various frequencies: $15kHz$ for main phase, $0.5Hz$ for circuit level, and $1Hz$ for plug-level. For the main phase, data were written to the log in buffers of one second and for the circuit and plug level once every three seconds. Figure 3.5 shows an example of power demand for devices over the course of a day.

Table 3.3: Data details for each house.

House Number	Days Span	Days	#Channels	#Devices
House 1	36	35	18	11
House 2	34	15	9	9
House 3	44	23	20	13
House 4	48	30	18	12
House 5	44	9	24	15
House 6	23	18	15	11

Table 3.3 shows characteristics for the collected data across the individual houses in the dataset. The *Days Span* column represents the total number of days between the start and ending date in the time series, whereas *#days* is the total number of days with at least one hour of available data. Similarly, *#channels* represents the number of data collection points (plug and circuit) in a house and *#devices* is the number of unique devices available in the house.

6) Market Data: In this thesis, the financial potential of the proposed methods are evaluated on the Nord Pool spot and regulating market in the DK1 (West Denmark) region, the region which produces the largest part of the wind power production in Denmark. We obtained the time series dataset from

Table 3.4: Sample Market Data

Date	Hour	Up Regulation Price	Down Regulation Price	Up Regulation Volume	Down Regulation Volume	Spot Price	Wind Power Production	Energy Demand
1/1/2014	0	222.43	113.00	200	0	113.01	1709.2	1893.8
-	-	-	-	-	-	-	-	-
-	-	-	-	-	-	-	-	-
1/1/2014	9	189.83	137.91	0	0	189.84	1359.6	1896.9
1/1/2014	10	183.71	164.6o	0	-84	183.72	1413.3	2027

the Danish TSO Energinet.dk[1]. For the financial evaluation, we utilize the hourly resolution market data set of a length corresponding to the experiment (detailed in individual chapters). The data set consists of 9 different attributes as shown in Table 3.4.

3.2.2 Device Categorization

Our analyses of the devices in a household pertain to the possibility extracting the operational flexibility, which is evaluated based on the cost and benefit of utilizing it under the TotalFlex scenario. In that respect, we define *cost* and *benefit* as follows:

- *Cost:* The loss of user-perceived quality caused by accepting flexibility.

- *Benefit:* The available time and energy flexibility for the device.

According to this cost and benefit trade-off for flexibility, we have classified type of devices into three different *flex-categories*.

- *Fully-flexible* : High benefit at low cost
  For example, a refrigerator exhibit repeating patterns in energy consumption, which allows higher flexibility in its operation without any loss in user experience (temperature).

- *Semi-flexible* : Benefit and cost are comparable.
  For example, the flexibility in shifting activation time for the oven is associated with a cost of users' willingness to delay their cooking.

- *Non-flexible* : Low benefit and high cost
  For example, the operation of devices such as lighting or television, is not flexible or comes with high loss in user experience.

Our flex-categorization of all devices from the datasets is shown in Table 3.5. This categorization may be somewhat arguable, but minor changes will not have a significant impact on our flex-detection analyses. Further, we also categorize

---

[1]//www.energinet.dk/EN/El/Engrosmarked/Udtraek-af-markedsdata/Sider/default.aspx

Table 3.5: Device Categoraization

Fully-Flexible	Semi-Flexible	Non-flexible
Dishwasher	Furnace	Bathroom_gfi
Electric_heat	Microwave	Miscellaneous
Refrigerator	Stove	Electronics
Washer_dryer	Oven	Kitchen_outlets
Heat Pump		Lighting

the operation of devices into three different *operation-states* as rendered in Figure 3.5:

- *Inactive State* : The device is in a non-operating mode.

- *Operating state* : The device is functioning or performing some task.

- *Steady State* : The ideal or low power consumption state between two peak consumption of a single operation.

3.2.3   Operation of Different Device Types

This thesis focuses only on three different flexible device types.

Wet-device: A device used for washing and cleaning, i.e., one that operates with water, is considered a *wet-device*, e.g., dishwasher, washer dryer, etc. For wet-devices, the *Switch-on* action from Figure 3.1 represents the event of pressing a ready button by a user, usually after loading the device. In a typical household wet-devices are used 1-2 times a day, and for a short time duration (1-2 hours), thus they are assumed to provide a substantial flexibility to the EC on selecting the timestamp to trigger the *Activate* action between EST and LST, where LST $\geq$ EST. For example, a dishwasher is usually activated after dinner and not used until the next morning or even later. Though wet-devices may have varying energy demand during the heating and washing cycles, they are usually operating at a fixed power level (depending on the selected program). Thus, the maximum and minimum energy bounds, which are parameters of the *Consume* action, are equal and describe the specific energy amount required by the device, i.e., $e_t \neq e_{(min,t)} = e_{(max,t)}$. The *Deactivate* action represents the completion of all tasks such as heating, washing, spinning, etc.

Electric Vehicle (EV):   After the daily driving, a user plugs in an electric vehicle (EV) for charging which generates the *Switch-on* action shown in Figure 3.1. EVs are usually used plugged in for a longer duration than the actual charging time. Thus, as in the case of wet-devices, an EV also provides a larger time margin to the EC for triggering the *Activate* action. The EC

executes the Consume actions that specify the energy amount $\langle e_1, \ldots, e_d \rangle$ assigned for charging the battery. Depending on the driving requirements, the user may desire to entirely or partially charge the battery, which provides the $e_{(min,t)}$ and $e_{(max,t)}$ bounds for e_t, i.e., $e_{(min,t)} \leq e_{(max,t)}$. For example, if the user set the minimum required charge level to 70%, then $e_{(min,t)} = 0.7 * e_{(max,t)}$.

Heat Pump: A heat pump is a device used to transfer thermal energy from a heat source to the heat sink, usually for in-door heating/cooling. If we consider an air-source HP, it has to operate continuously to maintain the internal room temperature. Thus, the EST and LST are equal and set to the beginning of the heating period, e.g., 12 AM. The EC has to trigger the *Activate* and *Consume* actions immediately after the *Switch-on* action from the user. The value of e_t, determined by the EC for each *Consume* action, depends on various factors such as internal and external temperature, the thermal capacity of the house, etc. Further, the minimum and maximum bounds for e_t also depend on the comfort (temperature) range θ_{max} and θ_{min}. A wider comfort ranges yield wider $e_{(min,t)}$ and $e_{(max,t)}$ bounds for e_t. Further, it might also generate a situation where for some of the time units $e_{min} = 0$, i.e., the HP has to consume no energy in maintaining the room temperature at the user defined range. Thus, the amount of energy consumed in each time unit can be varied from 0 to the maximum capacity of the HP.

3.3   Nordic Power Market

In this section we present detailed discussion on the Spot and Regulation markets, major components of the Nordic power market.

3.3.1   Spot Market

The Spot market is a wholesale electricity pool market where the selling and buying bids for energy are settled a day ahead of the actual dispatch. The market price/MWh for each hour in the next day is settled by a competitive auction that usually occurs 12-36 hour ahead of the actual supply and demand occurs. The bidding on the market closes at 12:00 AM and the final spot prices for the next day are broadcasted already at 12:45 PM. The spot price for each hour represents an intersection point between the aggregated curves for demand and supply for each hour, considering various market and transmission constraints. Figure 3.6 illustrates the price settlement using a 2 sided (demand and supply) auction mechanism in the spot market.

3.3.2   Regulation Market

The regulating power market is activated shortly before the time of the actual delivery and purchase of the power, when the market is anticipated to have any
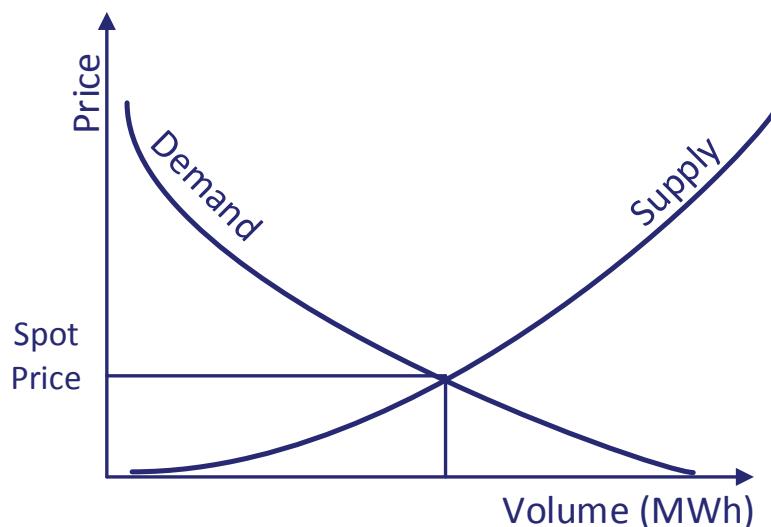
Figure 3.6: Spot Market Price Settlement Demand and Supply Curve

imbalance in supply or demand. The regulating power could be activated for any duration of time. In this thesis, we assume that the duration of activation of regulating power is in unit of an hour. This assumption is not essential for our analysis and could be changed if a more fine-grained control should be desired. Regulating power can be either up or down as a consequence of the following situations. If the supply is less than the demand, the BRP has to buy up-regulating power – at up-regulating power price – in order to maintain the energy balance in the market. The required amount of *up-regulating* power is fulfilled by other energy suppliers or by decreasing the demand by an amount equivalent to the difference. On the other hand, if the supply is greater than the demand, the BRP has to sell *down-regulating* power – at down-regulating power price – to maintain the energy balance in the market. The down-regulating power is sold to the reserve energy market or the demand is increased by an amount equivalent to the difference. The regulating power prices differ from the spot price, thus the BRP suffers financial loss when using the regulating power. The BRP may eventually settle the regulating loss with the energy suppliers that did not fulfill their commitment or the cost is transferred to the customers.

Regardless of the market situation, the regulating power market closes two hours before the actual deliveries and purchases take place. However, the clearance for the regulating power market is done only if needed to balance the market and take place 15 min before the actual deliveries and purchases of energy. The suppliers or buyers of energy in the regulating power market must therefore fulfill their bids within 15 min of being given notice. Here, we define various parameters associated with the regulating power market.

- Spot price, $p_s(t)$: Energy price at the spot market.

- Up-regulation volume, $v_u(t)$: The amount that is less than the actual

demand in the spot market.

- Down-regulation volume, $v_d(t)$: The amount that exceed the actual demand in the spot market.

- Up-regulating power price, $p_u(t)$: Price paid for the up-regulating power.

- Down-regulating power price, $p_d(t)$: Price received for down-regulating power.

At any point in a time, one of the regulation volumes in the pair $(v_u(t), v_d(t))$ will be zero. For notational convenience, we will in the following represent the regulation volumes with a single notation.

- Up/Down-regulation volume, $v_{u/d}(t)$: denotes the non-zero regulating element, or zero if both elements in the pair are zero.

3.4   Summary and Discussion

In this chapter, we detailed the flexibility and flex-offer concepts. Further, we mapped the attributes of flex-offer to the operation sequence of smart devices, namely Wet-devices, Heat Pumps, and Electric Vehicles. Further, we presented the device-level data collection method and datasets used in the experiments of the thesis. The household devices were categorized based on the possibility of extracting flexibility from their operations and the associated loss of user comfort, i.e., user acceptance of loss in perceived quality due to delay in device operation for a given financial incentive. Finally, we discussed the Nordic energy market and the applicability of demand flexibility to solve the demand and supply problems in the market. Specifically, we analyzed the spot and regulation markets. In the next chapter, we focus on evaluation of utility of the flexibility concept to validate the financial viability of the proposed flexibility-based DR.

Chapter 4

Utility of Flexibility Concept

> *What is the economic value of flexibility-based demand response to the market?*

A higher integration of RES into the grid system often creates challenges in demand management, which often results in market imbalances and financial loss. For example, in Denmark, the total volume of the regulating power market for the DK1 (West Denmark) price area was 395.4 GWh in 2013, with a maximum power imbalance of 750 MW [84]. Since the trading of this high amount of energy in the regulation market reduces the revenue, a BRP has always been interested in a dynamic market scheme that reduces their deviation and minimizes their losses. Hence, in this chapter, we perform an assessment of the utility of flexibility-based demand response scheme in perspective to financial benefits to market players. Here, the utility is measured in terms of the reduction in regulation cost and regulation volume that a BRP (market) can achieve utilizing demand flexibility. A positive result from experiments would motivate additional market operators such as Transmission System Operator, Distribution System Operator, etc. to implement, contribute, and adopt the flex-offer concept. With the overall goal of assessing the utility of the device-level flexibility, this chapter focuses on the following topics. First, it proposes and evaluates a number of structural models to select a model which best capture the relationship between the market power prices and regulation volumes. Then, the selected model is used to quantify the effect on the regulating power prices that are caused by the fluctuation in the market balance as a consequence of shifting flexible demand. Henceforth, it evaluates the financial benefit obtained with various types of energy flexibility and market objectives. The overall financial benefit is further analyzed by decomposing it into the direct benefit and the benefit due to the side effects of shifting a flexible demand. Finally, the results of the various experiments are compared to determine the

size and type of energy flexibility that maximizes the benefits of integrating flexibility in the market.

4.1 Relation between Energy Market and Flexibility

The utilization of flexibility for demand management involves shifting of some portion of demand from an originally planned timestamp to a new timestamp within a given *time flexibility*. This concept of demand management comes with side effects that change the market dynamics. A Lack of detailed information and dataset constraints the modelling of an experiment that could address all possible side effects. For example, the available dataset do not tell us detail information regarding the type of the device consuming the specific energy, this limits a model that can address the change in overall demand due to shifting of flexible demands. However, the thesis will now go on to carefully model some of these side effects on the market in order to simulate a real world scenario of utilizing the energy flexibility and its side effects on the market. We will in the following use the convention that t refers to the original timestamp of an available flexible demand and use t' to denote the timestamp whereto the flexible demand is shifted. A t' can be greater, or smaller than t depending on the type of time-flexibility, i.e., forward, backward, or bi-directional.

4.1.1 Modeling The Effect of Flexibility on Energy Markets

We analyse the displacement in the market states (at both timestamps t and t') and the corresponding changes in the regulation prices as a consequence of shifting flexible demand. In addition to evaluating the models in the literature, we propose various new structural models to capture the dynamic pricing mechanism of the current electricity market, i.e., the relationship between the market power prices and regulation volumes. We evaluate the models and select the model with best performance as the final model to be used in the experiments.

Displacement of market balance

The shifting of flexible energy from one timestamp to another will displace the anticipated market balance for both timestamps. This displacement will change the regulation volumes in the market and might also reverse the market balance state (e.g., from *demand > supply* to *supply > demand*). At any timestamp the market will be in one of three different states $S = \{$*up-regulated, down-regulated, balanced*$\}$, and, as such, the configuration (t, t') of the two timestamps may be in any of the market state configurations from the Cartesian product $S \times S$, as detailed in Table 4.1. A shifting of flexibility may affect the anticipated market states at both times t and t'. Let us capture these

Table 4.1: Possible market state at timestamps t and t' (represented as pairs).

Original timestamp

	Up-regulated	Balance	Down-regulated
Up-regulated	*(up, up)*	*(up, balance)*	*(up, down)*
Balance	*(balance, up)*	*(balance, balance)*	*(balance, down)*
Down-regulated	*(down, up)*	*(down, balance)*	*(down, down)*

New timestamp (row label)
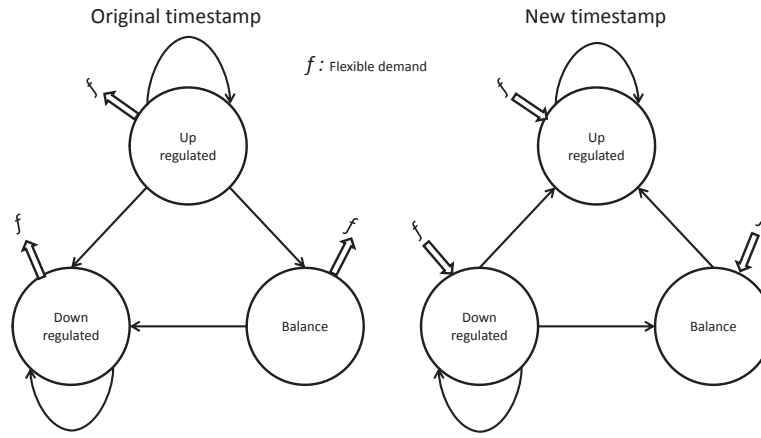
Figure 4.1: Possible market states transition for t and t' (arrows represent addition and removal of flexible demand).

state transitions by defining a shift of a flexible demand f as

$$a(t, t') \xrightarrow{f} b(t, t'),$$

where $a(t, t') \in S \times S$ and $b(t, t') \in S \times S$ are the market state configurations before and after the shift, respectively.

With $|S \times S| = 9$ market state configurations both before and after the shift, we would naively have to consider 81 different situations, when analyzing the effect of a shift of energy from t to t'. However, there are logical constraints that reduce this number of possible situations considerably. Namely, at time t, only the up-regulated state may shift to any of the three states in S, while for the remaining two states the shift of energy will result in a down-regulated state. For example, shifting of flexible demand f from an up-regulated market at t may change the market to down-regulated or balanced condition or can continue to be in up-regulated condition. However, shifting of flexible demand f from a balanced or down-regulated market will always change the market to the down-regulated condition. Similarly, at time t', only the down-regulated state may result in any of the three states in S, while the remaining two states
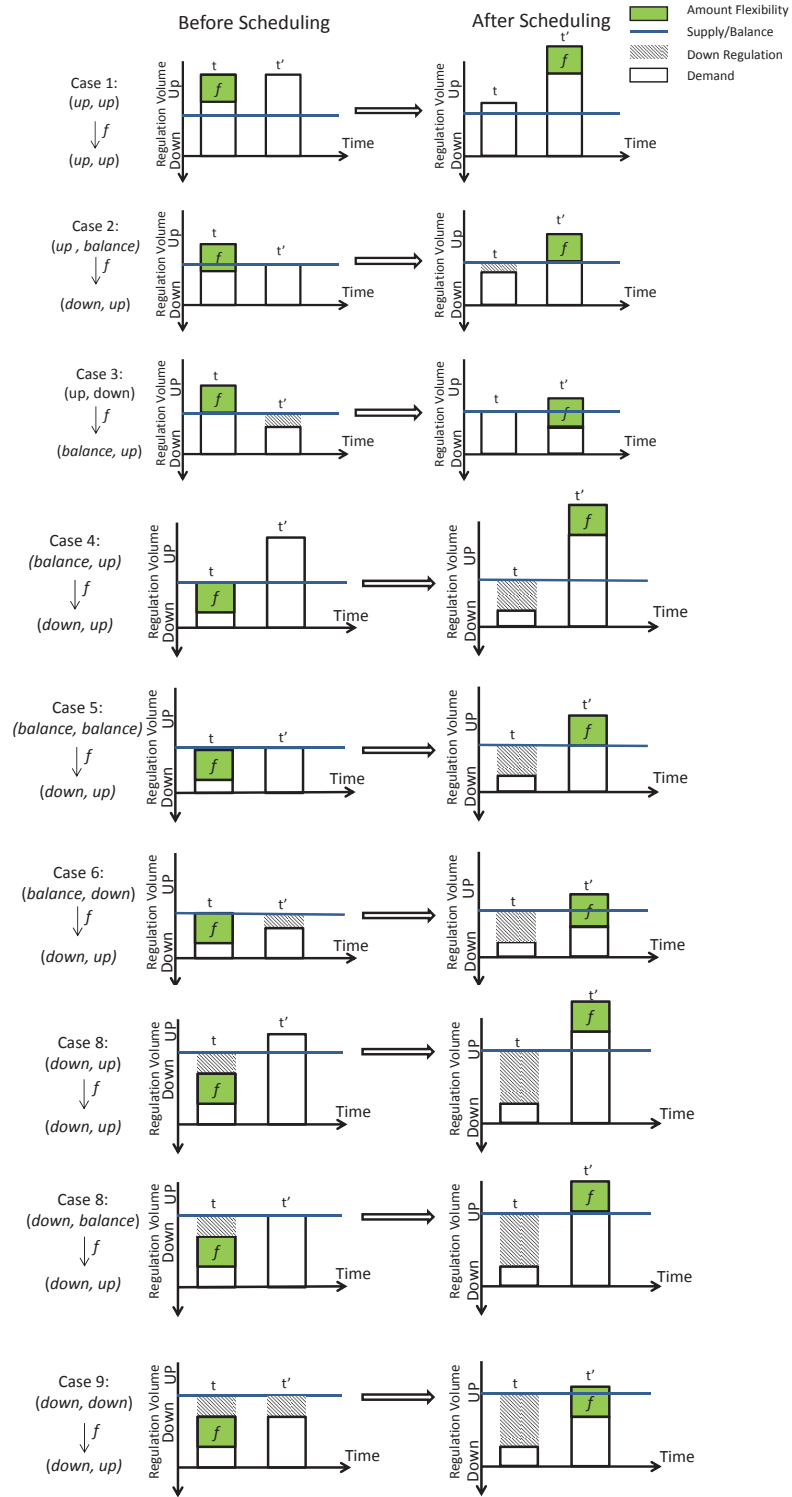
Figure 4.2: Example of changes in market states due to utilization of energy flexibility.

will be in the up-regulated state after the shift. For example, the initial pair of market states (up, down) can shift to one of many possible pairs such as (down, up), (balance, up), (down, down), etc., whereas, the (down, up) pair
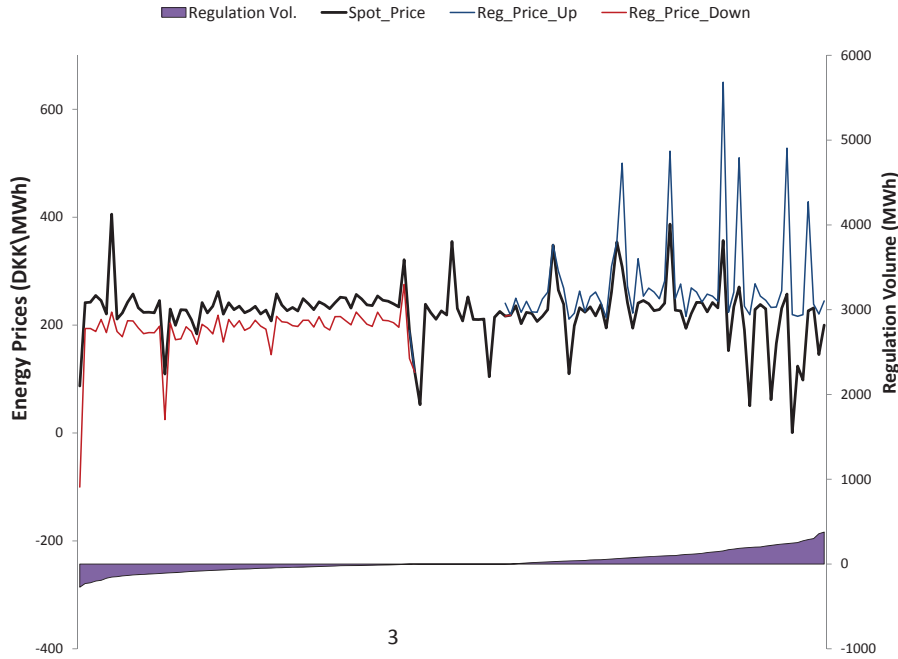
Figure 4.3: Dependency of energy prices on the regulation volumes

will always remain (down, up) irrespective of the size of the shifted flexible demand. These constraints reduce the number of possible situations to only 25 out of the 81, as illustrated in Figure 4.1.

Examples for 9 out of 25 possible situations, one example for each possible pair of initial market states, are shown in Figure 4.2. In each of the examples, the left–hand side of the figure shows the anticipated market states $a(t, t')$ before utilizing the flexibility, and the right–hand side represent the new market states $b(t, t')$ after utilizing flexibility, i.e., shifting the flexible demand from t to t'.

Changes in regulating power prices

Experiments on the relationships between the market power prices and regulation volumes, and the cost associated with a market imbalance has been discussed in [40]. Further, the effect of the level of the spot price and the volumes of regulation bid on regulating power prices has been analysed in [40] [85], respectively. The regulating power prices are generally affected by the market balance, i.e., supply and demand. A displacement in the market balance, due to the utilization of flexible energy, will consequently affect the regulating power prices in the market. Our economic analysis of flexibility incorporates a model for this relationship between regulation volumes and regulating power prices that we have inferred from historical data (described in further detail below). All experiments are performed using the obtained model to estimate updated prices at the timestamps t and t' affected by a shift in demand.

Figure 4.3, illustrates the dependency of regulating power prices on the reg-

ulation volumes in the market. The regulating power prices are clearly seen
to follow the spot price trend with a margin. The size of the margin depends
on the regulation volume in the market, with a few outliers where the margin
does not reflect the regulation volume. These outliers may occur due to the
dependency of regulating power prices on additional exogenous factors such as
hydrology - reservoir levels and inflow, temperature, wind speed, nuclear avail-
ability, etc. [84]. The figure further reveals some difference in the relationship
between spot and regulating power prices according to the type of regulation
(up or down) and, arguably, there seem to be some dependence on the regula-
tion volume as well. In [40], it was found that the most predictive features of
the regulating power price were, in fact, the spot price, the regulation volumes,
and the regulation type (up or down). Based on data from the Nordic energy
market in 1999, the paper inferred a model that defines a linear relationship on
spot price and regulation volume with a conditioning effect given by the type
of regulation. The parameterized model in [40] (in the following referred to as
Model 1) is as follows

$$
\begin{aligned}
p_{u/d}(t) = {} & 1 \cdot p_s(t) \\
& + 1_{v_d(t)<0}(-0.069 \cdot p_s(t) + 0.023 \cdot v_d(t) - 4.3) \\
& + 1_{v_u(t)>0}(0.028 \cdot p_s(t) + 0.042 \cdot v_u(t) + 13.07)
\end{aligned}
\tag{4.1}
$$

Here, $1_{a<b}$ denotes the indicator function for the predicate $a < b$, and $p_{u/d}(t)$
is the predicted up-regulating power price $p_u(t)$ in case of up-regulation and the
predicted down-regulating power price $p_d(t)$ in case of down-regulation. That
is

$$
p_{u/d}(t) = \begin{cases}
p_u(t) & \text{for } v_u(t) > 0 \\
p_d(t) & \text{for } v_d(t) > 0 \\
p_s(t) & \text{otherwise (i.e. } v_u(t) = v_d(t) = 0)
\end{cases}
$$

Since this analysis, the market may have changed in certain aspects. To account
for these potential changes, we consider two alternative models. The first model
maintains the same structural relations as in [40], but with a parameterization
that is re-estimated with our more current 2014 data. This model (in the
following referred to as Model 2) is as follows

$$
\begin{aligned}
p_{u/d}(t) = {} & 1 \cdot p_s(t) \\
& + 1_{v_d(t)<0}(-0.5101 \cdot p_s(t) - 0.0324 \cdot v_d(t) + 55.8372) \\
& + 1_{v_u(t)>0}(0.0657 \cdot p_s(t) + 2.6157 \cdot v_u(t)) - 12.281)
\end{aligned}
\tag{4.2}
$$

In order to account for the dramatic changes in the energy market since
1999, we created 14 different structural models which exhaustively defines all
possible linear combinations of main effects and multiplicative interactions be-
tween the three predictive features: spot price, regulation volumes, and reg-
ulation type (up or down). Out of the 14 constructed structural models, we

selected the model that best quantify the effect on regulating power prices that
are caused by the fluctuation in the market balance. The search for the best
structural model was performed as follows. We partitioned the data from 2012
- 2014 into three parts, with the data from Jan 2012 - June 2013 in a training
set, data from July 2013 - Dec 2013 in a validation set, and the latest data from
Jan 2014 - Feb 2014 set aside as a test set used for later evaluation/experiment.
The alternative model structures – each with individually optimized parame-
terization obtained by using the Matlab curve fitting toolbox on the training
dataset – are compared based on the Mean square Error (MSE) achieved on
the validation set. The model with the least (MSE) is selected as the final
structural model. Finally, with the model structure in place, the combined
training and validation data was used for estimating the final parameterization
of the model. The final learned model (in the following referred to as Model
3) resulting from the structural model selection and the associated parameter
estimation is as follows:

$$
\begin{aligned}
p_{u/d}(t) = {} & 1 \cdot p_s(t) \\
& + 1_{v_d(t)<0}(-0.3362 \cdot p_s(t) + 0.0005 \cdot (p_s(t) \cdot v_d(t))) \\
& + 1_{v_u(t)>0}(0.2378 \cdot p_s(t) + 0.0034 \cdot (p_s(t) \cdot v_u(t)))
\end{aligned}
\tag{4.3}
$$

As for the previous two models, we see that the prediction of the regulat-
ing power prices exhibits a direct relation on the spot price in the first term,
with the following terms accounting for the price adjustment. In contrast, the
price adjustment terms differ structurally from the previous models. Not sur-
prisingly, all three models show that in the condition, where there is neither
up- nor down-regulation the spot price equals the regulating power prices, i.e.,
$p_s = p_u = p_d$. Furthermore, the negative coefficients for the down-regulated
market and positive coefficients for the up-regulated market constraint the mar-
ket price to $p_u > p_s > p_d$, which is similar to the price trend obtained by [40].

4.1.2 Modeling Financial Aspect of Flexibility

At any point in time, the loss due to regulation is computed as the regulated vol-
ume times the price difference between the regulating and spot prices. Hence,
under the normal energy market condition, i.e., without utilizing a flexibility
shift from t to t', the combined regulation cost at the two time points is

$$
\begin{aligned}
R(t, t') = {} & v_{u/d}(t) * |p_{u/d}(t) - p_s(t)| \\
& + v_{u/d}(t') * |p_{u/d}(t') - p_s(t')|
\end{aligned}
\tag{4.4}
$$

Now, consider a flexible load $f(t)$ and let us define the resulting regulation

volumes after this load has been shifted from time point t to t' as follows

$$\underline{v_{u/d}}(t) = v_{u/d}(t) - f(t)$$

$$= \begin{cases} v_u(t) - f(t) & \text{for } v_u(t) \geqslant f(t) \\ f(t) - v_u(t) & \text{for } 0 < v_u(t) < f(t) \\ v_d(t) + f(t) & \text{for } v_u(t) = 0 \text{ (and } v_d(t) \geqslant 0) \end{cases}.$$

$$\overline{v_{u/d}}(t') = v_{u/d}(t') + f(t)$$

$$= \begin{cases} v_d(t') - f(t) & \text{for } v_d(t') \geqslant f(t) \\ f(t) - v_d(t') & \text{for } 0 < v_d(t') < f(t) \\ v_u(t') + f(t) & \text{for } v_d(t') = 0 \text{ (and } v_u(t') \geqslant 0) \end{cases}.$$

Notice the notation, where underbar and overbar denotes a shift to, respectively, lower and higher volumes of the market demand due to a shift of the flexible load. In section 4.1.1 we saw that a change in regulation volume affects the regulating power price. After the flexible load is shifted, the expected combined regulation cost $E(t, t')$ is therefore computed in a similar way as in Equation 4.4, but with estimated prices taking the changed volumes into the account. That is,

$$\begin{aligned} E(t, t') = {}& \underline{v_{u/d}}(t) * |\underline{p_{u/d}}(t) - p_s(t)| \\ & + \overline{v_{u/d}}(t') * |\overline{p_{u/d}}(t') - p_s(t')|, \end{aligned} \tag{4.5}$$

where $p_{u/d}(t)$ and $p_{u/d}(t')$ are updated regulation price at timestamp t and t', respectively, calculated by using regulating power price prediction Model 3 from Section 4.1.1.

The expected change in regulation cost due to shifting a flexible load is then

$$\Delta R(t, t') = R(t, t') - E(t, t') \tag{4.6}$$

A positive value for $\Delta R(t, t')$ represents a savings, i.e., decrease in the regulation cost, and a negative value represents an increase in a regulation cost. The decision regarding the shifting of flexible demand is therefore made based on the value obtained for $\Delta R(t, t')$, and is shifted only if $\Delta R(t, t')$ is positive. The details regarding the market objective and methods for selecting the best time for shifting the flexible demand are discussed in the next section.

4.2   Market Objectives of Utilizing Flexibility

Let us define a data set $X = \{f_1, f_2, \ldots, f_n\}$ of n flexible demands. To ease notation, we will assume the same fixed time flexibility for all the demands, but this assumption is easily generalized to varying time flexibilities across demands. Let T denote the set of possible time flexibilities and $\tau \in T$ be a

specific given flexibility. We will in the following analyses investigate flexibility ranges of all units (hours) of a day. That is, $\tau \in \{0, 1, \ldots, 24\}$ for forward and backward time-flexibility and $\tau \in \{0, 1, \ldots, 12\}$ for bi-directional time-flexibility (see Section 3.1), where, in particular, $\tau = 0$ corresponds to in-flexible demands.

Recall that a flexible demand $f_i(t)$ at time t can shift its load to any unit (hour) within the flexibility of τ. Hence, the new time t' for the flexible load of this demand can be any of

$$t' \in \begin{cases} t, \ldots, t + \tau \\ t - \tau, \ldots, t \\ t - \tau, \ldots, t + \tau \end{cases}.$$

where, $t' \geq t$ for forward time flexibility, $t' \leq t$ for backward time flexibility, and $t - 12 \leq t' \leq t + 12$ for bi-directional time-flexibility.

Let $C(t_i, t'_i; \tau)$ denote the benefit (or negative cost) of moving the i'th flexible load $f_i(t)$ at most τ time units to the new point in time t'. The objective is to maximize the total benefit of utilizing flexibility for all flexible demands in the data set X. In other words, we are optimizing the objective benefit criterion

$$C(X; \tau) = max_{(t'_1, \ldots, t'_n)} \sum_{i=1}^{n} C(t_i, t'_i; \tau) \tag{4.7}$$

We will be using a greedy procedure that optimizes the above benefit criterion for one flexible demand at a time and in this way lower bound the total benefit [1]. With a more sophisticated procedure the benefits could therefore be even bigger than what we demonstrate in the experimental section. Based on the requirement of a regulation market, the energy flexibility can be used to achieve various objectives such as minimizing regulation cost and volume, or even more elaborate objectives such as minimizing loss from underutilized wind energy. To demonstrate our approach, we will focus on minimizing regulation cost and volume, as follows

Minimizing regulation cost: The benefit criterion in Equation 4.7 will in this case maximize the savings that can be obtained in the regulation costs by time shifting the flexible demands. This benefit can be expressed by using the regulation change from Equation 4.6 on the right-hand side of Equation 4.7. That is,

$$C(t_i, t'_i; \tau) = \Delta R(t_i, t'_i; \tau),$$

---

[1] The lower bound will be tight if the optimal shifts for individual flexible loads are independent, which is, however, rarely the case. Consider, for example, $C(t_i, t'_i; \tau) = \Delta R(t_i, t'_i)$, the change in regulation cost from Equation 4.6. Here, the move of a flexible demand may affect the benefit associated with shifting of other flexible demands and, therefore, the order of greedy optimization may matter.

where the τ in the regulation change expresses the maximal shift considered during the optimization. In more detail, the greedy optimization procedure proceeds as follows. We use Equation 4.6 to calculate the regulation cost of utilizing a flexible demand $f_i(t)$ at each possible load shift under the given time flexibility τ for the demand, and then select the most cost-saving shift as the optimal t'_i for this flexible energy demand. Following, the regulation volumes at the two affected time points are updated, as described in Section 4.1, and new estimated regulating power prices are calculated by using prediction Model 3 from Section 4.1.1. For example, if the flexible demand $f_i(t)$ is shifted forward by a duration of 3 units (hours), then the regulation volume and associated predicted regulating power prices are updated at both time points t and $t' = t+$ 3. The next step in the greedy procedure will now consider one of the remaining flexible energy demands in X, but with the updated regulation volumes and prices.

Minimizing regulation volume: The benefit criterion in Equation 4.7 will in this case optimize the balance between the energy demand and supply. This benefit can be expressed by simply using the difference in regulation volume before and after each shift on the right-hand side of 4.7. That is,

$$C(t_i, t'_i; \tau) = (v_{(u/d)}(t) + v_{(u/d)}(t')) - (\underline{v_{(u/d)}}(t) + \overline{v_{(u/d)}}(t'))$$

The greedy optimization procedure will in this case proceed exactly as in minimization of the regulation cost above, except that there is no need for updating expected regulating power prices.

4.3 Experimental Analaysis

4.3.1 Minimizing Regulation Cost (First Experiment)

The first market objective we consider is to minimize the regulation cost paid by the BRP due to imbalance in the market. We analyze the reduction in the regulation cost that a market can achieve for each duration of time flexibility and a given amount flexibility.

Forward Time Flexibility: The average daily savings that can be achieved by utilizing the forward time flexibility in the energy demand is shown in Figure 4.4a and 4.4b. The figures show that with 24 hours of time flexibility and 100 MWh of amount flexibility available for each hour in a day, the market can achieve the highest average daily saving of 107K DKK. Further, we can see that savings in regulation cost generally grow with increasing time flexibility. However, a few drops in savings between the corresponding time flexibility can also be seen, such as in the case of 200 MWh of amount flexibility the saving for 22 hours is less than for 21 hours. This is mainly due the greedy approach we adopted to optimize the shifting of flexible demand (discussed in Section 4.2). In addition, the varying relationship between the market power
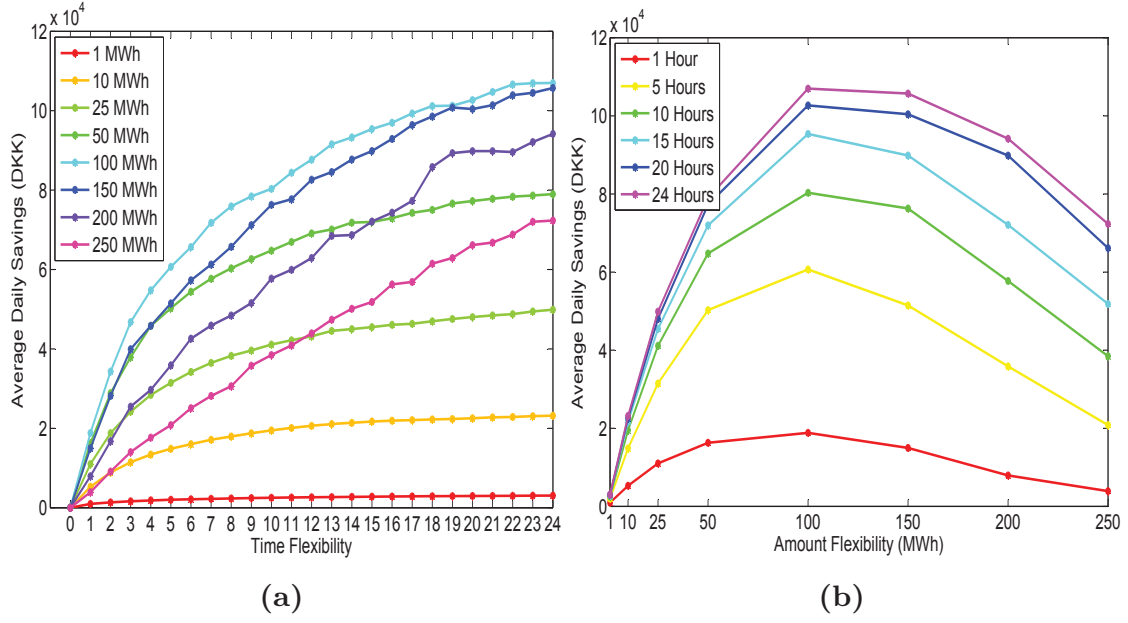
Figure 4.4: Savings in regulation cost - averaged across days (forward time flexibility).

prices and regulation volumes affects the savings based on the timestamp of the shifted flexible demand, e.g., if we have flexible demands at time t and $t+1$ and only one of them can be economically shifted to timestamp t', then the savings might differ based on the flexible demand we choose to shift. Further, we can see diminishing returns for the larger time flexibilities. We can see that the savings gradually grow with increasing amount flexibility up to a certain limit, in this case up to 100 MWh, after which it gradually decreases. This behavior is more clearly demonstrated by the bell shaped structure in Figure 4.4b. The saving decreases because for this particular market (DK1) with its average regulation volume of 63MWh, shifting of larger flexible demands (>100MWh) creates a higher fluctuation in the market and requires higher regulating power to balance the market. As a result, the cost paid for the side-effect is greater than the saving generated from the shifting of flexible demand, i.e., regulation cost increases and the shifting of flexible demand becomes uneconomical as it in some sense does more harm than good. This effect can be further seen in Figure 4.5, where we see a gradual decrease in the average percentage count of flexible demands that are shifted. Recall that in this chapter a flexible shiftable demand represents a macro flex-offer created by aggregating flexibilities from a number of smaller demands. Thus, it can only be shifted as a whole, not in parts.

The daily relative savings/MWh of flexible energy demand is shown in Figure 4.6. The curve is similar to that of average daily savings, showing a gradual growth in the savings with increasing time flexibility. The maximum saving/MWh is achieved for 1 MWh of flexible demand with 24 hours of time flexibility, after which savings/MWh gradually decreases with increasing amount flexibility, which is further demonstrated in Figure 4.7. The savings

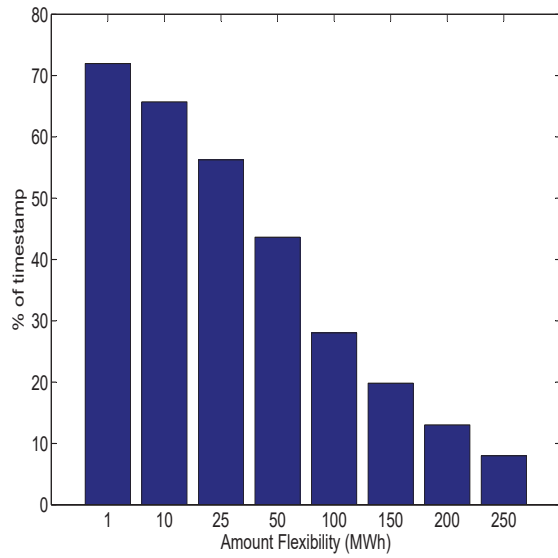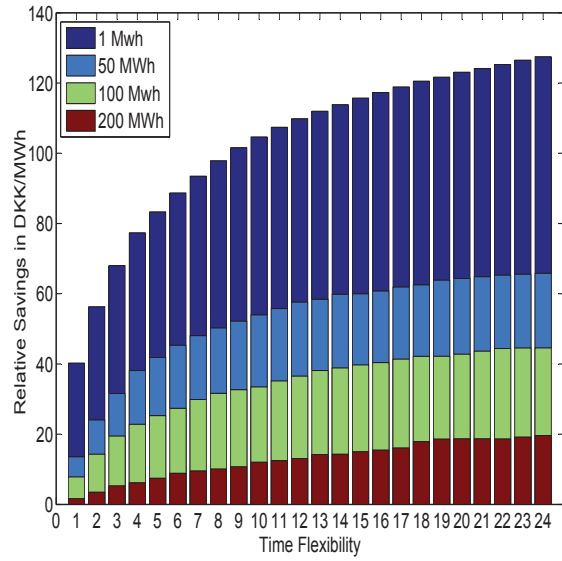**Figure 4.5:** Percentage of shifted flexible demand - averaged over the entire time flexibility.

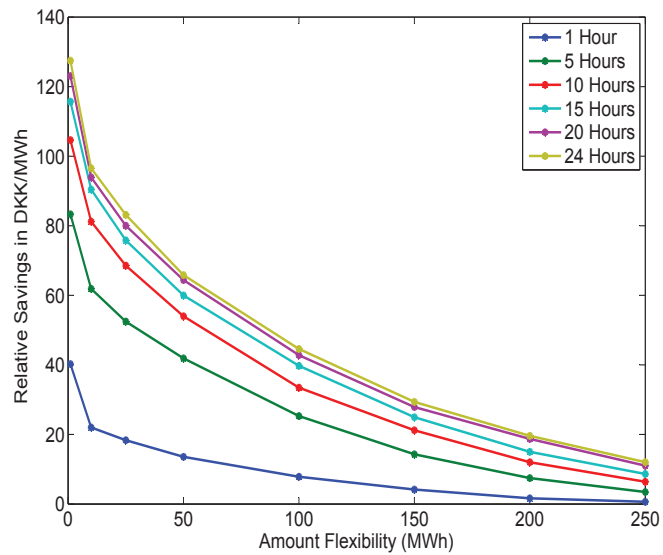**Figure 4.6:** Savings in regulation cost per MWh - averaged over the days.

Figure 4.7: Savings in regulation cost per MWh - averaged over the days.
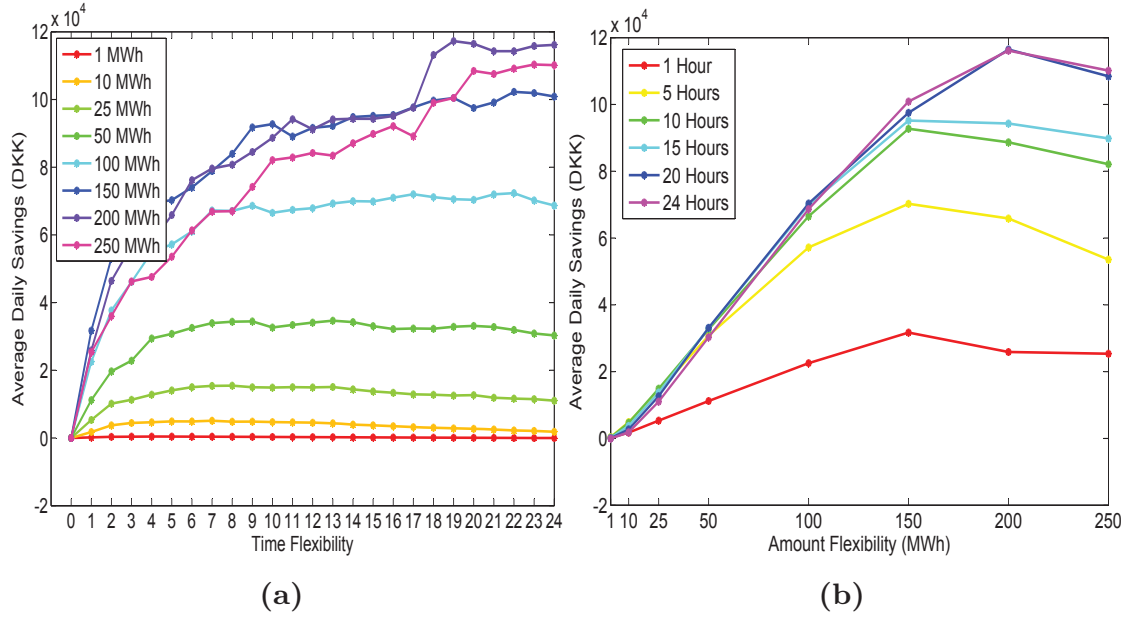
Figure 4.8: Direct savings from the volume of flexible amount - averaged across days (forward time flexibility).

in regulation cost come from two sources, first from the difference in the regulation cost paid for the volume of flexible demand at the original and new timestamps, i.e., direct savings. The second from the side effects, the change in the regulation cost for the remaining part of the regulation volumes at both timestamps. It is interesting to analyze the savings from these two sources separately. Thus, we disaggregate the average savings into direct savings and the savings from side effects, as shown in Figures 4.8a and 4.8b, respectively. The curves in Figure 4.8a show slightly different patterns in savings compared to that of total savings, i.e., for lower amount flexibilities (up to 100 MWh) direct savings are lower, whereas the savings are higher for larger amount flexibilities (> 100 MWh). Similarly, Figure 4.9b shows that the savings from side effects are positive for lower amount flexibilities, but negative for higher amount flexibilities. These differences are due to the effect of difference in market power prices on the overall savings. The lower amount flexibilities has a dual benefit because they could be shifted to an up-regulated timestamp, if the difference in the market power prices is high enough to cover the loss due to changes in regulation volumes.

On the other hand, for the higher amount flexibilities the difference in market power prices cannot cover the loss due to a higher fluctuation in regulation volumes, which results in negative savings for side effects and lowers the total savings as shown in Figure 4.4a. Further, compared to Figure 4.4b, Figure 4.8b shows an increase in threshold values of amount flexibility by 50 MWh and 100 MWh for lower and higher time flexibilities, respectively. The threshold value of the amount flexibility that generates positive savings for the side effects of shifting flexible demand depends on the energy demand and the average
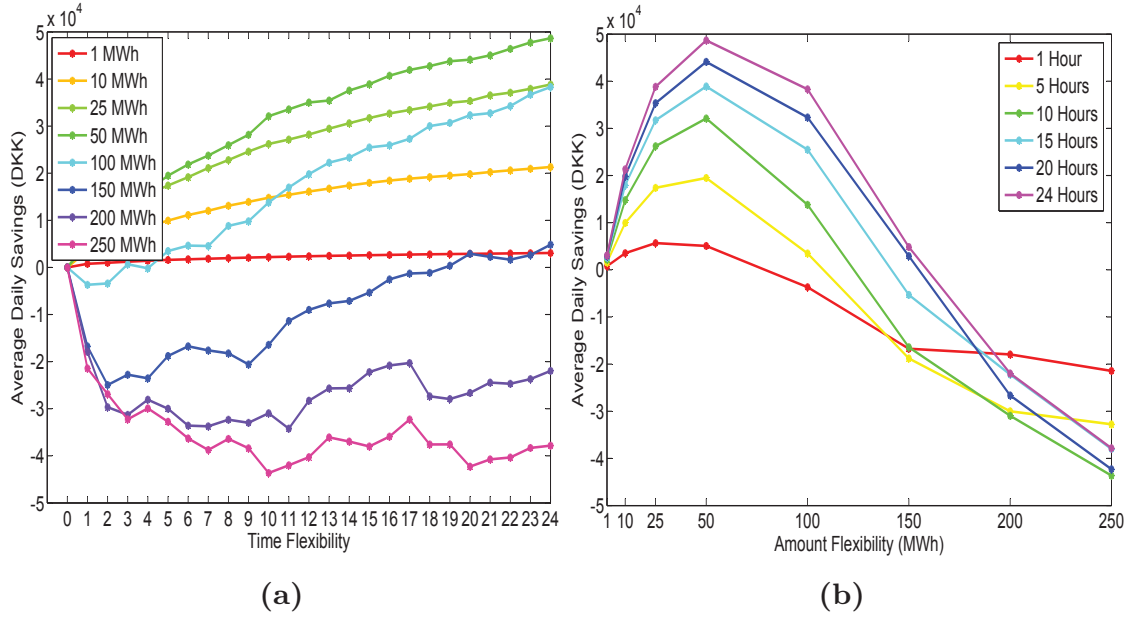
Figure 4.9: Savings from side effects of shifting flexible demand - averaged across days (forward time flexibility).

regulation volumes of the market.

Backward Time Flexibility:

The average daily savings that can be achieved by utilizing the backward time flexibility in the energy demand is shown in Figure 4.10a and 4.10b. The trends are similar to that of forward flexibility, but with slightly lower values, e.g., the best average daily saving is 6.1% less. This is due to the occurrence of a few consecutive up-regulations at the beginning of the time series, which decrease the possible shifting of flexible demand. Further, we see similar trends for direct savings and savings from side effects. Similarly, the relative savings/MWh also gradually decreases with growing amount flexibility. A detailed comparison of savings between various types of time flexibility and market objectives is shown in Table 4.2.

Bi-directional Time Flexibility: The average daily savings and trends obtained for the bi-directional time flexibility are very similar to that of forward flexibility, but with slightly lower values, e.g., the best average daily saving is 11% less. A comparison of savings between various types of time flexibility and market objective is shown in Table 4.2.

4.3.2   Minimizing Regulation Volume (Second Experiment)

The second market objective, we consider, is to minimize the volume of energy traded in the regulating market, i.e., regulation volume. In addition, we also analyze the reduction in regulation cost in this case. As seen above, the result for forward, backward, and bi-directional flexibility are similar, so we now describe them jointly.
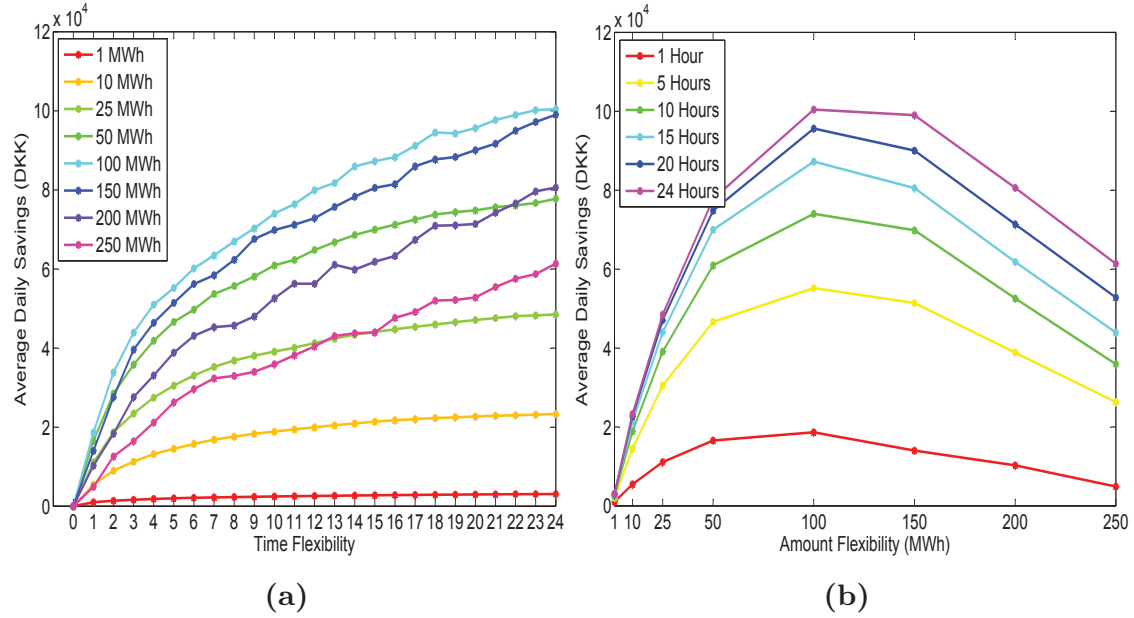
Figure 4.10: Savings in regulation cost - averaged across days (backward time flexibility).

Amount Flexibility (MWh)	Regulation Cost						Regulation Volume					
	Forward		Backward		Bi-directional		Forward		Backward		Bi-directional	
	Direct	Side Effect	Direct	Side Effect	Direct	Side Effect	Direct	Side Effect	Direct	Side Effect	Direct	Side Effect
1	-10	3070	-6	3071	-181	2939	741	1600	642	1563	571	1450
10	1844	21327	2571	20739	-326	23003	6692	10997	5291	10597	4574	10172
50	30313	48671	37585	40194	22377	52457	23858	27511	13406	33578	14624	30051
100	68685	38303	78114	22351	54459	40803	41970	40026	26913	37416	28214	30187
200	116136	-21969	111981	-31369	84037	-12009	66018	1540	55600	-3370	48125	-7669

Table 4.2: Relative average daily savings (in DKK) for 24 hours time flexibility: compared over various market objective.
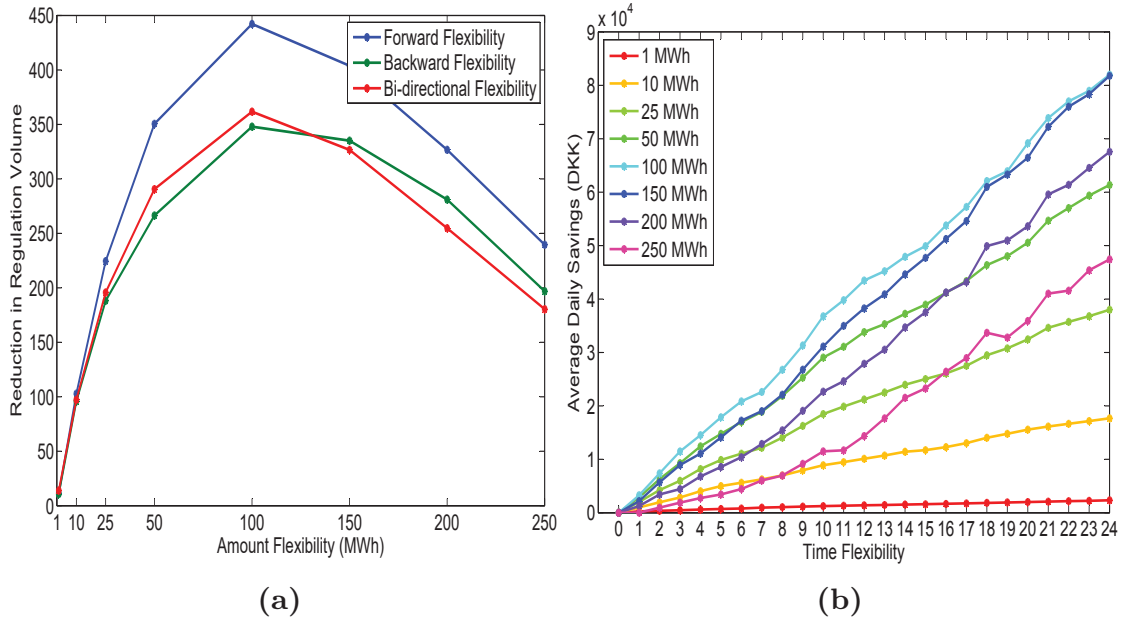
Figure 4.11: Savings in regulation volume (a) and regulation cost (b) - averaged across days (forward time flexibility).

Figure 4.11a shows the average daily reduction in the regulation volume utilizing the maximum time flexibility (24 hours). The figure shows that the regulation volume can be reduced by up to 442 MWh on average, which accounts for 29.4% of the average daily imbalance in the market. Further, the curves show that benefits for the market grows up to a certain value of amount flexibility, and then gradually decreases. In addition to the regulation volume, the regulation cost also reduces by up to 82K DKK on average, which is $\approx 24\%$ less than for the first experiment. The most interesting of all, is the disaggregation of the total savings into direct savings and savings from the side effects, as shown in Figure 4.12a and 4.12b, respectively. For the small amount flexibilities, the trends are different from the first experiment: the direct savings are negative and become positive with the increasing time flexibility. On the other hand, for higher amount flexibilities the savings are positive from first flexible hour and increases gradually. Similarly, the savings from the side effect are positive for small amount flexibilities and negative for higher amount flexibilities. This contrasting trend is explained by the market objective, where the shifting of flexible demand is only possible from an up-regulated market to a down-regulated market. In addition, it is also due to the varying effects of flexibility on the market balance and market power prices, based on the size of shifted flexible demand, as discussed in Section 4.3.1. The complete comparison of savings can be seen in Table 4.2.
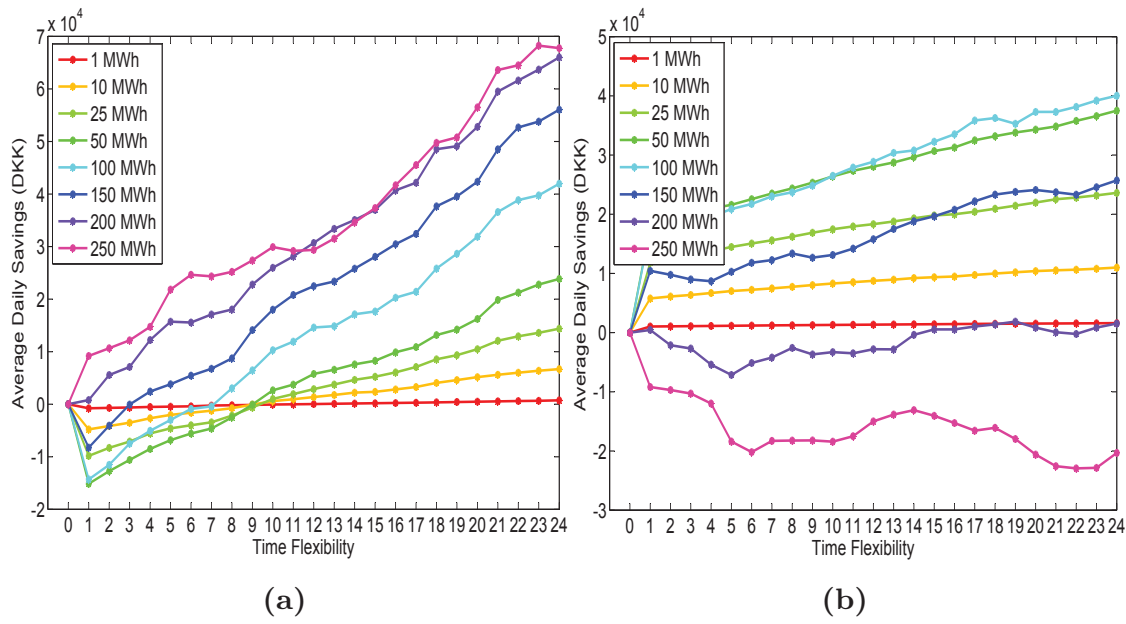
Figure 4.12: Direct Savings (a) and savings from side effect (b) - averaged across days (forward time flexibility).

4.3.3 Analysis

The results above show that in general a market can increase its savings in regulation cost with increasing time flexibility, but the specific trend depends on the market objective. For the first experiment, the market has diminishing returns for increasing time flexibility. For example, for 100 MWh of amount flexibility, 8 hours of forward time flexibility give 71% of the benefit of 24 hours. In contrast, for the second experiment the savings grow steadily with increasing time flexibility, where 8 hours of forward time flexibility give approximately a third of the benefit of 24 hours. On the other hand, the size of the amount flexibility plays an important role in determining the benefits of flexibility in the market. The financial benefits of the market grow with the increasing amount flexibility up to a certain limit, after which it decreases and can be negative, e.g., the highest benefit with 100 MWh of amount flexibility is almost 48% higher than that for 250 MWh. Further, the relative savings/MWh gradually decreases with increasing amount flexibility, e.g., the best possible relative saving/MWh (127 DKK/MWh) for 1 MWh of amount flexibility is twice that for 50 MWh and six times that for 200 MWh.

In addition, until the threshold value of amount flexibility (100 MWh) is reached, the market has a dual benefit of utilizing the flexibility, i.e., the market benefits from both direct savings and savings from side effects, whereas, above the threshold value the market loses a huge amount in side effects due to higher imbalance in the market. This loss reduces the overall savings and diminish the benefit of utilizing flexibility in the energy market, e.g., in the first experiment, savings from side effects for 50MWh is 62% of the best possible saving, whereas

for 200MWh it reduces to -52%. The higher the amount flexibility, the lower the possibility of gaining financial benefit from shifting it. This argument is supported by the fact that, on average flexible demand of size 10 MWh is shifted 65.2% of time, which reduces to 8% for 250 MWh. These results indicate the maximum size of energy flexibility that can be traded in an energy market with profitability. In addition, it also provides the guidelines for aggregating micro flex-offers to macro flex-offers.

A market with an objective to reduce regulation volume can achieve a best possible average daily reduction of 442 MWh along with a 37.5% reduction in regulation cost. These results show that, the time shifting of flexible demand can generate a substantial benefit regardless of the types of energy flexibility or market objectives. However, the geographical location, size of the market, and the type of RES will determine the optimal size of time and amount flexibility that maximizes the benefits, e.g., demand management for solar energy need flexible load to be shifted to day time, which requires higher time flexibility to maximize benefit. For the market in this study, forward time flexibility and objective to reduce regulation cost generates the best possible benefits for the market. Here, the market can achieve up to 49% (107K DKK) reduction in the average regulation cost, with 24 hours of forward time flexibility and just 3.87% of average gross demand (2.58 GW) being flexible. Further, with just 1 hour of time flexibility, the market can achieve 17.6% of the savings for 24 hours.

4.4   Summary and Discussion

In this chapter, we presented the answer to the research question *"What is the economic value of flexibility-based demand response to the market?"*. In this regard, we evaluated the utility of demand flexibility to the market. Specifically, we quantified the reduction in the regulation cost and volume that a market can achieve utilizing demand flexibility. The experiments on the Nordic energy market showed that regardless of the type of market objective and flexibility, a market can generate substantial economic benefit. Further, a market can trade-off between the available time and amount flexibility to maximize its benefit by better mapping energy demand and supply. Indeed, with just 4 hours of time flexibility and 100 MWh of amount flexibility, the market could reduce the regulation cost by 24.9. Finally, we can conclude that *"flexibility-based demand response is economically viable to a market"*. However, the extent of financial benefits is specific to a market and depends on the market demand and supply conditions and types and volume of RES integration. In the next chapter, we focus on the viability of extracting flexibilities from users' daily routines.

Chapter 5

User Behavior Analysis

Which flexibilities do users have in their daily routine and are those enough to support the concept?

In the previous chapter, we quantified the utility of the demand flexibility to the energy market. However, flexibilities come from users and should be enough to be effectively utilized in the market. The vision of the flexibility-based DR scheme is that for users having a flex-offer contract with an energy supplier, their flexibility is not specified by the user, but instead predicted based on past users' behavior. Hence, a comprehensive analysis of users' device usage behavior is fundamental for evaluating the existence of usable flexibilities in their daily routine. This chapter focuses on the device-level analysis of energy consumption data, which will form the foundation for accurate flexibility detection (Chapter 6), flexibility prediction (Chapter 6), and automated generation of flex-offers (Chapter 7). In this chapter, we first introduce various device operation properties that provide insights on the device usage preference of a user. Henceforth, we present a number of specific pre-processing steps to clean device-level data. Then, it performs state-of-the-art analysis of device-level energy consumption, including patterns and periodicity in device operation and the correlation between operations of different devices. Subsequently, we show the existence of detectable time and energy flexibility in user's daily routine and device operations.

5.1 Device Operation Properties

We will now introduce the important device operation properties that we investigate in order to support the challenges of deriving flexibility information about energy demand. Statistical analyses for these properties will provide significant information for generalized as well as user-specific device operation

patterns. Further, they provide information regarding existence of flexibility in the device operations and correlation between various devices.

1. There exists detectable Intraday and Interday patterns in device operation.

   (a) Weekend and Weekdays patterns are different.

   (b) Houses exhibit general and specific intra-day and inter-day patterns.

2. There exist time and energy flexibility in device operation.

   (a) A major percentage of energy consumption comes from flexible devices.

   (b) An alteration in device energy profile is feasible.

   (c) Device activation time can be shifted by some duration.

3. Some devices are correlated

   (a) Highly correlated device are operated simultaneously or just after one another.

   (b) There is some fixed sequence of device operation.

5.2   Device-level Data Preprocessing

In this section, we will discuss the preprocessing steps and challenges associated with the device level analysis. The complete sequence of the steps taken during the preprocessing of raw input data before allowing the statistical analysis is visualized in Figure 5.1, and details for each step are described in the following subsections:

5.2.1   Spike Removal

We define noise as both being the effect of unwanted artifacts and white noise influencing the time series data in a way that generates abnormal patterns in the recorded device operation. Noise may be introduced through different sources, e.g., error during data collection, abnormal behavior of the device, or mistakes by the users (mistakenly switching device on and off). We have considered a very high consumption value for very short duration, up to 2 data points (corresponding to 6 sec), as a noise spike. The first preprocessing step replaces the spikes in energy consumption with its preceding neighboring value, which eliminates the obvious artifacts created by the installed measurement devices or user mistakes in device operation. The data, in general, contains few noise spikes of this type. To give an impression only 17 consumption spikes were removed for the oven device-data in *house 1*, which accounts for 0.004% of total high consumption values for this device. However, removing these noise
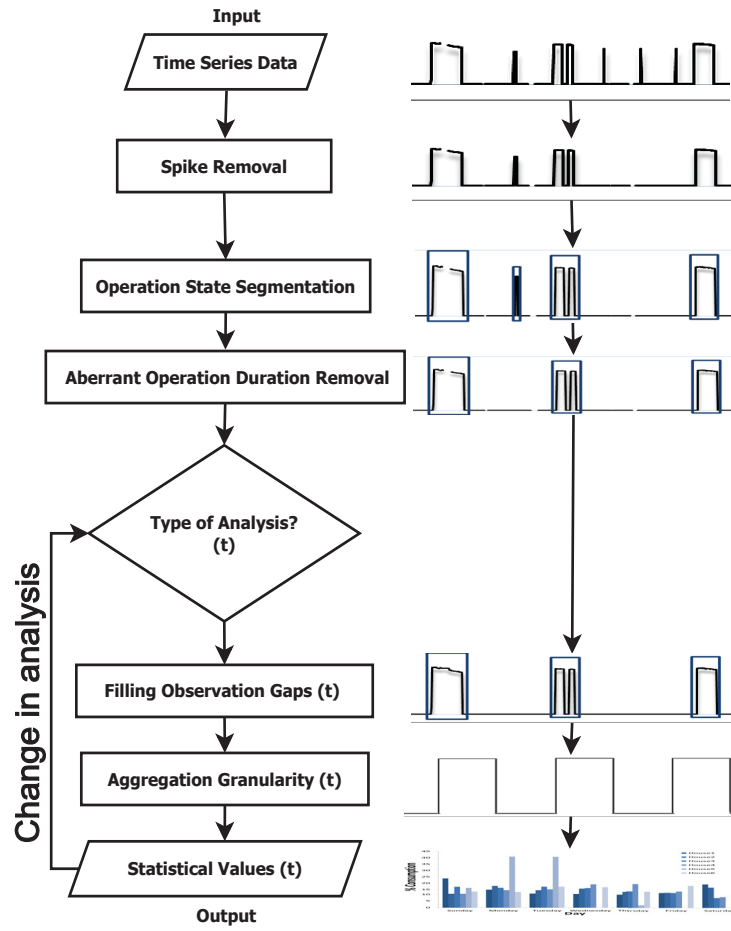
Figure 5.1: Preprocessing steps.

spikes significantly effects the statistical analysis, as each removed spike would otherwise be considered as an individual device operation, escalating the device operation frequency.

5.2.2 Operation State Segmentation

Before proceeding with device-level statistical analysis, we annotate the data according to the three different states of device operation, as discussed in Section 3.2.2. The presence of an unknown, and for some devices, a high number of intermediate power levels between the minimum and maximum power consumption makes it challenging to segment the data into their respective phases of operation. Furthermore, variability in the power consumption value within a given stage of device operation creates ambiguity in defining the correct stage. We rely on a simple segmentation approach, where we manually inspected the consumption time series for a device in order to determine decision thresholds between the different device operation states, e.g., 50 watts for microwave, 20 watts for electric heat, etc. This manual approach is only possible because we are dealing with a reasonably small number of different devices. In future work,

we will investigate the possibility of replacing this manual setting of thresholds with robust and optimal automated threshold detection techniques.

5.2.3   Aberrant Operation Durations Removal

Each device has its own functionality and, furthermore, the user can typically operate the device with varying parameters to accomplish an objective. For example, a cooking objective has a high number of parameters according to the meal being prepared. Despite variation in functionalities for different devices, their durations in operation are usually constrained by the device objectives, e.g., usually a dishwasher is not operated for a few seconds or more than a few hours, and a microwave is not operated for more than a few minutes. These abnormal behaviors in device operation were filtered out by replacing the consumption values in these periods with the inactive state values. In an extreme case, we discovered four instances in *house 1* with durations of oven operation that were less than *3 minutes*. When removed, it reduced the operation frequency of oven in the house by 20%. The rationale behind filtering out aberrant operation is that these are often caused by user's mistake, e.g. user forgetting to turn off the oven when the food is cooked or mistakenly switching on the device, and these kinds of behaviors are very rare and almost impossible to capture in analysis and prediction models.

5.2.4   Filling Observation Gaps

Observation gaps in the time series is a major challenge that we face for analyzing the dataset used in this thesis. We have implemented three simple approaches for handling the missing information in observation gaps. The approaches differ according to the amount of missing data in the gaps. In the first case, if an observation gap stretches over a full day, that day is discarded from the analyses. Second, for observation gaps up to at most 6 seconds (2 consecutive data points), this data is filled-in by simple extrapolation according to the last observed value. Finally, for observation gaps of more than 6 seconds and up to a day, the imputation for the gap is based on a computation of the expected value given time-matching observed values across all days in the analysed data of the given device (and house). For gaps less than an hour, the observed data values are averaged across the hour, and for gaps greater than an hour the observed data values are averaged over a block of time for each of the five distinct periods that are shown in Table 5.1. In this way, we fill in $\approx 750(1.67\%)$ missing hours out of the total of $\approx 45,000$ hours for the entire dataset.

    To illustrate our data imputation methods, let us first assume that we are interested in the operation state for a device and discover a gap of 20 minutes in a particular hour. Each of the three operation states (see Sec. 3.2.2) are then for the unobserved time ticks filled-in with the average observation value for each state, as computed for the considered hour across days with full observations

Table 5.1: Division of 24 hours into different block periods. The last column shows the expected value (pobability) of operation for the stove in *house 1*.

Periods	Hours of the day	Category	Probability of operation
Night	23.00 - 05.00	Off-peak period	0.03
Morning	06:00 - 08:00	On-peak period	0.67
Noon	09:00 - 13:00	Mid-peak period	0.28
Afternoon	14:00 - 17:00	Low-peak period	0.12
Evening	18:00 - 22:00	Mid-peak period	0.32

on that hour. In another example, say, that we are still interested in operation states, but now the observation gap is a three hour period from 7:00 to 09:59. The imputation of the operation states will now proceed as in the previous example with the minor difference that the imputed values are averaged across the hours for a block period (see Table 5.1). Notice that in this way, the imputed values for the first two hours may be different from the imputed values for the last hour, because they belong to different block periods. Finally, to illustrate how imputed values may be involved when computing statistics in an analysis, let us consider the oven, which has imputation values for the 'operating' operation state, as shown in the last column of Table 5.1. Say that we observe that the oven is turned on once in the evening and we have the same three hour observation gap as above. In this case, the expected frequency for the usage of the oven during that day would be 1+0.67+0.28=1.95 operations. Notice that to simplify the illustration, we have here assumed that the oven can only be turned on once during a block period and that operations do not cross blocks.

5.2.5 Aggregation Granularity

Finally, we aggregate the high frequency data in to the granularity of time that we target for analyses, e.g. hourly or daily. An example for aggregation of time series data into hourly resolution is shown in Figure 5.2.

5.3 Data Analysis

In this section, we will discuss statistical behaviors regarding energy profile, usage pattern, and correlations in the device operations.

5.3.1 Device Energy Profiles

In Figure 5.3, we show a disaggregaion of the total energy consumption for the different individual devices that are measured in the REDD dataset. It shows a significant variation in device consumption across the different houses. However, by aggregating according to the type of device flexibility (Fully/Semi/Non,
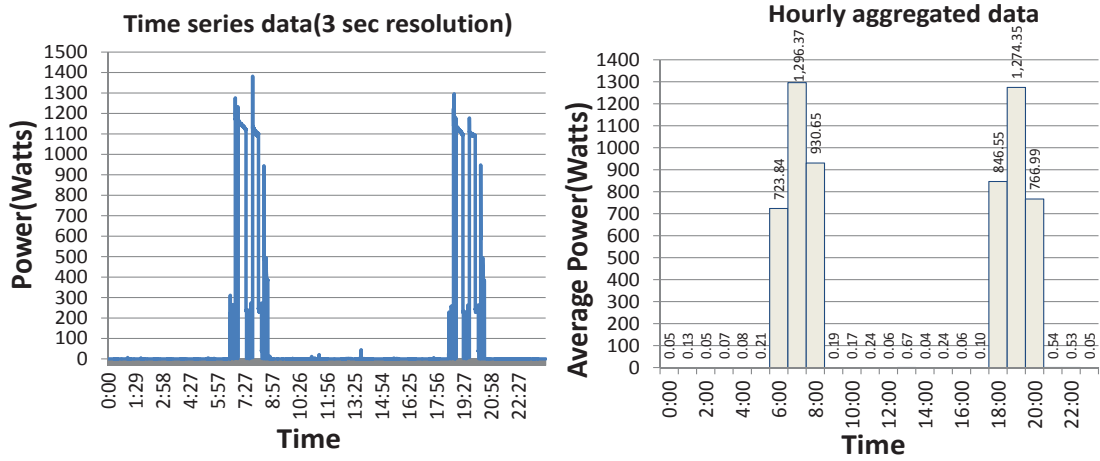
Figure 5.2: Hourly aggregation of time series data for *dishwasher*.

see Table 3.5), as shown in Figure 5.5, we see a more consistent pattern in the shares of the total energy consumed by flexible and non-flexible devices. The figure suggests that on average 50% of the total energy demand from a house can be considered flexible, supporting Property 2(a). The low percentage of consumption from flexible, energy-intensive devices such as heating, contributed to the seasonal behavior in the device operation (late Spring to early Summer).

Figure 5.4 shows the minimum, maximum, and average power during operation for different devices in *house 1*. We see many devices with average consumption far from the extremes, which for all these devices indicate a potential for device designs that support altering of the energy profile during operation and, hence, creating flexibility and supporting Property 2(b). Looking at the potential for flexibility due to users' behavioral changes, additional experiments (not reported in detail here) have shown a high deviation in total energy consumption, due to duration of operation and power level, across operations for an individual device. Also, as we will see in Section 5.3.2, there can be significant variation in consumption at the time an operation is initiated. These findings indicate a flexibility potential of shifting the energy profile or activation time for the device, if users are willing to –or compensated for–behavioral changes supporting Property 2(c).

5.3.2   Use Patterns

We will now look more into the patterns and (ir)regularities that device operations exhibit in individual households. If we first consider the interday distribution of total energy comsumption for the individual houses' in Figure 5.6, we see a general pattern of a somewhat evenly distributed consumption across the week, with a slight tendency towards lower consumption during the weekend. Exceptions are *house 5* where most of the consumption occurred on Monday and Tuesday, and *house 1* with higher consumption during weekends.
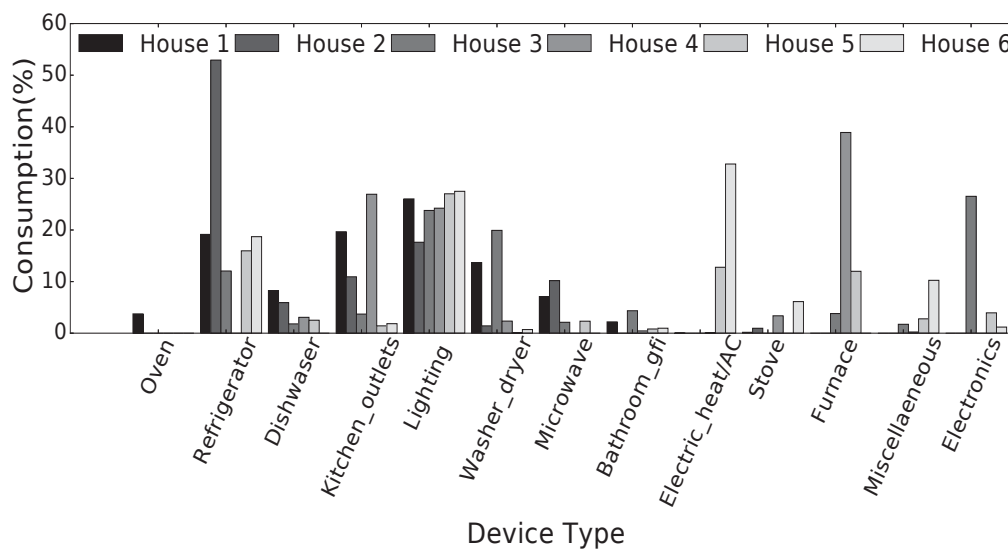
Figure 5.3: Distribution of energy demand over various devices (and house).
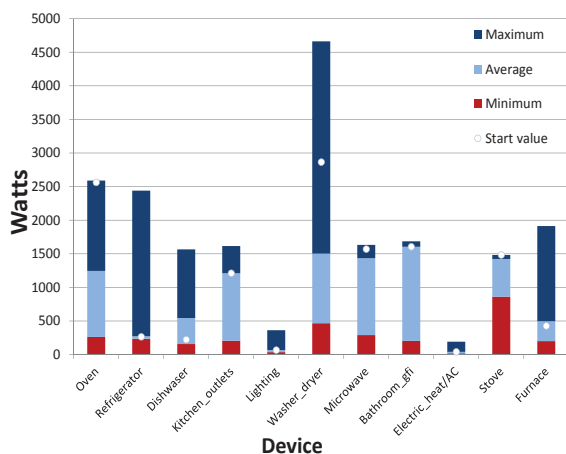
Figure 5.4: Min, Avg, and Max power consumption (watts) and power consumption during the device activation (shown by dot) for selected devices
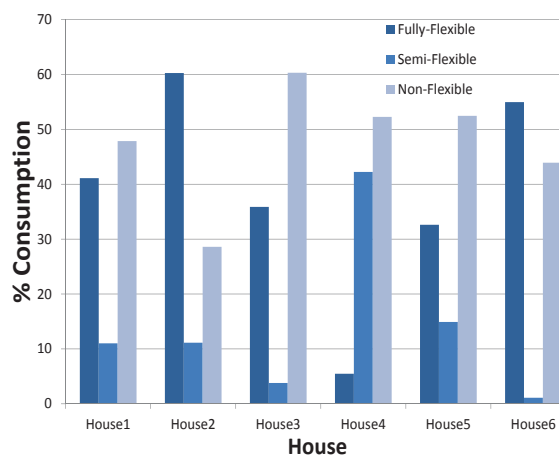
Figure 5.5: Percentage distribution of total energy consumption by flexibility type from individual houses.

The weekday versus weekend difference is further emphasized by the corresponding aggregation, as shown in Figure 5.7. For all but *house 1* we clearly see a higher energy consumption during the weekdays than in the weekends. This difference supports Property 1(a), but the pattern is, however, surprising in that it contrasts the common belief that people use most of their high energy consuming devices, such as washer and dryer, during weekends and holidays. It should be noted that the REDD data has a high rate of missing data during weekends and we are therefore collecting additional data to support this surprising preliminary result.

To analyze intra-day variation for individual devices within and across households, we aggregated device operations by the the hours of the day in
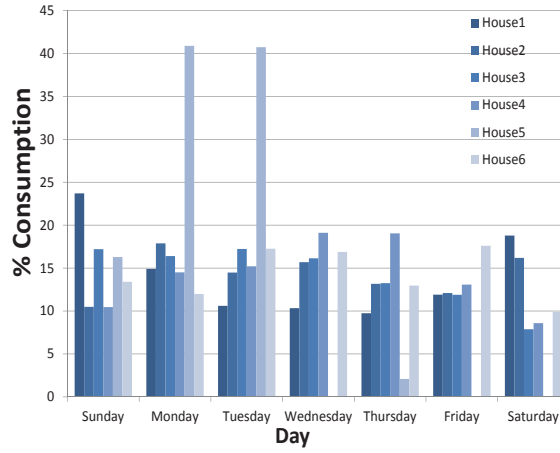
Figure 5.6: Relative daily energy consumption across households.
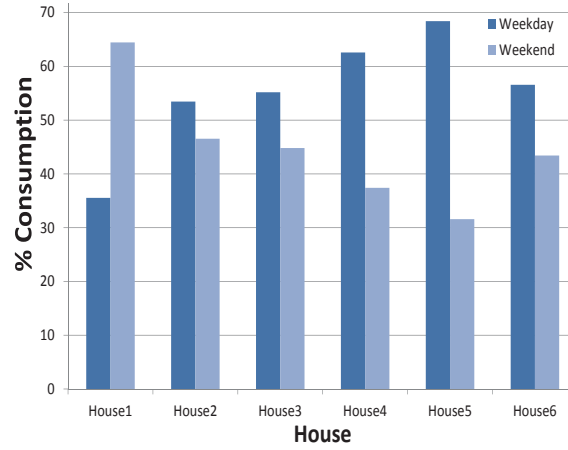
Figure 5.7: Relative weekdays versus weekends energy consumption across households.

order to compute the percentage of days that a device has been in operation during a particular hour. Figure 5.8 shows the results obtained for the *dishwasher* operation in the six houses. We can clearly see a similarity in operation across the houses, with a higher percentage of operation during the day and a much smaller percentage during the late night (00.00 am - 06.00 am). Hourly operation patterns were observed for most devices, but patterns were not necessarily similar across different households. For the *stove* and *microwave*, for example, we observed similar peak operation periods in the morning and evening hours corresponding to the typical hours for preparing meals. In contrast, for the *washer dryer*, we found varying patterns in operation. Some houses have evenly distributed hours of operation, whereas operations in other houses were concentrated at certain hours of the day. For example, operation of the *washer dryer* in *house 3* was highly concentrated in the hours between 5.00pm - 11.00pm, whereas a fairly even distribution of operation was observed in *house 1*. In some cases, these peak periods represent very high percentages of device operation, as e.g., for *dishwasher* usage in *houses 2,3,5* seen in Figure 5.8.

Results (not reported in detail here) also verify obvious variation in the typical duration for operation of different devices, reflecting their objectives and functionalities. For example, we see shorter operation durations for the microwave and high (continuous) duration for the refrigerator as justified by their different operational objectives. Looking at the aggregated daily frequency[1] of device operations, conclusions are household dependent. For some households we see a fairly even distribution across days, whereas for other households the frequency varies dramatically. For example, Figure 5.9 shows the daily frequency of *microwave* operations for different households (*houses 4,6* do not operate this device). We see a somewhat stable usage pattern for *houses 3,5*,

---

[1]More precisely, the *expected* frequency, as described in Section 5.2.4.

whereas *houses 1,2* have large daily variations.
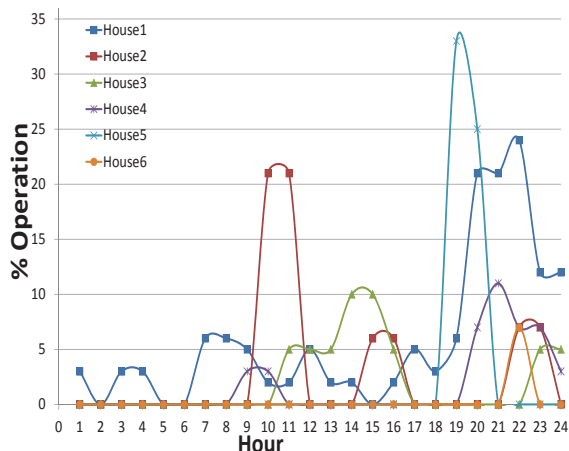
Figure 5.8: Distribution of hourly *dishwasher* operations – averaged across all days in households.
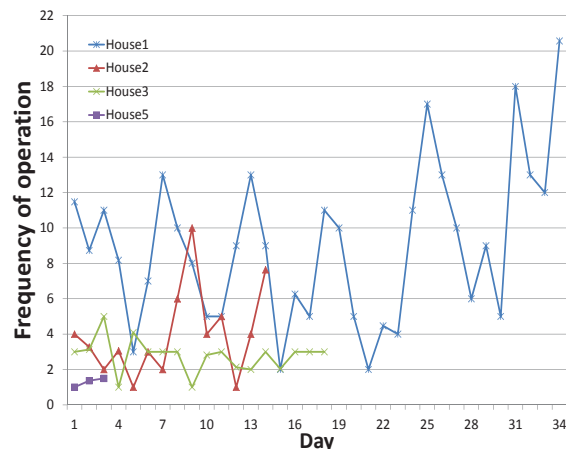
Figure 5.9: Daily operation frequency for *microwave* across households.

Summarizing the intra-day results from this section, we can support Property 1(b), that device dependent intra-day patterns do exist and repeat across days. On the more challenging side, however, these patterns seems to be highly household dependent. In addition, small peak periods for operation, as we see in Figure 5.8, may suggest a potential for shifting activation time, in further support of Property 2(c).

5.3.3 Device Correlations

Finally, we aim to provide information regarding the coherence and correlation between device operation occurring due to the usage patterns. We will further explore the sequence in which devices are operated, and there frequencies. Table 5.2 shows the number of times a given pair of devices were operated together within a one hour interval in *house 1*. The ordering of the pairs represents the sequence of activation of the devices. The high degree of correlation between devices such as the *stove* and *microwave* occur due to them supporting a joint activity, cooking. However, that devices such as the *washer dryer* and *microwave* exhibit a high degree of concurrent operation is more due to the user's behavior, preferences, and presence in the house. This conclusion is further justified by data from another house(not shown here), where we have very infrequent concurrent operations between *washer dryer* and *microwave*. This shows that devices can have a higher probability of concurrent operation without sharing similar functionality/purpose. Thus, we can say that occurrence of high simultaneous operation is specific to a house and does not always guarantee high correlations across households, thus providing somewhat conflicting arguments for Property 3(a). The sequential behavior in the device operations can be seen for the *microwave* and *dishwasher*, where the *microwave-dishwasher*

Table 5.2: Operation sequence for pairs of devices (*house 1*).

Device1	Device2	Frequency	Device1	Device2	Frequency
Oven	washer dryer	2	washer dryer	oven	2
Oven	microwave	8	Microwave	Oven	9
Oven	electric heat	1	electric heat	Oven	1
dishwaser	Oven	1	Oven	dishwaser	0
dishwaser	washer dryer	2	washer dryer	dishwaser	3
dishwaser	Microwave	2	Microwave	dishwaser	10
dishwaser	stove	1	stove	dishwaser	0
washer dryer	microwave	12	Microwave	Washer dryer	10
washer dryer	electric heat	1	electric heat	washer dryer	1
Microwave	electric heat	8	electric heat	Microwave	4
Microwave	stove	6	stove	Microwave	2
electric heat	stove	4	stove	electric heat	2
stove	Oven	1	Oven	stove	0

sequence occurred 10 times in 36 days, but *dishwasher-microwave* sequence occurred only twice. This shows that the correlations in device operations vary according to their sequence of activation, supporting Property 3(b).

5.4 Summary and Discussion

In this chapter, we presented the answer to the research question *"Which flexibilities do users have in their daily routine and are those enough?"*. In this regards, we demonstrated various repeating inter-day and intra-day, house-specific or general patterns of energy distribution and device operation across individual houses'. We also discovered some interesting correlations and sequences between device operations, which further provide valuable information regarding activation times of the correlated devices. Further, we showed the existence of a peak operating period and a longer idle period for some of the devices signifying the potential of extracting time flexibility (shifting activation time) from their operations. Similarly, we demonstrated the existence of variation in total energy consumption, due to the varying duration of operation and power level, which support the existence of energy flexibility in the device operations. We showed that a significant percentage of the total energy demand for a house could be considered to provide flexibility. Finally, we can conclude that even though user exhibits stochastic behavior in device usage, *"users have significant time and amount flexibilities in their daily routine"*. In the next chapter, we will focus on the automated prediction of device-level energy demand and associated flexibility.

Chapter 6

Device-level Energy Demand Forecasting

How can we predict the future device-level energy demand and associated flexibility?

Previous chapters have shown that there exist substantial flexibilities in users daily routine with economic values in the market. However, the flexibility has to to be automatically predicted from the historical device usage behavior. The stochasticity associated with device-level demand makes forecasting a difficult task, and the situation worsens in the absence of context information. In this chapter, we focus on assessing the accuracy and usability of general forecast models for device-level demand and flexibility forecasting. First, we present a number of features that reliably capture usage patterns and address the requirements of a device-level demand forecast. Then, we discuss three different prediction models to predict the device states for various data granularities, namely Logistic Regression, Logistic Regression with weighted class importance, and Pattern Matching as a baseline model. The models are evaluated based on Area Under the precision-recall Curve (AUC) and F1-score. A market player is always more interested in the utility that device-level demand flexibility brings to the rather than the intrinsic model-level quality (accuracy) of the forecast model. Thus, we evaluate the financial viability of device-level flexibility-based DR in relation to the achievable forecast accuracy, i.e., evaluate whether the forecasts are good enough to make money in the market. In this regards, we formulate a set of equations for quantifying the financial benefits of flexibility in energy demands and the loss due to forecast errors. The overall benefit and loss are decomposed and analyzed based on types of prediction categories represented by contingency table. Finally, we investigate the forecast model, forecast horizon, and data granularity best suited for the device-level flexibility analysis and flex-offer generation that maximizes the benefit of flexibility-based DR.

6.1   Device-level Forecasting

In this section, we discuss device-level demand forecast models that can achieve satisfactory results even in the absence of context information. The device-level demand forecast is done in two steps: i) predict the probability of device activation, and ii) predict the demand and operation duration for the activation. Any device at a particular timestamp could be in one of the three possible states, i) *idle-* switched off, ii) *activation* - switched on, or iii) *operating*. The *idle* and *operating* state of a device is represented by 0 and activation state by 1. Further, a *threshold value* represents the minimum power (watts) demand for a device to be in the *activated* or *operating state*.

6.1.1   Data Resolution

Given the historical device-level demand timeseries at a 15-minute resolution, we aggregate the dataset at various granularities to analyse the correlation between the data granularity and the forecast accuracy. Let us define a time series dataset X of device activation profiles for $d-1$ days.

$$X = \{a_1, a_2, \ldots, a_{d-1}\} \tag{6.1}$$

where a value of each "a" depends on the aggregation level discussed in details below. We will consider aggregation into the 3 most commonly analyzed data granularities.

   Hourly Resolution: Here, the forecasting problem is to predict the hourly device activation probability for the next 24 hours. Thus, the device-level consumption data is aggregated to an hourly resolution with the energy consumption replaced by a binary activation value. Specifically, if a reading is above a *threshold value* and represents an activation state then the reading is replaced by 1 else by 0. In this case, the dataset X in Eqn.(1) represents an hourly device activation profile. Hence, each $a_i \in \{0, 1\}^{24}$ is a vector composed of 24 hourly profiles corresponding to certain day i.

   Group Resolution: Here, the forecasting problem is to predict the device activation probability for each group in next 24 hours. The groups are created by distributing the 24 hours of a day into m groups, depending on the distribution of the device activation times in the historical data. For example, we create a set of group $G = \{g_1, g_2, g_3\}$, where $g_1 = \{1, \ldots, 7\}, g_2 = \{8, \ldots, 15\}$, and $g_3 = \{16, \ldots, 24\}$. The hourly resolution dataset is here aggregated to the group resolution. Specifically, if any hour in the group has value 1, then the group gets value 1 otherwise it is set to value 0. In this case, the dataset X in Eqn.(1) represents the device activation profile for the groups. Hence, each $a_i \in \{0, 1\}^m$ is a vector composed of activation profile for m groups corresponding to a certain day i.

   Daily Resolution: Here, the forecasting problem is to predict the probability of the device activation in the following day. Thus, the dataset is here

S.No.	Features	Notation	Description	Example	Feature Used		
					Hourly	Group	Daily
1	Last 24 States, $L(x^{\{24\}})$	if $h=0$ then $\{a_{i-1}^0,\ldots,a_{i-1}^{23}\}$ else $\{a_{i-1}^h,\ldots,a_i^{h-1}\}$	24 Binary Features	$\{0,0,1,\ldots,0,1\}$	X	X	
2	Last 7 States, $L(x^{\{7\}})$	$\{a_{i-7},\ldots,a_{i-1}\}$	7 Binary Features	$\{1,0,1,1,0,1,1\}$			X
3	Hour of a Day $H(x^{\{24\}})$	$\chi^{hour}(x_i)=\begin{cases}1 & \text{if hour of the prediction point}\\0 & \text{otherwise}\end{cases}$	24 Binary Features hour $\in\{0,\ldots,23\}$	For h = 2 $\{0,0,1,0,\ldots,0,0\}$	X	X	
4	Days of the Week, $D(x^{\{7\}})$	$\chi^{day}(x_i)=\begin{cases}1 & \text{if the prediction day}\\0 & \text{otherwise}\end{cases}$	7 Binary Features day $\in\{mon,\ldots,sun\}$	For Thursday $\{0,0,0,1,0,0,0\}$	X	X	X
5	Is weekend, $W(x^{\{1\}})$	$\chi(x_i)=\begin{cases}1 & \text{if weekend}\\0 & \text{otherwise}\end{cases}$	1 Binary Feature	If $sat,$ or sun than 1 else 0	X	X	X
6	Last Operation, $LO(x^{\{7\}})$	$\chi^n(x_{i-7},\ldots,x_{i-1})=\begin{cases}1 & \text{if } x_{i-n}=1 \text{ and } \forall\{x_{i-n+1},\ldots,x_{i-1}\}=0\\0 & \text{otherwise}\end{cases}$	7 Binary Features n $\in\{1,2,3,4,5,6,\geq 7\}$	If last operation was 2 days before, than $\{0,0,0,0,0,1,0\}$	X	X	X
7	Season, $S(x^{\{4\}})$	$\chi^s(x_i)=\begin{cases}1 & \text{if season of the prediction point}\\0 & \text{otherwise}\end{cases}$	4 Binary Features s $\in\{winter,spring,summer,autumn\}$	If spring, than $\{0,1,0,0\}$	X	X	X

Table 6.1: Features Description

aggregated to the daily resolution. Specifically, if any single reading in the day is greater than the *threshold value*, then the day gets value 1 otherwise it is set to 0. In this case, the dataset X in Eqn.(1) represents daily device activation profile and each $a_i \in \{0,1\}$ is the activation state corresponding to a certain day i.

6.1.2 Feature Extraction

We analyse the collected device-level dataset with an aim to extract features that can reliably capture the device activation patterns and energy demand. More specifically, we generate additional derived values from the initial measured data to enhance the information on the device activation and usage patterns. The descriptions of some of the extracted features is described in the Table 6.1.

The present state of a device is highly dependent on its previous states, i.e., a device with no recent activities has a higher probability of activation than the devices recently activated. Thus, we extract the device states in the *previous 24 hours* as 24 binary features and an additional 7 binary features to represent the time since the *last operation* (1 and 6 in Table 6.1). For the daily forecast, we extract the device activation patterns for the past l days, where l is the window size (2 in Table 6.1). We assume that the uses of devices have some temporal patterns, e.g., an oven is mostly activated during the morning and evening, and the dishwasher is mostly operated after the lunch or dinner, etc. Further, we can notice a variation in device activation patterns during the days of the week. Therefore, we generate 24 binary features to represent each *hour of the day* and 7 binary features representing the *day of the week* (3 and 4 in Table 6.1).

To capture the influence of seasonal factors on the usage patterns, we include four binary features representing the four *seasons* of the year (7 in Table 6.1)). In addition, we create various additional features as a multiplicative interaction between the above-extracted features. We will in the following use $x_i = x_i^1, x_i^2, \ldots, x_i^m$ to represent m features corresponding to a data point i in X, and use the convention that x^m refers to the mth feature and $x^{\{m\}}$ refers to a set with m features.

6.1.3   Learning Models

In this thesis, we present a simple pattern matching model as a baseline method for performance comparisons. We further present the logistic regression model with various extracted features and learning parameters. In our experiment, the forecast is made at the end of the day (24:00 hour). However, the same method can be used to build a model to forecast starting at any hour of the day.

Pattern Sequence Matching: The Pattern Matching (PM) algorithm works under the premise that days with the same intra-day pattern are likely to continue to behave somewhat similar the following day. Therefore, to predict the hourly device activation probability of the day d, the PM algorithm first searches the days with device activation pattern that matches the pattern of day $d-1$ in the hourly resolution dataset X. This search outputs a set of indexes for matching days M, defined as $M = \{i \mid a_i = a_{d-1}\}$. The probability of device activation for each hour of day d is calculated as:

$$p(h) = \frac{1}{size(M)} \sum_{i \in M} a_{i+1}^h, \tag{6.2}$$

where size(M) is the number of elements in the set M and h is the hour of the day. If set M is empty, we calculate the probability over the complete dataset. To predict at a group resolution, we follow the same procedure with group resolution dataset X, and h replaced with g. The hourly forecast model is described in Algorithm 1:

Algorithm 1 Hourly Forecast

Input: hourly resolution dataset $X = \{a_1, \ldots, a_{d-1}\}$.
Output: $p(h)$ for all hours of day d.
  1: Initialize $M = \emptyset$ and $p(h) = 0$
  2: H = hour with highest device activation frequency in X
  3: **for all** i such that $a_i \in X$ **do**
  4:      **if** $a_i = a_{d-1}$ **then** Append i to M
  5: **for all** hours h **do**
  6:      **if** $M = \emptyset$ **then**
  7:          **if** $h = H$ **then**
  8:              Update $p(h) = 1$
  9:          **else**
 10:              Update $p(h) = 0$
 11:      **else**
 12:          **for** each $j \in M$ **do**
 13:              Update $p(h) = p(h) + a_{j+1}^h$
 14:          Update $p(h) = \frac{p(h)}{size(M)}$
       return $p(h)$

Similarly, device activation has some repetitive patterns over the days, such as devices might be activated every second day, some daily, some once a week, etc. Let $S_i^l = [a_{i-l+1}, \ldots, a_i]$ be the sequence of a daily device activation pattern for l consecutive days, from day i backward. In our experiments, we use the $l = 7$ to capture the weekly behavior of device activation patterns. Now, the PM algorithm for prediction of the device activation probability of the day d first searches the sequence of device activation patterns that exactly matches the sequence of pattern S_{d-1}^l in the daily resolution dataset X. The search output a set of indexes for the last day of matching subsequences M, defined as $M = \{i \mid S_i^l = S_{d-1}^l\}$. If set M is empty, the algorithm searches for the sequence of device activations which are equal to S_{d-1}^{l-1} and thus successively until $l = 1$. The probability of device activation on day d is calculated as

$$p(d) = \frac{1}{size(M)} \sum_{i \in M} a_{i+1}, \tag{6.3}$$

The daily forecast model is described in Algorithm 2:

Algorithm 2 Daily Forecast

Input: daily resolution dataset $X = \{a_1, \ldots, a_{d-1}\}$.
Output: device activation probability $p(d)$ of day d.

1: Initialize $M = \emptyset$, $p(d) = 0$, and $l = 7$
2: **while** $M = \emptyset$ and $l > 0$ **do**
3:     **for all** i such that $a_i \in X$ **do**
4:         **if** $S_i^l = S_{d-1}^l$ **then** Append i to M
5:     Update $l = l - 1$
6: **for** each $j \in M$ **do**
7:     Update $p(d) = p(d) + a_{j+1}$
8: Update $p(d) = \frac{p(d)}{size(M)}$
9: **return** $p(d)$

Logistic Regression: The standard logistic regression [86] model has been used extensively in the literature for various binary classification problems. The model defines the relationship between a set of explanatory variables and a dependent classification variable, and provides the probability or likelihoods of the possible outcomes. Let $Y = \{y_1, y_2, \ldots, y_n\}$ be binary dependent variables where each $y_i \in \{0, 1\}$ represents the class label for the feature vector x_i. Let, z_i represent a linear function of the explanatory variables:

$$z_i = \theta^0 + \sum_{j=1}^m \theta^j x_i^j,$$

where $\theta = (\theta^1, \theta^2, \ldots, \theta^m)$ are the regression parameters associated with the explanatory feature vector $x_i = (x_i^1, x_i^2, \ldots, x_i^m)$ The probability that a new data point belongs to the class label 1 is represented as:

$$p(y_i = 1 | x_i; \theta) = \frac{1}{1+e^{-z_i}},$$

Let us use $\pi_i = p(y_i = 1 | x_i; \theta)$ to simplify the notation, the probability of the possible outcome to be class 0 is represented as: $p(y_i = 0 | x_i; \theta) = 1 - \pi_i$.

Given the training data $((x_1, y_1), \ldots, (x_n, y_n))$, the optimal regression coefficients θ can be estimated by maximizing the log-likelihood function:

$$L(\theta) = \sum_n (y_n \, ln\pi_n + (1 - y_n)ln(1 - \pi_n)),$$

We are considering a high number of features in the model and therefore introduce L1 regularization in order to counter overfitting the model to the training data. The L1 regularized log-likelihood function is:

$$L(\theta) = \sum_n (y_n \, ln\pi_n + (1 - y_n)ln(1 - \pi_n)) + \lambda \sum_{j=1}^{m} |\theta^j|$$

where λ is the regularization parameter. In this thesis, we use two versions of logistic regression i) standard implementation using *statsmodel* in Python, ii) *Vowpal Wabbit* [87] with weighted class importance - where positive and negative instances are assigned *importance weight* relative to the class ratio, i.e., instances of minority class are assigned a higher weight.

6.1.4   Demand Forecast

We need to estimate the demand at the point of forecasted device activation and also the duration of the operation. Depending on the device current state and configuration, demand from a device can range from a few watts to thousands of watts. This high sparsity in energy demand from a device, 705 unique demand values for washer dryer operations, creates challenges in estimating the demands for a particular operation. Thus, to reduce the sparsity in the data, we discretize the demand values into various bins, with an interval of 500 watts. Further, we average the demands value in each bin to generate a single value that represent the bin. The next task is to determine the duration of the device operation, i.e., the time between device activation and transition back to an idle state. We assume that the device always operates for the duration of 2 hours, i.e., the lower bound of average operation duration of 2.27hrs. Although this approach seems simplistic at first, we emphasize that it represents 64% of the actual *washer dryer* operations duration in the experiment.

The next step to demand forecast includes estimation of energy demand for the activation and subsequent durations of device operation. The process starts with the extraction of the energy profiles for all previous activation's triggered at the hour h, i.e., the hour to predict. The extracted profiles are comprised of demand for each unit time of the device operation. The multiple micro-profiles are aggregated into a single macro-profile by averaging the profiles values at the respective time unit of operation. The macro-profile value for each time unit is compared with the bins ranges. The value of the bin whose range includes the macro-profile value is selected as the estimated demand for the corresponding time unit. Figure 6.1 shows an example of a day-ahead hourly demand forecast, with 2 predicted device activation in the day.
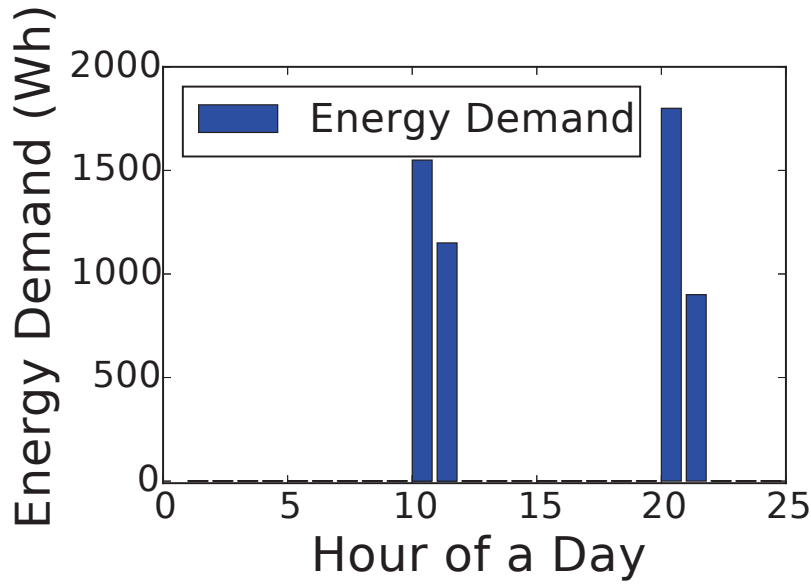
Figure 6.1: Demand Forecast - for next 24 hours

6.1.5 Model Evaluation

Precision, Recall, and Receiver Operator Characteristics (ROC) are commonly used in the literature for the binary decision problems. However, for the class imbalanced dataset, the ROC curve does not provide the real picture of the performance of the model due to the slower increasing rate of the dominant class (96.9% of instances in our dataset), i.e., false positive rate. Therefore, in our experiment, we evaluate the performance of the classifiers on Area Under the precision-recall curve[1] as discussed in [88]. Let $\hat{F} = \{\hat{f}(1), \hat{f}(2), \dots, \hat{f}(24)\}$

Actual $f(t)$	Predicted $\hat{f}(t)$	Type	Benefit	Loss
0	0	True Negative (TN)		
0	> 0	False Positive (FP)	X	X
> 0	0	False Negative (FN)	X	X
> 0	>0	True Positive (TP)	X	

Table 6.2: Categories of forecast result: based on actual and forecasted demand.

be 24 hourly forecasted demands, i.e., flexible demands for the day d, and $F = \{f(1), f(2), \dots, f(24)\}$ be the actual demands at the time of delivery. Based on the forecasted and actual demand the results of a forecast model can be divided into four categories, shown in Table 6.2. The table further shows the consequences, in terms of benefit and loss, that a market experience for each category of the result, detailed in Section 6.2. Therefore, a market player is always interested in quantifying the pros and cons of each result categories

---

[1]Precision-Recall curve is a plot of the Precision against the Recall at various threshold values.

individually, which leverage them in selecting a forecast model with the best performance on the desired category. For example, a market can desire a model with a higher precision or a higher recall, or have a trade-off between precision and recall.

6.2   Experimental Setup and Mathematical Formulations

In this section, we will define the savings that can be generated from the energy flexibility and analyze it in relation to the forecast error.

6.2.1   Analysis on Forecast Performance

In a flexibility market, a precise demand forecast model is crucial for the market to follow the planned schedules. However, at the device-level, the demand at a particular time highly depends on various factors, such as user availability, preference, weather condition, device settings, etc. Therefore, the forecast model suffers from stochastic user behaviors and external factors that are hard to capture, leading to a higher imbalance in the market due to the scheduling of false and unplanned energy demand. A market player is always interested to know the maximum limit of forecast error that can be handled in the flexibility market without any further financial loss. In this regard, we will analyse the effect of each type of forecast result described in Table 6.2.

For the TN results ($f(h) = \hat{f}(h) = 0$), the market neither has flexible demands to schedule nor experience any unexpected demands at the time of actual delivery. Thus, no financial loss or benefit comes with the TN result. In the case of FP results ($f(h) = 0$ and $\hat{f}(h) > 0$), the loss depends on the market balance at the time of actual deliveries. For example, the up-regulated market at the scheduled timestamp can achieve financial gain by a reduction in regulation volume. On the other hand, the FP will increase the anticipated total demand due to inaccurate estimation, which in turn causes the financial loss due to the change in the market prices, discussed in Section 6.2.3. Similarly, in the case of FN results ($f(h) > 0$ and $\hat{f}(h) = 0$), an unscheduled demand could increase the up-regulation volume causing financial loss or decrease the down-regulation volume generating financial benefits. Finally, for the TP results, the market generates financial benefit by pre-scheduling the flexible demand to reduce the regulation cost. In the next section, we will quantify the benefit that can be achieved by shifting of forecasted flexible demand and the corresponding loss due to the forecast error.

6.2.2   Scheduling of Flexible Demand

The extent to which a forecasted demand, i.e., flexible demand can be shifted is constrained by the time flexibility associated with the demand. Let $\tau \in$

$\{0, 1, 2, \ldots, 24\}$ be time flexibility associated with each forecasted demand in $\hat{F}$, where, in particular $\tau = 0$ corresponds to inflexible demands. Now, $\hat{F} = \{(\hat{f}(1), \tau_1), (\hat{f}(2), \tau_2), \ldots, (\hat{f}(24), \tau_{24}\}$ represents a vector of tuples, where τ_i is the time flexibility for $\hat{f}(i)$. To ease notation, we will assume the same fixed time flexibility for all the demands, but this assumption is easily generalized to varying time flexibilities across demands. Let, $V = \{v_{u/d}(1), v_{u/d}(2), \ldots, v_{u/d}(k)\}$ represents regulation volumes for the next k hours from the time of forecast, where $k = 24 + \tau$. Now, the scheduling task is to assign the flexible demands $\hat{F}$ to V such that the market minimizes their total regulation volume $\sum_{i=1}^{k} |v_{u/d}(i)|$ without violating the constraints. The shifting of flexible demand is subject to the following constraint: i) $s_i \in \{i, i + \tau\}$ where s_i is the scheduled time for $\hat{f}(i)$.

The scheduling problem has a finite solution space and could be solved by a greedy approach or a brute force algorithm. The greedy approach can find the best schedule for a single flexible demand, but does not guarantee the optimality for a large number of demands. On the other hand, even for the current problem the brute force algorithm takes significant time to find the optimal solution and creates a problem for generating schedules for a large number of devices. Therefore, we present the scheduling task as an optimization problem and solve it using the GLPK solver with PuLP in Python. The worst case running time for the solver, to find an optimal solution, is < 3ms.

6.2.3 Change in Regulation Price due to Scheduling

The inaccurate estimation of demand changes the anticipated regulation volume. Since, the regulation prices in the market depend on the volume and type of regulation [40], the change in volume affects the regulating power prices in the market. Thus, to evaluate the change in regulation price due to an error in the forecast, we use the hypothetical relationship between energy prices and regulation volume as proposed in [89].

$$
\begin{aligned}
p_{u/d}(i) = {}& 1.p_s(i) \\
& + 1_{v_d(i)<0}(-0.334 \cdot p_s(i) + .0005 \cdot (p_s(i) \cdot v_d(i))) \\
& + 1_{v_u(i)>0}(0.238 \cdot p_s(i) + .0034 \cdot (p_s(i) \cdot v_u(i))) \qquad (6.4)
\end{aligned}
$$

Here, $1_{a<b}$ denotes the indicator function for the predicate $a < b$, and $p_{u/d}(i)$ is the predicted up-regulating power price $p_u(i)$ in case of up-regulation the predicted down-regulating power price $p_d(i)$ in case of down-regulation.

6.2.4 Savings in Regulation Cost due to Scheduling

For each hour in V, the loss due to the market imbalance is computed as a product of the regulation volume times the price difference between regulating

and the spot price. Hence, the total regulation cost for V is calculated as:

$$R = \sum_{i=1}^{k} v_{u/d}(i) * |p_{u/d}(i) - p_s(i)| \tag{6.5}$$

where $p_{u/d}(i)$ and $p_s(i)$ are regulation price and spot price, respectively.

Given, the regulation volume and forecasted flexible demand $\hat{F}$, the market generate a demand schedule that minimizes the regulation volumes. Let the new expected regulation volumes be $\bar{V}|\{\forall i, \overline{v_{u/d}(i)} \leq v_{u/d}(i)\}$, where the over-bar denotes the change in regulation volume due to shifting of flexible demand. Hence, the total expected regulation cost E is given by:

$$E = \sum_{i=1}^{k} \overline{v_{u/d}}(i) * |p_{u/d}(i) - p_s(i)| \tag{6.6}$$

The objective of scheduling the flexible demands is to reduce the regulation cost of the market. Thus, the expected regulation cost E is always less than or equal to R, i.e., $E \leq R$. Therefore, savings in regulation cost due to shifting of flexible demand is given by: $\Delta R = R - E$

At the time of actual delivery, if the demand deviates from the previously forecasted demand the market player, i.e., BRP that caused the specific imbalance is financially responsible for the deviation. The total financial loss due to the error in demand forecast is calculated as:

$$L = \sum_{i=1}^{k} |f(i) - \hat{f}(i)| * |p_{u/d}(i) - p_s(i)|$$
$$+ 1_{f(i) \neq \hat{f}(i)}(v_{u/d}(i) * |\overline{p_{u/d}(i)} - p_s(i)|) \tag{6.7}$$

Notice the notation, where overbar denotes the updated regulation price calculated using the update regulation volume, i.e., $\overline{v_{u/d}(i)} = v_{u/d}(i) \pm f(i)$, i.e.,updated regulation volume due to the inaccurate estimation of demand. Recall the classification of forecast results in Table 6.2, ΔR and L represents the benefit and the loss for the table, respectively. The $\Delta R - L$ gives the total benefits of shifting flexible demands.

6.3   Experimental Analysis

6.3.1   Evaluation of Device-level Forecast Models

In this section, we will analyze the performance of the three different classifiers, i) Logistic Regression (LR), ii) Logistic Regression with weighted class importance(LR-W), and iii) Pattern Matching (PM), for various data granularities. The experiments are performed using the demand timeseries data for a *washer dryer*.
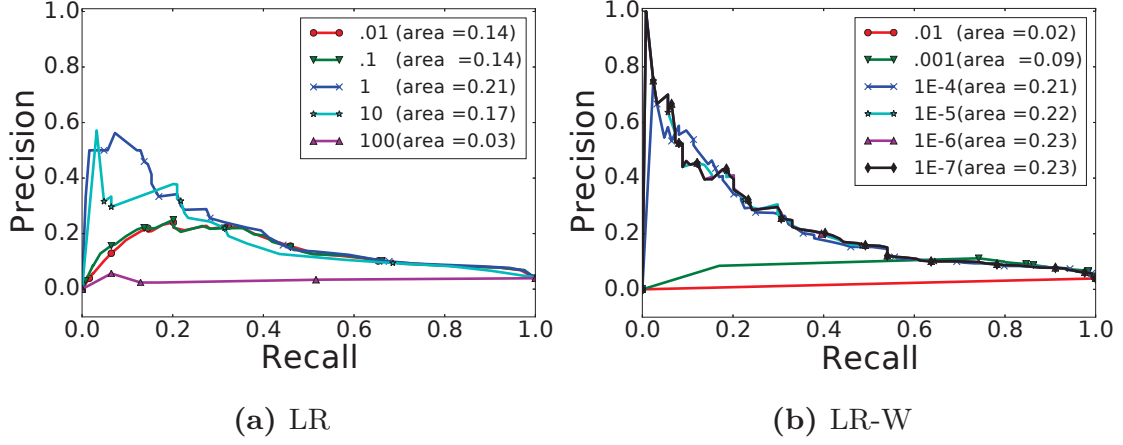
(a) LR

(b) LR-W

Figure 6.2: PR curve - for various λ values (hourly)
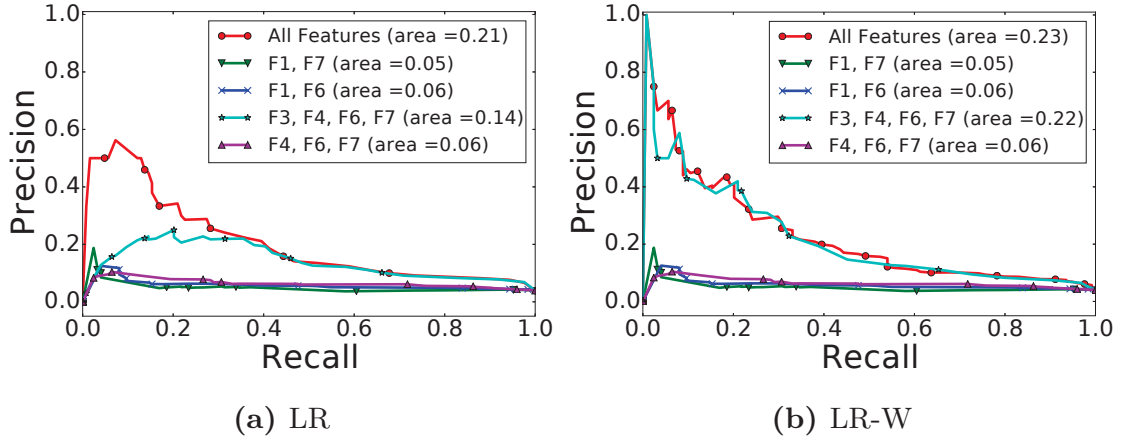

(a) LR

(b) LR-W

Figure 6.3: PR curve - for various feature sets (hourly)

First, we analyze the performance of the classifiers for the hourly data granularity. To compare the classifiers, first we evaluate the variation in the performance when changing the regularization value λ, shown in Figure 6.2a and 6.2b. The performance of both models (LR and LR-W) degrades with increasing λ values, mainly because increasing the penalty drives parameters θ to zero and deselect most of the features in x (feature vector). The best regularization parameter for the classifiers is estimated via cross-validation over each λ value. The cross-validation gives the best average performance with $\lambda = 1$ and $1E - 6$ for LR and LR-W, respectively, and the models achieve an AUC of 0.21 and 0.23 for the respective λ values. Further, Figures 6.3a and 6.3b show the performance of the classifier for various sets of features $x^{\{m\}}$. Here, both LR and LR-W achieve the best performance with the complete set of all extracted features. Thus, we argue that the best strategy is to feed a classifier with all the features and tune the model correctly so that it self-selects the most relevant ones.

The precision-recall curve for all 3 classifiers is shown in Figure 6.4a, and
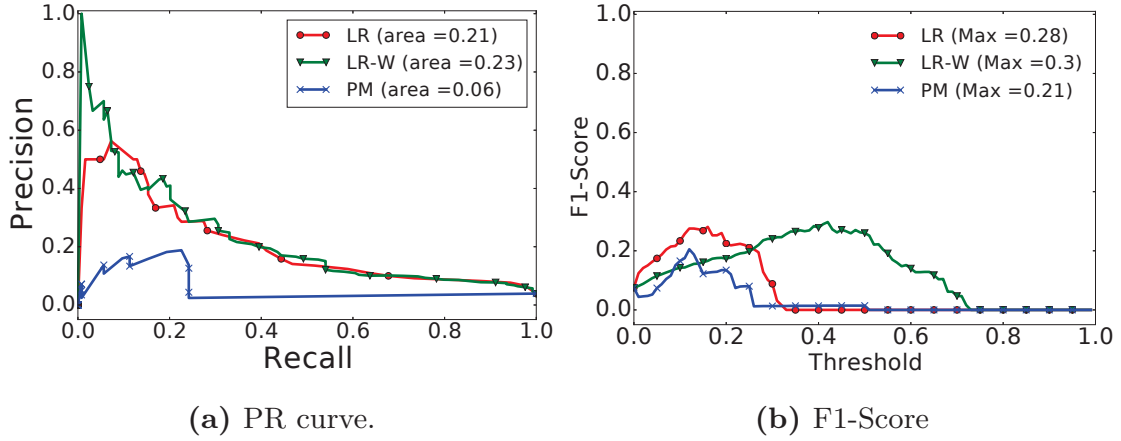
(a) PR curve.   (b) F1-Score

Figure 6.4: Performance of classifiers (hourly).
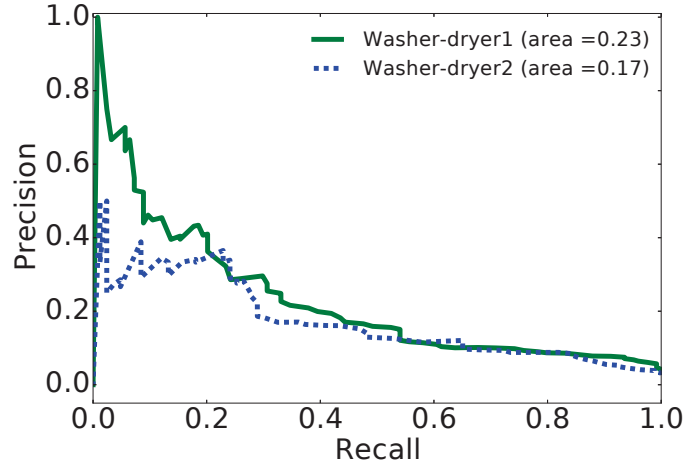
Figure 6.5: Performance of LR-W across washer dryers.

LR-W model has the highest area under the curve of 0.23 compared to 0.21 and 0.06 for LR and PM, respectively.

Further, the results show that the predicted class probabilities for LR and PM are clustered in a small region, i.e., have lower prediction confidence. On the other hand, LR-W predicts the positive and negative class with a higher confidence that gives smooth precision-recall curve and the model achieves a precision of 1, i.e., 0 unexpected demands, for some threshold values. The lower confidence in prediction is due to the class imbalance and stochastic behaviors associated with the device-level demand. The lower prediction confidence creates fluctuating (nonlinear) precision and recall curve as shown in 6.4a. Thus, to select the threshold value that gives the best performance of the model, we analyze the F1-scores of the classifiers as shown in Figure 6.4b. LR, LR-W, and PM achieve the best performance at a threshold value of 0.16, 0.42, and 0.12, respectively. However, the best threshold value depends on the objective of analysis, discussed in Section 6.3.2.

(a) LR

(b) LR-W

Figure 6.6: PR curve - for oversampling (hourly)

(a) LR

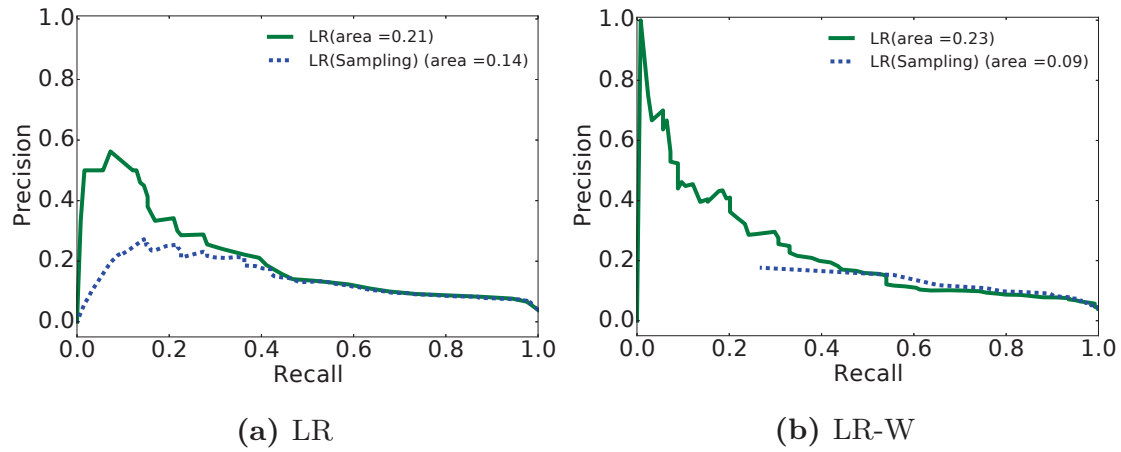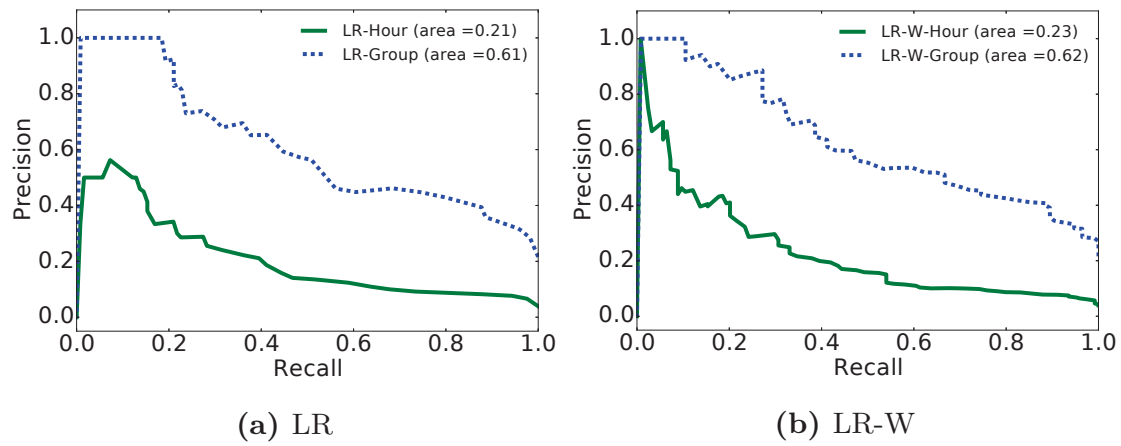**(b)** LR-W

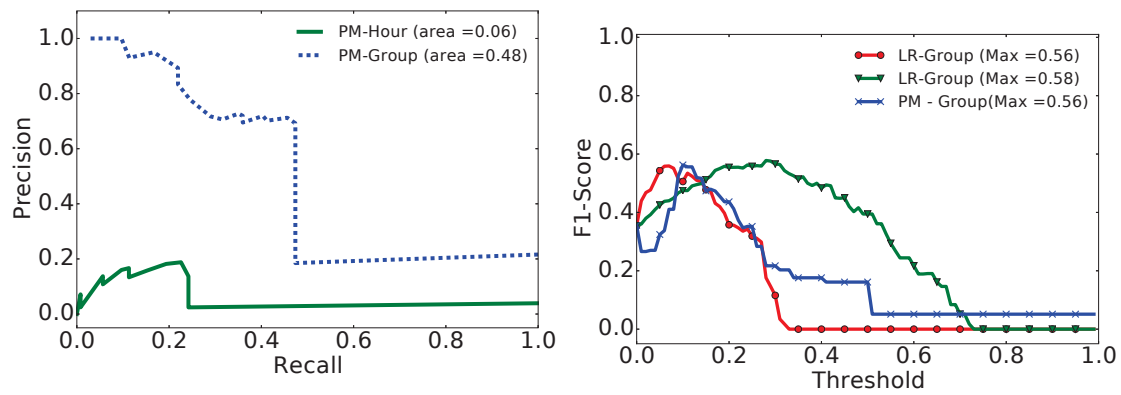**Figure 6.7:** Performance of classifiers (hourly versus group).

Figure 6.8: Performance of PM (hourly versus group).

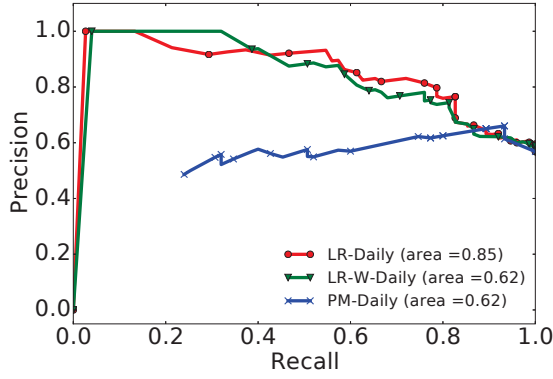**Figure 6.9:** F1-Score of classifiers(group).

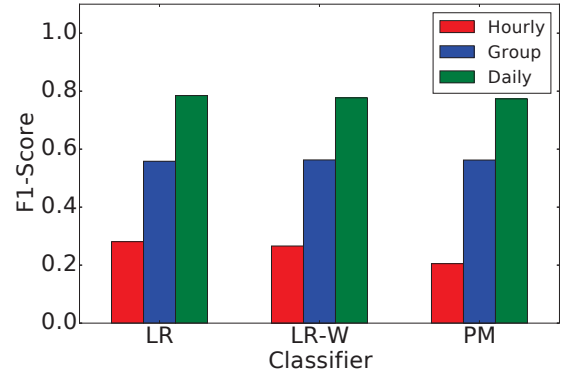**Figure 6.10:** Performance of classifiers (daily).

Figure 6.11: Comparison of classifiers - across all resolution.

From the figures, we can see that none of the classifiers show a good performance at an hourly resolution, and a simple model such as PM has a performance comparable to a complex model (LR). Further, as shown in Figure 6.5, the quite low AUC achieved by LR-W for 2 different *washer dryers*, shows the stochasticity associated with device-level demand forecasts. Nevertheless, the comparable performance across the two devices illustrates that the proposed device-level forecast model is generalizable. The lower performance of the classifiers is typically also due to a very high percentage of the majority class. Thus, we evaluate the performance of the classifiers with oversampling of the minority class, shown in Figures 6.6a and 6.6b. However, oversampling increases the sensitivity of the classifier towards the minority class which further degrades the performance giving more FPs. Therefore, for the device-level demand forecast, where the minority class is the target of the classification, weighted class importance yields better performance compared to over-sampling approach.

Figures 6.7a, 6.7b, and 6.8 compare the performance of the three classifiers on hourly and group resolution. The figures show that a classifier has a significantly better performance at the group resolution, with LR and LR-W achieving an improvement in AUC of 0.40 and 0.39, respectively. Especially, PM reports the best performance improvement with an increase in AUC of 0.42. These results suggest that device usage patterns are more repetitive in a cluster of hours, e.g., the user frequently activates a *washer dryer* in group g_3 (4 PM-12 AM) depending on his/her presence at home. Figure 6.10 compares the performance of the classifiers at a daily resolution, and Figure 6.11 compares the best F1-score. The figures clearly demonstrate that the classifiers achieve the best performance at a daily resolution with AUC of 0.85, 0.84, 0.62 for LR, LR-W, PM, respectively. At the daily resolution, the imbalance shift towards the positive class and the weighted measure does not contribute to the performance gain. Therefore, the performance of LR surpasses LR-W. Moreover, we can conclude that, at an device-level, the predictability increases with the data aggregation level.

The above results exhibit the stochasticity associated with an device-level
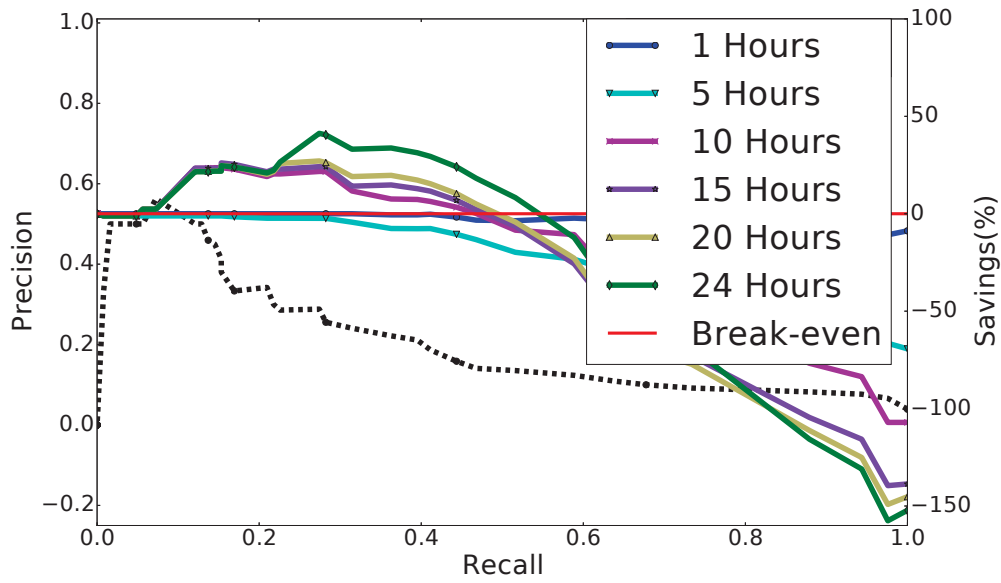
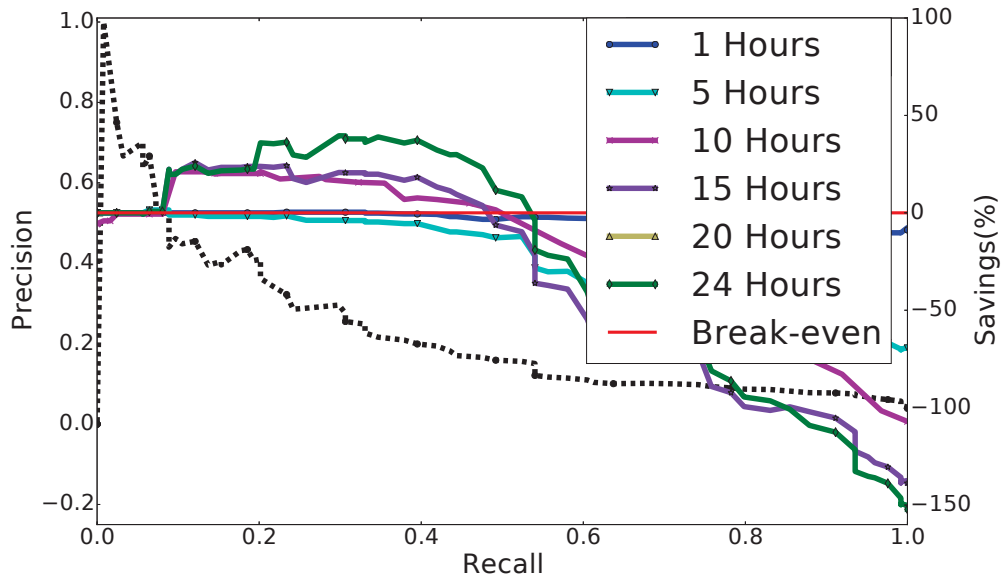**Figure 6.12:** Saving for varying time flexibility: LR($\lambda : 1$, Features: All)

Figure 6.13: Saving for varying time flexibility: LRW($\lambda : 1E - 6$, Features: All)
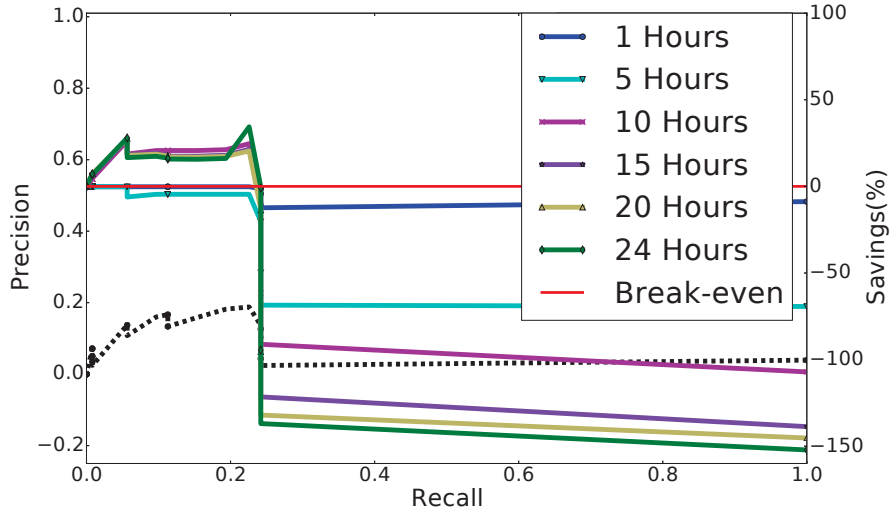
Figure 6.14: Saving for varying time flexibility: PM

demand where it is hard to capture any patterns at an hourly resolution. More-over, in the absence of context information, the unusual behaviors in the usage patterns are wrongly represented by forecast models which degrade their overall performance. However, our main interest is to evaluate the viability of flexibility market utilizing a device-level demand forecast in a stochastic environment. Therefore, in the next section, we will analyze the financial implication on the regulation market relative to the performance level of our forecast models (classifier with demand appended).

6.3.2   Evaluation of Utility of Forecast Models

Here, we will quantify the financial benefit (savings in regulation cost) of demand flexibility in relation to the achievable device-level demand forecast accuracy. First, we evaluate the theoretically maximal savings in regulation cost for a hypothetical 100% accurate demand forecast model. The experiment on a demand timeseries for a *washer dryer* shows that with the hypothetical model, a market (BRP) can generate the maximum savings of 6.1 DKK over the test period. Henceforth, we evaluate the performance of the forecast models relative to the percentage of the maximum savings they can achieve.

Percentages savings for hourly resolution demand forecast models at varying time flexibilities are shown in Figures 6.12, 6.13, and 6.14. The figures show savings for the LR and LR-W models configured for the best performance (discussed in Section 6.3.1). We can see that best savings grows with increasing time flexibility but has a diminishing return for larger time flexibility. The models achieve the best savings for the 24 hour time flexibility. However, for shorter time flexibilities, the models always have losses (negative savings) due to a decrease in the number of genuine demands (TPs) that can be shifted. Thus, all further experiments are performed considering 24 hours of time flexibility.

(a) LR

(b) LR-W
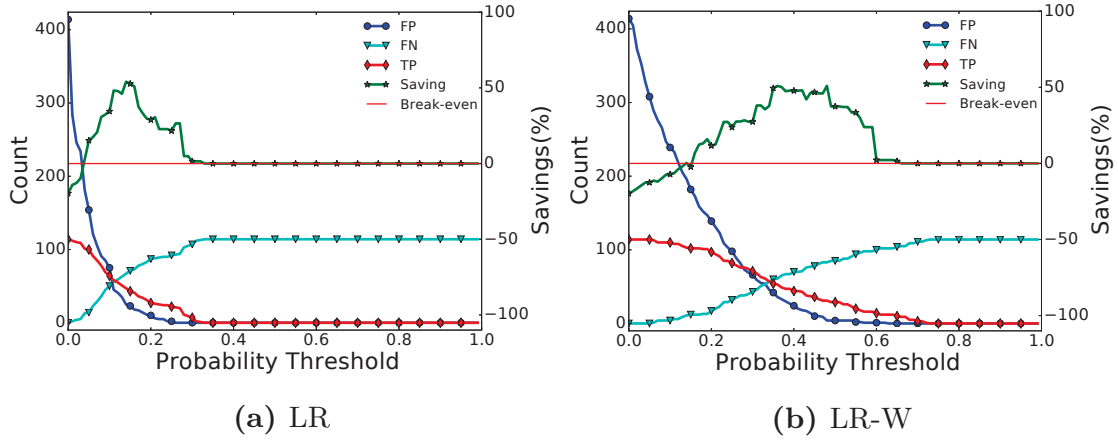
Figure 6.15: Savings from demand flexibility of a device- relative to forecast category.(Group)

(a) LR($\lambda : 1$, Features: All).

(b) LRW($\lambda : 1E - 6$, Features: All)
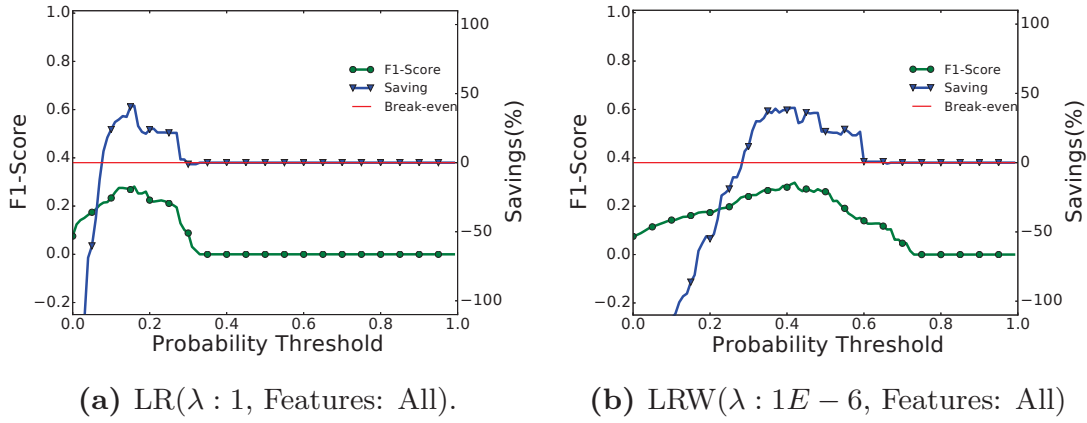
Figure 6.16: Savings from demand flexibility of a device- for various F1-score. (Hourly)

The Figures 6.12, 6.13, and 6.14 show that, given 24 hour time flexibility, the LR has the best savings of 42% compared to 40% and 35% for LR-W and PM, respectively. This contrast the hourly level forecast result where the LR-W has better AUC than LR and PM.

In general, the savings is positive for a higher precision value, but the corresponding recall value determines the size. For example, in Figure 6.12 the savings with a precision of 0.56 is 40% less than with 0.28. The savings decreases due to the lower *recall* where the loss due to the FP cases is higher than the gain due to the TP cases. The figures illustrate that savings are positive only for a portion of the precision-recall curve and negative for the rest. The positive region represents precision-recall values with accuracy enough to cover the losses due to FN and FP cases, and these regions are specific to the forecast model, e.g., LR-W has positive savings for more points than PM.

Figures 6.15a, 6.15b show the savings for various probability thresholds relative to the forecast categories, i.e., TP, FP, and FN. We can clearly see that at lower thresholds, the number of FPs is too high to generate any savings,
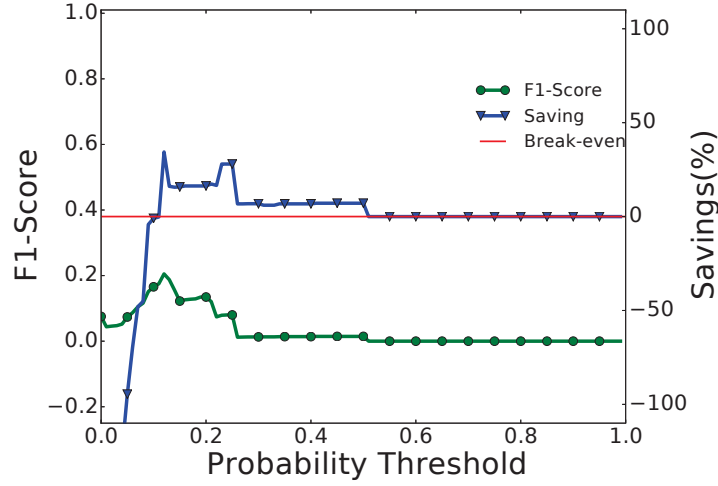
Figure 6.17: Savings from demand flexibility - PM (Hourly)

e.g., at a probability threshold of 0.02, LR and LR-W have FPs equivalent to 47% and 83% of the total instances, respectively. However, with an increase in the probability threshold, the FPs and TPs decreases as does the loss. The figures illustrate that a flexibility market can generate substantial savings even in the presence of a vast number of FPs. For example, a savings of 24% for LR and 11% for LR-W can be achieved even with 84% of total forecasted flexible demand being false. These savings are mainly attributed to the change in regulation prices, where for some FPs the loss due to increase in up-regulation price at i is lower than the gain due to a decrease in down-regulation price at $i + \tau$. With a further increase in threshold, the number of TPs drops to 0 and so do the savings.

Since the savings depend on the timestamps of the correctly forecasted instances and the number of FPs, the *precision* and *recall* values for positive savings are inconsistent with the models. For example, LR has a savings of 7.5% for the *(precision, recall)* value of (0.13, 0.52) , whereas, LR-W has a loss of -9.8% for (0.13, 0.54). In addition, the savings does not follow the patterns of the precision-recall curve. This behavior creates difficulty in selecting a probability threshold value for a model that guarantee the positive savings, i.e., the probability threshold that gives the desired *precision*. To this end, we evaluate the savings relative to the *F1-score* at various probability threshold values, shown in Figures 6.16a, 6.16b, and 6.17. The figures show that for all the models, the savings follows the respective *F1-score*, and the model achieves positive savings at a point of the highest *F1-score*. Thus, the problem of setting the optimal probability threshold can be solved by selecting the threshold value with the highest *F1-score*. This rule of thumb is valid for all cases with significant savings, i.e., all experiments with best positive savings of $> 1.1\%$. The highest *F1-score* achieved by the LR, LR-W and PM are 0.28, 0.3, and 0.21, respectively.

Figures 6.18a, 6.19a, and 6.20a illustrate the savings from the demand

(a) Group   **(b)** Daily

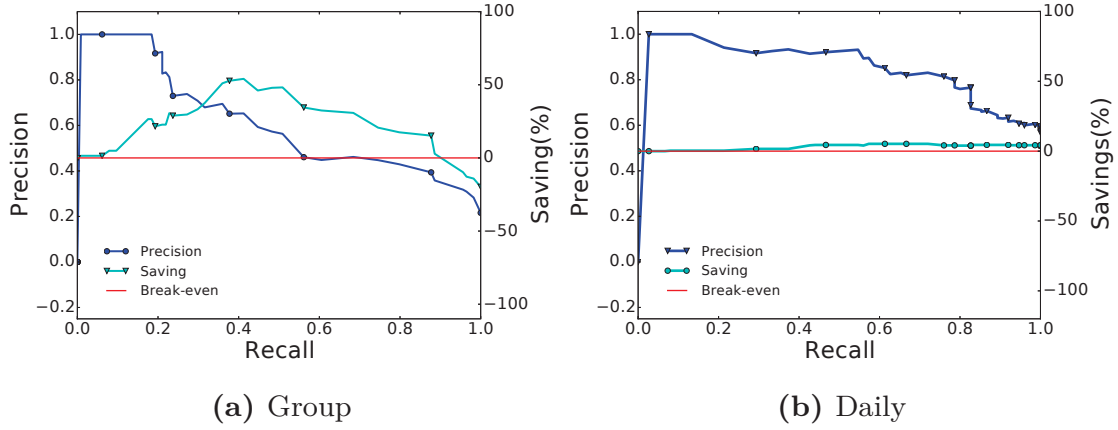**Figure 6.18:** Savings from demand flexibility of a device- relative to precision-recall (LR).

flexibility utilizing the demand forecast at the group resolution. The figures demonstrate that the savings increases with the data aggregation, however, the increase is not proportional to the improvement in forecast accuracy. For Example, the best *F1-score* for LR is double of hourly resolution, but at the same time, the maximum savings increase only by 29%. Similarly, we can see a 93% improvement in the best *F1-score* for LR-W, but the savings increase only by 27%. Because the best savings for PM lies at the point of the best *F1-score*, PM surpasses the savings from LR by 10% despite having lower prediction accuracy (AUC). For other models best savings does not coincide with the best *F1-score*, thus, they have a comparatively lower precision and recall value at the point of the best savings. These results show that in the flexibility market, a model with a better forecast performance (AUC) does not guarantee higher savings. Further, a comparable savings from the baseline model (PM) supports our argument that, at the device-level, a simple model can compete with a complex one. In addition, the figures demonstrate an extended positive region for all the forecast models, which indicates that a market can generate savings for almost all values of the precision-recall curve. Moreover, best savings are obtained from the *(precision, recall)* values of (0.65, 0.41), (0.6, 0.46), and (0.69, 0.47) for LR, LR-W, and PM, respectively. Further, a market (BRP) has a better confidence in the forecasted demand flexibility at a group resolution and is guaranteed to obtain savings from it.

The savings from the demand flexibility utilizing the demand forecast at the daily resolution is shown in Figures 6.18b, 6.19b, and 6.20b. For the daily resolution, the savings is drastically reduced due to a decrease in the number of available flexible demands to be scheduled, e.g., the maximum savings for LR is only 13% and 10% of hourly and group resolution, respectively. Though, the savings is comparatively less, a market will never have a loss due to FP or FN flexible demands.

The above results show that the highest saving in regulation cost is achieved with a demand forecast model at the group resolution, nevertheless a market

(a) Group

(b) Daily
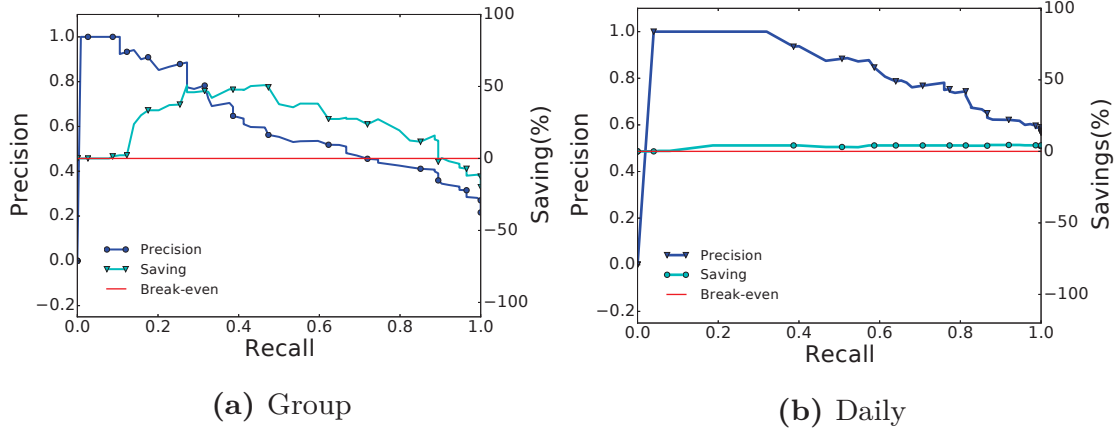
Figure 6.19: Savings from demand flexibility of a device- relative to precision-recall
(LR-W).

(a) Group

(b) Daily
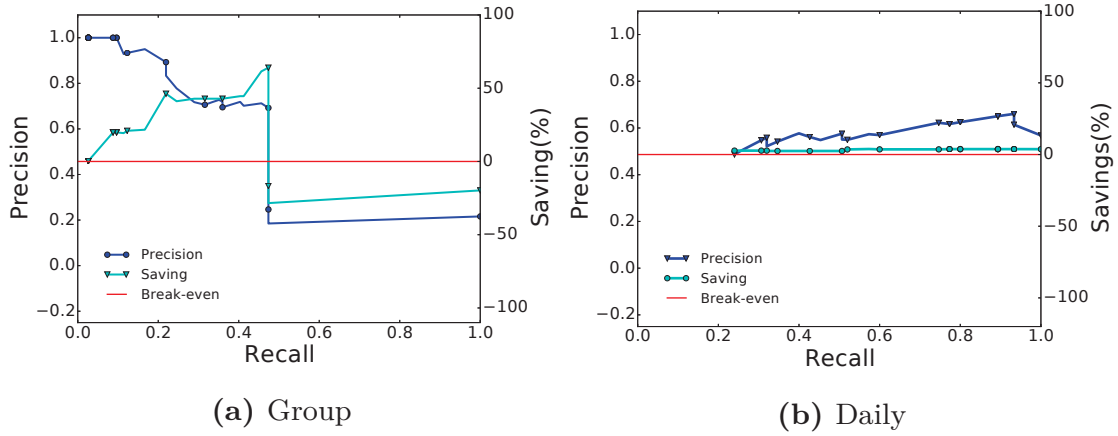
Figure 6.20: Savings from demand flexibility of a device- relative to precision-recall (PM).

can also generate substantial savings at the hourly resolution. However, even
though the accuracy of a model improves with data aggregation, a model has
the least savings at the daily resolution due to a corresponding decrease in
the number of available flexible demands. These results show that device-
level flexibility-based DR can be a promising tool to confront the challenges of
integrating RES into the grid system. Moreover, an energy market can extract
the benefit of device-level demand flexibility even in the presence of a large
number of false predictions, i.e., FPs. However, the maximum proportion of
FPs and FNs, i.e., the lower bound of precision and recall, that a market can
sustain are specific to the market.

6.4   Summary and Discussion

In this chapter, we presented an answer to the research question *"How can we
predict the future device-level energy demand and associated flexibility?"*. In

this regards, we assessed and compared the performance of the three different forecast models, namely two variants of L1-regularized logistic regression and one pattern sequence matching to evaluate their effectiveness and usability for device-level demand prediction. We presented various device-level features that could reliably capture usage patterns and address the requirement of device level forecast. We compared the performance of the forecast model at various data granularity and forecast horizon and identified the granularity and the horizon best suited for device-level flexibility analysis. The prediction models were evaluated based on the Area Under the precision and recall Curve and F1-score. Further, we quantified the effect of forecast error on the flexibility based demand response. Though the experiments, we showed that benefit of demand flexibilities could be maximized utilizing a forecast model at the group resolution. Further, we demonstrated that for the device-level demand forecast, financial gain for a market is much better than implied by the error metrics. Hence, even with a low accuracy at the hourly resolution, a market can still yield positive financial benefits. Finally, we can conclude that *"traditional forecast model does not provide higher accuracy at device-level, but the accuracy can be improved by tuning the data granularity and even with a lower prediction accuracy a market can achieve an acceptable utility"*. In the next chapter, we focus on the generation and evaluation of flex-offers from various flexible devices.

Chapter 7

Generation and Evaluation of Flex-offers

How can we effectively extract and generate flex-offers from household devices?

In previous chapters, we have shown that there exists significant time and amount flexibilities in user's daily routine and these flexibilities can be predicted at a level that is economically viable to the users and market players. However, the main question of how flexibilities from various device types can be extracted and modeled as flex-offers, for efficient aggregation, trading, and scheduling, is still unanswered. In this regards, this chapter presents a generalized Flex-offer Generation and Evaluation Process (FOGEP) that extract and model flexibilities from wet-devices (e.g. dishwashers), electric vehicles, and heat pumps. We first present a generalized FOGEP that analyzes past consumption behavior of a device (using real-world dataset) to capture flexibility in the device usage behavior and automatically predicts the future energy demands and associated flexibilities for the device. Henceforth, we present a method for modeling the predicted flexibilities from all device types into a unified format represented as flex-offers that facilitate, e.g., aggregation and trading. We utilize device-level forecasting techniques and algorithms (discussed in Section 6.1) to capture various attributes and temporal patterns required for flexibility extraction. Further, we perform an evaluation of the performance of FOGEP regarding the accuracy of the extracted flexibility and an economic assessment to identify the device-specific best market to trade flexibility. The economic evaluation of FOGEP is performed on the spot and regulation markets, discussed in Section 3.3.

7.1 Flex-offer Generation and Evaluation Process (FOGEP)
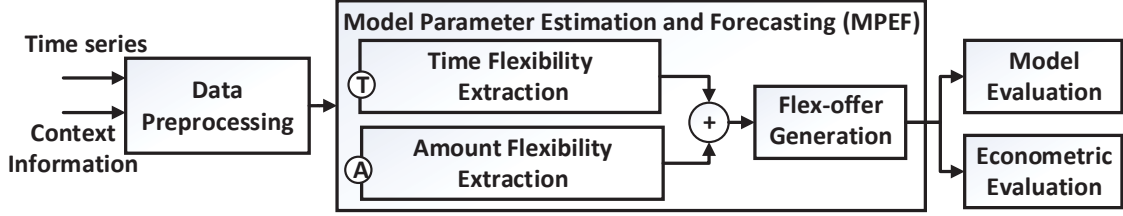
Figure 7.1: The General FO Generation and Evaluation Process.

Figure 7.1 shows the general steps involved in the generation and evaluation of FOs. The FO generation process starts with the gathering of the energy demand time series and available context information such as the description of house occupants, house insulation parameters, etc. The next step includes the preprocessing of the raw information into a format required for analyzing and predicting timestamps and values for the actions captured by FOs. The *Model Parameter Estimation and Forecasting (MPEF)* step includes the application of forecast models required for the generation of FOs. Specifically, this step comprises of two main tasks. The first task is the *Time Flexibility Extraction* which involves the prediction of timestamps of various actions such as Switch-on, Activate, Consume action, etc. The second task is the *Amount Flexibility Extraction* which includes the prediction of the number of slices in FOs, i.e., the number of Consume actions required for the completion of the defined task and the minimum and maximum energy bounds for the individual slices. Finally, this step combines the outputs of the two sub-steps to generate FOs for the forecasted device operations.

The last two steps evaluate the accuracy and the econometric benefit of the generated FOs. The accuracy provides the correctness of the predicted time and amount flexibility of the proposed FO generation process, and econometric benefit provides the financial viability of the generated FOs, e.g., in the regulation and spot market. The general process, shown in Fig. 7.1, is suitable for generating and evaluating FOs of different types of devices, including the aforementioned wet-devices, EVs, and HPs. The next Sections 7.2 and 7.3 present the key inputs and implementation of all the steps in the FOGEP in the case of wet-devices, EVs, HPs.

7.2 Flexibility Extraction

This section details our implementation of the MPEF step focusing on the *Time Flexibility Extraction* and *Amount Flexibility Extraction* sub-steps. The sub-steps rely on a number of forecast models required to predict the timestamp

and values captured by FOs. Hence, this section first discusses the feature set extracted from the preprocessed input data and presents our forecast models. Henceforth, it discusses the specific use of the forecast models and algorithms in the case of wet-devices, EVs, and HPs.

7.2.1 Input Data

In all cases, FOGEP requires energy demand time series and relevant context information (if available) as an input to perform various tasks such as model parameters estimation, training of forecast models, etc. The time series consist of a device energy consumption profile for $k-1$ days which will be used to detect the flexibility of the device on day k. The context information includes other information such as house area, family size, age, etc. The FOGEP also requires weather information for FOs generation for the temperature dependent devices such as HPs. FOGEP utilizes the first 80% of the time series to build various models required for FO generation process and the remaining 20% as the test set to evaluate the performance of the process. Though the proposed FOGEP works for time series of any resolution, this thesis focuses on an hourly resolution. Thus, let us define a time series $X = \{e_1, e_2, \ldots, e_T\}$ of device energy consumption profile for $k-1$ days, where $T \in \mathbb{N}_{>0} = (k-1) * 24$.

7.2.2 Forecast Models and Feature

This subsection presents the implementations of the general forecast models required in the *Time Flexibility Extraction* and *Amount Flexibility Extraction* sub-steps. As the base models, we use Logistic regression model along with the respective features sets (discussed in Section 6.1) and Pattern Sequence Matching model discussed below. The use of the models in the case of different device types is discussed in the next subsection.

Pattern Sequence Matching (PSM): Pattern Sequence Matching (PSM) is used to predict values for various attributes of FOs, e.g., the number of slices, energy profile, etc. required during the MPEF step. The device activation causes a noticeable increase in power consumption. First, all the changes in consumption values in the historical time series X are detected and are transformed into energy consumption patterns. The PSM algorithm (Algo. 3) works under the premise that these patterns are correlated to the time of activation, e.g., a dishwasher activated at 20:00 always operate for two-time units and has an average energy profile of $\langle 1.2, 1 \rangle$ kWh. Therefore, to estimate the energy profile for a predicted device activation at hour h of day k, the PSM first searches device activations triggered at hour h in the time series X. Then, for each activation the algorithm extracts the energy demand e_t for the duration of the device operation. This search outputs a set of indices of the device activation timestamps and profiles $P = \langle p_1, \ldots, p_n \rangle$, where each p_i is an energy profile of the device activation at the timestamp i and n is the number of device activations at the hour h. This algorithm returns an array of energy profiles

for matching device activations, and the specific use of the output profiles are discussed in the next subsections.

Algorithm 3 Pattern Sequence Matching (PSM)

Input: $= X$ - $\{e_1, \ldots, e_T\}$ a time series.
         h - a predicted device activation hour.
Output: $= P$ - a list of demand patterns starting at h.
        I - a list of index for the patterns.

1: **function** DEMANDPATTERN(X, h)
2:     $P \leftarrow \emptyset; p \leftarrow \emptyset; I \leftarrow \emptyset; active \leftarrow false$
3:     **for** $t \leftarrow 1 : T$ **do**
4:        **if** $e_t \geq thres$ **then**
5:           **if** $t\%24 = h$ **then**
6:              $p \leftarrow p \cup \{e_t\}$; $active \leftarrow true$; $I \leftarrow I \cup \{t\}$
7:           **else if** $active = true$ **then**
8:              $p \leftarrow p \cup \{e_t\}$
9:        **else**
10:          **if** $active = true$ **then**
11:             $P \leftarrow P \cup \{p\}$
12:          $p \leftarrow \emptyset$; $active \leftarrow false$
13:     Return P, I

The next subsection details the *Time Flexibility Extraction* and *Amount Flexibility Extraction* sub-steps in the case of different device types that utilizes the forecast models.

7.2.3   Device Type Specific Case of MPEF Step

This subsection presents the algorithms and the use of the forecast models to realize the *Time Flexibility Extraction* and *Amount Flexibility Extraction* sub-steps in the case of wet-devices, EVs, and HPs. Let us consider the forecasted device activation hour to be h and f to represent the final generated FO.
Wet-device: In the case of wet-devices, the LR algorithm is used to predict
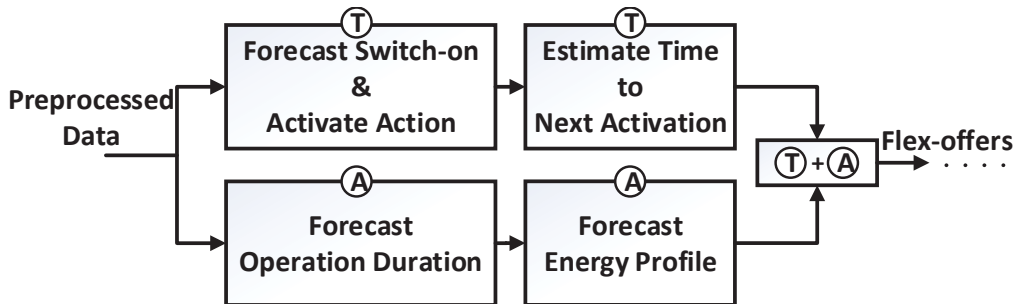

Figure 7.2: FO Generation sub-steps in the wet-device case.

Switch-on and *Activate* actions on the next day k (see Fig 7.2), utilizing wet-devices specific feature set.

Users exhibit some temporal behavior in their device activation patterns [56,90], i.e., they have specific device usage patterns during weekdays, weekend, or in some other temporal resolutions. Therefore, on day k, the user is assumed to follow a pattern similar to the patterns in the historical data X. Thus, after the prediction of *Switch-on* (using the LR algorithm), the output of the PSM (Algo. 3) is used to estimate the time gap between the subsequent *Switch-on* action, i.e., the next Switch-on action after the forecasted one. The time gap is calculated by averaging the gaps between the historical device *Switch-on* actions at the hour h and its subsequent *Switch-on* actions, as shown in Algo. 4. This time gap gives the EST (t_{es}) and LET (t_{le}). For example, if a device is predicted to be activated at 16:00 PM and all the historical device activations at hour 16:00 is followed by a subsequent activation at hour 23:00. Then, the average time gap is calculated as 7 hours and (t_{es}) and (t_{le}) is set to 16:00 and 23:00, respectively.

Algorithm 4 Estimation of Time Flexibility

Input: $= I$ - indices for matching patterns from Algo. 3.
  X - a time series.
  h - a predicted device activation hour.
Output: $= [t_{es}, t_{le}]$ - the earliest start and latest end time.

1: **function** ESTIMATETIMEFLEXIBILITY(I, X, h)
2:     $tf(f) \leftarrow 0; n \leftarrow length\ of\ I$
3:     **for** $i \leftarrow 1 : n$ **do**
4:         $idx \leftarrow I[i] + 1$
5:         **for** $t \leftarrow idx : T$ **do**
6:             **if** $X[t] > 0$ **then**
7:                 $tf(f) \leftarrow tf(f) + (t - idx)$
8:                 **break**
9:     $t_{es} \leftarrow h; t_{le} \leftarrow h + \frac{tf(f)}{n}$
10:     Return $[t_{es}, t_{le}]$

To estimate the energy profile for the forecasted activation, we search the historical time series X to extract profiles for all previous activations of the device at hour h, using PSM. Next, the operation duration d, i.e., number of slices for the forecasted device activation is estimated as the ceiling of the average number of slices in the historical profiles P, $d = \left\lceil \frac{1}{n} \sum_{i=1}^{n} length\ of\ p_i \right\rceil$, where $p_i \in P$. Then, the $[e_{min}, e_{max}]$ energy bound for each slice in the profile p is estimated as described in Algo. 5.

Electric Vehicles: Similar to wet-devices, LR algorithm is used to predict *Switch-on* and *Activate* actions for an EV on the next day k (see Fig 7.3) utilizing the EV-related feature set. The predicted *Switch-on* action represents the plug-in event of an EV.

Algorithm 5 Estimation of Energy Profile for a wet-device

Input: $= P$ - the extracted demand patterns from Algo. 3.

      d - an operation duration.

Output: $= p$ - an energy profile for forecasted activation.

1: **function** ESTIMATEPROFILE-WET(P, d)
2:     $p \leftarrow \emptyset; n \leftarrow length\ of\ p$
3:     **for** $j \leftarrow 1 : d$ **do**
4:         $e_{(min,j)} \leftarrow \frac{1}{n}\sum_{i=1}^{n} p_i.e_{(min,j)}$
5:         $e_{(max,j)} \leftarrow \frac{1}{n}\sum_{i=1}^{n} p_i.e_{(max,j)}$
6:         $s_j \leftarrow [e_{(min,j)}, e_{(max,j)}]$
7:         $p \leftarrow p \cup \{s_j\}$
8:     Return p
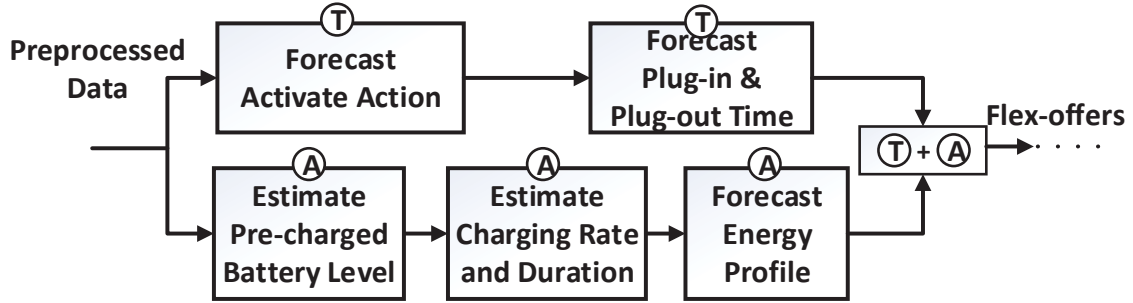
Figure 7.3: FO Generation sub-steps in the EV case.

The charging event of an EV is enclosed between the plug-in (t_{es}) and plug-out (t_{le}) events, i.e., between connection and disconnection to the charging port. The plug-in and plug-out events usually occur with some time gap from the actual charging event. For example, a user plug-in the EV at 5 PM after returning from work, the actual charging starts at 8.00 PM and ends at 12.00 AM, and the user plugs-out the car at 7.00 AM in the morning to drive to work. The time distance between these two events provides the time bound for an EC to trigger the *Activate* action and start charging, i.e., time flexibility $tf(f)$ for the generated FO. However, if the EV is charged using a quick charging station, these events may coincide with first and the last *Consume* action, i.e., $tf(f) = 0$. These events are predicted using the LR algorithm discussed in Section 6.1 with the dependent binary variable Y representing the plug-in and plug-out events, respectively.

The next step includes the estimation of the total (base and flexible) energy demand for the predicted EV charging event. It is essential to estimate the initial battery level of the EV before proceeding with the calculation of total energy demand. The initial battery level depends on the total distance (km) traveled, i.e., driven by a user on the day $k - 1$. In the absence of actual driving information, the initial battery level c_{init} is estimated by averaging the initial level for all historical charging events starting at hour h in the data

---

Algorithm 6 Estimation of Energy Profile for an EV

Input: $= e_b, e_f$ - a total base and flexible demand.
   ϕ - an energy distribution ratio.
Output: $= p$ - an energy profile for forecasted activation.

1: **function** ESTIMATEPROFILE-EV(e_b, e_f, ϕ)
2:     $p \leftarrow \emptyset$
3:     **for** $j \leftarrow 1 : d$ **do**
4:         $e_{min} \leftarrow \phi_j \cdot e_b$
5:         $e_{max} \leftarrow e_{min} + \phi_j \cdot e_f$
6:         $s_j \leftarrow [e_{min}, e_{max}]$
7:         $p \leftarrow p \cup \{s_j\}$
8:     Return p

X. Further, depending on the driving requirement, the user may fully or partially charge the EV. Let c_{min} be the minimum charge level required by the user. Then, the base energy demand for the EV charging is calculated as $e_b = (c_{min} - c_{init}) * r$, where r is the charging rate, i.e., the energy consumed per unit increase in charge level. Similarly, the energy demand flexibility $af(f)$ is calculated as $e_f = (100 - c_{min}) * r$. To distribute the base and flexible energy demand into e_{min} and e_{max} bounds for the individual slices of FOs, first, the charging duration d is estimated using a similar procedure as for wet-devices. Then, the energy distribution ratio $\phi = (\phi_1, \ldots, \phi_d)$, i.e., the portion of total energy consumed at each unit time of charging, is calculated by averaging the energy distribution in the profiles from Algo. 3, where $\phi = \sum_{i=1}^{d} \phi_i = 1$. The Algo. 6 calculates the profile p for the predicted EV charging.

Heat Pump: For a heat pump, the energy consumed at hour 1 effects the energy allowed to be consumed at hour ≥ 2, i.e., the size of slice s_t depends on energy consumed for slices $s_1, \ldots, s_{t-1}$. The actual amount of energy assigned at a particular timestamp is unavailable until the final schedule that creates an infinite combination of feasible energy (flexibility) allocations over the day (24 hours). This dependency problem has been explained and addressed in [91] using a Dependency-based Flex-offer (DFO) where an FO is represented as a series of two-dimensional polyhedrons, instead of just minimum and maximum energy bounds. However, the DFO is more complex and more challenging to deal with (compared to FOs considered in this thesis), which is often undesirable and thus not considered in this thesis. Thus, this thesis presents two different flexibility allocation approaches, discussed in detail later.

The first sub-step specific to the MPEF for a HP is the estimation of the model parameters of a house (see Fig 7.4). This thesis considers the flexibility of (air source) HPs for domestic heating use. Thus, we discuss a simple linear time-invariant (LTI) state space model of the house to estimate the internal temperature of the house given the input energy from the HP. The internal temperature of the house is dependent on the ambient temperature and the

thermal energy supplied by the HP. Further, the amount of thermal energy required to maintain the internal temperature is determined by the heat retaining capacity of the house. The first order LTI model of house is given by

$$\theta_0^{in} = \theta_0$$
$$For\ t = 1 : T$$

$$\theta_t^{in} = (1 - \frac{1}{R \cdot C}) \cdot \theta_{t-1}^{in} + (\frac{1}{C}) \cdot u_t + (\frac{\theta_t^{amb}}{R \cdot C})$$
$$e_t = \frac{u_t}{\eta}$$

where C and R are the thermal capacitance and thermal resistance of a house, respectively. Further, η is the Coefficient of Performance (COP) of the HP, θ_0^{in} is the initial room temperature, $\theta_1^{in}, \theta_2^{in}, \dots, \theta_T^{in}$ are the room temperatures at time intervals $1, \dots, T, \theta_1^{amb}, \theta_2^{amb}, \dots, \theta_T^{amb}$ are the ambient temperature at $t = 1 \dots T$, $u_1, u_2, \dots, u_T$ are the thermal delivery of the HP at $t = 1 \dots T$, and $e_1, e_2, \dots, e_T$ are the consumed energy amounts at $t = 1 \dots T$.

The optimized parameterization of the LTI model is obtained using the curve fitting technique, where the best numerical value for the parameters is estimated fitting a polynomial line to the consumption time series X. Table 3.2 in Section 3.2 depicts the parameter values for HPs and contextual information for some of the houses used in the experiments.
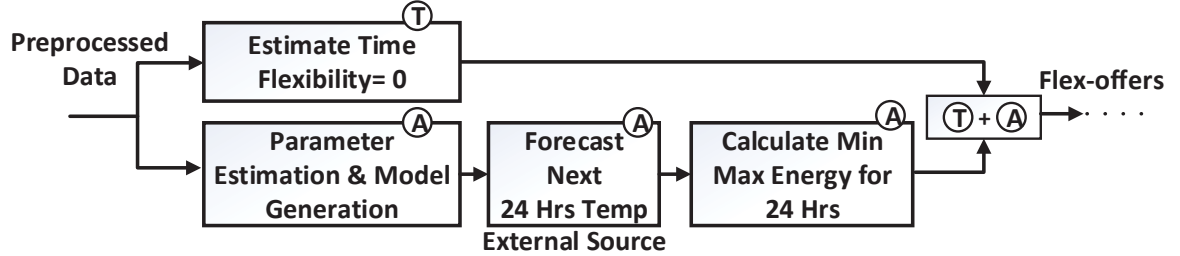
Figure 7.4: FO Generation sub-steps in the HP case.

To estimate the ambient temperature for the next 24 hours of the day k, the freely available weather forecast from [92] is used. Let us consider the *Switch-on* action for the HP operation is triggered at the first hour of the day, i.e., h=1 and f be the final generated FO. The air source HP has to operate continuously to maintain the indoor temperature, providing zero time flexibility (See Fig. 7.4). Thus, the *Switch-on* action and the *Activate* action always occur simultaneously, i.e., $t_{es} = t_{ls} = 12 : 00AM$ and an FO has an energy profile of 24 slices, i.e., $p = \langle s_1, \dots, s_{24} \rangle$. As discussed earlier, an HP has to maintain the indoor temperature within the comfort constraint defined by the minimum θ_{min} and maximum θ_{max} temperature bound. This comfort constraint provides substantial flexibility in energy required for each hour of the day k, i.e., e_{min} and e_{max} bound for the individual slices of the profile p. The constraint is estimated by averaging the temperature range across all

houses in the dataset. The temperature variation within a household support
our concept of flexibility; where a user is already comfortable and accept certain
degree changes in the indoor temperature. Recall, e_{min} and e_{max} for a slice
depend on previous slices. Thus, we present two different flexibility allocation
approaches:

1) Constraining Amount Flexibility (CAF): This approach imple-
ments the simple strategy that if you consume more for slice s_i then you have
to consume less for slice s_{i+1}, and vice versa. To avoid a situation where select-
ing a value from the given flexibility range violates the user comfort boundaries,
the amount flexibility of slice s_{i+1} is reduced so that any allocation of energy at
$t = 1$ in the range $[e_{(min,1)}, e_{(max,1)}]$ both fulfills the comfort constraints and
allows energy allocation at $t = 2$ be in the range $[e_{(min,2)}, e_{(max,2)}]$. Although
this approach favors the flexibility at the beginning and penalizes later periods,
it guarantees that the user comfort limit is always satisfied.

2) Weight-based Flexibility Distribution (WFD): There might be
cases where users can accept a few violations of the comfort constraints. Hence,
instead of penalizing the flexibility to satisfy the user constraints, this approach
supports assigning weights for each hour representing the amount of flexibility
desired at a particular hour. There are different ways to calculate weights,
and one way is using the spot prices, i.e., the hour with the lowest price gets
the lowest weight and vice versa. Compared to CAF, this approach gives
more control to the user or Balance Responsible Party (BRP) in deciding the
distribution of flexibility.

Moreover, the house can have zero occupancy for some hours in a day, the
so-called *relax* period (usually dictated by the user). During the *relax* period,
the HP can ease the comfort constraint, i.e., maintain the temperature above
or below the constraint boundary. For example, if the user has a comfort limit
of 20-23 °C, then the range could be extended to 18-25 °C. However, the new
limit depends on the capacity of the HP and the heat loss rate of the house.
The temperature should rebound to the normal limit at the end of the *relax*
period. The slices for the *relax* period has extended e_{min} and e_{max} bounds that
increases the amount flexibility $af(f)$. The profile p for the HP is calculated
using Algo. 7, where w_j is the weight assigned to timestamp j and $e_{(sch,j)}$ is
energy scheduled for the timestamp.

Flex-offer Generation Finally, the last step of the MPEF combines the
output from the various device-specific algorithms to generate the FOs. The
generalized MPEF step of the FOGEP is summarized in Algo. 8.

7.3 Mathematical Formulation for FOGEP Evaluation

In this section, we present details of the implementation of the statistical and
financial evaluation steps of the FOGEP, the final steps of the process shown in
Fig. 4. The statistical evaluation measures the accuracy of the predicted time

---

Algorithm 7 Estimation of Energy Profile for HP

Input: $= R, C, \eta, \ldots$ - LTI model parameters; w - weights
    $\theta_{min}, \theta_{max}$ - a comfort constraint.
    $\theta_0^{in}, \theta^{amb}$ - an initial indoor and ambient temperatures.
Output: $= p$ - energy profile for forecasted activation.

  1: **function** ESTIMATEPROFILE-HP$(R, C, \eta, \ldots)$
  2:      $\alpha \leftarrow 1 - \frac{1}{R \cdot C}, \beta \leftarrow \frac{1}{C}, \gamma \leftarrow \frac{1}{R \cdot C}$
  3:      $p \leftarrow \emptyset$
  4:      **for** $j \leftarrow 1 : 24$ **do**
  5:          $e_{min} \leftarrow \frac{\theta_{min} - (\alpha \cdot \theta_{j-1}^{in}) - (\gamma \cdot \theta_j^{amb})}{\beta \cdot \eta}$
  6:          $e_{max} \leftarrow \frac{\theta_{max} - (\alpha \cdot \theta_{j-1}^{in}) - (\gamma \cdot \theta_j^{amb})}{\beta \cdot \eta}$
  7:          **if** CAF **then**
  8:             $\theta_j^{in} \leftarrow \alpha \cdot \theta_{j-1}^{in} + \beta \dot{e}_{min} \cdot \eta + \gamma \cdot \theta_j^{amb}$
  9:          **else**
 10:             $e_{(sch,j)} \leftarrow (e_{min} + (1 - w_j) \cdot (e_{max} - e_{min})$
 11:             $\theta_j^{in} \leftarrow \alpha \cdot \theta_{j-1}^{in} + \beta \cdot e_{(sch,i)} \cdot \eta + \gamma \cdot \theta_j^{amb}$
 12:          $s_j \leftarrow [e_{min}, e_{max}]$
 13:          $p \leftarrow p \cup \{s_j\}$
 14:      Return p

---

and amount flexibilities, and the financial evaluation quantifies the viability of the generated FOs in the regulation and spot markets.

7.3.1   Statistical Evaluation

The generated FOs are compared with the actual available time and amount flexibility on the test dataset. The actual time flexibility $\overline{tf(f)}$ is calculated as the distance between the consecutive device operations on the day k (or k and $k+1$). Though the consecutive activations might not necessarily represent the actual time flexibility, e.g., a dishwasher can be opened before next activation; it provides the best estimation of flexibility. Further, the test dataset does not contain information about the amount flexibility. Thus, the accuracy of the FOGEP is estimated as deviations between the sum of forecasted e_{max} values and the corresponding energy demands in the test set (X) for the flexibility range, i.e., from t_{es} to t_{le}. We evaluate the time flexibility accuracy of the proposed FOGEP using Mean Absolute Error (MAE) where:

- MAE for time flexibility is the average over-estimation of flexible hours:

$$MAE = \frac{1}{n} \sum_{i=1}^{n} (\overline{tf(f_i)} - tf(f_i)) \cdot 1_{tf(f_i) \geq \overline{tf(f_i)}} \tag{7.1}$$

- MAE for amount flexibility is the average difference between actual and

---

Algorithm 8 Flex-offer Generation

Input: = $type$ - device type.

　　　　X - a time series.

Output: = f - flex-offer.

1: **function** GENERATE FLEX-OFFER($X, type$)
2: 　　**if** $type =$ 'wet-device' or 'EV' **then**
3: 　　　　$h \leftarrow$ predicted device Activation hour (using LR)
4: 　　　　$P, I \leftarrow$ DEMANDPATTERN(X, h)
5: 　　　　$d \leftarrow \lceil \frac{1}{n} \sum_{i=1}^{n} length\ of\ p_i \rceil,\ p_i \in P$
6: 　　　　**if** $type =$ 'wet-device' **then**
7: 　　　　　　$[t_{es}, t_{le}] \leftarrow$ ESTIMATETIMEFLEXIBILITY(I, X, h)
8: 　　　　　　$p \leftarrow$ ESTIMATEPROFILE-WET(P, d)
9: 　　　　**else**
10: 　　　　　　$t_{es} \leftarrow$ predicted Plug-in time (using LR)
11: 　　　　　　$t_{le} \leftarrow$ predicted Plug-out action (using LR)
12: 　　　　　　$c_{init} \leftarrow$ pre-charged level, $r \leftarrow$ charging rate
13: 　　　　　　$c_{min} \leftarrow$ desired minimum charge level
14: 　　　　　　$[e_b, e_f] \leftarrow [(c_{min} - c_{init}) * r, (100 - c_{min}) * r]$
15: 　　　　　　$p \leftarrow$ ESTIMATEPROFILE-EV(e_b, e_f, ϕ)
16: 　　**else if** $type =$ 'HP' **then**
17: 　　　　$R, C, \eta \leftarrow$ estimate HP parameters (curve fitting)
18: 　　　　$[t_{es}, t_{le}] \leftarrow [1, 24]; d \leftarrow 24$
19: 　　　　$p \leftarrow$ ESTIMATEPROFILE-HP($R, C, \eta, \dots$)
20: 　　$t_{ls} \leftarrow t_{le} - d;\ f \leftarrow ([t_{es}, t_{ls}], p)$
21: 　　Return f

---

forecasted total energy demand:

$$MAE = \frac{1}{n} \sum_{i=1}^{n} \left(\sum_{t=t_{es}}^{t_{le}} e_{(max,t)} - \sum_{t=t_{es}}^{t_{le}} e_t \right) \quad (7.2)$$

where $e_t \in X$ is the actual consumption per time unit. Further, we would like to emphasize that the proposed FOGEP is not conservative and does not under estimates the time flexibility to achieve higher accuracy.

7.3.2 Financial Evaluation

A BRP could utilize the flex-offers to compensate their deviation in the spot market and a financial gain by avoiding the regulation market. Similarly, an optimal scheduling of the flex-offers' Consume actions can achieve a cost-efficient consumption profile for the household. Thus, FOGEP evaluates financial benefits that can be achieved by i) a BRP with a cost-effective scheduling of the FOs in the spot market and ii) a BRP with optimal scheduling of FOs to avoid the regulation market.

Spot Market Optimization: The financial benefit in the spot market is the reduction in total energy cost that can be achieved by shifting the consumptions into hours of low energy prices. The key idea is to schedule the slices of the predicted flex-offer(-s) (for a device) for the following day based on spot prices such that the total cost of operating the device is minimized. The possible shift and variation in the value of a consume action are defined by the time and amount flexibility of the flex-offer. The bidding on the Danish spot market closes at 12:00 AM on the day-ahead and the final spot prices for the next day are broadcasted already at 12:45 PM. Hence, the FOs generation and scheduling is done starting 01:00 PM until 11:59 PM on the next day. Let $S = [s(t_{es}), \dots, s(t_{le})]$ represents the hourly spot prices between EST and LET of an FO slices. The overall objective of FO scheduling is to minimize the total cost of purchasing the energy required for the device operation which is given by $C(f) = \sum_{i=1}^{d} e_{(sch,i)} \cdot s(t + i - 1)$ where $e_{(sch,i)}$ is the energy scheduled for slice s_i, $s(t+i-1)$ is the spot price, and $t \in [t_{es}, t_{ls}]$ is the scheduled timestamp of the Activate action. The optimization problem is formulated as:

$$\text{minimize} \quad C(f) = \sum_{i=1}^{d} e_{(sch,i)} * s(t + i - 1)$$

$$\text{subject to} \quad e_{(sch,i)} \geq e_{(min,i)}, e_{(sch,i)} \leq e_{(max,i)}$$
$$t \geq t_{es}, t \leq t_{ls}$$

The financial benefit of an FO in the spot market is given by: $\Delta s = C - C(f)$ where $C = \sum_{t=h}^{t+d} e_t \cdot s(t)$ is the cost of energy for the default operation schedule, i.e., without any flexibility.

Regulating Market Optimization The regulating market is a part of the intra-day market, which is activated if a market is anticipated to have any imbalance in supply and demand, i.e., a BRP deviates from its previous commitment to the spot market. The deviation of the portfolio is handled by trading an equivalent amount of energy in the regulating market, and the BRP is responsible for the financial implication caused due to the differences in the spot and regulating energy prices. The financial benefit that a BRP can generate by utilizing FOs is the reduction in the regulating cost by minimizing the energy volume to be traded in the regulating market. The key idea is to schedule the slices of the predicted flex-offers (for a device) based on the anticipated deviation in the portfolio such that the total energy traded in the regulating market is minimized. The regulating power prices are not available at the time of the flex-offer generation. Therefore, a historical dataset is used to estimate the regulating power prices using the hypothetical model proposed in [93].

Let, $V = \{v_{u/d}(t_{es}), \dots, v_{u/d}(t_{le})\}$ and $P = \{p_{u/d}(t_{es}), \dots, p_{u/d}(t_{le})\}$ represent regulation volumes and prices between EST and LET of FO slices, where $v_{u/d}(t)$ denotes the nonzero element of regulating volume, i.e., up- or down-regulation and $p_{u/d}(t)$ is the predicted up-regulating power price $p_u(t)$ in case
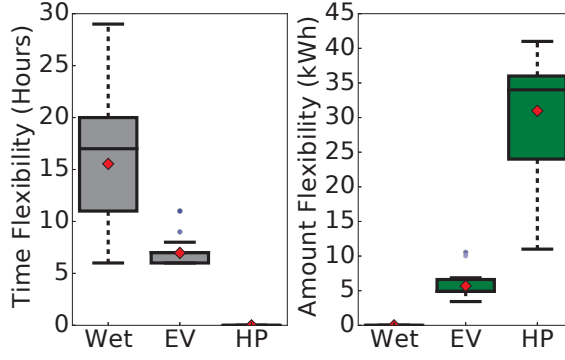
Figure 7.5: Average time and amount flexibility: over various device type.
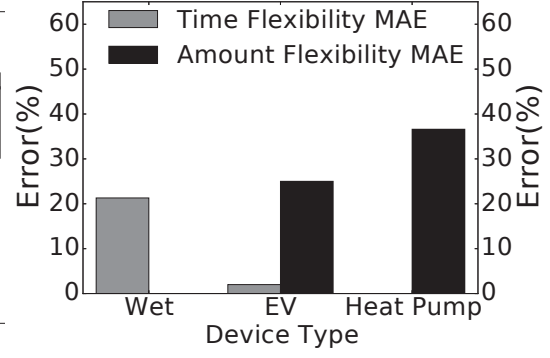
Figure 7.6: Time and amount flexibility error: over various device types.

of up-regulation and the predicted down-regulating power price $p_d(t)$ in case of down-regulation. Now, FOs scheduling task is to assign the FOs slices to V such that the BRP maximizes the total reduction in the regulation volume. Then, the optimization problem becomes:

$$\text{maximize} \quad \sum_{t=t_{ls}}^{t_{le}} |v_{u/d}(t)| - \overline{|v_{u/d}(t)|}$$

$$\text{subject to} \quad |v_u(t)| > e_{(sch,t)}, |v_d(t)| > e_{(sch,t)}$$

where the overbar in $\overline{|v_{u/d}(t)|}$ denotes the change in regulation volume due to time shifting of $e_{(sch,t)}$. Hence, the total financial benefit from an FO is given by:

$$E = \sum_{t=t_{es}}^{t_{le}} |v_{u/d}(t) - \overline{v_{u/d}}(t)| \cdot |p_{u/d}(t) - s(t)| \tag{7.3}$$

7.4 Experimental Analysis

We perform a number of experiments to analyze the effectiveness and efficiency of the proposed FOGEP, and evaluate the viability of the generated FOs in the spot and regulating markets based on real-world demand measurements. Hereby, we get a benchmark for the FO generation process and an experimental evaluation of markets for trading flexibility. The experiments uses the Zense-Home, INTrEPID, EVnetNL, and Heat Pumps dataset discussed in Section 3.2.1.

7.4.1 Evaluation of FOGEP Performance

The box plot from Figure 7.5 illustrates the distribution of time and amount flexibilities extracted using the proposed FOGEP on various device types. On
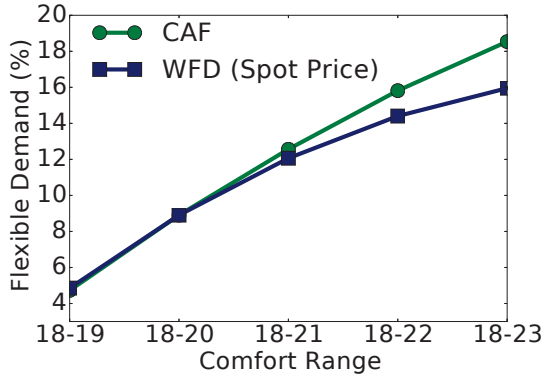
Figure 7.7: Flexible demand: over various comfort ranges during winter.
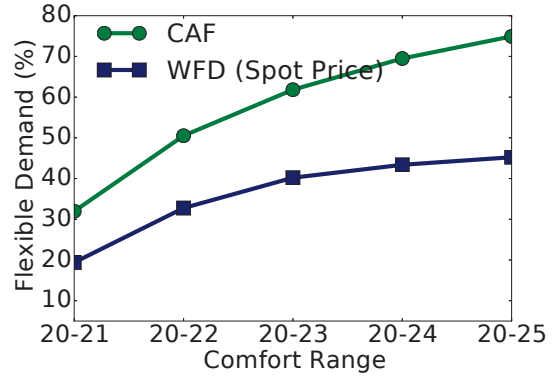

Figure 7.8: Flexible demand: over various comfort ranges during summer.

average, wet-devices provide 15.31 hours of time flexibility, 2 times the 7 hours average from EVs. On the other hand, HPs provide 33kWh of amount flexibility (daily average for 5th of January, for winter), 5.7 times greater than the 5.7 kWh from EVs. Accounting the continuous operation of the HPs throughout the day (24 slices), the average hourly flexibility drops to 0.92 kWh/h compared to 1.9 kWh/h for the EVs. The ranges of flexibility as shown by the box plot affirm that a significant flexibility can be extracted from individual device operations.

Figure 7.6 captures the accuracy of the proposed FOGEP on the extracted time and amount flexibility. The highly repetitive pattern in EV operations helps the proposed process to achieve 98% accuracy in the extracted time flexibilities. On the other hand, the FOGEP has a time flexibility error of 21% and amount flexibility error of 36% for wet-devices and HPs, respectively. The higher forecast error for wet-devices is mainly due to uncertainty in the usage patterns resulting in a significant number of falsely predicted Switch-on and Activate actions. Similarly, the error in the forecasted weather conditions and HP parameters estimation reduces the performance of the FOGEP for HPs. Thus, the accuracy could be significantly improved by incorporation of additional contextual information and using actual HP parameters. Nevertheless, the level of accuracy achieved by the proposed FOGEP is sufficient to increase the market confidence in utilizing FOs for balancing the deviations in the portfolio; this argument is further supported by the financial benefit of the FOGEP discussed in the next section. Further, the high degree of multidimensional (both time and amount) flexibility provided by EVs prove they are a major component of the flexibility-based DR.

The measurement data for HP includes the internal room temperature of households, which are set manually. Hence, here we assume a range of HP models, where a comfort range is set by a user and the HP automatically maintain the temperatures within the range. Figures 7.7 and 7.8 compares the average amount flexibility from 50 HPs across different flexibility distribution schemes and seasons at various comfort ranges. Specifically, we analyze the flexibility during the summer and winter seasons where the minimum comfort
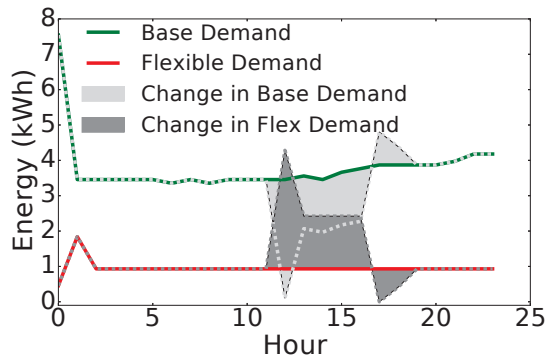
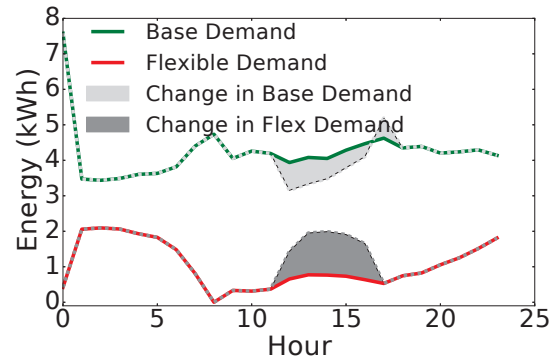**Figure 7.9:** Effect of relaxed hours: for CAF.

Figure 7.10: Effect of relaxed hours: for WFD.

boundary for the winter is set to 18 °C and maximum to 23 °C, increasing one °C at a time (estimated from the dataset). The temperature during the summer is usually high. Thus, we increase the minimum comfort boundary to 20 °C and maximum up to 25 °C. During the winter, on average 4.71% of total energy demand from HPs are flexible at 1°C comfort band and gradually increases to 18.5% at 5°C comfort band. On the other hand during the summer, even at 1°C comfort band 31.7% of total demand is flexible, which increases to 74% at 5°C.

A better flexibility during the summer is mainly attributed to the higher ambient temperature where an HP has to supply zero or very little energy to maintain the lower temperature bound. However, the absolute average amount flexibility during the summer is 20 kWh less than in winter. Comparing the flexibility distribution schemes, CAF gives 16% and 65% higher flexibility compared to WFD for winter and summer, respectively. The WFD distributes the energy demand based on the spot price, i.e., assign higher base demands to the points with a lower spot price, whereas CAF sets base demand to the minimum energy required to satisfy the lower temperature constraint. Thus, the assignment of higher base demand than required (to maintain the minimum comfort bound) reduces the amount flexibility for WFD. Further, the figure clearly shows that for both assignment schemes flexibility increases with the comfort range.

Figure 7.9 and 7.10 illustrate the effect of the relaxed hours on energy demands during the winter, for the respective optimization. The relaxed hours are set between 10:00 AM - 4:00 PM when users are usually away and the comfort range for the period is increased by 6 °C (3 °C on both directions). The experiments show that for each relaxed hour the amount flexibility increases by 6.9% and 4.3% for the CAF and WFD, respectively. The relaxed hours have better effect with CAF because the optimizer always tries to maintain the base demand to a minimum level, whereas WFD follows a weight based flexibility distribution.
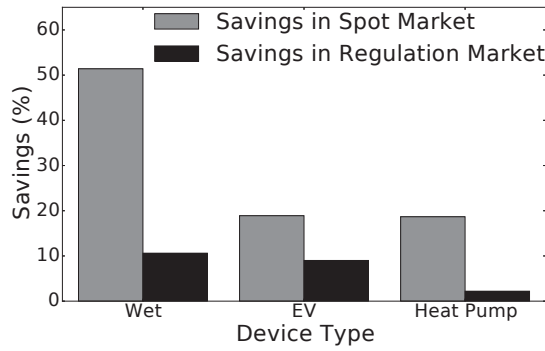
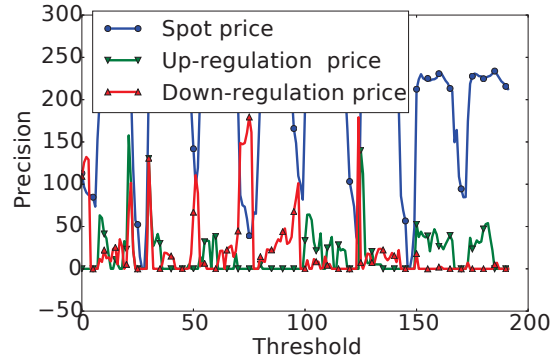**Figure 7.11:** Savings from flexibility: over various markets.

Figure 7.12: Energy price variation: for different regulation type.

7.4.2   Evaluation of FOGEP Utility

Here, we analyze the financial benefit that can be achieved by utilizing the extracted time and amount flexibilities and further compare the best market to trade the flexibility, i.e., spot market or regulating market.

A market player is always interested in knowing the market that gives best return for the available flexibility. Figure 7.11, compares the savings from flexibility at two different markets, i.e., spot and regulating market. Flexibility for all three device types generates better savings on the spot market compared to the regulating market. Specifically, for wet-devices and HPs the savings in the spot market are 5 and 8.4 times higher than the regulating market, respectively, whereas the savings is only 2.1 times higher for the EVs. The figure emphasizes that the flexibility has a greater impact on the spot market providing higher savings compared to the regulating market. As shown in Figure 7.12, a lower price variation/MWh in the regulating market decreases the potential benefit of flexibility in the market compared to the spot market. Further, the constraint (in Section 7.3.2) of shifting flexible demands only from up-regulated market to down-regulated market drastically reduces the number of flexible demands that are re-scheduled.

The savings in the regulating market also depend on the reduction in the price gap between the markets caused due to the shifting of flexible demand [40, 93]. The decrease in price gap from a small amount flexibility is not substantial enough to generate considerable savings, whereas larger amount flexibilities cause higher reductions making substantial savings. This behavior is demonstrated by EVs (Fig.7.5) where amount flexibilities are significant enough to generate comparable savings on both markets. However, even with a larger amount flexibilities, the unavailability of time flexibility makes FOs from HPs unsuitable for the regulating market.

Evaluating the savings in the Spot market, wet-devices give the highest relative savings of 51% compared to 18.9% and 19% for EVs and HPs, and the average daily absolute savings are 0.8 DKK, 4.5 DKK, and 3.85 DKK, respectively. Figure 7.13 demonstrates the distribution of savings into the
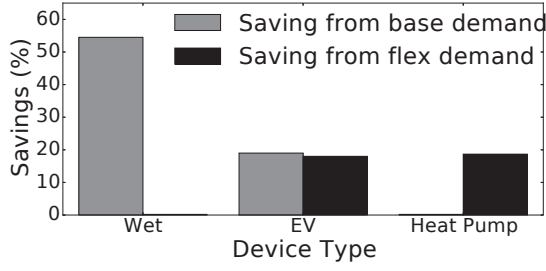
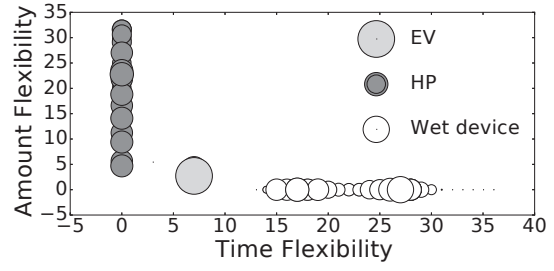**Figure 7.13:** Savings from base and flexible demand(winter, spot, CAF)

Figure 7.14: Distribution of savings over individual device operation.
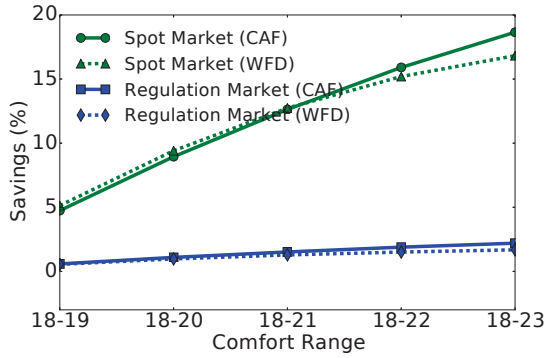
Figure 7.15: Savings from HP in winter: for different optimization.
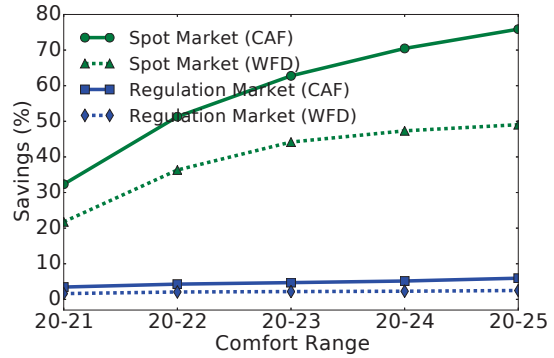
Figure 7.16: Savings from HP in summer: for different optimization.

base and flexible demand. The figure shows that all savings for wet-devices come from base demands. Similarly, for HPs entire savings come from flexible demands, whereas, the savings for EVs come from both demands. Figure 7.14 demonstrates the relationship between time flexibility, amount flexibility, and total savings. The x-axis represents the time flexibility; the y-axis represents the available amount flexibility, and size of the circle represents the total relative savings/kWh. Although the absolute savings/day from HPs is higher than the EVs and wet-devices, the saving/kWh of flexible demands are much lower for HPs compared to others. The higher saving for wet-devices and EVs is due to the availability of larger time flexibility, allowing a market to schedule the base demands to timestamps with lower spot prices. On the other hand, HPs provides no time flexibility generating no savings from the base demand (ref Fig. 7.5).

Figures 7.15 and 7.16 compare the savings during summer and winter for two different flexibility distribution schemes (CAF and WFD) and markets. The WFD is penalized for assigning higher base demand and maintaining a higher temperature to the point with lower energy prices resulting in lower savings compared to WFD at both markets. This behavior can be further inferred from the figures where the difference in savings increases with the comfort range, reaching 20% at 5 °C comfort band. Further, the temperatures during the summer are high, resulting in a larger percentage of total demand

Figure 7.17: Energy flexibility and spot price: over various season.

Figure 7.18: Effect of relaxed hours on savings: over various seasons.

being flexible (4 times more than in winter, see Figure 7.8). For HPs, relative percentage savings are higher during the summer, 75% in the spot market and 5% in the regulating market. However, average daily absolute savings is 0. 4DKK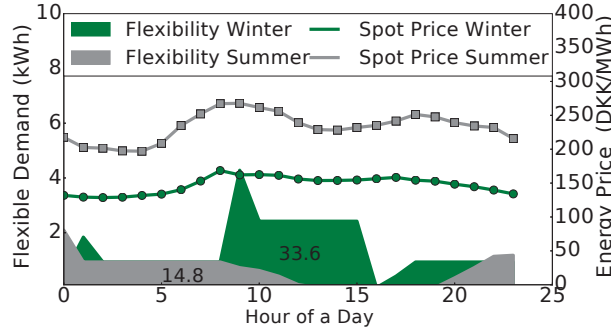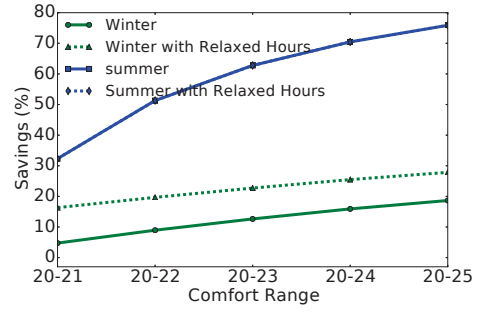 higher in winter then summer, 3.85 and 4.25 DKK in summer and winter respectively. Figure 7.18 illustrates the effect of relaxed hours during two different seasons using WFD. The energy demands during the winter are high even during the midday (office hours) and the relaxed hours increase the amount flexibility by 6.9% generating 1.5% increases in savings per relaxed hours. However, in summer the temperatures are usually high. Thus, the relaxed hours during the midday (office hours) have minimum effect on the savings, i.e., < 0.1% increase.

7.4.3   Analysis

The operating uncertainty for wet-devices gives relatively higher error for the FOGEP, whereas repetitive patterns for EVs result in a very high accuracy. The size and type of flexibilities are usually determined by device properties, operational objectives, settings, and other external factors. The results demonstrate that the preferred market for flexibility depends on its source and size. The device types with only one-dimensional flexibility (either time or amount) generates higher benefits in the spot market, whereas types with two-dimensional flexibilities (both time and amount) can generate significant savings in both spot and regulating markets. Specifically, the flexibility performs better in the regulating market when it has a right blend of time and amount flexibility (EVs). The device types with relatively lower energy demand, but with a substantial time flexibility (wet-devices), can have savings greater than devices with a larger amount flexibility (HPs). Therefore, the time flexibility shows a potential of generating higher savings, which is a valuable input to aggregators on deciding which dimension should be retained during aggregation of flex-offers.

The seasonal factor plays a significant role in determining the amount of the flexibility from HPs. HPs have a higher relative flexibility during the summer,

whereas the absolute flexibility is far better in winter. Further, HPs are not used for most of the time during the summer (as seen in the test dataset) and winters are usually longer in Nordic countries (the dataset). This behavior makes HPs one of the main components to consider for flexibility markets and energy optimization. Nevertheless, due to the temporal behavior of the spot price the average daily savings for both the seasons are almost equal. A higher flexibility with consistent savings makes CAF the best approach to handle the dependency between the flex-offer slices (from HPs).

7.5 Summary and Discussion

In this chapter, we presented an answer to the research question *"How can we effectively extract and generate flex-offers from household devices?"*. In this regards, we proposed a state-of-the-art flex-offer generation and evaluation process (FOGEP) for flexibility of household devices, namely wet-devices, EVs, and HPs. We have detailed various sub-steps involved in FOGEP and utilized two different forecast models and various algorithms required to predict the timestamps and values captured by FOs. We utilized real-world device-level time series to quantify the actual flexibility potential of devices and evaluated the performance of FOGEP. Further, we quantified the financial value of the flex-offers in the spot and regulation markets and identified the best market to trade flex-offers relative to flexibility type and size. The experimental results show that the proposed FOGEP is general enough to generate flex-offers from a variety of household electrical devices with a higher accuracy. The results showed that the proposed FOGEP could extract device-level flexibilities with up to 98% accuracy. Finally, we can conclude that *"we can extract and model flexibilities from various device types into a unified flex-offer format with good utility"*. In the next chapter, we focus on the user-comfort oriented scheduling of flex-offers.

Chapter 8

Scheduling of Flex-offers

How can we economically schedule flex-offers with minimal loss of user comfort?

In this chapter, we focus on the scheduling of the device-level flex-offers. Most of the previous solutions to flexible load scheduling (Discussed in Chapter 2) focus on the financial aspect of the load shifting potential of devices. We argue that the scheduling strategy should also consider the social aspect of a user of accepting the proposed schedule and should minimize the loss of user perceived quality induced by load shifting. In this regard, we propose a user-comfort oriented prescription technique to prescribe flex-offer schedule that maintains a balance between the economic and social aspect (user-perceived quality) of demand shifting. We present a novel Flexibility-aware Error (FAE) measure to evaluate the performance of prediction models in relation to demand scheduling. The proposed measure is evaluated with two demand prediction models and extensively compared with standard error measures. Finally, we quantify the financial potential of the demand shifting optimized on the spot and regulation markets, utilizing the proposed user-oriented scheduling technique.

8.1   Prescription of Flex-offer Schedules

In the previous chapters, we have described how flexibility can generate balance in the energy market. However, flexible demands need to be scheduled in a way that is accepted by the end-users. In the presented scenarios, customers' device usage preference information is often unavailable without surveys or direct knowledge. Nevertheless, the use of flex-offers as a modeling tool can provide insight on the user behavior. DR strategies for flexible demand scheduling of flexible devices involve multiple phases. Our proposal comprehends three main components: device-level load forecasting, flex-offer modeling, and flex-offer scheduling. The flex-offermodeling of the flexibility has already been discussed
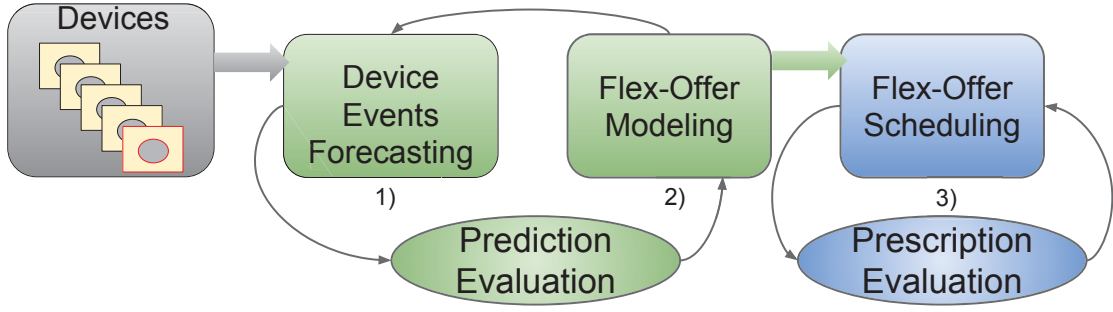
Figure 8.1: Prescriptive workflow of Flex-offer scheduling

in Section 3.1. Hence, in this chapter, we only discuss the demand forecast models with the focus on flex-offerscheduling and our scheduling strategy.

8.1.1   Prescriptive Strategy for Flex-offer Scheduling

Our proposed flex-offer scheduling can be described by the workflow shown in Fig. 8.1. The objective is to prescribe to the end-users a time schedule for the usage of their devices. Hence, our method first identifies the future energy needs from user's devices, model them as flex-offers, and then finds an optimal demand shifting schedule as a trade-off between the user's device usage preference and the savings in the energy cost and regulation volumes.

A preliminary step is to transform the data collected from the devices, a load consumption time series data describing the individual device usage, into an *event* time series. Events are the *activations* of a device, associated with the respective timestamps. To identify the device events we start by finding the device *operations*, regions of time in the time series, longer than a minimum threshold, with high deviation in load consumption. Operations are associated with a set of hours for the length of the operation, each corresponding to the respective energy load consumption. Finally, the starting timestamps of these operations are the activation event, from which we build the event time series.

Then, utilizing the event time series, we use forecasting models to predict the probability of device's activation events in the future (step 1 of Fig. 8.1). The result is a set of predicted activation timestamps for the device.

Third, we generate flex-offers utilizing the predicted events (step 2 of Fig. 8.1). In Fig. 8.2 we show an example of a flex-offer where two consecutive predicted activation events become the Earliest-Start-Time (EST) and Latest-Start-Time (LST) of the predicted flex-offerrespectively, and the interval between these two the time flexibility range. The flex-offer is also associated to a device *signature*, representing the device operation in terms of both duration and energy consumption for each hour of operation. A device signature is defined as:

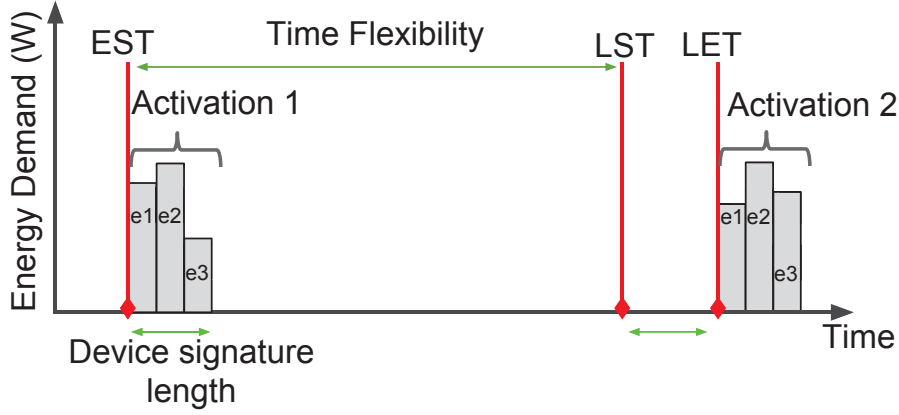$$\sigma = [e_1, e_2, \ldots, e_k] \tag{8.1}$$

Figure 8.2: Flex-offer model for an activation/deactivation event of a device

where e_i is the energy demand per time unit of the average operation length k. We simplify the device signature by averaging the operation duration and load consumption (at hour resolution in the thesis) over all the historical device operations, respective to the day of prediction.

The final phase consists of the prescription of flex-offers schedules (step 3 of Fig. 8.1). Considering, for example, a predicted flex-offer f_i, with $EST = 10 : 00$ and $LST = 24 : 00$, we have a 14 hours flexibility range for rescheduling the device activation event. The decision on the time to prescribe the activation to depends on both financial (reduction of energy price in spot and regulation markets) and user specific factors. In order to describe the prescriptive process, we start by defining the flex-offer schedule as follows:

Definition 2 (Flex-offer Activation Schedule). *The scheduling of a flex-offer is*

$$S : FO \to \mathbb{R}_{EST,LST}$$

a function that maps a flex-offer to a new timestamp for the device activation. The new timestamps is included between the time flexibility interval $[EST, LST]$ of the flex-offer.

Our current definition of flex-offer scheduling depends only on the flex-offer prediction. However, any proposed schedule has to be accepted by the device's owner. Direct information from the user on the device usage is often unavailable. Nevertheless, insight into the user behaviour is a main factor in the success of the flex-offer scheduling, e.g. users would not want a dishwasher to be scheduled before the dishes are ready to be cleaned, or after the dishes are needed again. Therefore, to enforce the dependency of the flex-offer scheduling on the end-users' preferences for the device usage, we introduce an assumption on which we base our scheduling approach:

User Assumption 1 (User Comfort). *Users tolerate their devices to be rescheduled in return for a financial benefit, as long as the scheduling does not restrict*

*their preferred behaviour. Their tolerance threshold to the device scheduling is
proportional to the financial benefits received, and inversely proportional to the
amount of shift of the scheduling.*

A user would, all else being equal, prefer to maintain the control over the
usage of the devices, i.e. to not reschedule the activation times. However,
according to our user comfort assumption, we can describe the scheduling pro-
cess as a trade-off between the financial benefits, i.e. the ability to schedule
device activations over a wide time flexibility interval, and the *user comfort
loss*, i.e. the distance between the user's preferred activation time and the new
scheduled time. In the next sections we present more details on the phases and
evaluation of the scheduling workflow.

8.2   Evaluation Metrics for Flex-offers Forecasting Models

In this section, we discuss the first phase of our proposed flex-offer scheduling
process in Fig. 8.1, the forecasting of device events. We describe two simple
forecast models and as well as our novel performance measure to evaluate the
models, specifically tuned for flex-offer scheduling.

8.2.1   Most Frequent Timestamp (MF)

The Most Frequent Timestamp (MF) model predicts n activation events in a
given time window at the n most frequent activation hours. n is the average
(rounded down) operating frequency of the device, i.e. the average number of
activations per time window (e.g. 24 hours) we observe in the historical usage
of the device.

For example, if Figure 8.3 represents the distribution of device activations
and the device has an average operating frequency of 2 per 24h, then the MF
model predicts 2 activations each day, the first at hour 9 and another at hour
21. In the cases, where two hours have same operating frequency, we select the
hour which is the farthest from the first selected activation. Similarly, if the
distances between the most frequent hours are less than the average operation
duration of the device, one activation is selected from those hours, and another
is selected from the next frequent operating hours.

8.2.2   ARIMA

We utilize a standard ARIMA [94] model on the event time series, for which the
parameters have been fitted by cross validation. The model, given a training
set, predicts the next activation event p to be the EST of a flex-offer . The
prediction p is then used as additional training point for predicting the LET of
the same flex-offer. Finally, p is replaced in the training set by the respective
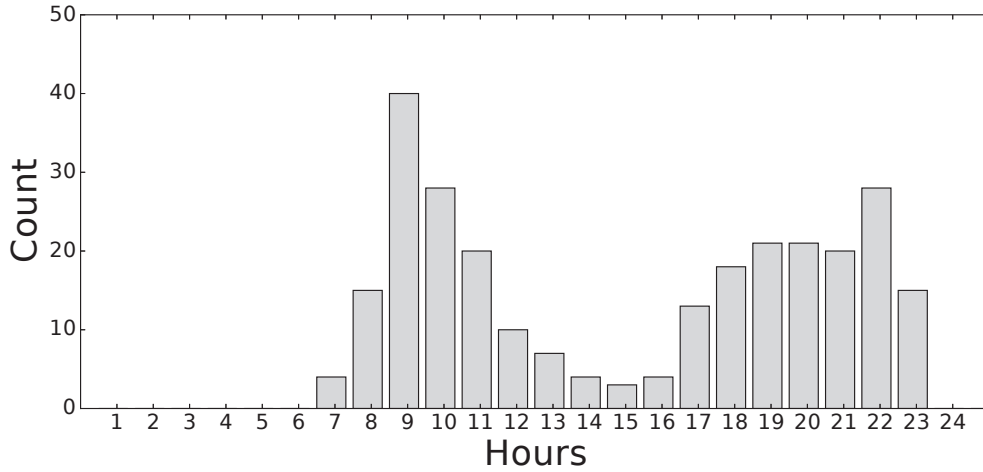
Figure 8.3: Distribution of device activations over hours

observation, i.e. the next event in the test set, and the model is used to predict a new flex-offer.

8.2.3   Flexibility-Aware Evaluation Metrics

The applicability of the predicted flex-offer models for demand scheduling is dependent on the probability that such flex-offers correctly capture the device usage. An intuitive approach for evaluating the predicted flex-offers is to measure the (time) deviation between *predicted* and *observed* device events. Traditional forecasting measures, such as RMSE, MAPE, precision, recall, etc., perform a point-wise comparison between the predicted and observed values. However, flex-offers present a more complex structure, that simple point-wise comparisons do not capture. Moreover, we would like to emphasize that although the correctness of the flex-offermodeling influences the quality of the demand scheduling process, device event prediction is only a middle step in the process, where the ultimate goal is to schedule the demands in a way that is both financially viable and acceptable by the user.

For example, an error of +2 hours in a predicted dishwasher activation event is largely acceptable, as long as the operation is scheduled before the flex-offer latest-start-time, i.e. next dishwasher usage. As long as this constraint is fulfilled, the schedule preserves the user behaviour. To capture this correlation between prediction error and user specific error, we present a *Flexibility-Aware Error* (FAE) measure of the performance of the forecast model with a focus on flex-offer scheduling. FAE is defined as follows

$$FAE(f, \overline{f}) = 1 - \frac{min(\overline{f_l} - \overline{f_e}, f_l - f_e)}{max(f_l, \overline{f_l}) - min(f_e, \overline{f_e})}, \tag{8.2}$$

where f and $\overline{f}$ are the observed and predicted flex-offers, and f_e, f_l the EST
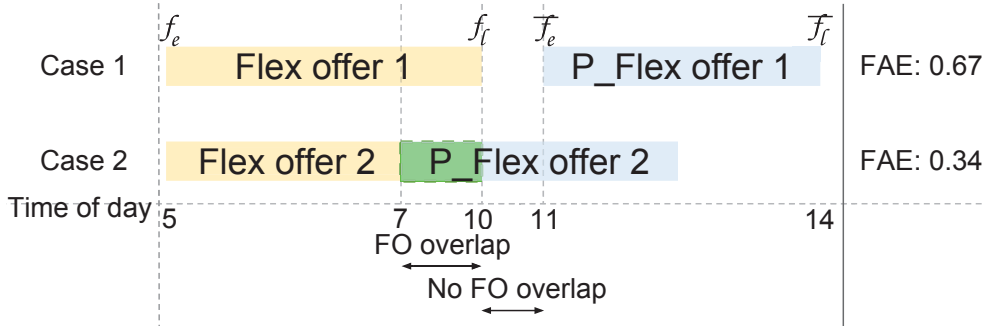
Figure 8.4: FAE for two different cases of actual and predicted flex-offers.

and LST, respectively. Intuitively, FAE describes the overlap between the time intervals $\langle EST, LET \rangle$ of the predicted and observed flex-offers. The larger the overlapping region between the two flex-offers intervals, the smaller is associated FAE. Further, the smaller FAE represents a better forecast regarding flex-offers scheduling. Alternatively, the farther away the two intervals are from each others, the higher the FAE. The nominator represents the smallest interval between the observed and predicted flex-offer flexibility intervals, while the denominator is the entire time region encompassing both intervals. If the predicted interval is on a different length scale than the observed one and/or with no overlapping region, the denominator becomes larger, with a resulting higher FAE.

For example, as shown in Figure 8.4, the comparison between an observed time interval $\langle 5, 10. \rangle$ and a one predicted $11, 14$ leads to the FAE of 0.67. The same observed interval compared to an overlapping predicted interval $\langle 7, 11 \rangle$ gives the lower FAE of 0.34. In Sec. 8.4 we will show how FAE compares to other traditional error measures for flex-offer scheduling. In the next section, we present the evaluation techniques for flex-offers scheduling.

8.3   Prescriptive Model For Flex-offer Scheduling

In this section, we describe the mathematical models to quantify the benefits that our flex-offer scheduling brings to BRPs and end-users, the last phase shown in Figure 8.1. The schedule for the predicted flex-offers is prescribed based on the analysis of three factors. The first and the second are the financial gain that can be achieved in the spot and the regulation markets respectively by scheduling the flex-offers. The third is the loss of user comfort due to the delayed activation imposed by the schedule. In these measures we reflect the trade-off between financial gain and user comfort and thus select a new device activation scheduling that optimizes the combination of these factors. The next sections describe these factors in detail.

8.3.1 Spot Market Savings

Spot market savings (SpMS) describes the total financial savings of energy demands and the corresponding flex-offers at the spot market for the predicted device activation. To maximize this factor, a flex-offer is scheduled such that the cost of purchasing the energy required for the device operation is minimum. Let $P = [p_s(est), \ldots, p_s(let)]$ represent the hourly spot prices between EST and LET. The energy cost for each timestamp of the device operation is calculated as the product of energy demand and the respective spot price given by

$$S = \sum_{i=0}^{k-1} e_i \cdot p_s(t_a + i), \tag{8.3}$$

where k is the duration of the device operation a starting at timestamp t_a and e_i is the energy demand for each operating time unit.

To calculate the spot market cost after the scheduling, we simply modify Eq. 8.3 with the new activation timestamp t_{new}

$$\bar{S} = \sum_{i=0}^{k-1} e_i \cdot p_s(t_{new} + i). \tag{8.4}$$

Therefore, the SpMS in spot market cost given by the activation schedule is given by

$$\Delta S = S - \overline{S}. \tag{8.5}$$

8.3.2 Regulation Market Savings

The mathematical model for regulation market savings has been discussed in Section 6.2.4, where the saving is given by

$$\Delta R = R - \overline{R}$$

where R is the actual regulation cost and $\overline{R}$ is the expected regulation cost after scheduling of flex-offer.

8.3.3 User Comfort

In Sec. 8.1.1 we have introduced our assumption of *user comfort*. Here, we describe how to utilize the device usage history to extract user information, and how do we use to define the user comfort.

User comfort can be described as the difference in time of activation between the original user preference for the device activation, and the new scheduled activation time. To cope with the absence of exact user behavior information, we start by looking at the *idle* time periods between device events. These *idle intervals* give us an, albeit partial, description of the device usage, letting us
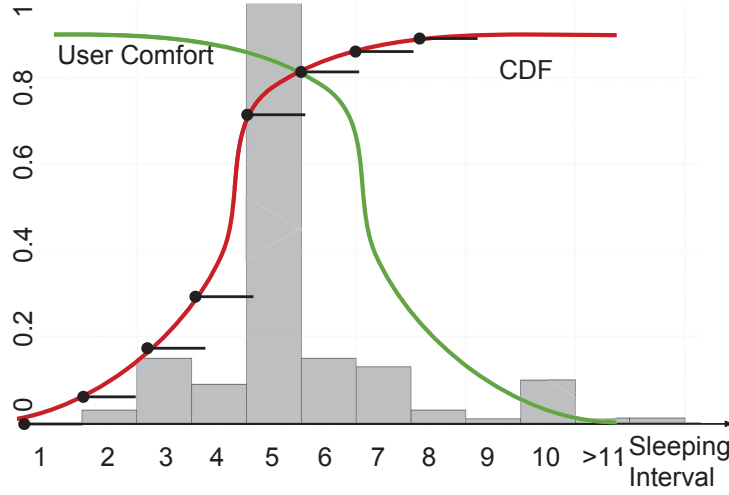
Figure 8.5: Relation between User Comfort and cumulative density function of unit's sleeping times

capture the user preference. Intuitively, a device activation scheduled extending over the average idle interval would result in an interference with the user activities, as it is more probable that the result of the device's operation (e.g. clean dishes in case of the dishwasher) is needed before that time.

Therefore, we define the user comfort of scheduling a device activation a from the original timestamp t to a new timestamp t_{new} as

$$U(t_{new}, t) = \begin{cases} 1 - \mathrm{cdf}(t_{new} - t), & \text{if} \quad t_{new} \geq t \\ 0, & \text{otherwise,} \end{cases} \qquad (8.6)$$

where *cdf* is the cumulative density function of the device's idle intervals. The user comfort is then inversely proportional to the distance $t_{new} - t$. We define the case in which $t_{new} < t$, i.e. the device operation has been scheduled before the observed activation time, as a scheduling fail, hence a user comfort value of 0. Fig. 8.5 shows an example of a histogram of idle intervals, and the associated cumulative density function. The user comfort is symmetric to the cumulative density function curve.

The performance of a flex-offer scheduling can now be evaluated as a combination of $\Delta S + \Delta R$ and preservation of user comfort. Fig. 8.6 shows an example of these three components. Currently, the device's operation starting at $4:00$ lies within a high spot/regulation market price time slot. To cope with this problem, we can utilize the flex-offer information to shift the device's activation to a lower energy price time, between EST and LET. However, the farther we shift the activation time, the lower the probability the user will accept such a scheduling, as the user comfort diminishes.

The scheduling process can therefore be seen as an optimization problem of opposite objectives: increasing flexibility to lower energy prices and to avoid the regulation market (this might lower user comfort), while maintaining a desirable
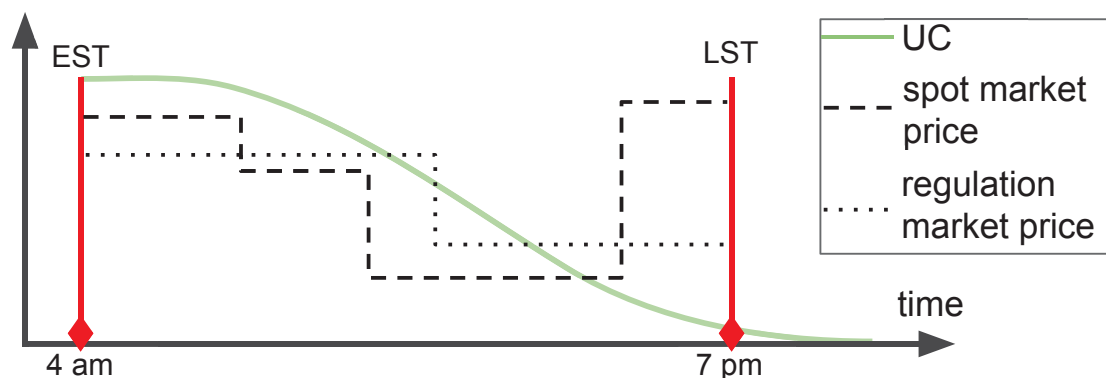
Figure 8.6: Example of the device scheduling as optimization of spot and regulation market price and user comfort.

comfort level for the user (this might lower flexibility). This combination can be defined as the scheduling Gain of a flex-offer f:

$$G(f) = U(t_{new}, t) \cdot (\Delta S + \Delta R). \tag{8.7}$$

After having found a schedule that optimizes these objectives, it can be submitted to the end-user. If assumed to be accepted, the actual improvement that results from the prescription of the flex-offer schedule can be simply quantified as the increase SpMS and ReMS due to the scheduling of the device demand, compared to the original price for such demand. A positive decrease will affect both BRPs and end-users, that share the benefits of lower energy prices.

In the next section we present the results of our experimental evaluations, where we analyze the applicability of our proposed techniques.

8.4 Experimental Analysis

We performed a number of experiments to analyse the effectiveness and financial impact on both spot and regulation markets of the proposed flex-offer prescription scheduling. First, we present a benchmark of the forecasting models we have described in Sec. 8.2. Then, we show the evaluation of the flex-offer scheduling process.

8.4.1 Dataset

We utilize device-level energy consumption and market datasets from INTREPID and Energinet.dk, respectively (discussed in Section 3.2.1).

	Average Shift	UC loss	SpMS	ReMS
ARIMA	12 hours	39%	18%	44%
MF	7.1 hours	25%	14%	43%

Table 8.1: Summary of the average demand shift (in hours), User Comfort (UC) loss, SpMS, and ReMS.

8.4.2 Prescription Scheduling Results

Table 8.1 shows a summary of the flex-offer scheduling results. In the table we show the average results for all the device datasets on which we have performed the evaluation. The learning phase is performed by applying *prequential evaluation* (or interleaved test-then-train evaluation), where each observed event in the time series is used first for the test set, and then for the training set. With both forecasting models it is possible to achieves positive decrease in the energy price, calculated as a percentage of the initial price. Specifically, the MF model leads to a percentage of SpMS and ReMS of 14% and 43% respectively, while keeping the loss of user comfort under 25%, resulting in an average demand shift of approximately 7 hours from the original user preferences.

To achieve these results we have utilized the scheduling gain defined in Eq. 8.7. The applicability and quality of a scheduling prescription can be evaluated at the forecasting time. In this evaluation we take an agnostic approach to the problem, utilizing distance aware measures (FAE and RMSE), as well as classification error measures (F1 score and accuracy), in order to evaluate which one can better describe the flex-offer scenario. We expect a negative correlation between FAE and RMSE, as these two measures capture prediction error, and positive for accuracy and F1 score. The measures are calculated for each flex-offer , comparing the predicted event timestamps (EST, LET) with the respective observed event timestamps. The results are then averaged over all the devices. In Table 8.2 we show a summary of the correlation between these measures and the scheduling gain. The scheduling gain is calculated as a percentage of the gain we achieve with the predicted flex-offers, compared to the gain we would achieve with fully observed flex-offers (obtainable if we had knowledge of the future device events, e.g. from the end-user). As we can see in the table, for both models our proposed FAE measure provides the *highest (negative) correlation* between the flex-offer forecasting and the flex-offer scheduling gain. With the ARIMA model, the predicted flex-offers have a correlation between the FAE measure and the scheduling gain of 78%. In both models, the classic accuracy measure does not provide any insight on the scheduling Gain, with a correlation of < 0.01. RMSE and F1 score perform poorly in both models, although being standard distance-aware measures in other scenarios.

In Fig. 8.7 we also show the correlation, for the individual devices, between the scheduling gain and the same 4 error measures of Table 8.2. These results are obtained only from the MF model. We sorted the devices by of the 4

Model	FAE	RMSE	Accuracy	F1 Score
ARIMA	**-0.78**	-0.03	-0.007	0.05
MF	**-0.69**	-0.13	-0.006	0.33

Table 8.2: Comparison of correlation between scheduling gain and forecasting performance measures, using the Pearson correlation coefficient
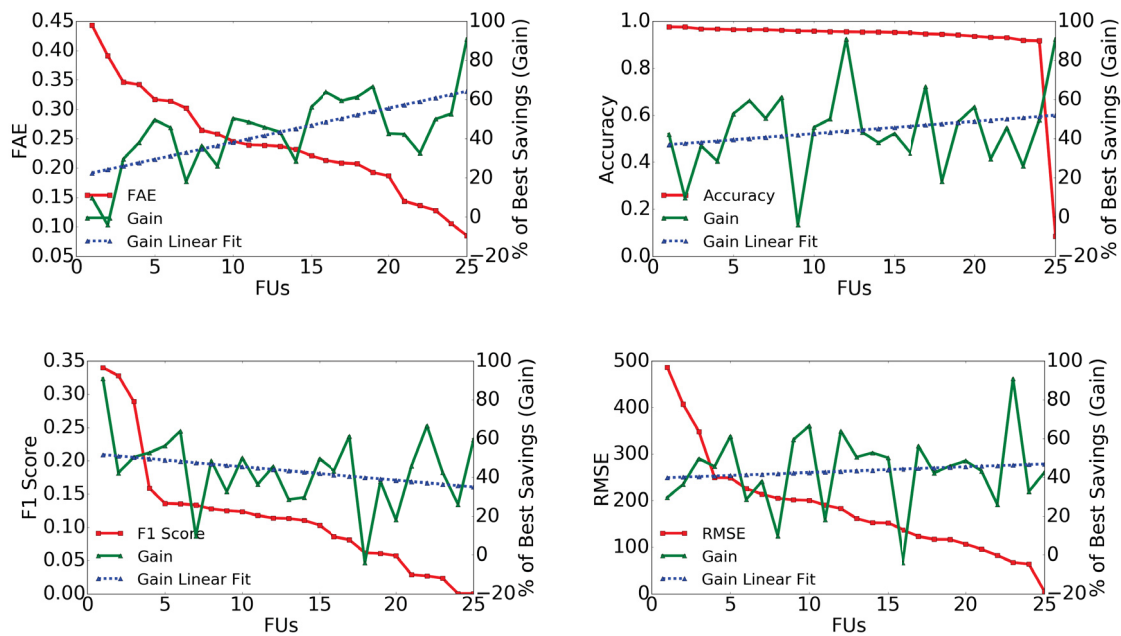
Figure 8.7: Correlation between forecasting measures and user gain percentage.

different measures, separately. In the case of the FAE and RMSE, the device have been sorted in descending order, ascending otherwise. From these plots it can be noticed how the FAE measure gives the *highest correlation* between forecasting measure and flex-offer gain, described by a steadier increase in the Gain curve as the FAE score decreases. Moreover, we can observe that despite the complexity of device-level load forecasting, approximately 50% of the device obtain a scheduling gain that is at least 45% of the optimal gain.

8.5 Summary and Discussion

In this chapter, we presented an answer to the research question *"How can we economically schedule flex-offers with minimal loss of user comfort?"*. In this regards, we presented a user-comfort orient scheduling of flex-offers that accounts both the social and financial aspects for demand shifting. We defined the energy demand prediction task in the context of flex-offer scheduling and presented the Flexibility-Aware Error (FAE) measure that quantifies the actual performance of forecast models designed for the flexibility market. We evaluated the proposed measure with two demand prediction models and ex-

tensively compared with standard error measures. We have demonstrated that
the standard point-wise forecasting measures, e.g. RMSE, F1 score and accu-
racy, do not provide us with a correct overview of the actual performance of the
prediction models. Instead, the proposed FAE measure gave a better guideline
on the quality of the prediction models. Through the experimental results, we
showed that with the proposed scheduling technique, it is possible to achieve
significant financial gain on both spot and regulation market with a minimal
loss of user-comfort quality. Hence, we can conclude that *"flex-offers can be
economically scheduled minimizing loss of user comfort"*. In the next chapter,
we focus on the designing a device-level forecast evaluation and benchmark
platform.

Chapter 9

Device-level Demand Forecast Platform

The selection of an effective device-level load forecast model for a variety of devices, data granularity, and forecast horizon is a challenging and resource-intensive task, mainly due to the (1) the diversity of the models, (2) the lack of proper tools, similar to [95], and (3) the unavailability of proper datasets that can be used to validate all these models. Further, the forecast model selection and validation process includes a number of steps which are often time-consuming (see Figure 9.1). In this process, researchers have to spend an enormous amount of time, especially, in data preprocessing, i.e., *data extraction*, *cleaning*, *transformation*, and *handling outliers and observation gaps*, and feature generation, where a set of features is generated repeatedly until a sufficient model accuracy is obtained. Here a *feature* is a variable, derived based on input dataset values or additional external information, that is assumed to be helpful for improving forecasting accuracy, e.g., temperature, wind speed, the day of the week, potentially, influencing demand and supply.

We can find a number of works dedicated to forecasting and analyzing device-level demand [70, 71, 90, 96]. However, work on device-level forecasting is still limited, because experiments are typically fine-tuned for a particular dataset and usage patterns, and are hardly reproducible with the reported level of accuracy. Furthermore, efficient and precise extraction of all relevant device-level data is a challenging and still ongoing research [97–99]. Consequently, there is a lack of proper datasets containing high-resolution measurements of a large number of devices that include all the relevant external influences. The experiments are typically performed with private datasets containing measurements that are collected within the scope of a project and are not freely available. Even the freely available datasets only include measurements for limited (short) time durations and are often too noisy to perform any detailed analysis [47]. Lastly, there are no effective tools designed specifically for tuning and
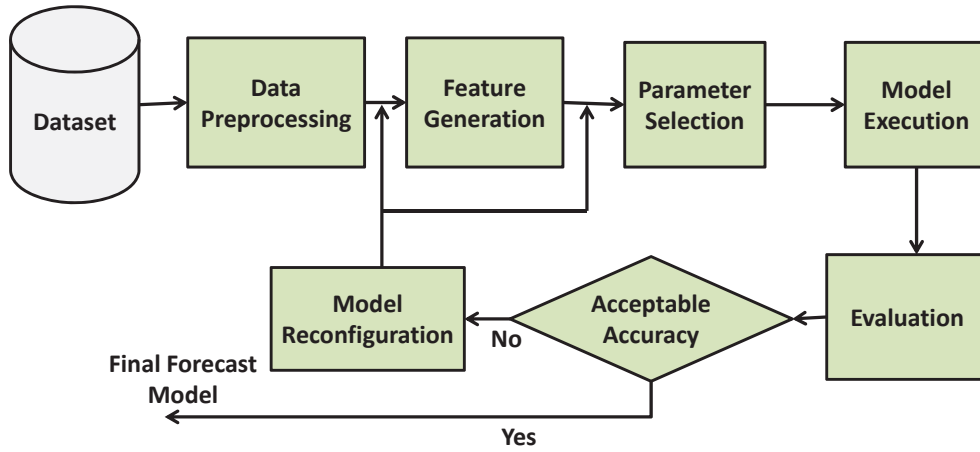
Figure 9.1: The process of forecast model selection and validation
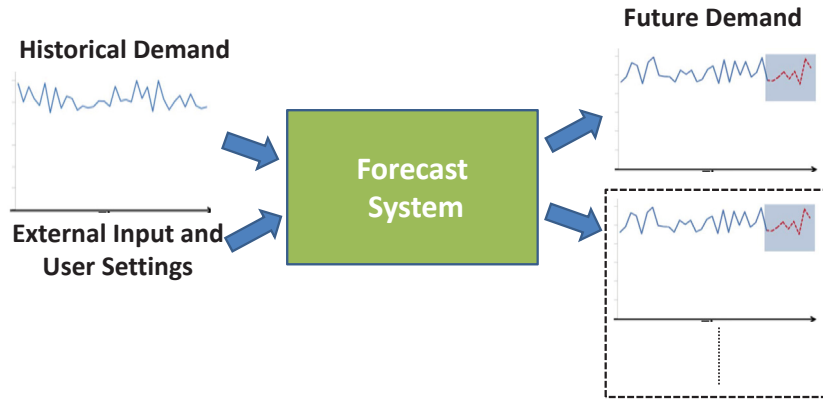
Figure 9.2: Inputs and outputs of the DeMand system

validating various device-level demand and supply forecasting models based on real measurements.

In this chapter, we present the DeMand system that allows the user to analyse and validate various forecasting models using a number of provided datasets and built-in or user-defined functions. The system is designed to automate most of the preprocessing steps. At the same time, it provides flexibility for the user (a researcher or energy market player) to use either existing system modules or plug-in custom user-defined modules. Figure 9.2 shows the inputs and outputs of the DeMand system. The system offers the following features and functionality: i) a repository of available device-level datasets for evaluating and comparing forecast models, ii) access to existing (standard) forecasting algorithms, iii) dynamic generation of features for various forecast horizons and data resolutions, iv) support for various experimental configurations and generation of multiple forecast models, v) functionality to compare experiment results, and vi) easy integration of external features and learning algorithms.
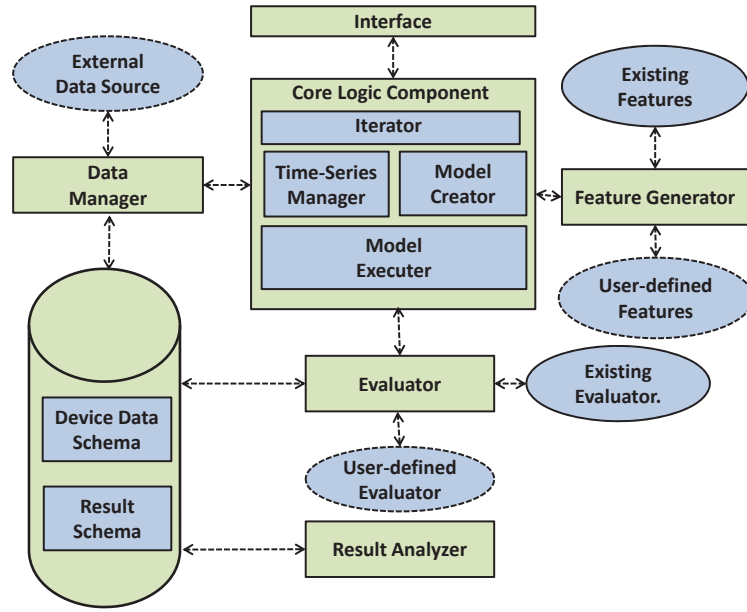
Figure 9.3: System Components

9.1 DeMand System Overview

In this section, we present the architecture and the functionality of the De-Mand system. The system is designed as a tool to automate experiments on device-level forecasting and to facilitate the comparison and re-(evaluation) of the existing experiments. Further, the system provides flexibility in using the available resources or uploading user-defined resources such as datasets, forecast models, evaluation metrics, etc. The user can define all necessary parameters and system configuration using the user interfaces. Once the user selects the timeseries and configures the experimental setup, the system provides all the available suggestion for the experiments such as a list of features, evaluation metrics, etc. Furthermore, the DeMand system is also envisioned to provide an open repository of device-level datasets that will be accessible to the research community for further experiments. Therefore, in addition to the graphical display in the interface, the experimental results and datasets are stored in the system database.

The DeMand system with the most essential components (rectangles) and their dependencies is shown in Figure 9.3. Here, the use of independent components for feature extraction, evaluator, and data management, etc. allows adding and removing system features in the plug-and-play fashion, making the system highly flexible and customizable for specific use-cases. We now present each of these components individually.

9.1.1   Interface Component

This components is designed to simplify and speed-up the process of setting up an experiment. Specifically, it offers a graphical user interface (GUI) and allows selecting a data source, a predictor, the values of configuration parameters, features, and error measures to use. It also provides visualization of time-series data. The user can set multiple values for the parameters to submit multiple tasks in a single execution. It also offers the visual representation of outcomes of the experiment, which plays a significant role in the forecast model analysis. Thus, in the end, the interface plots graphs for the experimental outcomes, according to the selected settings. For the multiple sets of parameters, the interface shows detailed plots for each configuration. If the user utilizes all the existing modules and functionality, the interface can reduce model analysis time drastically.

9.1.2   Core Logic Component

Core Logic is the central component in the system, and it orchestrates data manipulations and data exchange between other components according to user settings. The Core Logic component includes four different sub-components, shown in Figure 9.3, that automate the data preprocessing and parameter selection steps, shown in Figure 9.1.

Iterator The Iterator sub-component is responsible for parsing all the input parameters and determining the number of tasks to be executed. Here, a task is an independent execution of a forecast model with a particular predictor, dataset, parameters, etc. For example, if the user has selected two values of the probability threshold, e.g., in a classification task, then the iterator executes the same forecast model for each threshold value with the other parameters unchanged.

Model Creator The Model Creator sub-component creates an object that contains all the required parameters for the tasks. Then, all other (sub-)components fetch the parameters from this object. For multiple tasks, the sub-component updates object elements (parameters) that take more than a single value. The Model Creator sub-component is also responsible for persistent storage of model parameters until the completion of the current task.

Time Series Manager The Time Series Manager sub-component is responsible for the automation of all data preparation and manipulation tasks, such as data aggregation, noise filtration, filling observation gaps, etc. Further, it also acts as a communication bridge between the Data Manager and the Feature Generator components.

Model Executer The Model Executer sub-component executes the task with the provided feature set and configuration. If the experimental configuration already exists in the database, i.e., the experiment is not unique, the Model Executer terminates the current task and extracts the prediction results from the database. Further, the Model Executer handles errors in user-defined

predictors/classifiers by replacing them with the default predictor; else it terminates the current task and requests the Iterator sub-component to delete all remaining tasks in the queue. At the end of the execution, it passes the predicted values to the (model) Evaluator component.

9.1.3 Data Manager Component

The Data Manager component is responsible for extracting timeseries data from a user-defined source, i.e., a database or files in a user-specified location, and feeding extracted data to the System Core component. The Data Manager includes a data parser that transforms the raw data into the required format. After the data is successfully parsed, it is then persistently stored in the database for further analysis. This gives the user the comprehensive repository of datasets (e.g., for different device types) for future experiments, where the validation of forecast models in homogeneous environments becomes possible. Although the database is confined to the energy domain, there are no methods to validate the domain of the dataset except the required format. Thus, it is possible to upload timeseries from any domain without getting an error. However, a user can manually validate the timeseries before submitting the task by using the graphical interface (plot) provided by the system.

9.1.4 Feature Generator Component

Script 9.1: Python code for user defined feature

```python
#INPUT: time series with date column at first
#OUTPUT: binary feature representing weekdays
# or weekend for each data point
from datetime import datetime as dt
def is_weekend(timeseries = None):
    if timeseries == None:
        raise TypeError('Data can not be null')
    elif type(timeseries[0]) is not dt.date:
        raise TypeError('First column must be a
                        datetime.date, not a %s'
                        % type(timeseries[0]))
    else:
        feature_series = []
        for item in timeseries:
            day_of_the_week = item[0].isoweekday()
            if day_of_the_week <=5: # not weekend
                feature_series.append(0)
            else: # is weekend
                feature_series.append(1)
    return feature_series #binary features
```

The Feature Generator component is responsible for the generation of all features chosen by the user. As discussed earlier, *features* are variables (e.g., temperature or the day of the week), derived based on the input dataset values (time series) or external information, specified to, potentially, improve forecasting accuracy. The features are generated by using pre-defined functions from the feature repository. Additionally, the user can define new functions for the specification of custom features.

For example, Script 9.1 shows a simple user-defined function given in Python, which, for each data point in the original time series, identifies if a specific data point has been recorded during the weekend (value 1) or during a weekday (value 0). The component calls this user-defined function with the time series dataset as the input parameter. The function computes the features for each row in the time series and returns an output. The feature generator calculates the size of the output, i.e., the number of attributes, and increases the size of the feature vector by the respective number. In some cases, the device-level measurement dataset contains some sensitive context information, such as location, family size, age, group, occupation, etc. Therefore, this information is only accessible through the feature generator and is never revealed to the user. This approach facilitates the user when using sensitive information as features without requiring to disclose such information.

9.1.5   Evaluator Component

The Evaluator component receives as input the output of the predictor/classifier and evaluates its performance according to the chosen error measure and parameters. Further, it writes the experiment attributes and results to the database in the Result Schema for future reference and queries. The evaluator also invokes the functions to plot the graphical representation of the experimental results. The user can select existing error measures (out of many available ones), or define custom measures using a Python function, similarly to feature specification.

9.1.6   Result Analyser Component

The Result Analyzer is a component that processes the user queries and fetches the requested dataset from the system database. The user can write simple SQL queries to extract all the relevant data and results of an experiment satisfying certain conditions, such as type of predictor, dataset, forecast horizon, etc. Further, the user can write queries to compare experiments using a particular error measure. For example, let us consider a binary classification task where the objective is to predict a device state, i.e., idle (0) or active (1), at a particular hour in the future. Using the SQL query shown in Script 9.2, a user can query the results of the entire classification tasks based on the logistic regression model, ordered in the decreasing order of the Area Under the Curve (AUC) value. Further, from the list of available results, the user can select

up to two experiments and compare their performance graphically. Finally, if needed, the earlier experiments can be re-executed on a new dataset using the same or new configuration of parameters.

Script 9.2: SQL Script to extract data of earlier experiments

```sql
Select a.*, AUC from Experiment a
left join
(select experimentID, AUC from Result) b
on a.experimentID = b.experimentID
where predictor = 'Logistic_Regression'
order by AUC desc
```

9.2 DeMand Use-case Example

In this section, we briefly walk-through the DeMand functionality using the real-world use-case of the device-level forecast model analysis. Here, we continue with the binary classification problem of predicting a device state for the next hour.

9.2.1 Execution of Forecasting

First, we start by making a choice of the dataset to use for the experiment. In our case, we select the consumption time series of a washer dryer from the existing database. Alternatively, we can also select a new dataset using the window shown in Figure 9.5. The system then automatically plots the selected dataset in the main window and also provides some general statistics on the dataset, such as minimum, average, and maximum demand (see Figure 9.4). This information is helpful in deciding the threshold value (in watts) for the segmentation of data based on a device state (active or ideal). In our example, we have selected two values 10 watts and 100 watts as the threshold (see Figure 9.4). As result, the system creates two experiments to generate a classification model for each of the threshold values.

Next, we select the percentage of time series to use as a test set (20% in our example) using the main DeMand system window, shown in Figure 9.4. The timeseries is sequentially split into training and test sets based on the selected split ratio. Then, we also choose a predictor to be used for classification, which, in our example, is the *Logistic Regression* with *L1 regularization* (LR-L1). Further, we configure an hour-ahead forecast model by selecting the forecast horizon and data granularity of 1 hour. Afterwards, we select the set of features which we think will help to improve the performance of the classifier. As shown in Figure 9.6, the system automatically provides a list of available features that can be used with the selected dataset. In our example, as seen in Figure9.6, we select all the available features. Further, we choose
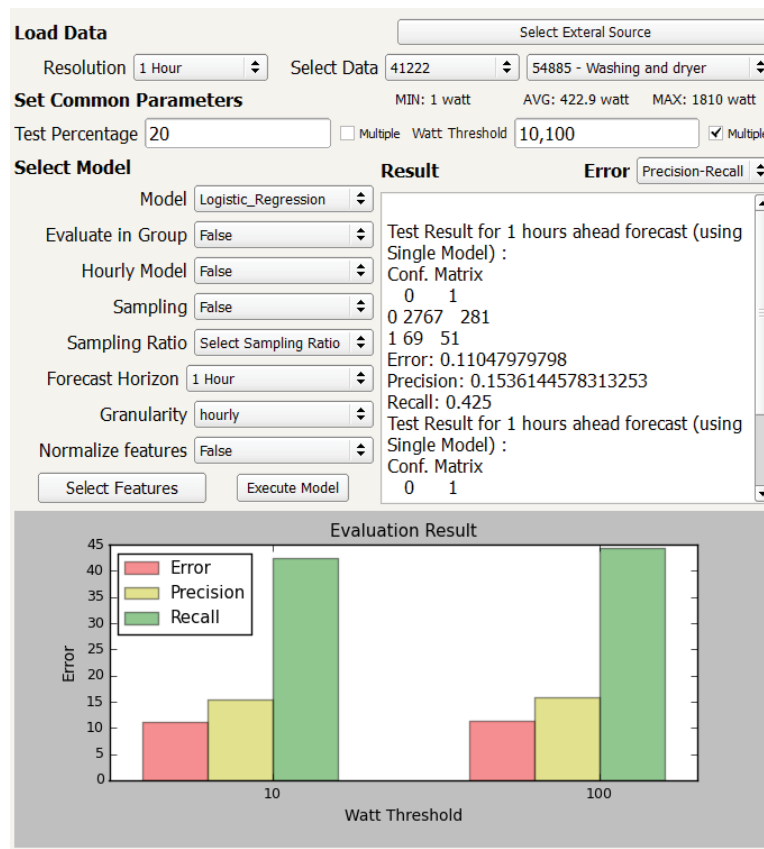
Figure 9.4: The main window of the DeMand system

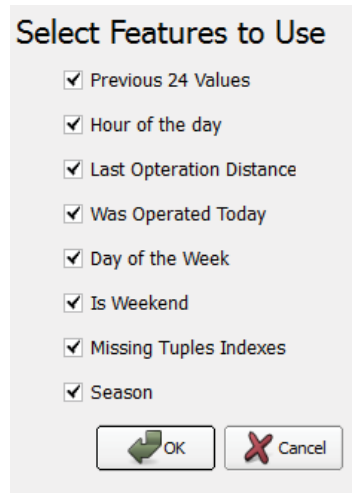**Figure 9.5:** Defining external source in DeMand

Figure 9.6: Feature selection in DeMand

the precision-recall curve as the error measure for evaluating the classifier (see Figure 9.4).

Finally, we execute the experiments from the completely configured experiment. The system might still request some additional parameters specific to the predictor. In the case of user-defined predictors, all additional parameters have to be handled by the user. To set the parameter value, a user has to include a call to the *add_parameter* function with a parameter name and a value as input. In our example, the LR-L1 model requires the values of the penalty parameter λ and the probability threshold. After we provide the values of all required parameters, the system executes the experiment and stores results in the database along with all forecasted values.

9.2.2 Result Presentation

After the completion of all experiments, the results are presented in the main window, as shown in the Figure 9.4. The system plots the values of the classifiers according to the chosen parameters. In our example, the system plots performance values in terms of *error*, *precision*, and *recall* of two classifiers with different *threshold watt* parameter values. The system also provides a detailed description of the results in the main window as a textual description. Further, if we click on the individual plots, the system automatically shows precision-recall curves for each classifier, as seen in Figures 9.7a and 9.7b. For a detailed comparison of the results, we can also use the result analyzer, as illustrated in Figure 9.8. In this example, we query all experiments that has been performed with the same *deviceID* and *dataGranularity* parameter values, sorted based on AUC values. The output of the query can be seen in the lower section of the Figure 9.8. Here, the experiment at the top of the table has the best performance in terms of the selected error measure. Additionally, we can choose any two experiments for a visual side-by-side comparison, as shown in Figure

(a) 10 watt

(b) 100 watts
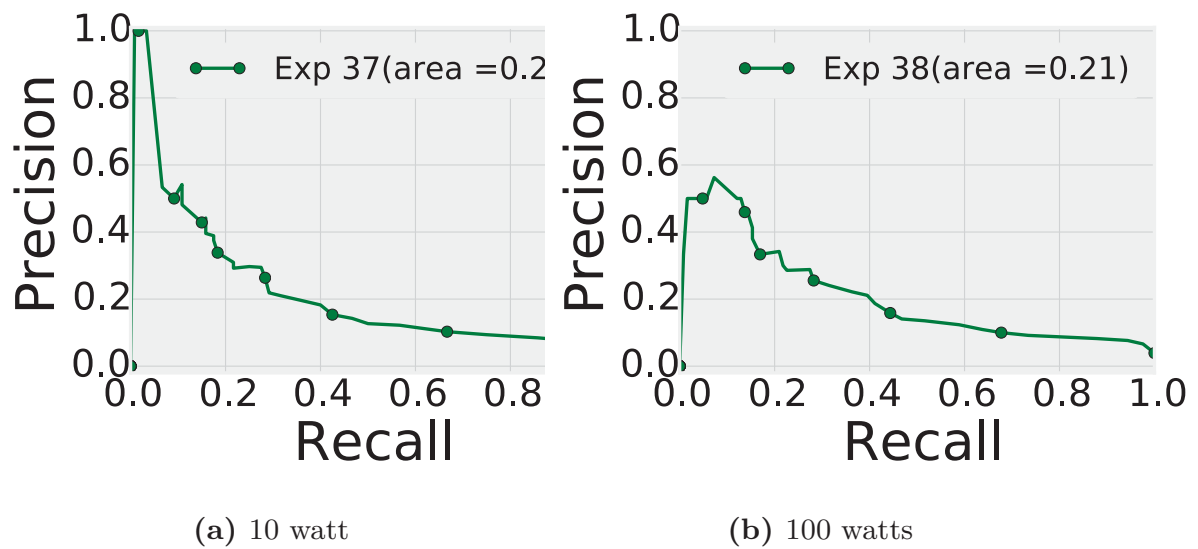
Figure 9.7: Precision-Recall curve: for various demand thresholds
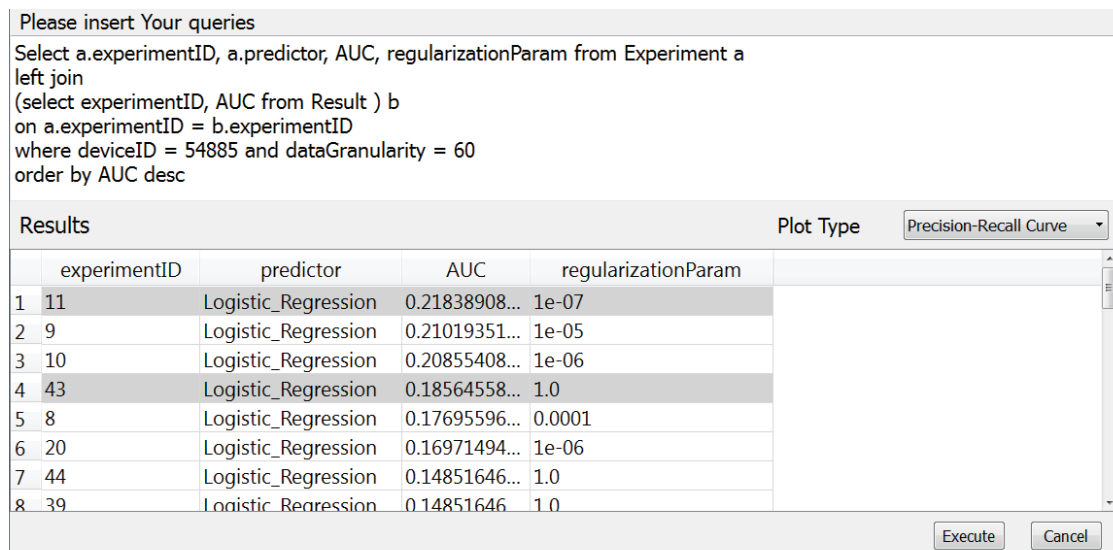
Figure 9.8: Result Analyzer Windows of the DeMand system

9.9.

As seen from the use-case example presented above, the DeMand system provides an analytical (decision-support) platform for comparing, validating, and choosing different forecast models and their parameter values. With the comprehensive model comparison information offered by the DeMand system, the user can decide on the best prediction model (algorithm) and select its parameters for a specific dataset or the given collection of datasets. As it can be seen from this use-case, the DeMand system with all its features and built-in functionality allows significantly reducing time needed to select and validate forecast models, compared to using general tools or hard-coded solutions requir-ing, typically, much more data pre-processing, system configuration, and result
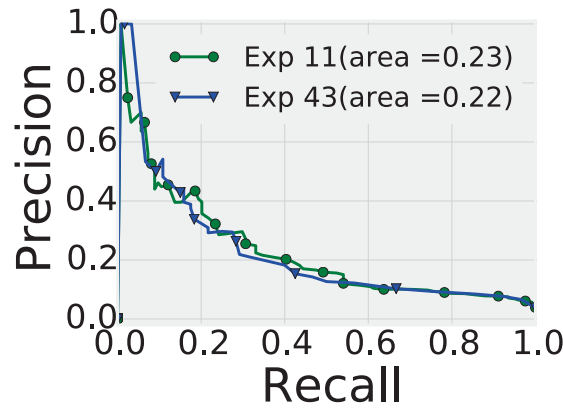
Figure 9.9: Comparison between experiments

processing time. The system can be accessed at `https://54.186.113.212`. However, due to the security reasons, a user should first whitelist his IP address in order to access the system. The request can be sent to the author at bj21.neupane@gmail.com.

9.3 Summary and Discussion

In this paper, we presented the DeMand system for fine tuning, analyzing, and validating the device-level forecast models. The system offers a number of features such as built-in device-level measurement datasets, forecast models, and error measures. The presented system allows users to evaluate and compare different forecast models based on various parameters, making device-level forecasting more accessible and efficient. Further, the chapter also provided the use-case example on how a forecast model for predicting a device state can be analyzed using the DeMand system. Thus, we showed that DeMand is an easy-to-use system automating most of the steps of the forecast model selection and validation process.

Chapter 10

Conclusion and Future Research Directions

10.1 Conclusion

The Ph.D. thesis presented a number of analytical techniques for analysis, extraction, and evaluation of flexibilities in energy demand from heat pumps, electric vehicle, and wet-devices, with a focus on filling the gaps for efficient integration of RES into a grid system. The Ph.D. thesis was supported by the TotalFlex project. The aim of the project is to develop a cost-effective, market-based system that utilizes total flexibility in energy demand, and provide financial and environmental benefits to all involved parties. In addition to analytics such as user behavior analysis, the thesis focused on designing models for forecasting device-level energy demands and extracting flexibilities from the forecasted demands. Further, as forecasting device-level demand is a very challenging task, the thesis performed a comprehensive evaluation of data granularity and forecast horizon that best suits the device-level forecasting. In addition, as the integration of a new concept to an established market always depends on the utility it brings to the market players, the thesis also modeled and evaluated the utility of the flexibility concept and financial viability of all proposed techniques. Similarly, as scheduling of flexibility is a complex task, the thesis took efficiency and user-acceptability into consideration in devising solutions for demand scheduling. The developed solutions and techniques were broad enough to be applied to various analytical systems and forecasting applications. In the following, the methodologies and outcomes for each chapter are summarized. Further, an overview of the contributions of the Ph.D. thesis is presented.

Chapter 3 detailed the flexibility and flex-offer concepts. The concepts were discussed analogous to the operation sequence of a smart device. The presented approach has opened an avenue of modeling flexibility from all de-

vice types into the flex-offer format proposed by the MIRABEL and TotalFlex projects, and the thesis demonstrated it for electric vehicles, heat pumps, and wet devices. Further, the presented strategy of modeling flexibilities from a variety of household devices into a unified flex-offer format facilitates the efficient aggregation and trading. The chapter presented a cost-benefit based method for categorization of household devices, where devices were categorized based on the trade-off between available flexibility and ease of extracting it. The chapter also discussed the spot and regulation markets and detailed the use of the demand flexibility by these markets to confront the challenges imposed by increasing RES integration. The chapter also presented the deployment architecture for device level data collection and provided a detailed description of datasets used in the experiments.

Chapter 4 evaluated the utility of demand flexibility in the energy market. The chapter presented a detailed analysis of the effects of incorporating demand flexibility in the existing energy market. The effects were analyzed in terms of changes the demand flexibility brings to the market balances and the corresponding changes in the market prices. The impact on market balances were modeled by analyzing all the possible changes in market states that can be caused by demand shifting. Similarly, a mathematical model to capture the relationship between various market prices was proposed to analyze the effect on the market prices. The mathematical model was derived by evaluating all possible linear and multiplicative interactions between the spot price, regulation prices, and regulation volumes. Thereafter, the chapter discussed the utility of the demand flexibility to the market players and proposed a model to quantify the utility. The presented quantification model captures the total increase or decreases in regulation cost as a consequence of demand shifting. Further, the chapter presented two optimization strategies, minimizing regulation cost and minimizing total imbalances, for evaluating the utility. The analyses were performed with demand flexibility of various sizes and types to identify the market specific threshold size of the amount flexibility that can be traded profitably. The experimental results showed that irrespective of the type of flexibility and market objectives, a market could always obtain a positive utility by adopting demand flexibility. Indeed, the market can achieve 5.2% reduction in regulation volume and 24.9% reduction in regulation cost with just 4 hours of time flexibility and 100 MWh of amount flexibility. The proposed mathematical models could be beneficial to the energy market players for assessing the applicability and utility of adopting demand flexibility in their market area.

Chapter 5 presented a descriptive analytics on the device-level energy consumption behavior. The primary goal was to find various patterns that support the existence of extractable flexibilities in device operations. The chapter introduced various properties representing flexibility in device operations, and statistical analyses were performed to validate these properties. It was found that analyses focused on the presented properties can effectively capture users' device usage behavior and provide significant information on the type and size of flexibilities available in device operations. The chapter also presented the

preprocessing steps and techniques to handle raw device-level data and convert it into a format required for higher level analysis. It was found that the proposed preprocessing steps helped to remove outliers and reduce stochasticity in the dataset. The experimental results supported the existence of significant amounts of extractable multidimensional (both time and amount) flexibilities in users' daily routine. The presented analytics revealed interesting patterns representing flexibility in device operation behavior. Thus, the proposed techniques can help energy market by quantifying the flexibility potential of any household devices.

Chapter 6 presented a detailed methodology for forecasting device-level energy demand. The primary goal was to evaluate the effectiveness and usability of widely used forecasting models for device-level demand forecast. Several forecasting techniques such as Logistic Regression (LR), Pattern Sequence Matching (PSM) were used to forecast device future activations and signature based pattern extraction was used to estimate the associated energy demand for the predicted activations. The chapter presented various device-level features to enhance the performance of forecast models (especially LR), and the experimental results showed that the proposed device-level features together with L1-regularization significantly improved the performance of the forecast models. The chapter also addressed the challenges of the class imbalance problem present in the real world device-level consumption data. Based on the Area Under the precision-recall curve and F1-score, we found that the L1-regularized Logistic Regression was the best device-level forecast model among the models used in the experiments. Further, it was found that over-sampling of the minority class did not improve the forecast model, whereas the weighted class importance helped to increase the forecast accuracy. The performance of the forecast models were evaluated for hourly, group, and daily data resolution. It was found that the forecast models achieve the best accuracy for daily resolution at the cost of significant loss of demand flexibility. Hence, group resolution was found to provide a good trade-off between forecast accuracy and available demand flexibility. The chapter quantified the financial benefits and losses of demand flexibility in relation to achievable forecast accuracy. The overall benefits and losses were decomposed and analyzed based on types of forecast categories represented by contingency table. The results showed that the financial gain from the presented models was much better than implied by the error metrics. Indeed, with a precision and recall of just 0.29 and 0.28, the market achieved regulation cost savings of 42% of the theoretically optimal. Further, the chapter presented a rule of thumb of selecting an optimal threshold for device-level forecast models. The proposed device-level demand forecast system can help the energy market for generating effective demand and supply schedules to reduce market imbalances and related financial losses. The system and analysis can also be applied to forecast demand for various other household appliances.

Chapter 7 presented a state-of-the-art flex-offer generation and evaluation process (FOGEP). The proposed process analyzes the past consumption pat-

tern of a device to extract available flexibility and encapsulates the flexibility for the device future operation into a flex-offer format. The FOGEP was divided into data preprocessing, Model Parameter Estimation and Forecasting (MPEF), and evaluation sub-processes, each with a distinct role. The MPEF sub-process predicts the attributes representing time and amount flexibilities for device future operations utilizing forecast models and device-specific algorithms. The predicted timestamp and values for the flexibilities were used to generate flex-offers for the device. The chapter proposed two different flexibility allocation approaches, Constraining Amount flexibility (CAF) and Weight-based Flexibility Distribution (WFD), to tackle the dependency problem present in amount flexibility distribution for heat pumps. Further, to capture the flexibilities during the zero occupancies a concept of the relaxed period was introduced. The chapter also presented models to estimate the savings in the spot and regulation markets that could be achieved by the efficient scheduling of the flex-offers generated using the proposed FOGEP. The chapter reported a comprehensive experimental study utilizing various real device-level consumption datasets and market data. The results showed that the proposed FOGEP could capture flexibilities in device usage behavior with an acceptable accuracy. The experimental results showed that household devices have up to 32% of reduction and 15 hours of shifting flexibility in their energy demand, and the proposed FOGEP can extract these flexibilities with up to 98% accuracy. Further, the experiments with three different device types showed the proposed FOGEP is flexible and general enough to be used for generating flex-offers from various other household devices.

Chapter 8 presented a novel user-comfort oriented prescription technique for scheduling of flex-offers. The proposed technique performs a trade-off between user comfort and financial gain as a consequence of shifting of flexible demand and prescribes a schedule that optimizes the combination of these two factors. Further, the chapter defined the energy demand forecasting task in the context of flex-offer scheduling and presented a novel Flexibility-Aware Error (FAE) measure that quantifies the actual performance of forecast models designed for a flexibility market. The proposed measure was evaluated with two demand prediction models, ARIMA and Most Frequent Timestamp, and was extensively compared with standard error measures. Through the extensive experiments with the real-world dataset, it was demonstrated that the proposed FAE measure gives better guidelines on the real, utility-based, quality of the prediction models compared to the standard point-wise forecast evaluation measures, such as RMSE, F1-score, and accuracy. Further, the proposed prescription technique was able to achieve a positive utility on both spot and regulation markets with a minimal loss of user-comfort quality. The proposed technique provided an efficient scheduling solution. Thus, the technique can help the energy market players to effectively utilize the device-level flexibility to solve the problems imposed by higher integration of RES.

Chapter 9 presented the DeMand system as a benchmark platform for device-level demand forecasting and flexibility analysis. The platform allows

users to evaluate and compare different forecast models based on different parameters, making device-level forecasting more accessible and efficient. The presented implementation offers a number of built-in device-level consumption datasets, forecast models, features, and error measures and also supports a user defined external features and learning algorithm. The chapter also provided the use-case example on how a forecast model for forecasting a device state can be analyzed using the DeMand system. An open repository of device-level datasets provided by the platform can help research community to perform extensive experiments on device-level demand forecasting.

In summary, predictive analytics of device-level energy demand is an integral part of the flexibility-based dynamic DR. The atomic level analysis of energy demands can be very useful in devising effective strategies to confront challenges imposed by integrating higher percentage of RES into the grid system. However, the raw device-level energy consumption data are often noisy and inconsistent, thus, making difficulties in further analysis. This thesis presented a number of analytical techniques for effective and efficient analysis of device-level data. Modeling of flexibility for various device types into a unified flex-offer format was a challenging task. The thesis presented the flexibility modeling approach that took this into consideration and solved it for important and typical devices. The adaptation of flexibility-based DR highly depends on the utility it brings to the market players. The thesis presented a financial quantification model that can help the energy market players in estimating the utility of adopting demand flexibility in their market areas. Further, to solve the response fatigue problem, the thesis presented solutions for automatically predicting flexible demands with minimal user intervention. The thesis presented a general flex-offer generation and evaluation process to automatically extract flexibilities and generate flex-offers for a variety of household devices. In brief, the thesis has presented various aspects of predictive analytics of device-level data with a special focus on extracting and modeling demand flexibility. The thesis resolved various challenges with the device-level data analytics. A number of novel techniques in the field of flex-offer modeling, generation, and scheduling were proposed that can significantly contribute to energy market to speed up the adaptation of flexibility-based DR.

10.2   Future Research Directions

Several directions for future work exist in the area of device-level data analytics and for the methods and techniques presented in this thesis. Here, we review future research directions for each of the presented chapters and overall new directions.

There are several directions for future work on modeling utility of demand flexibility presented in Chapter 4. Consideration of grid capacity constraint and other market constraints during the shifting of flexible demand is an interesting future research direction. In addition to the regulation volumes, it will

be interesting to include various other external features to capture the relationship between market prices. The presented demand shifting model adopts a greedy procedure; more efficient optimization techniques should be designed. Further, to better access viability of the flexibility-based DR, the utility should be evaluated for various other market objectives and market areas.

There are a number of future directions in device-level consumption behavior analysis presented in Chapter 5. In order to capture seasonal and long-term device usage pattern, an analysis should be performed with a longer consumption timeseries. New operation state segmentation techniques are required to supplement the manual inspection method. Demand flexibility analysis for smaller resolutions, e.g. 15 min or even 1 min, is an important direction for future work. It will also be interesting to analyze flexibilities potential in energy supply. Further, collecting user behavior and preference data through a survey to improve the confidence over a purely data-driven flexibility analysis will be an interesting approach.

There are several directions for future work on the assessment of forecast models presented in Chapter 6. There exists some correlation between devices and are operated in a specific sequence; the inclusion of correlated device states as features during forecasting will be a major research direction. The device-level consumption data is highly stochastic; it will be interesting to evaluate forecast models for a pool of similar devices clustered based on various market conditions. There should be a more comprehensive exploration and evaluation of forecast models suited for device-level forecasting.

For the case of flex-offer generation and evaluation process presented in Chapter 7, designing and utilization of more robust device-level forecast models during flex-offers generations sub–steps will be an important future research direction. Various hardware and grid constraints should be incorporated in FO-GEP. It will be interesting to consider the storage capacity of electric vehicles for a flex-offer generation. The flexibility-based market setup should also be established to validate the applicability of the proposed FOGEP.

In the case of Chapter 8, establishing statistical models based on capacity constraints to evaluate multi-device flex-offers scheduling will be interesting future research direction. In addition, designing of forecasting techniques that align the prediction to the scheduling task will be another future work.

For the DeMand system presented in Chapter 9, designing of an API-centric architecture is an important future research direction. It is also interesting to incorporate additional features such as model and parameter recommender system, more comprehensive data processor module, and support for ensemble learning, etc.

The thesis addressed various issues in device-level data analytics, especially in the demand flexibility extraction and modeling scenario. However, there are still various future research topics that we are interested in exploring further, especially related to device-level data analytics. First, the proposed preprocessing steps only consider the data at hourly resolution. Hence, we will explore the device-level data preprocessing and design a general solution that applies to

data of different resolutions. Second, for the class imbalance problem, we will further explore the problem and perform extensive statistical tests to validate the approach proposed in this thesis. Third, a flex-offer can represent flexibility for both energy demand and supply. Hence, we will extend the current FO-GEP to capture and model flexibilities from various sources of energy supply. Fourth, this thesis mostly evaluated regression and pattern matching models for predicting device-level demand. In the future, we will further explore the forecasting field to design and assess more sophisticated prediction models such as BayesNet, DeepLearing, HMM, etc., for device-level demand forecasting.

References

[1] Eurostat, 2016. [Online]. Available: http://ec.europa.eu/eurostat/tgm/table.do?tab=table&plugin=1&language=en&pcode=t2020_31

[2] M. Nikzad and B. Mozafari, "Reliability assessment of incentive-and priced-based demand response programs in restructured power systems," *International Journal of Electrical Power and Energy Systems*, vol. 56, pp. 83–96, 2014.

[3] J. Torriti, "Price-based demand side management: Assessing the impacts of time-of-use tariffs on residential electricity demand and peak shifting in northern italy," *Energy*, vol. 44, no. 1, pp. 576 – 583, 2012.

[4] F. Brahman, M. Honarmand, and S. Jadid, "Optimal electrical and thermal energy management of a residential energy hub, integrating demand response and energy storage system," *Energy and Buildings*, vol. 90, pp. 65 – 75, 2015.

[5] L. Huang, J. Walrand, and K. Ramchandran, "Optimal demand response with energy storage management," in *2012 IEEE Third International Conference on Smart Grid Communications (SmartGridComm)*, 2012, pp. 61–66.

[6] .-. The *EcoGrid 2.0* Project. [Online]. Available: http://ecogridbornholm.dk/nyheder/ecogrid-20/

[7] .-. The *iPower* Project. [Online]. Available: http://www.ipower-net.dk/

[8] TotalFlex. (2012) The *TotalFlex* project. [Online]. Available: http://www.totalflex.dk/Forside/

[9] .-. The *GoFLEX* Project. [Online]. Available: http://www.goflex-community.eu/

[10] M. Muratori and G. Rizzoni, "Residential demand response: Dynamic energy management and time-varying electricity pricing," *IEEE Transactions on Power Systems*, vol. 31, no. 2, pp. 1108–1117, 2016.

[11] . The *MIRABEL* Project. [Online]. Available: http://www.mirabel-project.eu.

[12] L. Dannecker, M. Böhm, W. Lehner, and G. Hackenbroich, "Forcasting evolving time series of energy demand and supply," in *Proceedings of AD-BIS'11*, 2011, pp. 302–315.

[13] A. G. Paetz, T. Kaschub, P. Jochem, and W. Fichtner, "Load-shifting potentials in households including electric mobility - a comparison of user behaviour with modelling results," in *2013 10th International Conference on the European Energy Market (EEM)*, 2013, pp. 1–7.

[14] E. Klaassen, C. Kobus, J. Frunt, and H. Slootweg, "Load shifting potential of the washing machine and tumble dryer," in *2016 IEEE International Energy Conference (ENERGYCON)*, 2016, pp. 1–6.

[15] C. B. Kobus, E. A. Klaassen, R. Mugge, and J. P. Schoormans, "A real-life assessment on the effect of smart appliances for shifting households' electricity demand," *Applied Energy*, vol. 147, pp. 335 – 343, 2015.

[16] W. Labeeuw, J. Stragier, and G. Deconinck, "Potential of active demand reduction with residential wet appliances: A case study for belgium," *IEEE Transactions on Smart Grid*, vol. 6, no. 1, pp. 315–323, 2015.

[17] A. Giannitrapani, S. Paoletti, A. Vicino, and D. Zarrilli, "Bidding strategies for renewable energy generation with non stationary statistics," in *World Congress*, 2014, pp. 10 784–10 789.

[18] X. Zhang, "Optimal wind bidding strategy considering imbalance cost and allowed imbalance band," in *Energytech, 2012 IEEE*, 2012, pp. 1–5.

[19] G. Bathurst, J. Weatherill, and G. Strbac, "Trading wind generation in short term energy markets," *Power Systems, IEEE Transactions on*, pp. 782–789, 2002.

[20] P. Siano, "Demand response and smart grids - a survey," *Renewable and Sustainable Energy Reviews*, pp. 461 – 478, 2014.

[21] Z. Zhao, W. C. Lee, Y. Shin, and K.-B. Song, "An optimal power scheduling method for demand response in home energy management system," *Smart Grid, IEEE Transactions on*, pp. 1391–1400, 2013.

[22] A.-H. Mohsenian-Rad and A. Leon-Garcia, "Optimal residential load control with price prediction in real-time electricity pricing environments," *Smart Grid, IEEE Transactions on*, pp. 120–133, 2010.

[23] A. Roscoe and G. Ault, "Supporting high penetrations of renewable generation via implementation of real-time electricity pricing and demand response," *Renewable Power Generation, IET*, pp. 369–382, 2010.

[24] M. Albadi and E. El-Saadany, "Demand response in electricity markets: An overview," in *Power Engineering Society General Meeting*, 2007, pp. 1–5.

[25] H. Aalami, M. P. Moghaddam, and G. Yousefi, "Modeling and prioritizing demand response programs in power markets," *Electric Power Systems Research*, pp. 426 – 435, 2010.

[26] B. Biegel, P. Andersen, T. Pedersen, K. Nielsen, J. Stoustrup, and L. Hansen, "Electricity market optimization of heat pump portfolio," in *Control Applications (CCA), 2013 IEEE International Conference on*, 2013, pp. 294–301.

[27] O. Sundstrom and C. Binding, "Flexible charging optimization for electric vehicles considering distribution grid constraints," *Smart Grid, IEEE Transactions on*, pp. 26–37, 2012.

[28] D. Papadaskalopoulos and G. Strbac, "Decentralized participation of flexible demand in electricity markets - part i: Market mechanism," *Power Systems, IEEE Transactions on*, pp. 3658–3666, 2013.

[29] O. Lünsdorf and M. Sonnenschein, "A pooling based load shift strategy for household appliances," in *24th International Conference on Informatics for Environmental Protection*, 2010, pp. 734–743.

[30] T. Bigler, G. Gaderer, P. Loschmidt, and T. Sauter, "Smartfridge: Demand side management for the device level," in *ETFA 2011*, 2011, pp. 1–8.

[31] H. T. Haider, O. H. See, and W. Elmenreich, "Dynamic residential load scheduling based on adaptive consumption level pricing scheme," *Electric Power Systems Research*, vol. 133, pp. 27 – 35, 2016.

[32] S. Nistor, J. Wu, M. Sooriyabandara, and J. Ekanayake, "Capability of smart appliances to provide reserve services," *Applied Energy*, vol. 138, pp. 590 – 597, 2015.

[33] S. A. Pourmousavi, S. N. Patrick, and M. H. Nehrir, "Real-time demand response through aggregate electric water heaters for load shifting and balancing wind generation," *IEEE Transactions on Smart Grid*, vol. 5, pp. 769–778, 2014.

[34] S. Feuerriegel and D. Neumann, "Measuring the financial impact of demand response for electricity retailers," *Energy Policy*, pp. 359 – 368, 2014.

[35] J.-H. Kim and A. Shcherbakova, "Common failures of demand response," *Energy*, vol. 36, pp. 873 – 880, 2011.

[36] J. Ma, H. H. Chen, L. Song, and Y. Li, "Residential load scheduling in smart grid: A cost efficiency perspective," *IEEE Transactions on Smart Grid*, vol. 7, pp. 771–784, 2016.

[37] N. Rotering and M. Ilic, "Optimal charge control of plug-in hybrid electric vehicles in deregulated electricity markets," *IEEE Transactions on Power Systems*, vol. 26, pp. 1021–1029, 2011.

[38] Z. Ma, D. S. Callaway, and I. A. Hiskens, "Decentralized charging control of large populations of plug-in electric vehicles," *IEEE Transactions on Control Systems Technology*, vol. 21, pp. 67–78, 2013.

[39] C. Shao, X. Wang, X. Wang, C. Du, and B. Wang, "Hierarchical charge control of large populations of evs," *IEEE Transactions on Smart Grid*, vol. 7, pp. 1147–1155, 2016.

[40] K. Skytte, "The regulating power market on the Nordic power exchange Nord Pool: an econometric analysis," *Energy Economics*, pp. 295–308, 1999.

[41] M. Alizadeh, A. Scaglione, A. Applebaum, G. Kesidis, and K. Levitt, "Reduced-order load models for large populations of flexible appliances," *IEEE Transactions on Power Systems*, vol. 30, pp. 1758–1774, 2015.

[42] M. Zeifman and K. Roth, "Nonintrusive appliance load monitoring: Review and outlook." *IEEE TCE*, 2011.

[43] J. Kelly and W. Knottenbelt, "Neural nilm: Deep neural networks applied to energy disaggregation," in *Proceedings of the 2Nd ACM International Conference on Embedded Systems for Energy-Efficient Built Environments*, ser. BuildSys '15, 2015, pp. 55–64.

[44] N. Batra, A. Singh, and K. Whitehouse, "Neighbourhood nilm: A big-data approach to household energy disaggregation," *CoRR*, vol. abs/1511.02900, 2015.

[45] J. Z. Kolter, S. Batra, and A. Y. Ng, "Energy disaggregation via discriminative sparse coding," in *Proceedings of the 23rd International Conference on Neural Information Processing Systems*, ser. NIPS'10, 2010, pp. 1153–1161.

[46] A. Zoha, A. Gluhak, M. A. Imran, and S. Rajasegarar, "Non-intrusive load monitoring approaches for disaggregated energy sensing: A survey," *Sensors*, vol. 12, no. 12, pp. 16 838–16 866, 2012.

[47] J. Z. Kolter and M. J. Johnson, "REDD: A Public Data Set for Energy Disaggregation Research," in *SustKDD workshop*, 2011.

[48] Source of dataset: Intrepid, intelligent systems for energy prosumer buildings at district level, funded by the european commission under fp7, grant agreement n. 317983.

[49] N. G. Paterakis, O. Erdinç, A. G. Bakirtzis, and J. P. S. Catalão, "Optimal household appliances scheduling under day-ahead pricing and load-shaping demand response strategies," *IEEE Transactions on Industrial Informatics*, vol. 11, pp. 1509–1519, 2015.

[50] J. J. Shah, M. C. Nielsen, T. S. Shaffer, and R. L. Fittro, "Cost-optimal consumption-aware electric water heating via thermal storage under time-of-use pricing," *IEEE Transactions on Smart Grid*, vol. 7, pp. 592–599, 2016.

[51] B. Biegel, P. Andersen, T. S. Pedersen, K. M. Nielsen, J. Stoustrup, and L. H. Hansen, "Electricity market optimization of heat pump portfolio," in *Control Applications (CCA), 2013 IEEE International Conference on*, 2013, pp. 294–301.

[52] K. Kouzelis, Z. H. Tan, B. Bak-Jensen, J. R. Pillai, and E. Ritchie, "Estimation of residential heat pump consumption for flexibility market applications," *IEEE Transactions on Smart Grid*, vol. 6, pp. 1852–1864, 2015.

[53] M. U. Kajgaard, J. Mogensen, A. Wittendorff, A. T. Veress, and B. Biegel, "Model predictive control of domestic heat pump," in *2013 American Control Conference*, 2013, pp. 2013–2018.

[54] M. Pipattanasomporn, M. Kuzlu, S. Rahman, and Y. Teklu, "Load profiles of selected major household appliances and their demand response opportunities," *IEEE Transactions on Smart Grid*, vol. 5, pp. 742–750, 2014.

[55] C. Develder, N. Sadeghianpourhamami, M. Strobbe, and N. Refa, "Quantifying flexibility in ev charging as dr potential: Analysis of two real-world data sets," in *2016 IEEE International Conference on Smart Grid Communications (SmartGridComm)*, Nov 2016, pp. 600–605.

[56] R. D'hulst, W. Labeeuw, B. Beusen, S. Claessens, G. Deconinck, and K. Vanthournout, "Demand response flexibility and flexibility potential of residential smart appliances: Experiences from large pilot test in belgium," *Applied Energy*, vol. 155, pp. 79 – 90, 2015.

[57] D. Kaulakienė, L. Šikšnys, and Y. Pitarch, "Towards the automated extraction of flexibilities from electricity time series," in *EDBT/ICDT Workshops*, 2013.

[58] D. Fischer, T. Wolf, J. Wapler, R. Hollinger, and H. Madani, "Model-based flexibility assessment of a residential heat pump pool," *Energy*, vol. 118, pp. 853 – 864, 2017.

[59] A. M. Foley, P. G. Leahy, A. Marvuglia, and E. J. McKeogh, "Current methods and advances in forecasting of wind power generation," *Renewable Energy*, pp. 1 – 8, 2012.

[60] X. Wang, P. Guo, and X. Huang, "A review of wind power forecasting models," *Energy Procedia*, pp. 770 – 778, 2011.

[61] A. Fabbri, T. Román, J. Abbad, and V. Quezada, "Assessment of the cost associated with wind generation prediction errors in a liberalized electricity market," *Power Systems, IEEE Transactions on*, pp. 1440–1446, 2005.

[62] G. E. P. Box and G. Jenkins, *Time Series Analysis, Forecasting and Control*, 1990.

[63] P. Brockwell and R. Davis, *Introduction to Time Series and Forecasting*, ser. Lecture Notes in Statistics, 2002. [Online]. Available: http://books.google.dk/books?id=VHB4OSAmwcUC

[64] J. Nyblom, "Testing for the constancy of parameters over time," *Journal of the American Statistical Association*, vol. 84, no. 405, pp. 223–230, 1989.

[65] D. Bunn, "Forecasting loads and prices in competitive power markets," vol. 88, no. 2, pp. 163–169, 2000.

[66] R. Cottet and M. Smith, "Bayesian modeling and forecasting of intraday electricity load," *Journal of the American Statistical Association*, vol. 98, no. 464, pp. 839–849, 2003.

[67] Z. Aung, M. Toukhy, J. Williams, A. Sanchez, and S. Herrero, "Towards accurate electricity load forecasting in smart grids," in *Proceedings of DBKDA'12*, Feb. 2012, pp. 51–57.

[68] L. Zhang and P. Luh, "Neural network-based market clearing price prediction and confidence interval estimation with an improved extended kalman filter method," *IEEE TPS*, 2005.

[69] A. Reinhardt, S. S. Kanhere, and D. Christin, "Predicting the power consumption of electric appliances through time series pattern matching," in *BuildSys (posters)*, 2013.

[70] A. Reinhardt, D. Christin, and S. S. Kanhere, "Can smart plugs predict electric power consumption?: A case study," in *Proceedings of the 11th International Conference on Mobile and Ubiquitous Systems: Computing, Networking and Services*, ser. MOBIQUITOUS '14, 2014, pp. 257–266.

[71] A. Barbato, A. Capone, M. Rodolfi, and D. Tagliaferri, "Forecasting the usage of household appliances through power meter sensors for demand management in the smart grid," 2011.

[72] N. Prüggler, "Economic potential of demand response at household level-are central-european market conditions sufficient?" *Energy Policy*, pp. 487 – 498, 2013.

[73] M. Vasirani and S. Ossowski, "Smart consumer load balancing:: state of the art and an empirical evaluation in the spanish electricity market," *Artificial Intelligence Review*, pp. 81–95, 2013.

[74] O. Erdinc, "Economic impacts of small-scale own generating and storage units, and electric vehicles under different demand response strategies for smart households," *Applied Energy*, pp. 142–150, 2014.

[75] D. Setlhaolo, X. Xia, and J. Zhang, "Optimal scheduling of household appliances for demand response," *Electric Power Systems Research*, pp. 24 – 28, 2014.

[76] L. Šikšnys, E. Valsomatzis, K. Hose, and T. B. Pedersen, "Aggregating and disaggregating flexibility objects," *IEEE Transactions on Knowledge and Data Engineering*, vol. 27, pp. 2893–2906, 2015.

[77] E. Valsomatzis, T. B. Pedersen, A. Abelló, K. Hose, and L. Šikšnys, "Towards constraint-based aggregation of energy flexibilities," in *Proceedings of the Seventh International Conference on Future Energy Systems Poster Sessions*, ser. e-Energy '16, 2016.

[78] E. Valsomatzis, T. B. Pedersen, A. Abelló, and K. Hose, "Aggregating energy flexibilities under constraints," in *2016 IEEE International Conference on Smart Grid Communications (SmartGridComm)*, 2016, pp. 484–490.

[79] E. Valsomatzis, K. Hose, and T. B. Pedersen, *Balancing Energy Flexibilities Through Aggregation*, pp. 17–37.

[80] N. Sadeghianpourhamami, M. Strobbe, and C. Develder, "Real-world user flexibility of energy consumption: Two-stage generative model construction," in *Proceedings of the 31st Annual ACM Symposium on Applied Computing*, ser. SAC 16, 2016, pp. 2148–2153.

[81] L. Šikšnys, M. Khalefa, and T. Pedersen, "Aggregating and disaggregating flexibility objects," in *SSDBM 2012*, 2012, pp. 379–396.

[82] *Zensehome.* [Online]. Available: http://www.zensehome.dk/da/privat. aspx

[83] Source of dataset: Evnetnl ev charging dataset, 2016. [Online]. Available: https://www.elaad.nl/

[84] N. E. R. (NordREG), "Nordic market report 2014." [Online]. Available: http://www.nordicenergyregulators.org/wp-content/uploads/2014/06/Nordic-Market-Report-2014.pdf

[85] I. Ilieva and T. F. Bolkesjø, "An econometric analysis of the regulation power market at the nordic power exchange," *Energy Procedia*, pp. 58 – 64, 2014.

[86] D. W. Hosmer Jr and S. Lemeshow, *Applied logistic regression.* John Wiley & Sons, 2004.

[87] A. Agarwal, O. Chapelle, M. Dudík, and J. Langford, "A reliable effective terascale linear learning system," *CoRR*, 2011.

[88] J. Davis and M. Goadrich, "The relationship between precision-recall and roc curves," ser. ICML '06, 2006, pp. 233–240.

[89] Anonymous, "Details omitted for double-blind reviewing," 2015.

[90] B. Neupane, T. Pedersen, and B. Thiesson, "Towards flexibility detection in device-level energy consumption," in *Proceedings of the Second ECML/PKDD Workshop, DARE'14*, 2014, pp. 1–16.

[91] L. Šikšnys and T. B. Pedersen, "Dependency-based flexoffers: Scalable management of flexible loads with dependencies," in *Proceedings of the Seventh International Conference on Future Energy Systems*, ser. e-Energy '16, 2016, pp. 1–13.

[92] Dark sky api, 2016. [Online]. Available: https://darksky.net/poweredby/

[93] Anonymous, "Details omitted for double-blind reviewing," 2015.

[94] W. W.-S. Wei, *Time series analysis.* Addison-Wesley publ Reading, 1994.

[95] R. Ulbricht, U. Fischer, L. Kegel, D. Habich, H. Donker, and W. Lehner, "Ecast: A benchmark framework for renewable energy forecasting systems." in *EDBT/ICDT Workshops*, 2014, pp. 148–155.

[96] A. Alhamoud, P. Xu, F. Englert, A. Reinhardt, P. Scholl, and R. Steinmetz, "Extracting human behavior patterns from appliance-level power consumption data," in *Wireless Sensor Networks*, ser. Lecture Notes in Computer Science, 2015.

[97] O. Parson, S. Ghosh, M. Weal, and A. Rogers, "An unsupervised training method for non-intrusive appliance load monitoring," *Artificial Intelligence*, pp. 1 – 19, 2014.

[98] D. Egarter, V. Bhuvana, and W. Elmenreich, "Paldi: Online load disaggregation via particle filtering," *Instrumentation and Measurement, IEEE Transactions on*, vol. 64, no. 2, pp. 467–477, 2015.

[99] S. Gupta, M. S. Reynolds, and S. N. Patel, "Electrisense: Single-point sensing using emi for electrical event detection and classification in the home," in *Proceedings of the 12th ACM International Conference on Ubiquitous Computing*, ser. UbiComp '10, 2010, pp. 139–148.

Statement of Authorship

I hereby declare that I have written this thesis individually and without the use of *documents* other than those referenced in this thesis and *inputs* others than those which I acknowledge in this thesis and explicitly mention in the presented co-author statements (available with the thesis). Up to now, the work is submitted only at TU Dresden and the IT4BI-DC partner university AAU. The work has not been presented in this or a similar form to another examination agency neither in Germany nor in any other country, and it has not yet been published either.

Bijay Neupane
August 10, 2017