Energy demand models for buildings in a smart cities context

Panagiota Gianniou

PhD Thesis

Department of Civil Engineering
2018

DTU Civil Engineering Report R-390

DTU Civil Engineering
Department of Civil Engineering

ENERGY DEMAND MODELS FOR BUILDINGS IN A SMART CITIES CONTEXT

PhD Thesis

by

Panagiota Gianniou

International Centre for Indoor Environment and Energy, Department of Civil Engineering

Technical University of Denmark

Title: Energy Demand Models for Buildings in a Smart Cities Context

Period: 15th September 2014 – 31st January 2018

Institute: Department of Civil Engineering, Technical University of Denmark

PhD student: Panagiota Gianniou

Main supervisor: Professor Carsten Rode

Supervisors: Associate Professor Alfred Heller, Senior Researcher Per Sieverts Nielsen

Declaration

I hereby declare that this thesis was written by myself entirely. All sources and materials I have used are hereby enclosed. The thesis has not been submitted further in this form or any other form, and has not been used to obtain any other equivalent qualifications at any other organization/institute.

Panagiota Gianniou

31 January 2018

Kgs. Lyngby, Denmark

‘All models are wrong, but some are useful.’

- *George E. P. Box*

Acknowledgements

This PhD thesis sums up the work carried out at the Technical University of Denmark (DTU), Department of Civil Engineering, International Centre for Indoor Environment and Energy (ICIEE), Kgs. Lyngby, Denmark, for the period September 2014 - January 2018. The work was composed under the DTU PhD program and was supported by the Danish research project CITIES (Centre for IT-Intelligent Energy Systems in cities) No DSF1305-00027B funded by the Danish Strategic Research Council.

Main supervisor during this PhD study was Professor Carsten Rode, from the Department of Civil Engineering, DTU. Co-supervisors were Associate Professor Alfred Heller from Department of Civil Engineering, DTU and Senior Researcher Per Sieverts Nielsen, Department of Management Engineering, DTU. I would like to express my sincere gratitude to my three supervisors and mentors for the constant inspiration to conduct research, guidance, motivation and knowledge they provided me with. Their overwhelming encouragement was important during this whole journey.

I would also like to thank my advisors Prof. Christoph Reinhart from Department of Architecture, Massachusetts Institute of Technology and Ass. Prof. David Hsu, Department of Urban Studies and Planning, Massachusetts Institute of Technology for hosting me during my external research stay and for the valuable knowledge and perspective on scientific research they shared with me.

I would also like to thank Adam Bruun, Aarhus AffaldVarme, who trusted me with the distribution of smart meter data and enabled the conduction of my research work. Special thanks to Rune Andersen, Department of Civil Engineering, DTU, for the productive discussions we had throughout my studies. I would also like to thank Snjezana Skocajic, who helped me with several practical issues and kept the administrative process running smoothly. Thanks to Prof. Geo Clausen, Head of the Section for Indoor Environment and Building Physics, DTU, for his helpfulness.

My sincere gratitude goes to my co-authors that helped me in accomplishing my research work and contributed to that. Thank you for the inspiring discussions and the exceptionally good collaboration we had. Special thanks goes to Dr. Xiufeng Liu, whose help on data analytics was significant. I would also like to thank Audrey Ryan for her kind assistance with part of the simulation analysis.

Friends and colleagues at the Section for Indoor Environment and Building Physics have made the efforts towards my degree infinitely more pleasant. Thank you, Kyriaki, Mariya, Eleftherios, Ongun, Emilie and Barbora. I couldn’t wish for better office mates!

I am indebted to my family and friends for their unlimited and unconditional support that incented me to strive towards my goals. Huge thanks to Christos for supporting me constantly. And most of all to my parents and sisters, Maria and Christina, for believing in me and always being by my side. *Σας ευχαριστώ!*

Abstract

Energy is one of the major drivers in smart cities along with smart environment and smart living. The role that buildings can play in the development and operation of smart energy cities is important, due to the large share of energy use they are responsible for and the smart energy solutions they can potentially integrate. The increasing number of smart cities initiatives and their focus on city level energy policy management has emphasised the need to move from the traditional micro-level building energy modelling towards the development of aggregated energy demand models. To accomplish that, methods that can be scalable to higher levels of aggregation, ranging from clusters of buildings to neighbourhoods and cities are needed.

The present thesis aims at providing enhanced modelling methodologies that provide building energy demand related insights in a high spatial and temporal resolution, which can help to evaluate energy policies and demand side management strategies. The main objectives of the thesis are to propose and investigate engineering-based approaches, statistical methods and data mining techniques that can contribute to the accurate building energy demand modelling at urban scale; to determine the potential that buildings can have on the stabilization of the energy grid and the flexible operation of the energy system -considering a smart energy cities context- both at building level and urban level; and finally, to indicate the suitability of each category of proposed modelling methodologies and provide guidelines for future investigations.

The most important findings of the research performed in this thesis are summarized in the following:

Aggregation of building energy demands is enhanced with archetypal approaches that significantly reduce the modelling and computation time. The minimum information level to model reliably a housing stock contains basic typological information including knowledge about building construction characteristics and floor areas, as well as current refurbishment state that greatly affects the overall heat transfer coefficient of the building envelope. Uncertainty in input parameters is decisive and may exceed uncertainties that are induced by model simplifications.

The availability of measured energy data at large scale enables the application of simplified thermal models on housing stocks to gain insights concerning building characteristics and indoor conditions. The heating setpoint temperature is key assumption of occupant behaviour that determines heating demand in cold climates. The use of an urban dataset to determine temperature setpoints and thermal transmittance for thousands of buildings is demonstrated. The distributions of these parameters at urban scale can be particularly beneficial for urban building energy models and for determining energy refurbishment measures that have been applied to the thermal envelope. The use of this thermal modelling approach at urban scale can be adopted by utilities and authorities that aim at estimating or validating basic characteristics of buildings at urban or national level.

The application of data mining methods to smart meter data unveils hidden correlations between energy demand and its influencing factors, as well as identified consumer behaviour patterns and variability over time. A clustering-based analysis is proposed that consists of three main modules: data preparation, clustering and load profiling analysis. District heating residential customers are classified into different groups based on their consumption intensity and representative patterns. Typological building characteristics such as building age and floor areas are found to be more correlated with heating data rather than family size and age of occupants, hence they can still be used in new classification schemes for housing stocks with regards to energy use intensity. The high energy intensive groups of the examined consumers have more predictable and regular behaviour than the low energy intensive ones and are therefore suitable for predictive analysis.

In the smart energy city context, where the increased penetration of volatile energy sources calls for balancing the supply and demand sides, buildings can actively contribute to the stabilization of the grid. The thermal flexibility potential that the building stock can offer to the grid varies significantly mainly based on the thermal transmittance of building envelope, heat capacity of internal wall mass, solar gains and ambient temperature. The quantification of thermal flexibility potential is done in terms of the degradation of indoor comfort that people perceive after a specified heating control strategy and the energy savings or peak load that are created after this event. Optimized scheduling of a grid-connected heat pump is proposed based on a price signal that reflects the lack or excess of renewables to the grid, which can lead to peak shaving and cost savings that are highly dependent on the input prices.

When optimizing the whole energy system operation, the utilization of the thermal mass of buildings for heat storage is introduced as an additional source of flexibility to the system. An archetypal approach to characterize the building stock according to the construction age and refurbishment level is employed to better represent the respective flexibility potential of each building category. A flexibility indicator representing the viability of the different scenarios of the system setup is introduced and shows that the utilization of the building thermal mass would be more beneficial to future energy systems than current district heating systems, due to the larger capacities of intermittent generation that can be successfully integrated in the district heating supply.

Resumé

Energi er et af de væsentligste omdrejningspunkter for smarte byer, og for byers smarte måder at indrette sig med hensyn til at blive miljørigtige og for borgernes smarte måder at indrette deres liv på. Den rolle, som bygninger kan spille i udvikling og drift af smarte byer, er vigtig på grund af den store andel af energiforbruget, som bygninger er ansvarlige for, og på grund af de smarte energiløsninger, der er mulighed for at integrere i bygninger. Det stigende antal initiativer om smarte byer og deres fokus på energistyring på byniveau har understreget behovet for at flytte fra den traditionelle energimodellering på et niveau omkring hver enkelt bygning, til udviklingen af aggregerede modeller for energibehov. For at opnå dette er der behov for metoder, der kan skaleres til højere aggregeringsniveauer, lige fra klynger af bygninger til kvarterer og byer.

Denne PhD-afhandling sigter på at tilvejebringe forbedrede modelleringsmetoder, der kan give forbrugsprofiler for bygninger med høj rumlig og tidsmæssig opløsning, som kan bidrage til at evaluere energipolitikker og strategier for forbrugsstyring. Et hovedformål med afhandlingen er at foreslå og udforske ingeniørbaserede tilgange, statistiske metoder og teknikker til datamining, som kan bidrage til nøjagtig bygningsmodellering på byskalaniveau for at fastslå det potentiale, som bygninger kan have på stabilisering af energinettet og på energisystemets fleksible drift. Herunder skal der tages hensyn til en kontekst for smarte byer - både på bygnings- og bydelsniveau; og endelig skal egnetheden af hver kategori af foreslåede modelleringsmetoder kunne angives sammen med retningslinjer for fremtidige undersøgelser.

De vigtigste resultater af den forskning, der udføres i denne afhandling, er opsummeret i følgende:

Bygningers samlede energibehov kan bedst vurderes ud fra betragtning af arketyper, der reducerer modellerings- og beregningstiden betydeligt. Et minimumsniveau af informationer, der skal til for pålideligt at udforme en bygningsmasse, indeholder grundlæggende typebaserede oplysninger, herunder karakteristika for bygningskonstruktioner og bebyggede arealer samt oplysninger om den nuværende renoveringstilstand, der i høj grad påvirker bygningers samlede varmetransmissionskoefficient. Usikkerhed i inputparametre er afgørende og kan overstige de usikkerheder, der knytter sig til modelforenklinger.

Tilgængeligheden af målte energidata i stor skala muliggør anvendelse af forenklede termiske modeller for en samlet bygningsmasse, så der kan opnås indsigt i bygningskarakteristika og indendørsforhold. Setpunktstemperaturen for opvarmning er en nøgleparameter for beboeradfærd, der har indflydelse på opvarmningsbehovet i kolde klimaer. Brugen af et datasæt for byer til bestemmelse af setpunktet for indendørs temperatur og varmetransmittans for tusindvis af bygninger er demonstreret. Fordelingen af disse parametre på byskala kan være særligt gavnligt for energimodeller for bygning i byer og til bestemmelse af foranstaltninger til energieffektivisering, som er udført på bygningernes klimaskærm. Anvendelsen af denne forenklede termiske modelleringsmetode på byskala kan benyttes af

forsyningsvirksomheder og myndigheder, når de skal estimere eller validere de grundlæggende egenskaber ved bygninger på by- eller nationalt plan.

Anvendelse af metoder til datamining ud fra data fra smarte målere kan afsløre skjulte korrelationer mellem energibehov og de faktorer, der påvirker det, samt til at identificere adfærdsmønstre blandt slutbrugere og deres variation over tid. En klyngebaseret analyse foreslås, der består af tre hovedmoduler: Databehandling, klyngedannelse, og analyse af belastningsprofiler. Fjernvarmekunder klassificeres i forskellige grupper baseret på intensiteten af deres forbrug og repræsentative mønstre for forbruget. Typeinddelte bygningskarakteristika, såsom bygningsalder og etageareal, er mere korreleret med opvarmningsdata end data såsom familiestørrelse og beboernes alder, og de kan derfor anvendes til nye klassifikationsmodeller for bygningsmassen med hensyn til energiintensitet. Blandt de undersøgte forbrugere har grupper med en høj energiintensitet mere forudsigelig og regelmæssig adfærd, og er derfor egnede til prognostisk analyse.

I sammenhæng med smarte byer, hvor den øgede anvendelse af fluktuerende energikilder kræver balance mellem udbuds- og efterspørgselssiden, kan bygninger aktivt bidrage til at stabilisere nettet. Det termiske fleksibilitetspotentiale, som bygningsmassen kan tilbyde til nettet, varierer væsentligt og hovedsageligt på grund af klimaskærmens varmeisoleringsevne, varmekapaciteten af de indvendige bygningsdele, soltilskud og omgivelsestemperatur. Kvantificering af potentialet for termisk fleksibilitet er gjort med hensyn til den mulige forringelse af den indendørs komfort, som folk kan opleve ved indførelsen af en bestemt strategi for regulering af opvarmningen og ved at vurdere energibesparelser eller spidsbelastninger, der opstår efter en sådan regulering. Optimal planlægning af en varmepumpe, der tilkobles nettet og styres efter et prissignal, der afspejler mangel på eller overskud af vedvarende energikilder i nettet, kan føre til spidslastreduktioner og omkostningsbesparelser, der er yderst afhængige af de forudsatte priser.

Ved optimering af hele energisystemets drift indføres udnyttelse af den termiske masse af bygninger til varmelagring som en ekstra kilde til fleksibilitet for systemet. En arketypisk tilgang til karakterisering af bygningsmassen efter bygningsalder og renoveringsniveau anvendes for bedre at repræsentere hver bygningskategoris potentiale for energifleksibilitet. En fleksibilitetsindikator er indført til at repræsentere egnetheden af de forskellige scenarier for systemopsætning, og det viser at udnyttelsen af bygningens termiske masse kan være mere gavnlig for fremtidens energisystemer end i de nuværende fjernvarmesystemer på grund af den større kapacitet af varierende energiforsyning, der med succes kan integreres i fjernvarmeforsyningen.

Table of contents

List of figures

List of tables

List of abbreviations

ACH	Air change per hour
ANN	Artificial neural networks
ASHRAE	American Society of Heating, Refrigerating and Air-Conditioning Engineers
BBR	Danish Building Register
BIC	Bayesian Information Criterion
CDA	Conditional demand analysis
CO	Heat cut-off
CO_2hPH	Heat cut-off following two hours of preheating
CO_4hPH	Heat cut-off following four hours of preheating
DR	Demand response
DSM	Demand side management
ELM	Electric load management
EUI	Energy use intensity
GHG	Greenhouse gas
GIS	Geographic information systems
HP	Heat pump
i.e.	That is
ICT	Information and communication technologies
IEA EBC	International Energy Agency, Energy in Buildings and Communities
IT	Information technology
KDD	Knowledge discovery in databases
MOBO	Multi-objective building performance optimization program
MPC	Model predictive control
OLS	Ordinary least squares
PNN	Probabilistic neural networks
SFH	Single-family houses
UBEM	Urban building energy model
UMI	Urban modeling interface

List of publications

The following list includes the list of journal and conference papers produced by the author in the context of this PhD thesis.

Journal articles

I. **Gianniou, P.**, Liu, X., Heller, A., Sieverts Nielsen, P., Rode, C. (2018). Clustering-based Analysis for Residential District Heating Data. *Energy Conversion and Management 165*; 840–850.

II. **Gianniou, P.**, Reinhart, C., Hsu, D., Heller, A., Rode, C. (2018). Estimation of Temperature Setpoints and Heat Transfer Coefficients Among Residential Buildings in Denmark Based on Smart Meter Data. *Submitted to Building and Environment*, January 2018. Accepted.

III. Franjo Dominkovic, D., **Gianniou, P.**, Münster, M., Heller, A., Rode, C. (2018). Utilizing Thermal Building Mass for Storage in District Heating Systems: Combined Building Level Simulations and System Level Optimization. *Energy 153*; 949-966.

Conference papers

IV. **Gianniou, P.**, Heller, A., Nielsen, P. S., Negendahl, K., & Rode, C. (2015). Aggregation of Building Energy Demands for City-scale Models. In Proceedings of *Building Simulation 2015*, Hyderabad, India.

V. **Gianniou, P.**, Heller, A., & Rode, C. (2015). Building Energy Demand Aggregation and Simulation Tools: a Danish Case Study. In *Proceedings of CISBAT 2015*, Lausanne, Switzerland.

VI. **Gianniou, P.**, Foteinaki, K., Heller, A., & Rode, C. (2017). Intelligent Scheduling of a Grid-Connected Heat Pump in a Danish Detached House. In Proceedings of *Building Simulation 2017*, San Francisco, United States.

VII. Sarran, L., Foteinaki, K., **Gianniou, P.**, & Rode, C. (2017). Impact of Building Design Parameters on Thermal Energy Flexibility in a Low-Energy Building. In Proceedings of *Building Simulation 2017*, San Francisco, United States.

VIII. Zilio, E., Foteinaki, K., **Gianniou, P.**, & Rode, C. (2017). Impact of Weather and Occupancy on Energy Flexibility Potential of a Low-energy Building. In Proceedings of *Building Simulation 2017*, San Francisco, United States.

IX. **Gianniou, P.**, Heller, A., Nielsen, P. S., & Rode, C. (2016). Identification of Parameters Affecting the Variability of Energy Use in Residential Buildings. In Proceedings of the *12th REHVA World Congress*, *CLIMA*, Aalborg, Denmark.

List of publications that are not included in the thesis

The following publications were produced as part of the current PhD project, but were not the main focus of it. Hence, they are only attached at the end of the thesis.

X. Heller, A., Liu, X., & **Gianniou, P.** (2017). A Science Cloud for Smart Cities Research. *Energy Procedia 122*; 679-684.

XI. Liu, X., Nielsen, P. S., Heller, A., & **Gianniou, P.** (2017). SciCloud: A Scientific Cloud and Management Platform for Smart City Data. In Proceedings of *28th International Workshop on Database and Expert Systems Applications (DEXA)* (pp. 27-31).

1 Introduction

Energy is central to most challenges the world is facing today, hence being an important goal in the sustainable development agenda of United Nations [1]. Buildings are one of the key sectors to be addressed towards achieving the environmental and energy goals set by International Directives. In particular, they can influence 42% of final energy consumption and 35% of total greenhouse gas (GHG) emissions. Furthermore, they can contribute to significant resource savings, such as water and extracted materials [2]. At the same time, the building sector plays an important role on energy hubs, which are defined as units where multiple energy carriers can be converted, conditioned, and stored. Energy hubs can be considered as interface between different energy infrastructures and loads [3]. Therefore, they are represented as complex and multi-scaled systems. This creates a challenge in modelling them and calls for advanced and computationally sophisticated methods [4].

Modelling building energy demands has received particular interest during the last decades. Many methods have been developed towards this from a physics, statistical and data-mining perspective. Some of them consider the building physics and dynamics taking place into buildings, while others investigate the correlations among building-related variables and extract hidden information from data.

The development of sustainable future energy systems calls for solutions that can potentially have a large impact on the energy system. In this context, integrated solutions at urban scale is seen as a promising approach, due to the density of demands and technical infrastructure that they can accommodate. It is expected that such efforts can be cost-effective due to the availability of technologies and the large density of population, although complexity and installation costs in urban areas can be impediments to their application. The current work has started at a time, where such issues were not fully raised and this rather new concept was put under the microscope. This is one of the topics that this thesis aims to investigate, being an early attempt in the subject of smart energy cities.

Energy is one of the major drivers in smart cities along with smart environment and smart living. The integration of energy systems and technologies into the buildings is described by the concept of smart energy cities. Therefore, smart energy cities refer to cities which can improve present systems and implement new technologies in a coordinated and optimal way, by optimizing the synergies among all energy solutions [5]. Buildings occupy a key place in the development and operation of smart energy cities, since they are responsible for a large share of energy use and represent a large potential to integrate smart energy solutions.

Smart cities initiatives and their focus on city level energy policy management has emphasised the need to move from the traditional micro-level building energy modelling towards the development of aggregated energy demand models. This calls for methods that can be scalable to higher levels of aggregation, ranging from clusters of buildings to neighbourhoods and cities. The need to explore these different levels of aggregation has led to research endeavours like IEA EBC Annex 51 (Energy Efficient

Communities), Annex 63 (Implementation of Energy Strategies to Communities) and Annex 67 (Energy Flexible Buildings). The current project is executed parallel to these efforts and has had some contribution to the latter.

This PhD Thesis aims at investigating the different approaches that characterize building energy demand modelling at large scale. Such models are focal for the operation of energy grids, electrical, gas and thermal. The models are also pivotal with respect to new establishment of energy infrastructure in planning and designing. The main objectives of the thesis are to present available modelling techniques, analyse their applicability on the aggregation level by applying them on real world cases, and give directives for the utilization of these models for different purposes. Hereby, the thesis is an early attempt to carry out research in a field, where path finding approaches were the most important findings with respect to setting directions for future research.

1.1 Research hypothesis and problem statement

The main hypothesis underlying this thesis is that *the development of building energy demand models at aggregate level is feasible and reliable and that it can be introduced into current practice for energy system planning and operation processes at a justifiable effort*. This hypothesis was addressed using models, information and data that enable the modelling at large scale as well as through a comparison of all the respective findings.

<u>Objectives</u>

1. The first objective of this thesis is to propose methodologies that prove how **dynamic and steady-state energy simulations** can contribute to the accurate building energy demand modelling at urban scale.
2. The second objective of this thesis is to test how **data-driven methods** can model housing stocks with regards to energy demand.
3. The third objective of the thesis is to model the **potential that buildings can have on the stabilization of the energy grid** and the flexible operation of the energy system - considering a smart energy cities context- both at building level and urban level.
4. The fourth objective is to **identify the strengths, limitations** and suitability of each category of proposed modelling methodologies and to provide guidelines for future investigations.

The four objectives collectively lead to the answer of the research hypothesis. A vision for an enhanced urban energy modelling and its applicability to real cases is provided. The relevance of this thesis is that enhanced modelling methodologies can provide building energy demand related insights in a high spatial and temporal resolution, which can help to evaluate energy policies and demand side management strategies.

1.2 Guide to the reader

This part presents the outline of the thesis, which consists of five different chapters. The first chapter, which is the introduction, presents the motivation, background, problem formulation leading to the general framework and the formulation of the research hypothesis of the thesis including the objectives that are going to be addressed. The second chapter provides the necessary background and state of the art to the reader to follow the project. In particular, the smart cities context is presented together with the building stock modelling approaches that brings the building energy modelling to the "aggregated" city level. Hereby, the role of data in smart cities, as well as the concept of energy flexibility are central concepts. The third and fourth chapters summarize the findings of the publications produced along the study. In particular, the third chapter includes the findings with regards to building energy modelling and corresponding aggregation methods, while the fourth chapter presents the findings on the contribution of buildings to energy flexibility, one of the promising solutions to stabilize energy grids in general. The fifth chapter discusses the overall conclusion of the thesis, as well as sets the outlook and new perceptions of this work. Detailed conclusions from the studies are given in the respective studies and in the publications. The sixth chapter concludes the thesis by proposing future investigations that would continue this work. The publications that were produced in the context of this work are included in the Appendix of the thesis.

Table 1 illustrates an overview of the publications produced during this research project and relates them to the objectives of the thesis. Furthermore, the scale of the analysis is presented and the indication of development of models. It should be noted that the fourth objective is mostly answered in the conclusion chapter taking into account all findings of the thesis. The last two publications presented in Table 1 do not address the objectives of the thesis, but present a framework and propose a platform for the integration of smart city data. Hence they are not included in the main body of the thesis.

Table 1. Overview of publications in the thesis

Paper	Main focus	Objective	Model development	Scale analysis
Paper I	Clustering-based Analysis for Residential District Heating Data	#2	Yes	Urban
Paper II	Estimation of Temperature Setpoints and U-values Among Residential Buildings	#2	Yes	Urban
Paper III	Thermal Mass Utilization for Storage in District Heating Systems	#3	Yes	Urban
Paper IV	Aggregation of Building Energy Demands for City Scale Models	#1	Yes	District
Paper V	Building Energy Demand Aggregation and Simulation Tools	#1	Yes	District
Paper VI	Intelligent Scheduling of a Grid-Connected Heat Pump	#3	Yes	Building
Paper VII	Impact of Building Design Parameters on Thermal Energy Flexibility	#3	Yes	Building
Paper VIII	Impact of Weather and Occupancy on Energy Flexibility Potential	#3	Yes	Building
Paper XI	Identification of Parameters Affecting the Variability of Energy Use	#1*	Yes	Building
Paper X	A Science Cloud for Smart Cities Research	**	No	Urban
Paper XI	A Scientific Cloud and Management Platform for Smart City Data	**	No	Urban

* Introductory work for dynamic energy modelling

** Give a perspective towards the integration of data in smart cities

1.3 Motivation

This PhD thesis is motivated by the need of the energy community to understand the impact of housing stocks onto the energy system, especially with respect to transition towards smart energy cities, while accomplishing the energy and environmental goals. Furthermore, the balancing of the smart grid challenged by volatile energy sources and the contribution of buildings to achieving stable energy grids are main motives for this research. The focus of the project lies primarily on the residential building sector and specifically on single-family houses, which consist a large share of the building stock in Denmark. The aim is to understand the impact of clusters of buildings on energy infrastructures through modelling. Hereby, models will play a major role in the analytical part of the thesis. Models will be classified on theoretical basis and with regards to the application that is determined by our knowledge of the buildings, the objective of the modelling (e.g. planning or optimization) and systems in focus (information level). Heating demand is the main attribute of building energy demand in Denmark. Cooling demand is therefore out of the scope of the thesis.

Figure 1 presents the modelling flow of the thesis. The models that are developed are dependent on data and theory, and are used to run simulations and data-driven analysis. The outputs of these are used to conduct optimization, prediction and classification. The results of the thesis can be used to facilitate decision-making and planning processes.

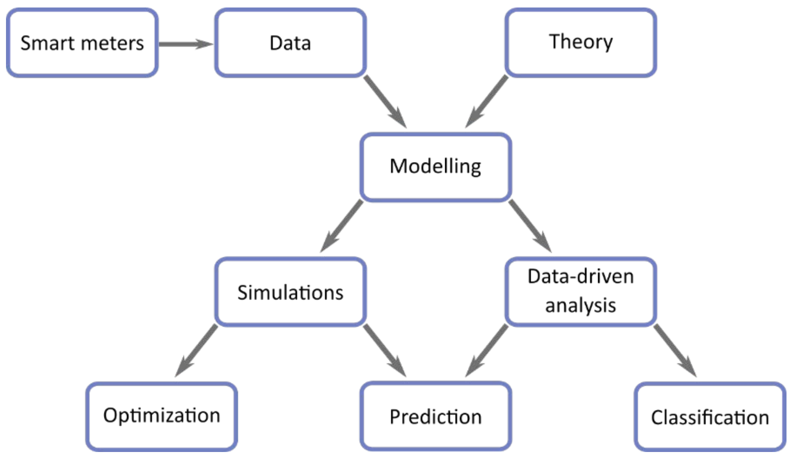


Figure 1. Modelling flow of thesis

2 Background

During the past decades, European and international directives have been introduced to support the reduction of greenhouse gas emissions. An overall reduction of 50% in global greenhouse gas emissions (GHG) by 2050 compared with 1990 values has been proposed by science [6]. The energy sector is crucial to meeting climate targets. According to the International Energy Agency, global energy consumption is expected to increase approximately 30% until 2035 [7]. Buildings are responsible for 40% of energy consumption and 36% of CO_2 emissions in the EU [8]. This results in the building sector playing a key role in effective climate policy.

2.1 Urbanization and smart cities

At the same time, with the rapid increase of urbanization, cities have become the pivotal centres of our lives. More than three-quarters of global energy production is used within cities today, while they are the main generators of GHG emissions, as well. In order to attain durable and liveable future cities, sustainability is seen as the only promising solution. Thus, apart from being seen as a part of the problem, urban environments should be treated as a part of the solution to achieve sustainable development [9]. Climate change mitigation measures are proven to be more cost-effective and efficient in large and compact cities than in low-density areas [10]. This makes cities an ideal platform, where sustainable solutions will be applied. More than four thousands cities and towns have signed the 'Covenant of Mayors', which is a climate action and renewable energy package that aims at reducing the CO_2 emissions by 20%, and increasing energy efficiency and utilization of renewable energy resources by 20% by 2020 [11]. New signatories even pledge to reduce CO_2 emissions by 40% by 2030.

To comply with the environmental policy goals and to make extensive use of information and communication technologies (ICT), new urban development concepts have been created in the last decades. One of them are smart cities. The term has been widely adopted to represent the '*application of complex information systems to integrate the operation of urban infrastructure and services, such as buildings, transportation, electrical and water distribution and public safety*' [12]. The increasing interest towards the smart cities development has originated from the Smart Growth Movement of the late 1990's. Smart cities exhibit a high potential to contribute significantly to the achievement of energy and climate targets [13]. The main characteristics of a smart city are smart environment, smart people, smart mobility, smart governance, smart economy and smart living [14]. In addition, a smart city should be able to optimize the use of its tangible and intangible assets [15]. Thus, buildings occupy a key place in the development of smart cities as they represent an important potential to integrate smart energy solutions.

2.2 Smart energy cities

The concept of smart energy cities refers to *cities which can improve present systems and implement new technologies in a coordinated and optimal way, by optimizing the synergies among all energy solutions*

[5]. In other terms, they may include smart energy systems [16] or smart energy networks [17]. Smart energy systems differentiate from smart grids in the fact that they integrate more sectors than just the electricity sector, such as heating, cooling, buildings etc. and promote the integration of all sectors and infrastructures. They incorporate energy solutions that can lead to future renewable and sustainable energy systems [18]. The integration of smart energy solutions and technologies at different spatial scale (building, block, district, city) can be achieved in sectors such as buildings, energy and ICT infrastructure, transportation, as well as data infrastructure. Examples of applied technologies in each domain can be found in [19]. The basic elements that a smart energy city consists of are presented in Figure 2.

Smart energy city projects have become particularly popular during the last years, due to support by both the European Union (e.g. Horizon 2020) and the private sector [20]. Such examples can be found in all over the world. In Denmark, smart city vision is very popular, where several Danish municipalities have defined relevant strategies and participated in smart city projects (e.g. Copenhagen, Vejle, Kalundborg, Sønderborg, Randers). A large number of these activities begun in the run up to the UNFCCC COP 15 meeting in Copenhagen in December 2009, when many cities launched their own voluntary climate strategies [21]. The EnergyLab Nordhavn is an example of smart energy city project based in Copenhagen, Denmark, which will provide home and work for 40,000 people within the next 50 years [22]. It is supported by the Energy Technology Development and Demonstration Program (EUDP), which is a public grant scheme. This scheme supports new technology in the energy field, which can help meet Denmark's energy and climate goals. In France, the EcoCité initiative aims at facilitating the transition towards more sustainable urban development, with 13 cities being part of it [13]. The Morgenstadt initiative in Germany includes 14 cities which have come together to realize the future of integrated, sustainable and resilient cities.

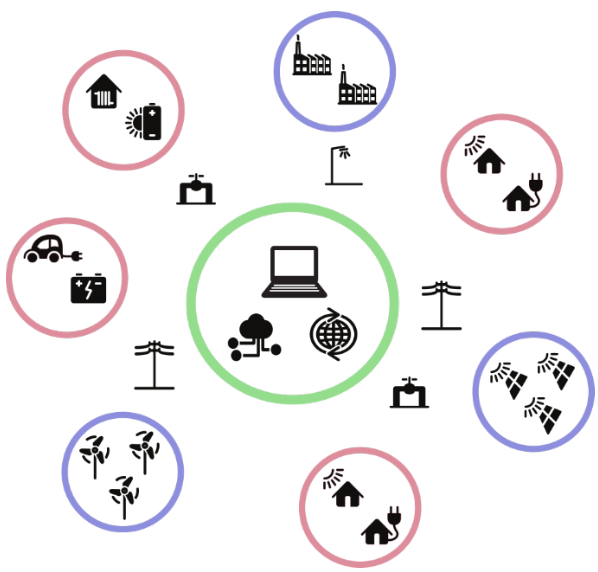


Figure 2. Elements of a smart energy city

2.3 CITIES project

This thesis is part of the Centre for IT-Intelligent Energy Systems in cities (CITIES) project, which is a research project launched in January 2014 as a 6-year research centre which has been financially supported by a range of industrial and academic partners, and the Danish Council for Strategic Research. CITIES aims at developing methodologies and solutions for the achievement of fully integrated urban energy systems, which will be entirely fuelled by RES. This thesis is linked to the activity of Working Package 3 entitled 'Intelligent Energy System Integration'. The objective of this Working Package is to optimize the interactions and synergies between individual system components, amongst buildings. Figure 3 shows the spatial scale and the complexity of the problems that are in the CITIES research agenda. The interactions among the thermal, power, biomass, fuel and gas sectors are modelled with the utilization of models, data, optimization techniques and communication flows. The modelling of building energy demands plays an important role since it reflects the central role of the building stock in an integrated urban energy system. The main focus of the third Working Package lies on the residential building stock, since it accounts for 64% of the total heat demand in Denmark. Cooling demand is not part of the investigation of the project.

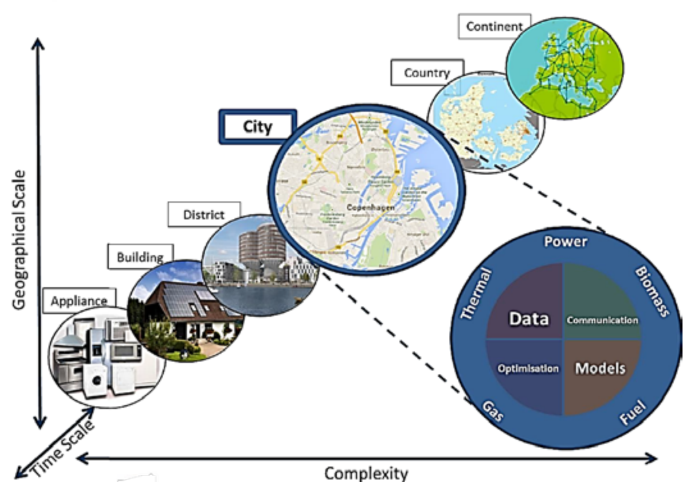


Figure 3. Spatial scale and complexity of CITIES focus [23]

2.4 Modelling approaches of building stock

A number of modelling approaches have been developed to obtain information on physical resource flows through the building stock [24]. These methods have been mostly used for energy demand prediction, characterization, analysis and benchmarking of housing stocks, as well as to evaluate the impact of specific energy renovation measures or installation of energy technologies. The modelling techniques for residential building stock can be classified into two categories: *bottom-up* and *top-down approaches*. Bottom-up approaches calculate the energy consumption of individual buildings or clusters of buildings and extrapolate these results to a wider region based on the representative weight of the modelled sample [25]. They include models that use input data based on a hierarchical level and account for individual end-uses. The main advantage of bottom-up approaches is the high level of detail with which they can model energy technologies. This enables the identification of targeted areas, which either perform well exhibiting a high potential for integration of specific energy solutions, or are problematic, so they need to integrate specific energy conservation measures. However, they require an extensive amount of input data to model accurately the building stock, which range from building-related parameters, such as floor areas, construction materials, insulation thickness and infiltration rates to occupant-related parameters. These are not always available and are quite difficult to be collected and processed especially when the analysis includes many buildings. Thus, many assumptions need to be made. Bottom-up approaches combined with geospatial data, which are becoming more commonly available, can provide useful tools to model building stocks [26, 27].

On the contrary, top-down approaches do not distinguish energy use due to individual end-uses [25]. Their main objective is to determine supply requirements by determining the effect of changes in the residential sector onto energy consumption. They are based on historical or statistical data and economic theory. Top-down approaches are mainly categorized into econometric and technological. Econometric models emphasize on price and income, while technological models correlate energy use with general characteristics of the building stock. Hybrid top-down models have also been used in literature. The main benefit of top-down models are the less detailed data that are required, usually at aggregate form, as well as their simplicity. Nevertheless, the historical data that are needed do not support modelling of discontinuous technological changes or planning processes, which is a drawback for this type of approaches. The main strengths and limitations of these approaches are summarized in Table 2.

Table 2. Characteristics of bottom-up and top-down modelling approaches

Approach	Benefits	Limitations	Categories
Bottom-up	- Model new or existing energy technologies in high detail - Determine end-use energy contribution - Identify targeted areas to apply specific strategies - Use of physically measurable data	- Require big amount of input data and information - Computationally intensive - Poorly describe market interactions	- Physics-based methods - Data-driven methods
Top-down	- Capture the economic and energy macro-scale - Consider the effect of economic variables on energy - Determine the impact of changes onto the energy sector - Make use of aggregated data - Enable the modelling of social cost-benefit policies - Simplicity of input information	- Require historical data - Project future trends based on past events - Cannot model specific technologic-related strategies - Cannot model specific end-users	- Econometric methods - Technological methods

2.5 Categorization of bottom-up models

Bottom-up approaches for housing stocks include a plethora of modelling techniques which have been studied extensively. These can be classified into two main categories: engineering methods and statistical techniques. Specifically, bottom-up energy models are characterized as *physics-based* (also known as white-box models) and *data-driven* (black-box or grey-box models). They have been also found in literature as forward or classical and inverse modelling [48]. All of the above terminologies differentiate the models based on their capability to describe physical phenomena and on their dependence on data. These categories are presented in the following. Evaluation methods for non-residential buildings follow a similar classification [47].

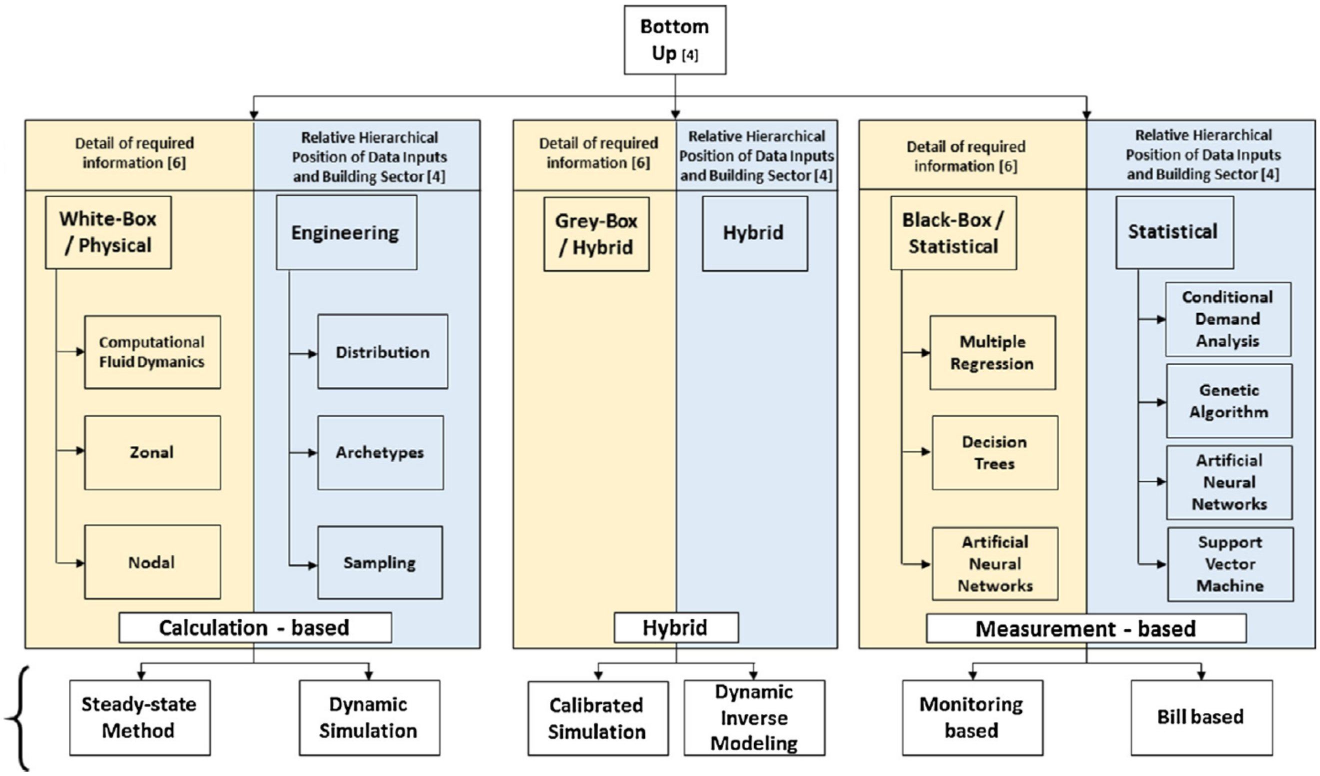


Figure 4. Categorization of bottom-up models [28]

2.5.1 Physics-based methods

Physics-based models predict the output variables of a specified model given specified input variables [48]. Thus, they explicitly consider energy use of end-uses based on use of equipment and systems or heat transfer and physics equations [25]. They consist of a physical description of the system or model and a core calculation engine. The input data that are required to describe physically these models regard the building envelope characteristics, thermal comfort conditions, number of occupants, energy systems installed in the buildings and their efficiency, operation of these systems, heating and cooling patterns and many more. Thus, white-box models rely on detailed housing information. The amount of the input data is defined by the complexity of the model and the method. These data can be collected from personalized visits to the buildings, technical drawings, nationwide building databases or housing surveys. Despite the amount of information available in these sources, some data are unavailable and specific assumptions need to be made. Nevertheless, the combination of physics phenomena and empirical data enable the building energy performance estimation for the past, present and future [29]. One of the main limitations of physic-based models is the uncertainty in input data. Occupant behaviour is a leading factor in uncertainty of such models especially when determining the applicability of building energy retrofits [30]. Neglecting the aspect of stochasticity can lead to severe miscalculations and inaccurate estimates of building energy performance [31].

Concerning the energy calculation, the physics-based models can be classified into dynamic and quasi-steady-state methods. According to the international standard on building energy performance (ISO13790) [32], three different methods of calculating thermal loads are proposed which calculate: i) monthly and seasonal results (quasi steady-state methods), b) simple hourly results (dynamic methods) and c) detailed simulation results (dynamic methods). Quasi steady-state methods calculate heat balance for a long time (e.g. monthly) considering dynamic phenomena through correlation factors. To estimate heating, a utilization factor is usually introduced, which considers that only a part of the internal and solar heat gains can reduce the heat demand, while the rest of the heat gains result in an increase of the room temperature above the setpoint. Quasi-steady-state methods are employed by steady-state building performance simulation tools that aim at providing estimations of energy demands at a low temporal resolution (i.e. monthly or annually). On the other hand, in dynamic methods heat balance is calculated with short time steps (e.g. hourly or variable) considering transient phenomena and the thermal inertia of the building. Dynamic methods model explicitly thermal transmission, internal and solar heat gains, heat flow by ventilation, as well as thermal storage. They differentiate from the rest of the modelling methods in the way that they handle changing surroundings by computing the system's characteristics in very short time-intervals [33]. Dynamic methods are usually employed by dynamic building performance simulation tools (e.g. IDA ICE, EnergyPlus, TRNSYS, IES-VE).

2.5.2 Data-driven methods

Data-driven methods mostly represent statistical techniques and data-mining approaches. Statistical methods make use of historical data to attribute household energy consumption to particular end-uses. Their main advantage is that they usually include macroeconomic and socio-economic effects, estimate end-use energy consumption and are easy to develop and be used [29]. However, their main limitation is that they require historical data and cannot interpret the effect of specific energy technologies or energy conservation measures. Data-driven methods that have been commonly applied to a sample of houses can be roughly classified into three categories: regression analysis, conditional demand analysis and neural networks [25]. A regression analysis refers to the prediction of a continuous dependent variable from a number of independent variables. These methods are usually based on long term load measurements and weather data [34]. They are evaluated on the basis of goodness of fit. Based on the selection of variables, the models may or may have not physical significance. Conditional demand analysis (CDA) refers to regression-based methods that are used to estimate energy consumption of households as a function of appliance holdings [35]. Thus, the total consumption can be broken down into constituent end-use components. Due to statistical considerations, the number of variables included in the models often has to remain small, so the effect of many factors cannot be evaluated, such as socio-economic parameters [44]. Furthermore, they are mostly less flexible than other statistical or engineering approaches. Artificial neural networks (ANN) include a set of computational units and a set of data connections joining units [36]. Their operation is similar to human brain, by retrieving or modifying information when some factors are missing. Thus, they can be considered as black-box models, which generate their outputs by hidden layers from the inputs, without including any information on the physical characteristics of the building [37]. Neural networks can be applied from individual homes to national building energy forecasting [38]. A plethora of neural networks has been developed and applied to various fields including feedback artificial neural networks, feed forward neural networks or probabilistic neural networks (PNN). ANN often outperform other traditional statistical methods in energy consumption prediction because of their ability to handle nonlinear patterns with high computing speed and high accuracy [39, 40].

The data-driven algorithms that have been most commonly applied to building stocks can be classified into two main categories with regards to their objective: prediction and classification [41]. The predictive methods include artificial neural networks, support vector machines, statistical regression, decision trees and genetic algorithms. The most common methods used for classification are *K*-means clustering, hierarchical clustering and self-organizing maps. Data mining approaches are further analysed in section 2.8.1.

2.5.3 Hybrid models

Hybrid methods have been developed to overcome the limitations of the above-presented methods. They generally belong to the category of data-driven methods. The main characteristics of them are that they combine the physical representation of the models with the employment of statistical methods to estimate and predict energy loads. Hybrid methods also rely on empirical data, such as hot water demand per person [29]. They have been used extensively on a wider scale ranging from clusters of buildings to nationwide analyses. Grey-box or stochastic state-space models can be considered as a type of hybrid methods. They consist of a set of stochastic differential equations describing the building dynamics in a continuous time and a set of discrete time measurement equations [42]. Opposed to black-box models, grey-box models often include fewer parameters, which are more meaningful and valid over wider ranges of state spaces. Other hybrid models may include adaptive time-series techniques or auto-regressive models, which can be automatized to a high degree [43]. Grey-box models are criticized for their limited ability to represent accurately transient heat transfer phenomena. Nevertheless, their fast computation times make them particularly attractive for determining control-based actions.

2.6 Heat transfer modelling for buildings

Building energy modelling should not be confused with heat transfer modelling that is the basic theory applicable to most energy domains, including buildings. When investigating heat transfer in buildings, two modelling approaches are found in literature: the mathematical and the physical ones. The mathematical approaches are based on the differential equation of an *n*-node model employing *Euler* or *Fourier* series methods to solve them [45]. The physical approaches include explicit solutions of the heat diffusion equations, model reduction techniques and model simplification techniques. The explicit solutions include methods such as finite difference, finite volume or transfer function. They often require extensive data collection and analysis, thus increasing the complexity of equations when involving more than one building model [46]. Model reduction techniques represent building performance with fewer sets of equations, which however result in repeating reduction procedures, when the definition of an integral energy system is changed. Model simplification techniques are employed by grey-box models, which were previously presented. In particular, they use analogies to represent physical phenomena in a similar way as network models. They include the resistance-capacitance models and admittance methods. Electrical analogies are used to represent heat transfer. Temperature nodes represent the conduction in building elements, resistors indicate heat transfer, while capacitance stands for the thermal inertia. The number of elements that is selected to represent the model affects its complexity and accuracy. Models that include a small number of elements are called reduced-order ones and are easier to be solved.

2.7 Aggregation

The study of urban energy systems calls for efficient modelling and simulation techniques. Aggregation is such a technique that provides an important step towards estimating building energy demands at city level. The first step is to analyse individual buildings' energy demands and then continue with an upwards aggregation, which can be used to determine the energy demand of the whole building stock [65]. Thus, it can be characterised as a bottom-up approach. The estimation of building energy demands and their extrapolation to a larger scale in engineering methods can be achieved with the use of distributions, archetypes and samples [25]. Archetypes are representative buildings that can describe a building stock. They can be also found in literature as reference or example buildings. They are designated as representing actual buildings of the existing stock [49]. They can be statistical composites of features that characterize specific building types, usually derived from available data for national building stocks [50]. These features can be classified into four broad categories: geometry, fabric, equipment and operation [51]. This is an intensive procedure which requires efforts in data collection, preparation and analysis. Capturing the variations in energy use is crucial in stock aggregation. According to the findings of IEA EBC Annex 53 [52], these variations are mainly attributed to the following areas: building envelope and equipment, operation and maintenance, occupant behaviour and indoor environmental conditions, as well as outdoor climate. Selecting the parameters upon which the stock aggregation is based can be a hard procedure, which many techniques have been employed for. Some of them include sensitivity analysis, parametric analysis, statistical methods and clustering analysis. In general, the archetypal approach simplifies the diversity of construction variations, thus reducing the variety of predicted building energy demands [53].

<u>Archetypal approaches</u> can be used for supporting policy development or implementation of energy-related programs [47]. Furthermore, they can facilitate decision-making in urban development projects, while they are not directly affected by poor data quality since they mostly make use of mean values [54]. Archetypes can be as detailed as desired, but this would correspond to a higher number of archetypes needed to represent the building stock [55]. However, the potential benefits coming from a more detailed description of the building stocks through archetypes should be weighed against the increased complexity of the data structure [55]. The reliability of the archetype-based urban-scale building energy models can be evaluated based on the deviation of the predicted results from measured results, similar to the

validation process of white-box models. This deviation or error ranges between 7% and 21% for heating loads when comparing aggregated annual simulated versus measured energy use as found in reported studies, since the inaccuracies are averaged out [65]. This error is significantly higher when using disaggregated results.

The approaches that use archetypes to model building stocks mainly comprise of three stages: segmentation, characterization and quantification [49]. During the segmentation step, the number of archetypes that represent the examined stock is being decided. The building stock is divided into categories and then, one archetype is selected to represent each category and modelled. The second step refers to the characterization process, according to which the archetypes are described by their technical characteristics. Afterwards, the archetypes are aggregated by considering the total number of buildings included in every category [54]. In the majority of aggregation studies, weighting factors are used for this purpose. When measured data are available, validation can conclude the building stock aggregation process.

The stock aggregation can be applied in conjunction with statistical and machine learning methods to facilitate the analysis especially at urban or national scales and to handle uncertainty. Most of these methods were traditionally used for micro-level simulations and were only recently employed for housing stock analysis due to their computational complexity. One way to maintain the simplification purpose of stock aggregation and to minimize the number of archetypes is to combine cluster values into parameters based on statistical analysis, such as frequency histograms of populations [56]. Bayesian-based methods have been used recently to facilitate the determination of residential archetypes [57, 58]. The combination of probabilistic sensitivity analysis with Bayesian calibration processes can quantify uncertain parameters more accurately and account for model inadequacies [59]. Monte Carlo simulations have been also successfully used to generate probability distributions of input parameters in conjunction with nationwide building databases towards energy demand prediction [60]. In addition, statistical learning algorithms and feature selection methods including linear regression, random forest and support vector regression have been fit to energy benchmarking data for cities in order to predict the energy demand. The majority of these models have proved to be more efficient for electricity consumption prediction though, due to inadequate information regarding heating systems [61].

2.7.1 Urban building energy modelling

Urban building energy modelling is a nascent field that receives growing attention during the last decade [64]. Bottom-up urban building energy models (UBEM) have been introduced to engage and support various stakeholders and evaluate the potential impact of different energy retrofitting scenarios in urban areas. UBEMs can accurately represent the impact of urban context on building energy demands [62]. Thus, they have the potential to become key planning tools for utilities, municipalities and urban planners [65]. In combination with geographic information systems (GIS), they can import city model geometry and visualize and map energy demand calculation results. Furthermore, they can be used for urban massing proposals for urban development sites. UBEMs originate from the development of urban lighting and climate simulations tools that evolved later on to also account for energy demand prediction [63]. They can be also found in literature as *micro-simulation*, since the micro stands for individual building level. Consequently, meso-scale refers to district scale, while macro-scale refers to urban or national scale. Key inputs of UBEM are building information of the examined building stock that cover a variety of characteristics, from occupant-related parameters to building fabric and systems description.

A plethora of <u>urban simulation tools</u> has been developed during the last years to conduct energy performance calculations, as well as additional analyses, such as daylighting, thermal comfort and mobility. They mostly consist of two modules, i) the visualization and modelling interface including the data inputs and ii) the calculation engine. These tools may incorporate steady-state or dynamic-based thermal models to calculate energy demands. Thermal modelling in urban simulation tools differs from building performance simulation tools in the simplification algorithms that are often applied. In particular, thermal zones can be clustered based on common construction or usage characteristics to decrease the amount of model variables. UBEMs can be coupled with CFD and sophisticated solar radiation algorithms to account for localized weather conditions and shading. The archetypal approach is often adopted in urban simulation tools, which reduces significantly computation time, but may introduce uncertainty leading to less accurate prediction results. For this reason, validation and calibration processes are required. The majority of UBEMs have been calibrated using monthly, seasonally or yearly measured data. Few of them have been validated against hourly energy data. Occupant behaviour is one of the main challenges to cope with in urban energy simulations, due to the variability that it is associated with. Both deterministic and stochastic approaches have been applied in UBEMs to model occupancy, with the stochastic ones being less frequent because of their complexity and computational intensity. Due to the extensive data quantity embedded in UBEMs, an automation of modelling workflows is required. Some of the most promising urban energy simulation tools that have been developed are *Urban Modeling Interface (Umi)* developed by MIT, *CityBES* by Lawrence Berkeley National Laboratory and *CitySim* by EPFL. For the collection and integration of city data, an international data standard has been developed and widely applied, *CityGML* that can facilitate the exchange and representation of 3D urban models [66]. In Figure 5, an example of the visualization of energy demand results for a Portuguese residential and commercial district is presented, as simulated with *Umi*.

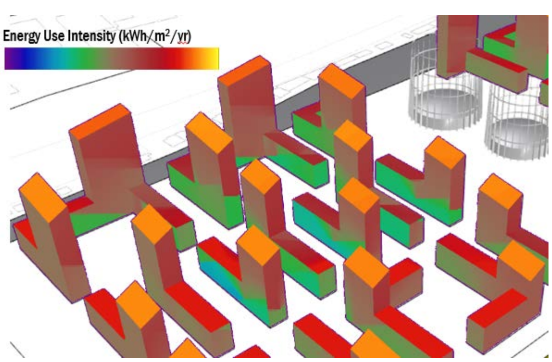


Figure 5. Energy simulation results for a district in Portugal made in Umi (P. Gianniou, S. Letellier-Duchesne, S. Nagpal)

2.8 Smart meter data

Due to the increasing penetration of ICT, conventional energy systems are being digitized. The deployment of intelligent or smart meter systems in buildings and district systems by utility companies is increasing rapidly during the last years, This is motivated by the *Third Energy Package* in the Electricity [67] and Gas Directive [68] issued by the European Commission in 2006. According to this, European countries plan to convert part of their legacy meter stock to smart by 2020 with a main focus on electricity. It is expected that almost 72% of European consumers will have a smart meter for electricity and about 40% will have one for gas by 2020 [69]. The enormous wealth of information included in big databases and inventories has led to a growing interest towards knowledge discovery [70]. The wide <u>use of smart meters</u> enables collection of a large amount of fine-granular time series, which can be used to understand end user behaviour, energy performance and building dynamics better. Furthermore, they can be very useful for smart cities operation and implementation of energy efficiency programs. Building-related data can reveal hidden correlations between energy consumption and its influencing factors, which can facilitate the application of control strategies to optimize building performance [71]. Therefore, they can be utilized to decrease uncertainty related to building energy performance and provide detailed information on energy monitoring. Furthermore, they can identify consumer patterns that can be utilized towards the optimal operation of smart grids and the implementation of demand-side management strategies. Smart meters can also provide valuable feedback to consumers, which can help them adopt efficiently demand response programs. There is also an expectation that smart metering or advanced metering 'infrastructure' can contribute to demand and cost reduction and to the adoption of domestic low- and zero-carbon technologies [72]. They can also help fraud reduction and accurate billing, while identifying anomalous behaviours. A significant drawback that hinders the availability of smart meter data is privacy. Smart meter data are mostly collected by utilities and energy suppliers and usually come with additional information about the users, including unique household codes that can be combined with address codes and therefore be identified. In fewer cases, census data are also included, such as age and number of occupants, income, education level etc. Anonymization processes can overcome the privacy issues, so that personal information about the users is hidden and their identification is not possible. Thus, the data acquisition and publishing process can be greatly facilitated. For this reason, most utilities anonymize their data before making them available.

2.8.1 <u>Data mining techniques for energy data</u>

Different data mining techniques have been developed and applied so far on measured energy data as mentioned previously. According to [73], data mining is defined as *"an interdisciplinary field bringing together techniques from machine learning, pattern recognition, statistics, databases and visualization"*. <u>Knowledge Discovery in Databases</u> (KDD) is a set of tools and theories that can extract useful information (knowledge) from digital data [74]. KDD typically consists of six steps as presented in Figure 6: data selection, data preparation, data transformation, application of data mining algorithms, interpretation of results and extraction of knowledge. The main advantage of data mining approaches against statistical methods is that the latter may fail in obtaining reliable mathematical models by over fitting or under fitting the data, while data mining can highlight common sense knowledge [75]. In addition, when applied to big data, they outperform traditional techniques by providing reliable models of consumer behaviours with fast legibility and high replicability. However, data mining techniques are entirely dependent on data quality, meaning that low-quality data lead to low-quality results.

Data mining is classified into two main categories of data analytics: supervised and unsupervised analytics [76]. Supervised analytics are powerful for predictive modelling, while unsupervised analytics are suitable for discovering correlations in data. The latter do not require high-quality training data and are more promising in realizing the true value of big data [77]. Clustering is one of the most common and efficient unsupervised techniques to identify groups and interesting distributions in the underlying data. The main reasons are that it can improve the predictive accuracy of subsequent energy models and that stable clusters are reproducible [78]. A clustering problem is about partitioning a given dataset into groups, classes or clusters such that the data points in the same cluster are more similar to each other than data points in different clusters [79]. The main steps of clustering include feature selection, choice of clustering algorithm, validation and interpretation of results. A plethora of clustering algorithms have been developed including partition-based –such as *K*-means and *K*-medoids-, hierarchical-based, density-based and grid-based algorithms [80]. These differentiate mainly upon the approach to arrange observations into clusters. Clustering has been applied widely to various fields, such as biology, web-mining etc. Due to the deployment of smart meters in the energy sector, clustering techniques are being widely applied on measured energy data. The majority of studies on cluster analysis found in literature using smart meter data regard electricity data. There are significantly fewer studies on heating data analysis. This is mainly due to the fact that smart meters are less commonly installed in heating sector, which is more challenging than other sectors due to high variability of heat production and demand of end users. Therefore, the collection of data on heat production and utilization is more difficult.

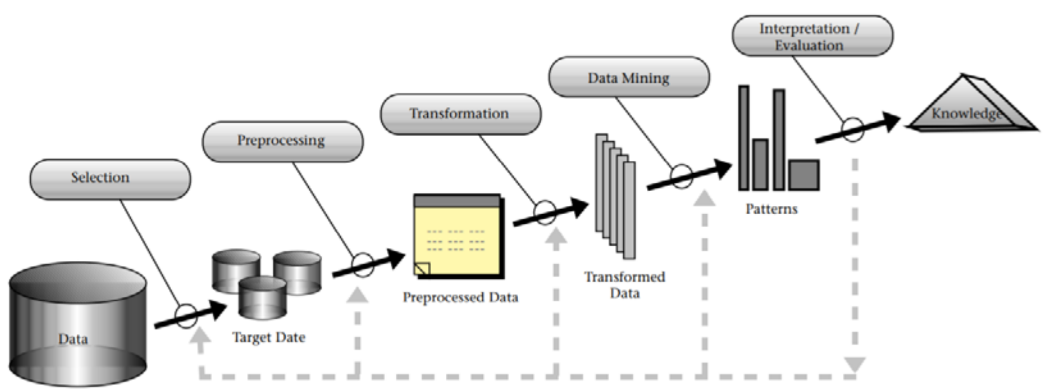


Figure 6. Overview of the processes in KDD [74]

2.8.2 Cloud computing for smart city data

Cloud computing is an emerging information theory paradigm that enables the access to shared pools of system resources and services, without requiring central management efforts. In other words, cloud computing describes the provision of information technology (IT) solutions as a service rather than as a product, mostly through the Internet [81]. Cloud solutions offer a great potential to integrate smart city data and provide solutions to different stake-holders involved in the smart city management and operation. They can be primarily used to deliver services that improve people's lives in a plethora of sections, such as energy efficiency and air quality. Furthermore, cloud computing applications in the area of smart grids, city management or even healthcare can predict future challenges and facilitate their solution. They also enable the sharing of open data and the citizen engagement that can be important towards monitoring the environment, energy savings and improvement of societal efficiency. The employment of smart meters leading to big data necessitates the development of cloud solutions. As a result, the data mining collected from smart meters opens up endless opportunities for smart city research. To take advantage of the cloud computing, specific platforms are required to collect, manage and analyse sensor data in a secure and scalable way. Recent advances include cyber physical systems or frameworks to support knowledge acquisition methods. The main challenges that need to be faced in cloud computing applications are standardization, data protection, user confidence and most importantly cloud security [82]. A proposed cloud computing framework and management platform for smart city data can be found in Paper X and Paper XI, included in the Appendix of the thesis.

2.9 Energy flexibility

The increased penetration of renewable sources to energy grids is a key solution towards the achievement of environmental and energy goals. The electrification of energy demand is expanding continuously worldwide through the adoption of promising technologies, such as electrical vehicles or heat pumps. Furthermore, district heating systems are dependent on several renewable sources. The volatility of renewable sources introduces a challenge to the energy grids, which needs to be addressed to ensure the stabilization of the grid without jeopardizing the security of supply. Towards this, the concept of energy flexibility has been introduced as a potential solution. Energy flexibility can facilitate the synchronization and balancing of supply and demand side. As far as Denmark is concerned, renewable energy is expected to cover about 80-85% of electricity consumption and up to 65% of district heating in 2020 [83]. In addition, approximately 63% of heating in private Danish houses is provided by district heating, which covers both space heating and domestic hot water demand [84]. One of the main goals is that the entire energy supply is covered by RES by 2050. That calls for strategies and techniques, which can facilitate the operation of the smart grid while enabling controllability of the electric loads.

Three main categories of these techniques can be identified: demand side management (DSM), demand response (DR) and electric load management (ELM). DSM characterizes any activity adopted on the demand side that ultimately changes the utility's system load profile. The concept of DSM originates in the 1980's. While the objective of any DSM activity is to create a load-shape change, successful implementation of DSM strategies achieve balancing of utility and customer needs [85]. DR describes techniques that are adopted to induce the customer to change their consumer behaviour and subsequently their energy demand, while ELM refers to policies that aim at managing a set of electric loads to obtain the desired goal. The high share of energy use that is attributed to the building sector makes it suitable to deliver significant flexibility services to the system by intelligent control of the thermal and electric energy loads [86]. DSM in grid-connected buildings has a high potential to provide flexibility to the energy grid and can be realized in different ways. It is classified in four distinct categories: reducing peak loads (peak clipping or shaving), shifting load from on-peak to off-peak (load-shifting), increasing the flexibility of the load (flexible load shape) and reducing energy consumption in general (strategic conservation).

Thermal energy storage has proven to be a technology that can be beneficial towards several purposes, such as minimization of fossil fuel utilization, reduction of GHG emissions, integration of renewable energy sources and energy savings. It can also actively facilitate the implementation of DSM strategies. Thermal storage options can be realized in several different ways. They can be central (closer to the supply side of the system) or decentral (close to the consumer side of the system). Another classification considers the thermodynamic nature of the way heat is stored; whether it is latent, sensible or thermochemical storage. Sensible heat storage includes liquid and solid storage mediums, such as water tanks and building structure. Latent heat storage includes phase change materials and eutectics, while thermochemical heat storage includes sorption and chemical phenomena materials. These systems are characterized by high thermal storage density, but also low efficiency, special consideration of safety and large initial investment costs [88]. Many different thermal storage options have been researched and some have already been implemented on a large scale. The zero investment cost along with the sufficient thermal insulation that is available in the majority of buildings in northern countries make the utilization of short term heat storage in the sensible thermal mass of the buildings a very promising solution. Opposed to the other storage methods, it can be entirely safe, applicable to several building types and cost-effective. On the other hand, active storage systems in the building envelope could be used when constructing new buildings [89]. The integration of active thermal storage in buildings should be planned during a design phase in order to overcome the problems of availability of space for installations.

The optimal operation or intelligent scheduling of systems in buildings is a topic gaining increasing interest in the last decade. The integration of energy flexible buildings in future energy systems requires holistic approaches that can combine building models, energy calculations, energy market design and occupant interaction [86]. The techniques that are mostly employed to conduct this are optimization or model predictive control (MPC). Optimization in building energy systems scheduling is conducted using mathematical techniques, such as linear programming, mixed-integer linear programming, nonlinear programming, and even mixed-integer nonlinear programming. The selection of the optimization algorithms depends on the nature of the variables in the problem. Model predictive control utilizes forecasts of future disturbances to determine optimal control actions minimizing a certain cost function over the prediction horizon. It has been widely applied towards HVAC optimal control. The main difference between MPC and traditional optimal control is that in MPC the optimization is performed repeatedly online (receding horizon) [87]. The main challenges associated with MPC are the forecast data and the compromise between the simplification and complication of the building thermal dynamic modelling in order to maximize its potential flexibility value. The model complexity affects significantly the MPC performance on buildings and has to be as complex as necessary to produce accurate predictions without leading to huge computation times. Thus, the models should require a minimum of states to model zones separately, as well as walls and floors. This minimum number of states should further increase with the building mass content [90].

3 Energy modelling of buildings at district and urban scale in a smart city framework

This chapter summarizes the findings from the engineering, statistical and data-driven analyses that were carried out for building energy modelling at district and urban scale in a smart city framework. These findings address the first and second objectives of the thesis. The following part is based on Paper I and Paper II, as well as Paper IV, Paper V and Paper IX. Detailed information can be found in those publications. All investigations refer to Danish case studies, hence heating demand being the main contributor to energy use. The overall context of the modelling is presented in Figure 1. The aspects that this chapter handles are first, the most influencing building envelope-related factors that cause variability in residential building energy demands, which is important to determine the aggregated load profile of a cluster of buildings. Second, the aggregation of building energy demands is investigated on a district level and the minimum information level about building parameters is determined that ensures reliable modelling. Third, the applicability of quasi-steady-state and dynamic energy calculation methods is discussed to conduct aggregation of building energy demands at district scale. Fourth, a combined engineering-based model with a statistical approach is proposed based on an urban dataset to gain insights into the building stock. Its suitability with regards to urban scale application is evaluated. Fifth, a data mining approach is introduced and applied on an urban-scale smart meter dataset that can reveal valuable information about the heat load profiles and consumption patterns. The strengths and limitations of such an approach are discussed.

3.1 Identification of parameters affecting the variability of energy use in buildings

Findings of this study have been published in the proceedings of CLIMA2014:

Gianniou, P., Heller, A., Nielsen, P. S., & Rode, C. (2016). Identification of Parameters Affecting the Variability of Energy Use in Residential Buildings. In Proceedings of the 12th REHVA World Congress, CLIMA, Aalborg, Denmark.

Variations in residential energy use intensity of buildings are significant. These variations are particularly critical when aggregating energy load profiles, as they determine the aggregated load profile. This preliminary study aimed at identifying the parameters that lead to the largest variations in space heating demand of residential buildings in Denmark. The identification of these parameters was important to understand the energy performance of housing stocks.

3.1.1 Method

In this study, a typical Danish single-family house from the 1940's was used as a reference building. A parametric or local sensitivity analysis was applied using a Brute-force algorithm. Brute-force is a very simple straight-forward approach that can sample the whole solution space. For this reason, it is computationally expensive and not very fast. However, it is very accurate and for the simple case examined in the present study the computation time did not exceed 15 minutes for each parameter. Main properties related to building envelope and occupancy were investigated, which included: orientation of the building, U-value of glazing, insulation thickness of the building components, infiltration rate and number of occupants. To quantify the changes in heating demand that the respective parameters caused, a dimensionless impact factor was defined as the related change of the extreme values of the heating use divided by the average value of the heating use. A dynamic energy model of the building was created in IDA ICE using a single-zone building model, which was coupled with an open-source software (MOBO) [93], designed to conduct parametric and optimization analysis.

3.1.2 Results

The parameters that were investigated in the sensitivity analysis, indicating their range and the values used in the reference model that was used for the parameter variation can be seen in Table 3. The results of the sensitivity analysis with regards to the dimensionless impact factor, *I*, are also presented in Table 3. The positive values of the impact factor indicate an influence to increase the heating demand, while the negative ones indicate a diminishing impact on the heating demand. The parameter that resulted to the greatest variations in space heating demand was the insulation thickness of external walls, followed by the insulation thickness of roof. The infiltration rate and number of occupants had significant although lower impacts on heating demand. These findings were highly dependent on the structure of the examined reference building, which in this case was poorly insulated. Different occupant profiles were not examined in the study, even though they might affect significantly the residential heating demand. The effect of outdoor climate was out of scope for this analysis, since the study aimed at investigating variations in energy use that can be observed among different buildings located in the same geographic region.

Table 3. Examined parameters and results of the sensitivity analysis

Parameters	**Min. value**	**Max. value**	**Ref. value**	***I*, Impact factor [-]**
U-value of glazing (W/m^2K)	1	3	1.5	0.058
Infiltration rate (ACH)	0.1	1	0.4	0.141
Number of occupants	0	6	3	-0.101
Orientation (deg)	0	350	50	0.013
Insulation thickness in roof (m)	0.001	0.32	0.095	-0.242
Insulation thickness in ext. walls (m)	0.001	0.32	0.01	-1.021
Insulation thickness in floor (m)	0.001	0.35	0.019	-0.030

3.1.3 Main findings and conclusions

- The impact of insulation was significant in an old, poorly insulated single-family house leading to large variations.
- Cold weather conditions intensified the impact of insulation due to higher thermal losses.

- Interactions among the building parameters should be taken into account through global sensitivity analysis, especially when the total energy use is in focus.

3.2 Aggregation of building energy demands and comparison between dynamic and steady-state methods

Findings of this study have been published in the Proceedings of BS2015 and CISBAT2015:

Gianniou, P., Heller, A., Nielsen, P. S., Negendahl, K., & Rode, C. (2015). Aggregation of Building Energy Demands for City-scale Models. In Proceedings of Building Simulation 2015, Hyderabad, India.

Gianniou, P., Heller, A., & Rode, C. (2015). Building Energy Demand Aggregation and Simulation Tools: a Danish Case Study. In Proceedings of CISBAT 2015, Lausanne, Switzerland.

The initiatives for smart cities and the focus on city level energy policy have emphasized the need for development of aggregated energy demand models. The objectives of these studies were to introduce two different modelling approaches to aggregate building energy demands and to test different energy simulation methods upon a district-scale building sample.

3.2.1 Method

In this study, the aggregation of building energy demands of a neighbourhood in Denmark was investigated. Two methods of aggregating energy demands were examined: the first method aggregated individual building energy estimates, while the second method treated the investigated housing stock with archetypes representing different building categories. The minimum information level to model accurately 16 single-family houses -located in southern Denmark- was identified after comparison of the predicted monthly energy demand results with measured energy data. Measured energy data were available for the specific houses, along with information gathered by questionnaires and personal visits to the houses (i.e. number of occupants, current refurbishment state of building, installation of PV systems). Four information levels were set and examined, taking into account the previous findings of the influencing building-related characteristics, which included: a) simple, predefined national typological data (i.e. TABULA [91]), b) information collected by national databases about building age, floor areas and construction characteristics (i.e. BBR [92]), c) detailed geometry information gathered from Google Maps and Street View, d) information acquired by personal visits (i.e. number of occupants, current renovation state). The combination of these information levels led to six investigated scenarios.

The modelling of the buildings was conducted in Termite, which is a newly developed parametric tool at DTU, based on architectural software (using Rhinoceros as its graphic interface and Grasshopper for its parametric properties) and using a Danish steady-state energy calculation tool (Be10, 2010 version) as its energy calculation core, which is based on EN ISO 13790 [94]. The main benefit of this tool was that all house models could be simulated with just one model setup. The different building parameters were included in inputs lists, which were all connected with a main controller that determined which house would be modelled every time. The automatic movement of the main controller from one building model to the other via a timer enabled the modelling of a large number of buildings. As a second step, a dynamic energy simulation tool (i.e. IDA ICE) was used and compared with the results of the steady-state calculations upon measured energy data. Conclusions were drawn on the suitability of each approach.

3.2.2 Results

In this study, a sensitivity analysis was applied to prioritize the impact of certain input parameters in the form of information levels on the energy performance. The results showed that the current renovation state of the houses was substantial, the existence of which decreased the error by 24% compared to the models which excluded it. Detailed information about the geometry (e.g. glazing area) and the number of occupants proved to be less significant to predict accurately building energy use, hence not essential. This was attributed to the fact that the assumption about occupancy made in the models was not far from reality, as the majority of the houses were occupied by 2 or 3 people. The effect of the exact envelope area and glazing area was not outlined in this study due to the steady-state calculations. The scenario that incorporated all available information (geometry, refurbishment state, typology, occupancy) showed the smallest deviation from the measurements, as expected, which was 12% on annual terms.

When implementing the aggregation method based on archetypes, five building types were identified based on their construction age, according to TABULA classification, as presented in Table 4. The results showed that the aggregated yearly energy result deviated only 1% from the measurements (Table 5). However, the monthly energy results showed a higher deviation, with the highest one being in May. It should be noted that the weather file used by the steady-state tool was based on the Danish Design Reference Year [95], which even though shared the same minimum ambient temperature as the year which the measurements were taken from, the rest of the weather conditions did not coincide throughout the year. This was one of the main factors of disagreement between predicted and measured energy demands. As a conclusion, even if the archetypes failed to represent well the individual heat demands, the aggregate annual prediction was very accurate. This can be explained by the fact that uncertainties and errors might be averaged out when scaling energy use up to a district scale.

Table 4. Characteristics and annual energy results for the archetypes

Building type	Construction period	No. of incl. buildings	Total floor area [m^2]	EUI [kWh/m^2]	Total floor area [m^2]	Energy demand [kWh]
A	1931-1950	2	238	82	238	19,516
B	1951-1960	2	180	91	180	16,380
C	1961-1972	10	1,530	158	1,530	241,740
D	1973-1978	1	117	118	117	13,806
E	1979-1998	1	122	104	122	12,688

Table 5. Comparison of aggregate heat demand results between measurements and archetypes using the quasi-steady-state approach

Calculation method	Aggregate heat demand [kWh]	Deviation
Measurements	300,858	1%
Archetypes	303,582	

The limitations introduced by the quasi-steady-state energy calculations called for a more dynamic approach, which could model internal gains, climate conditions and building physics more accurately. Thus, dynamic energy simulations in IDA ICE were carried out for the same building sample. The energy results were assessed both for each building category and for the aggregate total. It was found that IDA ICE performed better, having a lower deviation from the measured monthly energy data than Termite, as seen in Figure 7. It should be also noted that the created building models were simple box models, hence simplifications about envelope and glazing areas were made. The biggest deviation between predicted and measured energy was observed during the warmer months, for which IDA ICE overestimated the energy demands. This could be due to disagreement of the reference climate file (that was used in both tools) and the real weather conditions, as no historical weather data were available. It can be concluded that uncertainty in input parameters affects both dynamic and more simplified energy simulation tools and methods. This uncertainty may even exceed uncertainty induced by model simplifications. Parameter uncertainty mainly comes from assumptions that have to be made regarding building envelope parameters, indoor and outdoor climate conditions and occupant schedules, as well as values extracted from national building databases, which are not updated on a regular basis. Therefore, uncertainty in input parameters can counterbalance the accuracy of the implemented method.

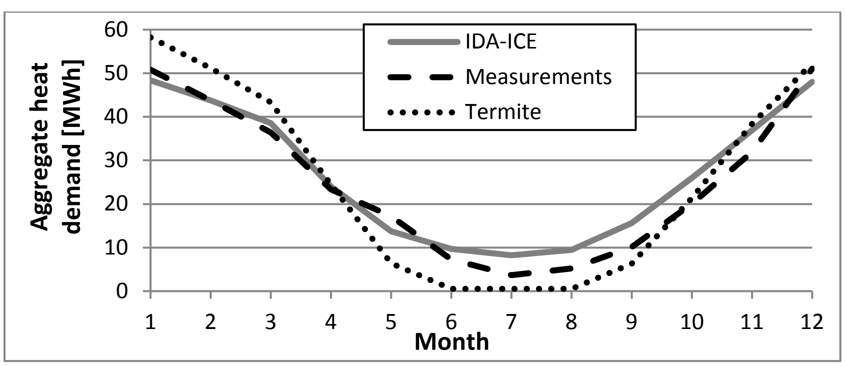


Figure 7. Monthly aggregated heat demand results from IDA-ICE (dynamic) and Termite (quasi-steady-state) compared to measured data

Table 6 summarizes the main characteristics of the two aggregation approaches implemented in this study. Their strengths and limitations with regards to their application are presented, as well as their suitability when applied to real cases. The data availability and the modelling or computational time, which is dependent on the scale of analysis, are decisive factors when selecting the appropriate aggregation approach.

Table 6. Characteristics of two aggregation approaches

Approach	Benefits	Limitations	Suitability
Individual building modelling	- Encompasses detailed description of physics, construction and systems - Leads to accurate predictions of energy loads when input information is available	- Requires large modelling effort - Computationally intensive - Requires large amount of input parameters - Highly dependent on uncertainty of input data	- When detailed information about buildings is available
Archetypal	- Facilitates the modelling for large building stocks - Gives a classification overview of the investigated buildings - Low computation times - Requires smaller modelling effort	- Smoothens dynamics and variations of energy performance - Dependent on national building classification schemes or statistical average values - Dependent on uncertainty of input data	- When average statistical values of the housing stock is available

3.2.3 Main findings and conclusions

- Even though dynamic simulations led to more accurate energy predictions than quasi-steady-state ones on monthly resolution, the error in their estimations mostly during warmer months was not insignificant.
- Quasi-steady-state models reduced significantly the computation times, but at the expense of accuracy.
- The archetypal approach used to aggregate building energy demands led to accurate prediction of the annual aggregated energy use of a cluster of buildings, hence being suitable for district system planning.
- The information level which was necessary to conduct building energy simulations at large scale included basic information of the building construction.
- The knowledge of the current refurbishment state of a building was critical information to model accurately its building envelope (i.e. U-value); hence its lack may introduce high uncertainty to the energy results.
- The archetypes that encompassed all available information about the type and construction of the building, geometry, occupancy and refurbishment level exhibited the lowest deviation from measured energy data.
- Uncertainty involved in input parameters may exceed the uncertainty induced by model simplifications.

3.3 Estimation of thermal comfort-related and building envelope properties among residential buildings in Denmark based on smart meter data

Findings of this study have been submitted to Building and Environment journal:

Gianniou, P., Reinhart, C., Hsu, D., Heller, A., Rode, C. (2018). Estimation of Temperature Setpoints and Heat Transfer Coefficients Among Residential Buildings in Denmark Based on Smart Meter Data. Submitted to Building and Environment, January 2018.

The increasing installation of intelligent metering systems in buildings and district scales has produced a vast amount of data that may be available at urban scale. These urban datasets combined with information derived from national building registers opens up new possibilities for gaining insights into the housing stock. Temperature setpoints are key assumption of occupant behaviour that determine heating demand in cold climates. Occupant behaviour and thermal comfort conditions are hard to estimate, thus are being assumed or estimated in the majority of building models. Several macro-level models assume the same constant heating temperature setpoints for the whole housing stock, while others estimate this based on the building typology and systems. The objectives of this study were to demonstrate the use of an urban dataset to determine temperature setpoints and thermal transmittance for thousands of residential buildings.

3.3.1 Method

In this study, a scalable approach to determine thermal properties of thousands of buildings was proposed and applied to measured energy data from circa 14,000 households in Aarhus, Denmark, over a full year. These data were provided by the local district heating supplier. Mean daily temperature setpoint distributions and overall heat transfer coefficients were calculated at city scale by correlating energy use to external weather conditions and building characteristics using a heat balance model assuming steady-state conditions that can be seen in equation (1). The approach is inspired by the theory of degree-days [100].

$$T_o = T_i - \frac{Q_H + Q_{SG} + Q_{IG}}{UA + c_p\, \rho\, n} \tag{1}$$

where Ti is the equivalent internal temperature in °C, T_o is the ambient temperature in °C, Q_H is the heating demand in W, Q_{SG} stands for solar gains in W, Q_{IG} represents the internal heat gains from occupants and equipment in W/m^2, U is the overall heat transfer coefficient across the building envelope in W/m^{2}K, A is the total envelope area in m^2, n is the ventilation rate in m^3/s, c_P is the specific heat capacity for air under constant pressure which equals to 1000 J/kg K and ρ is the density of air which equals to 1.2 kg/m^3.

Building characteristics were extracted from national building databases and according to national building regulations (i.e. BBR). The ordinary least squares (OLS) method was used and run for each household's data. According to [101], the OLS method is the most efficient for urban scale models having the lowest deviation from measured data.

Thus, in the current model, the ambient temperature T_o was modelled as an affine function of the total energy load, so that $T_o = T_i - \frac{load}{losses}$, where $load = Q_H + Q_{IG} + Q_{SG}$ and $losses = UA + c_p \, \rho \, n.$

The internal temperature and losses are expressed by the intercept and the slope of the linear model, respectively. The losses represent the envelope and ventilation losses. Q_H represents the space heating consumption of the households which was given on hourly resolution. However, in order to account for latent thermal mass effects in the buildings, it was decided to aggregate the hourly energy data on daily basis and report the aggregate value of energy consumption (heating). Thus, internal temperatures in equation 1 refer to the volumetric mean temperatures indoors, representing the relative heating profile of a dwelling. Specifically, the estimated temperature setpoints represent the volume-weighted mean temperature across all conditioned and unconditioned spaces in every household that is considered as constant throughout the heating season. In addition, the total heat transfer coefficients for each building were inferred extracting geometry information (i.e. envelope areas) from an open source GIS dataset of the city [98].

A set of considerations were made to run the analysis. The regression model was restricted to periods when the mean ambient temperature was lower than 15^{o}C, thus focusing only on heating season, when the solar gains were significantly lower and transient phenomena would be less dominating. Furthermore, all households were treated as single-zone models. Occupant profiles were assumed to be similar among the investigated households as no detailed information was available. Internal heat gains were defined according to the Danish Building Research Institute guidelines [99] as 5 W/m^2 (corresponding to external floor areas), which is a sum for people loads and equipment loads dependent on the size of residential building on an hourly resolution. This value was used as an average for all different spaces in households (kitchen, living room, bedrooms etc.), so that it was in line with the single-zone model approach. The infiltration rate of the buildings was calculated based on knowledge about the building age and construction characteristics. No mechanical ventilation was assumed in the examined housing stock.

3.3.2 Results

The above-presented method was applied to smart meter data corresponding to district heating consumption of 14,182 Danish single-family households. The dataset covered a big spectrum of single-family houses in Aarhus, ranging from buildings constructed in 1800 to 2015. The floor areas ranged from 25m^2 to 504m^2, with a median of 134m^2. All the examined buildings were single floor houses. Based on the above-described estimations, the mean value of infiltration rate across the investigated building stock was 0.5 ACH. The regression models showed a satisfactory fit, with a mean and median coefficient of determination (R^2) equal to 0.8 and 0.83, respectively (Figure 9). That confirmed the good fit of the linear regression models to the smart meter data, as well as a quite consistent model, considering the large amount of data used and assumptions made. The temperature setpoint distribution and the U-value distribution that were calculated are presented in Figure 8. The mean thermostat setpoint temperature was calculated to be 19.1°C with a standard deviation of 1.54°C. This was slightly lower than anticipated according to the Danish Building Research Institute guidelines [102], which would be close to 20°C, since it represented both conditioned and unconditioned spaces in a household and it represented an equivalent value over a heating season. So, unoccupied periods and hours with night-setback were also included in this estimation. The mean value of the equivalent U-values (corresponding to the total building envelope) for the examined residential building stock was calculated to be 0.58 W/m^{2}K with a standard deviation of 0.22 W/m^{2}K. According to the Danish Building Research Institute [103], the area-weighted U-values of Danish single-family houses -as calculated with regards to the national building regulations- mainly vary from 0.3 to 0.65 W/m^{2}K depending on the construction age. It should be noted that the majority of the investigated single-family houses were constructed in the 1960's, which should have an indicated area-weighted U-value of 0.52 W/m^{2}K according to [103]. The overall heat transfer coefficients of the buildings do not always comply with the ones defined in the national building databases or introduced by the building regulations. Therefore, the hereby calculated U-values can determine the level of energy refurbishment that might have been implemented to the buildings.

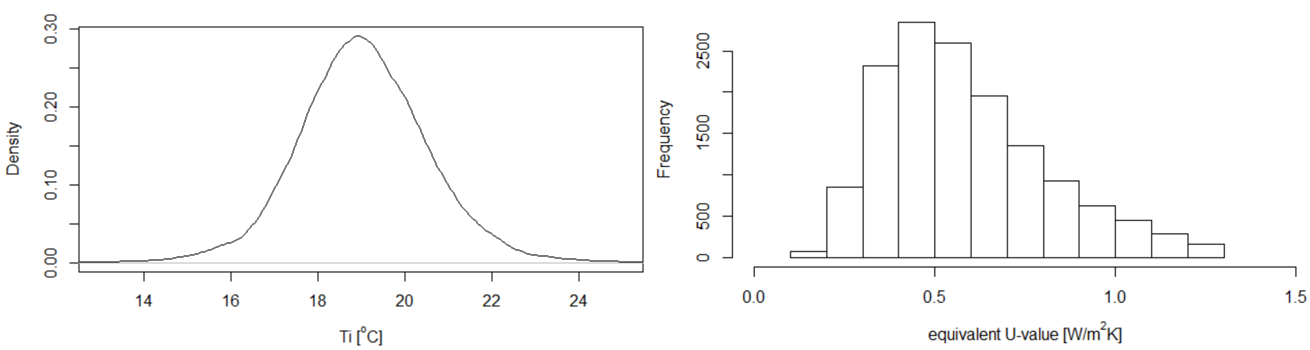


Figure 8. Temperature setpoint (left) and equivalent U-value (right) distributions

Figure 9 presents the regression fit for an exemplary house, as well as the distribution of the coefficients of determination across the building sample. The mean value of the errors of the regression models was equal to zero, as indicated by the Gauss-Markov theorem. In conclusion, the findings gave insight into the thermal comfort conditions across Danish single-family houses (SFH), as well as their current energy refurbishment state. These findings are applicable to urban building energy models where uniform temperature setpoints and U-values are adequate to be extracted for the whole heating season and the whole building envelope, respectively.

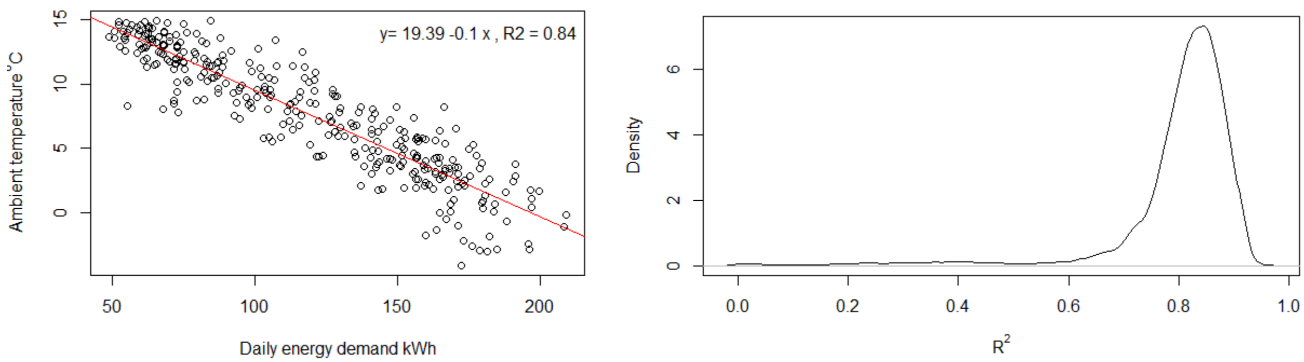


Figure 9. Regression fit for an exemplary household (left) and R^2 distribution for all models' fits (right)

3.3.3 Main findings and conclusions

- The proposed approach allowed to capture the full range of indoor temperature preferences among occupants.
- A good model fit was achieved in the majority of the households with a mean R^2 equal to 0.8, despite the steady-state conditions that were assumed.
- Results of equivalent thermostat heating setpoints and U-values were in line with previous findings.
- The estimated U-values based on the measured energy data determined the current energy refurbishment level of the buildings.
- A simplified approach for urban scale analysis was, thus, proposed that can be adopted by utilities and authorities to estimate basic characteristics of buildings based on big datasets.

3.4 Clustering-based analysis for residential district heating data

Findings of this study have been submitted to Energy Conversion and Management journal:

Gianniou, P., Liu, X., Heller, A., Sieverts Nielsen, P., Rode, C. (2018). Clustering-based Analysis for Residential District Heating Data. Submitted to Energy Conversion and Management, December 2017. Under review.

Due to the enormous wealth of information contained in big databases created by the digitization of conventional energy systems, which are mostly owned by utilities, there is a growing interest towards knowledge discovery. The wide use of smart meters enables collection of a large amount of fine-granular time series, which can be used to understand user demand behaviour and consumption better. This enables the use of data-mining methods, which can uncover hidden correlations between energy use and its influencing factors. The aim of this study was to propose a clustering-based knowledge discovery in databases method to analyse residential heating consumption data, examine load profiles of households and the consumption behaviour changes over time.

3.4.1 Method

District heating hourly consumption data from 8,293 households in Aarhus City, Denmark were utilized in this study. All data were gathered from single-family houses. These data were part of the dataset of the previous study, but were enriched with additional data about customer information, including addresses, family size, date of birth of residents and the dates of them moving in and moving out. The readings were collected from different starting dates, with the earliest being March 02, 2009, but all end on November 29, 2015. Anonymization techniques were applied to remove personally identifiable information such as names, addresses and social security numbers to ensure the sources' privacy.

The method applied for the heating consumption segmentation analysis was classified into three stages that can be seen in Figure 10. First, the heating consumption data were collected by the preparation module. Missing or erroneous values and outliers were cleaned to ensure good data quality for the subsequent clustering-based data analysis. In the second module, clustering was applied on the daily load profiles. The *K*-means algorithm was selected and applied to cluster the data in order to identify or segment the customers with similar load intensity and consumption patterns. Given a set of load patterns (daily load patterns in this paper), the clustering algorithm classifies the load patterns into *K* groups or clusters. The patterns are similar within the same cluster, but dissimilar to the other clusters, according to a distance metric. Clustering is an iterative process with the aim of minimizing the intra-cluster inertia criterion defined by equation 2.

$$C(P,\mu) = \sum_{i=1}^{n} \sum_{X_i \in P_k} \|X_i - \mu_k\|^2 \tag{2}$$

where $P = (P_1, P_2, \ldots, P_k)$ is the set of clusters, $\mu = (\mu_1, \mu_2, \ldots, \mu_k)$ is the set of the cluster centers, $\| . \|$ is the L_2 norm associated to the distance metric and X_i is the load profile of a customer.

To identify the <u>consumption intensity</u> groups, clustering was applied on the representative daily load profiles of all the customers without normalization. The representative load profile of a customer *i*, $\widehat{X_\imath}$, was estimated by the mean of the hourly consumption. The Euclidean distance was used as the distance metrics.

To further investigate the <u>representative patterns</u> in the load profiles, clustering was applied on all the normalized daily load profiles of the customers. The normalized consumption was expressed as the ratio of the raw consumption to the total consumption of that day. In this case, the KSC-distance was used [105], which is an improved distance metric to discover energy consumption patterns with a minor shift on the time scale. This means that the difference between the two patterns regarding a slight shifting along the time scale can still be classified into the same cluster. This can solve minor pattern shifting problems; for example, a person getting up late with the morning peak at 8–9 am, instead of his/her usual morning peak at 7–8 am.

The optimal number of clusters was determined according to the Bayesian Information Criterion (BIC) [104], which is based on the following considerations. In clustering, adding more clusters will decrease the variance (and increase the Bayesian likelihood). To avoid constantly adding centroids, the BIC works by penalizing the loglikelihood more when the complexity of the model (e.g. the number of parameters) increases. Therefore, this approach can search over different values of *k*, score each clustering model according to the BIC value, and determine the optimal number of clusters.

Different labels were assigned to the identified clusters. The analysis based on the clustering results was conducted in the third module, including customer segmentation, load profiling for individual customers, studying load transition over time, and consumption variability. The heating data was combined with building information contained in national building databases (i.e. BBR). The correlation of heating consumption with characteristics -that are often used in building classification schemes to categorize the building stock- was investigated through logistic regression. These were the building age and the floor area. Furthermore, occupant-related characteristics were also investigated to determine whether their correlation to heating consumption is significant and can thus, be associated with groups of consumers. These included the number of adults, teenagers and children in a household. The variability of consumer behaviour was quantified with the concept of entropy, which is often used in computer science and information theory to assess the uncertainty in data and as a measure of disorder [106, 107].

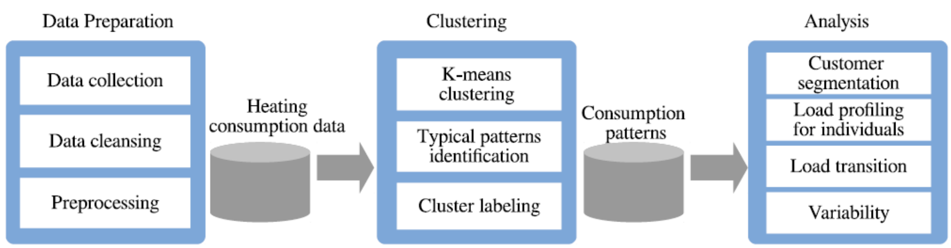


Figure 10. Overview of heating consumption analysis

3.4.2 <u>Results</u>

Based on their consumption intensity using the non-normalized data, customers were segmented into five consumption groups that can be seen in Figure 11. In each cluster, the bold red line shows the centre of clusters for each hour of the day which is the representative group load profile, while the other lines are the mean daily load profiles of the customers in the cluster. Based on the group load profile, the groups were ranked into five consumption levels labelled with the alphabets, *A-E*. The clusters were characterized by fairly constant load profiles, with two soft peaks during the day. Cluster B represented the majority of the households, namely 36%, having heating consumption higher than cluster A and lower than the rest of clusters. Cluster E which characterized the lowest share of households was the most energy intensive one, having also slightly more pronounced peaks during morning and evening time.

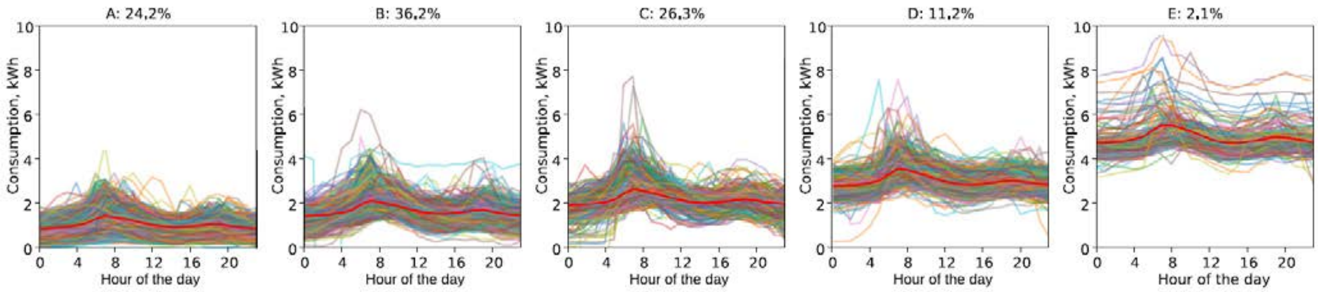


Figure 11. Customer segmentation based on their daily consumption intensity

In addition, Figure 12 represents the average hourly consumption of all households for weekdays and weekends during January. It can be observed that the morning peak load was shifted later during the weekends –by approximately 3 hours- in the majority of households.

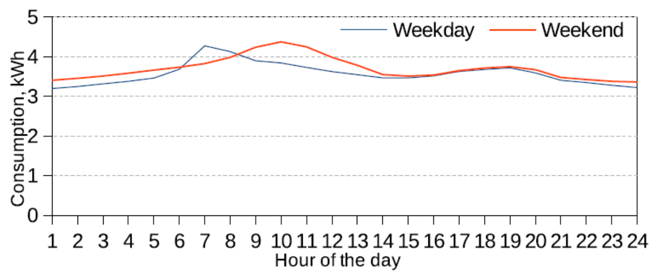


Figure 12. Average daily heating consumption patterns

The normalized energy data were used to determine the representative patterns in the consumer behaviour. According to that, nine clusters were identified and can be seen in Figure 13. According to this classification, the vast majority of the district heating customers, namely 81%, were represented by Cluster A, which was characterized by a constant load profile. The remaining clusters that represented 19% of the customers, exhibited pronounced morning or evening peaks that differentiated in appearance time during the day.

The logit regression analysis identified the variables with the highest correlation to the heating data. The results showed that the floor area and age of the building were the most significant variables. The number of teenagers was positively associated with heating use in some of the identified clusters.

Afterwards, the behaviour changes were investigated to identify the variability of heating demand over time with regards to the consumption intensity clusters. The quantification of the clustering results for the daily load profile of all customers differentiating between weekdays and weekends/holiday can be seen in Figure 14. All five segments showed a similar trend for weekdays, but not for weekends. The magnitude of heating consumption differentiated the five clusters in weekdays. All of the five clusters exhibited a remarkable morning peak, which took place around 7am in the three most energy intensive clusters, and an hour later in the two lowest heating demand clusters. The trends in load profiling for weekends were less evident, as occupants usually do not follow a steady schedule during weekends and holiday.

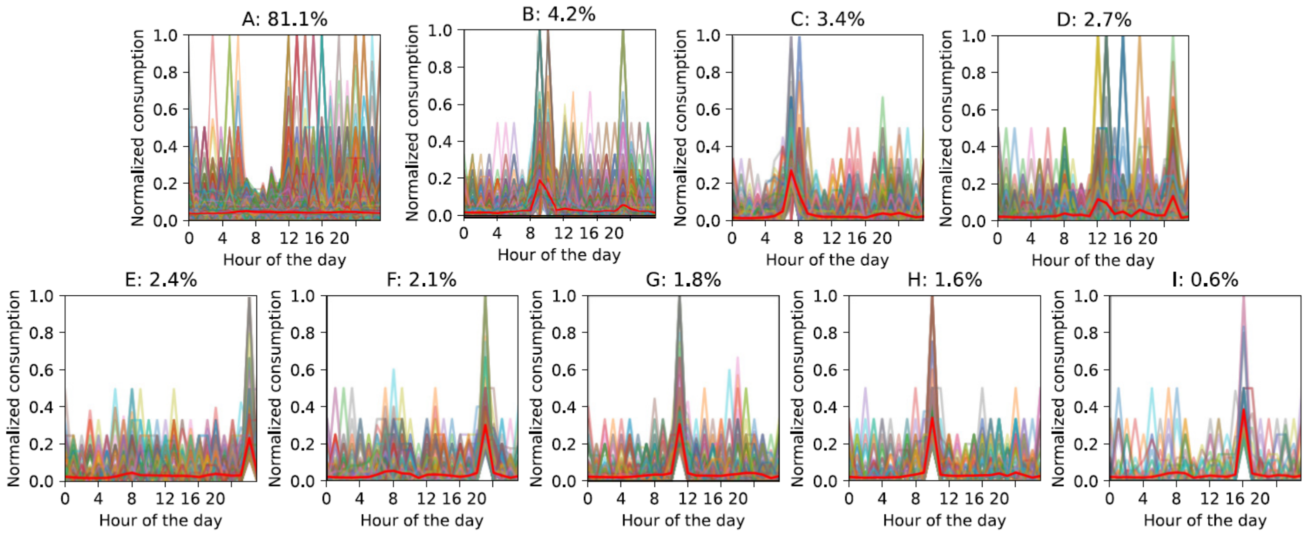


Figure 13. Clusters of normalized daily consumption patterns

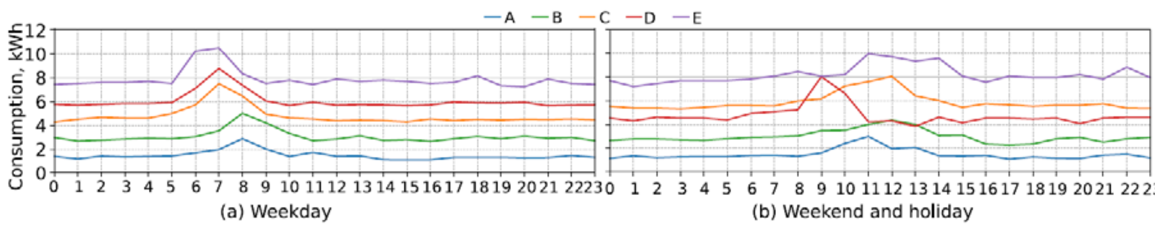


Figure 14. Clustering of daily load profiles into five groups for weekdays and weekends

The load transitions from one month to the other were also examined with regards to consumption intensity and are presented in Figure 15. The cluster distributions of all households can be seen for each month. It is evident that the most energy intensive clusters represented the majority of customers during winter season, while the clusters with the lowest heating demand were the most common in summer time. During transitional months from the heating season to the non-heating season and vice-versa, the medium energy intensive clusters gained ground. These seasonal load transitions could be interpreted as behaviour changes in consumers that were dependent on weather conditions and calendar events.

The consumption variability was quantified by computing the correlated and uncorrelated entropy based on the identified patterns and is presented in Figure 16 (right graph). It was found that the appearance of daily patterns was serially correlated and that the majority of customers had relatively low correlated entropy (i.e. 0.25), which indicated that their behaviour was quite predictable. The entropy was also calculated for the different consumption intensity groups. Results showed that the lower consumption intensity groups had higher variability in terms of the entropy values, while the higher consumption intensity groups had lower variability (Figure 16 left graph). According to the transition probabilities, in general, there was a high chance for a household to repeat the same pattern in the next few days.

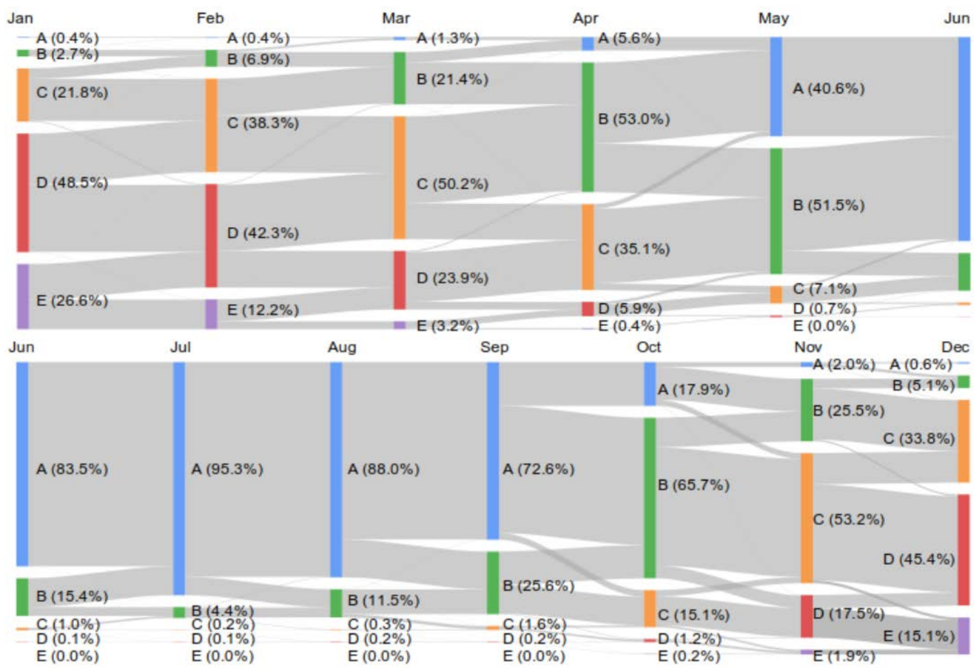


Figure 15. Load transition over the months in the year of 2014

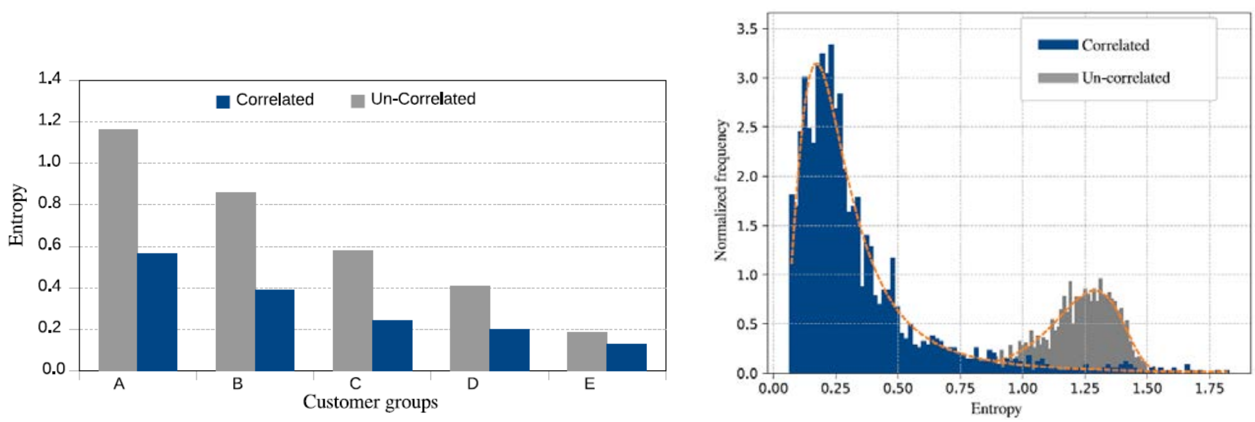


Figure 16. Average entropy of each consumption intensity group (left) and distribution of entropy (right)

3.4.3 Main findings and conclusions

- The proposed technique can easily be implemented to large datasets, since the *K*-means algorithm used for clustering the data is efficient and computationally attractive.

- Depending on the consumption intensity or the representative consumer patterns, five or nine clusters were identified for the examined residential consumers.
- The majority of the households was represented by fairly constant group load profiles.
- The sample was very homogeneous consisting only of single-family houses in the city of Århus, which might have affected the number of clusters.
- Building information, such building age and floor areas, was highly correlated with the heating consumption intensity groups, hence can be used in classification schemes to represent also load profiles.
- Occupant related information, such as family size and age of occupants had lower correlation with the heating consumption data.
- A large share of the examined consumers had regular and predictable consumer behaviours.
- The resulting load profiles can support a scaled analysis of buildings on urban level.

4 Energy flexibility provided by buildings

This chapter summarizes the findings from the studies carried out to determine the potential of buildings to provide energy flexibility to the grid. It is based on Paper III, Paper VI, Paper VI and Paper VIII. Detailed information on the models, methods and results can be found in these publications. All the following studies refer to Danish cases of existing or new buildings. First, the concept of thermal flexibility is defined. The impact of building design parameters, weather and occupancy on flexibility is investigated, to provide guidelines for building design optimization. Second, the intelligent scheduling of a grid-connected residential building is proposed that can contribute to the stabilization of the grid and specifically, to balancing the supply and demand sides. Third, the utilization potential of existing thermal building mass for storage coupled with energy system optimization is investigated for a Danish case. The results are evaluated with regards to different strategies.

4.1 Impact of Building Design Parameters, Weather and Occupancy on Thermal Energy Flexibility in a Low-Energy Building

Findings of this study have been published in the Proceedings of BS2017:

Sarran, L., Foteinaki, K., Gianniou, P., & Rode, C. (2017). Impact of Building Design Parameters on Thermal Energy Flexibility in a Low-Energy Building. In Proceedings of Building Simulation 2017, San Francisco, United States.

Zilio, E., Foteinaki, K., Gianniou, P., & Rode, C. (2017). Impact of Weather and Occupancy on Energy Flexibility Potential of a Low-energy Building. In Proceedings of Building Simulation 2017, San Francisco, United States.

The high penetration of renewable sources to energy grid can ultimately cause insecurity of supply. One solution to face that is demand side management, which aims at adjusting energy demand to match the fluctuating energy production, as defined in [85]. The objective of this study was to understand how energy efficient buildings can adapt perturbations in a city's thermal and power grids. The focus of this work was thus, on thermal energy flexibility, which is expressed as the capacity of a building to provide good indoor comfort in spite of changes in delivered heating power. This preliminary study aimed at determining the impact of building design parameters, weather and occupancy to a building's potential to adapt simple perturbations on the heating grid in which it is integrated. Passive flexibility solutions and specifically, heat storage into the thermal mass of the building was investigated. This study was conducted in the context of the EnergyLab Nordhavn project.

4.1.1 Method

First, a dynamic multi-zone model of a low-energy apartment building was created in IDA ICE. The chosen building was representative of the current and future constructions in Denmark complying with the strict building regulations. The focus was placed on a single apartment to get thorough understanding of the obtained results. Second, a parameter variation followed, which included a plethora of building-design variables: glazing-to-wall ratio, external wall insulation and concrete thickness, floor concrete slab thickness, window U-value and orientation of main façade. Regarding the internal gains and weather conditions, variations in the following variables were investigated: occupant schedules, number of occupants, heating setpoint, outdoor temperature and solar gains.

Two flexibility indicators were introduced to quantify the effect of the investigated variables to the flexibility potential of the apartment. These are represented by equation 3. The first indicator reflected the thermal autonomy potential of a building standing for the duration of the thermal comfort period, during which the operative temperature inside the room does not fall below 20°C. This indicator indicated the degradation of thermal comfort that occupants perceived after an experiment of heat cut-off was implemented. To avoid abrupt fluctuations in the indoor temperature, a tolerance factor was introduced, which made the indicator practically calculated as the duration of the time interval, starting from the cut-off, during which operative temperature in the living room has been above 20°C for 90% of the time since the cut-off. The second indicator was defined as the difference in peak heat load (i.e. space heating demand) of the building caused by the heating strategy implemented as opposed to the 'reference' peak load. Thus, it represented the pressure set on the heating system when performing a heating control operation. The two flexibility indicators are graphically presented in Figure 17.

$$Ind_1 = \min\{t|T(t) = 20^oC\} \qquad Ind_2 = \frac{Q_{max} - Q_{max,\ ref}}{Q_{max,\ ref}} \tag{3}$$

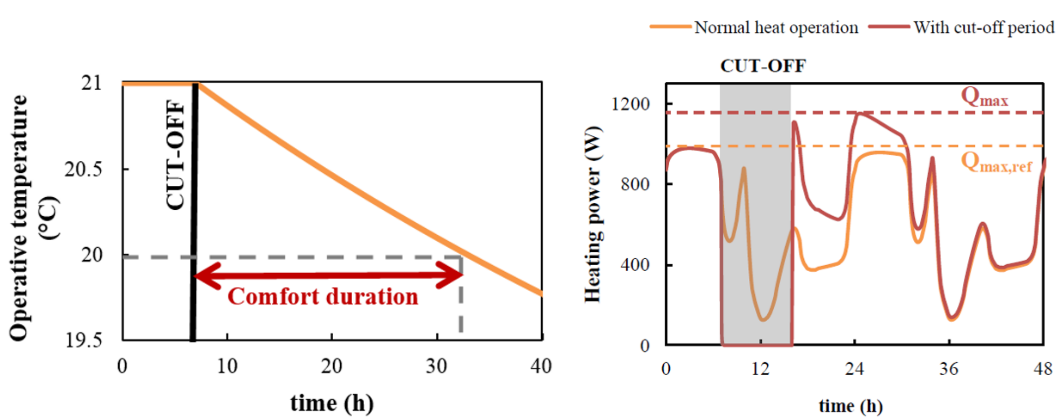


Figure 17. Definition of proposed thermal flexibility indicators

4.1.2 Results

The obtained results showed that the design parameters with the greatest influence on the temperature drop in the inhabited space after a heat cut-off were those related to the heat losses through the building envelope. Specifically, the impact of the external wall insulation on the flexibility indicators was significant and can be seen in Figure 18. The concrete layer in external walls and floor did not affect clearly the flexibility indicators. The importance of the window design, particularly the window-to-wall ratio and U-value was also highlighted due to its influence on heat losses and solar gains and can be also seen in Figure 18. Heat retention time strongly decreased when the glazing-to-external-wall ratio increased, showing that the losses due to a larger glazed surface impacted more the indoor temperature than the possible increased solar gains when the window area increased. The orientation of the apartment exhibited a moderate but clear impact on heat retention: any other orientation than the South significantly reduced the time in the comfort range after a cut-off.

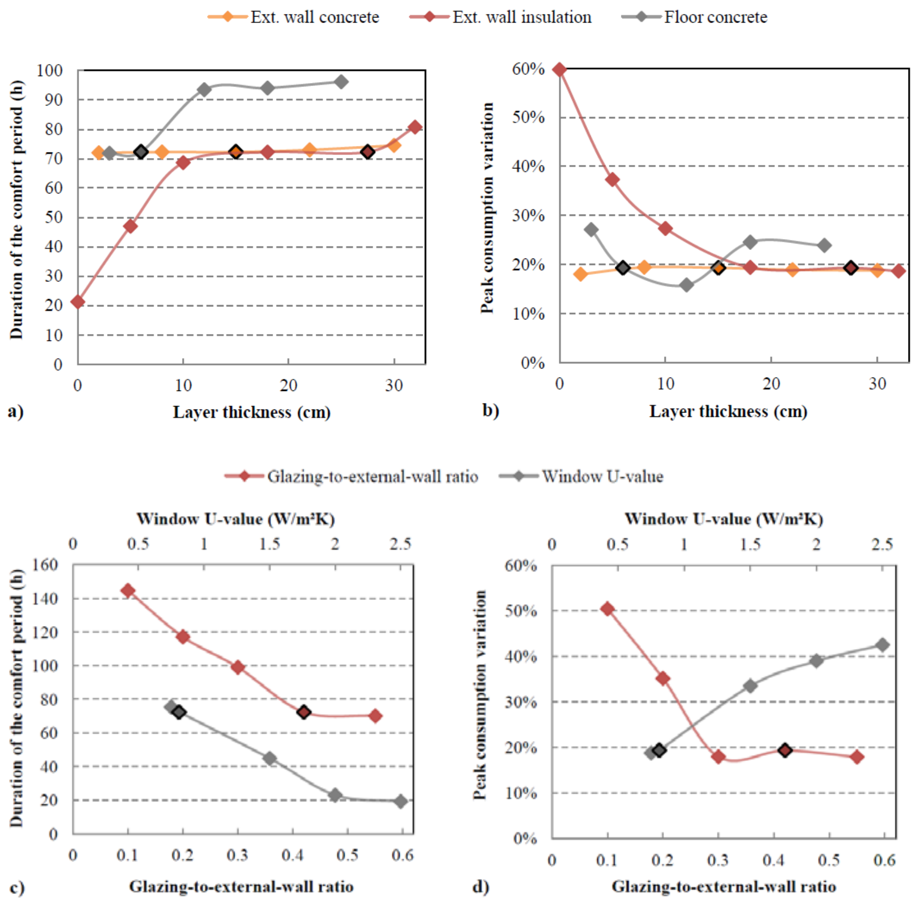


Figure 18. Parameter variation results for construction of building components

The results of the second part of the study regarding the influence of external factors showed that the influence that the outdoor temperature and solar gains had on the two flexibility indicators was significant (Figure 19). The results are presented for two different rooms of the apartment, since they had opposite orientations and different occupant schedules. Large solar gains for the living room led to 215h of thermal autonomy (indicator 1), indicating also the airtight envelope of the investigated building. The impact of the occupant schedule and the number of occupants was found to be lower. It was also concluded that increasing the number of occupants by one unit led to an almost linear increase in thermal autonomy time (in hours).

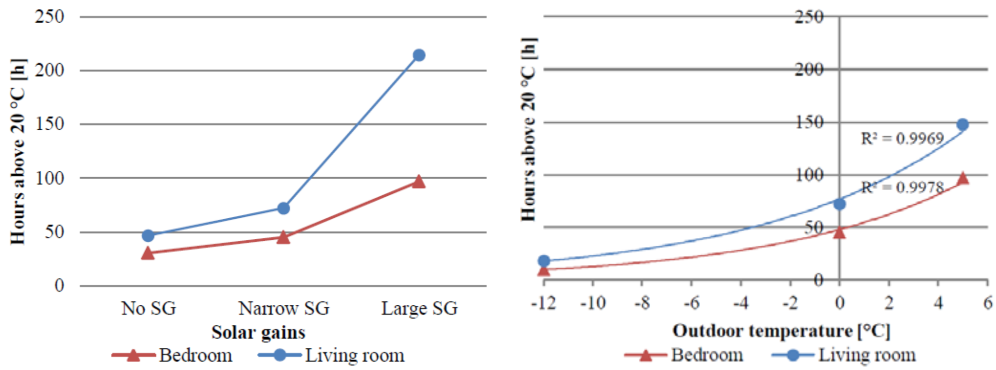


Figure 19. Parameter variation results for solar gains and outdoor temperature

4.1.3 Main findings and conclusions

- Two flexibility indicators were introduced that reflected the indoor comfort degradation and the peak heat load that were observed after a heat cut-off experiment.
- External wall insulation affected significantly the flexibility indicators.
- The impact of the concrete layers of different building components was less important than expected, with the floor slab being the most influential.
- The impact of solar gains and outdoor temperature on flexibility potential was significant.
- The heat power consumption peak that was created after a heat cut-off was greatly affected by the thermal resistance of the building envelope, and not by the thermal mass in the building structure.
- Newly-constructed, low-energy Danish apartment buildings could provide more than 200h of thermal autonomy to the grid after an instance of heat cut-off –as long as the weather conditions were beneficial-, in which the thermal comfort would not deteriorate greatly.

4.2 Intelligent Scheduling of a Grid-Connected Heat Pump in a Detached House

Findings of this study have been published in the Proceedings of BS2017:

Gianniou, P., Foteinaki, K., Heller, A., & Rode, C. (2017). Intelligent Scheduling of a Grid-Connected Heat Pump in a Danish Detached House. In Proceedings of Building Simulation 2017, San Francisco, United States.

The role of buildings in DSM is important, integrating different strategies to reduce power consumption during peak periods (peak-shaving) and/or to shift the power consumption from peak periods to off-peak periods (load-shifting). Price response strategies have a large potential in this context. This study aimed at proposing a method for optimally scheduling a heat pump (HP) according to the grid's requirements without compromising the occupants' comfort.

4.2.1 Method

The proposed method combined a dynamic building energy model with cost optimization that was solved in an integrated environment. A typical Danish single-family house (built in the 1960's) was used as reference, which was modelled in IDA ICE. The properties of the building envelope and systems were extracted from the TABULA database [96]. The house contained an air-to-water heat pump (HP) of 13 kW with *COP* equal to 3.5 and its building envelope had undergone extensive energy renovation. The material layers of the construction elements can be seen in Table 7. The heat emission system installed was hydronic with low-temperature water radiators. The controller of the hydronic system was a thermostat based on the operative temperature. The base heating setpoint was set to 20^{o}C. Internal blinds were drawn in cases of excessive solar radiation. It was also assumed that windows started to open when indoor temperature reached 25^{o}C and opened fully at 27^{o}C. The main façade of the house was assumed to be oriented towards the south. The house was modelled as a single-zone model. The wall model used was a finite difference model of multi-layered components. Each material layer was discretized into four nodes.

Table 7. Material layers of building model

Building component	Layer material	Thickness (m)	U-value (W/m^{2}K)
External walls	Brick	0.05	0.13
	Light insulation	0.25	
	Brick	0.05	
Roof	Light insulation	0.34	0.10
	Wood	0.03	
	Gypsum	0.013	
Floor	Wood	0.02	0.12
	Light insulation	0.28	
	Gypsum	0.02	

A price-based control strategy for the management of the heat pump was implemented that enabled the exploitation of dynamic electricity prices. An additional controller was added in the loop to allow for the system's flexibility when the electricity prices were low. In particular, a positive deadband of 3^{o}C was added to the base setpoint which was set to 20^{o}C, resulting in 23^{o}C upper threshold. The optimization was defined as minimization of the operating cost of the HP. This cost depended on the variable electricity prices and on the consumption of the HP. The horizon of the cost-based optimization was set to three days, for the following reasons: it would ensure cost savings in the electricity bill along with the desired peak shavings in the heat load, any phenomena of cumulative heat storage into the thermal mass would be observed and the computation time would be feasible. The optimization was conducted with the use of the open-source software MOBO, which was coupled to IDA ICE. Four different optimization algorithms and solver settings were tested. Three climate cases were selected out of the entire simulation period to present the results of the optimization. The optimization problem is presented below.

$$Minimize\ C_{HP} \tag{4}$$

$$C_{HP} = p_{el}\,.W \tag{5}$$

$$b_i\ \epsilon\ [0,1]\ for\ \ i = 1 \dots 30 \tag{6}$$

where C_{HP} was the operating cost of the HP, p_{el} was the total electricity price and W was the energy consumption of the HP. The optimization of the system operation was conducted through the boiler schedule which was characterized by ten variables per day, b_i, corresponding to the schedule of the part load operation of the HP. This means that a different schedule of the HP operation was optimized for each day.

The prices for electricity were set in the model through a stochastic profile reflecting real electricity prices in 2015 according to the Nordic electricity market Nord Pool for East Denmark. During this period, wind power generation accounted for 74% of the total electricity production, which explains the low electricity prices that comprised the price signal that was coupled to the IDA ICE model. These prices reflected the total electricity prices consisting of variable el-spot prices (commercial), which account for 32% of the total price in Denmark, while the remaining 68% are fixed taxes for local network, grid and system tariffs, public service obligation tariff and further subscriptions to electrical companies based on the Association of Danish Energy Companies [108]. The total electricity prices were used in the model so that they corresponded to the ones that electricity customers have to pay. The electricity spot prices during the 3-day investigated period can be seen in Figure 20. The negative prices indicate the surplus of wind power generation during these days. The hourly tariff concept was assumed to reflect the lack of or excess of renewable energy in the grid and thus, represent the stress on the grid.

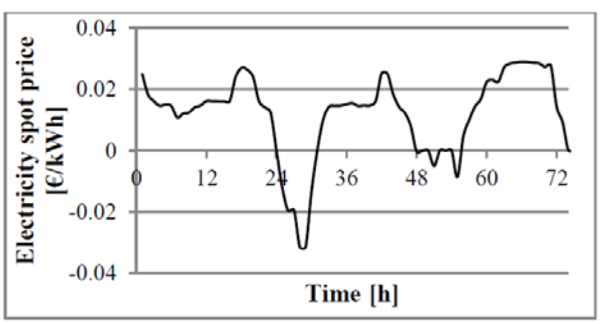


Figure 20. Electricity spot prices over 3 days

4.2.2 Results

Out of the four examined optimization algorithms, a hybrid genetic algorithm was selected that calculated the biggest cost reduction in all three weather cases. The increase in the heating setpoint according to the price signal enabled heat to be stored into the thermal mass during low price periods, and be released back to the room when prices were higher. The comparison of the optimized scheduling of the HP compared to its reference operation was made with regards to power peak shavings, cost savings, operative temperature and thermal comfort. Regarding peak shavings, which was the main goal of this demand side management approach, the results showed a clear decrease in peak consumption at the optimal case both for the cold and the intermediate weather that can be seen in Figure 21. Taking the highest peak in power for each case, there was a reduction of 27% (2.75 kW) for the cold weather and 21% (1.32 kW) for the intermediate weather. This was achieved due to the smoother operation pattern of the heat pump that shifted parts of the load in time, which led to reduced peak demand. The magnitude of the potential peak decrease would always depend on the price signal, but the main outcome was that the implemented control was an effective peak shaving strategy. In the warm weather case, the very low heating demand left no room for optimization.

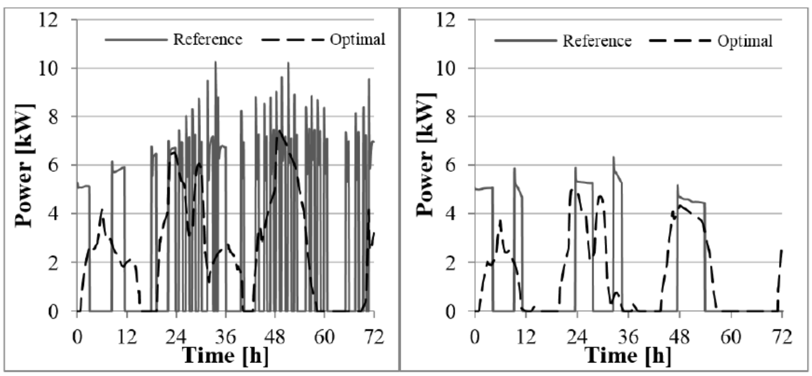


Figure 21. Power results for the optimized scheduling of the HP and the reference operation for cold weather (left) and intermediate weather (right)

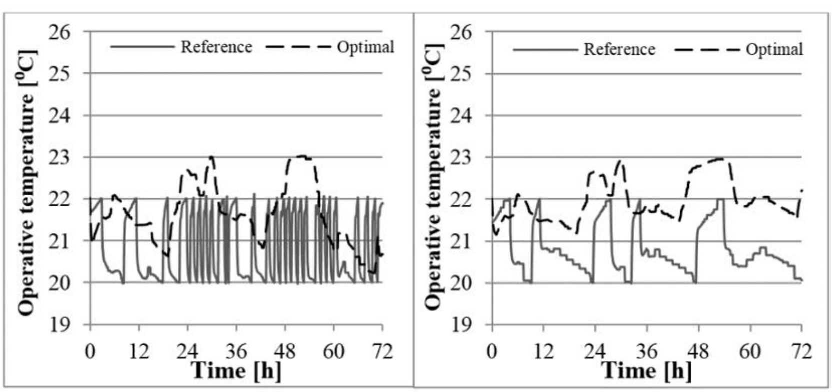


Figure 22. Operative temperature results for the optimized scheduling of the HP and the reference operation for cold weather data (left) and intermediate weather (right)

Regarding the operative temperature in the reference and optimized operation of the HP, results can be seen in Figure 22. Overall, there was a trend for the optimal cases to result in higher operative temperature than the respective of the reference cases at all climate cases. This can be explained by the control of the heating system, which was able to increase the setpoint up to 23ºC when the electricity prices were low. It has to be pointed out that the electricity prices according to which the optimal scheduling of the HP was conducted in the three weather cases were found to be low compared to the average monthly price, which resulted in utilization of the increased heating setpoint most of the time.

The results of the operating cost of the HP can be seen in Table 8. In the cold weather case, optimally scheduling the HP resulted in 42% decrease in the total operating cost of the HP for the examined 3-day period. In the intermediate case, the cost savings were 22% compared to the reference operation of the HP, while in the warm weather case these were significantly lower, approximately 13%, as it was expected.

The optimal scheduling of the HP maintained good thermal comfort of the occupants according to EN/DS 15251 with regards to the operative temperature. These results are presented in Table 9. If a looser comfort threshold was selected, the acceptable range would be further expanded so that lower operative temperatures than 20ºC should be allowed during the unoccupied hours. Furthermore, different price signals could have been applied to the case so that an estimate of the range of peak and power savings could be made and also a realistic correlation to the weather data would be possible.

Table 8. Operating cost results of the HP for 3 days

Weather data	**Total operating cost (€)**	
	Reference case	**Optimal case**
Cold	2.998	1.742
Intermediate	0.979	0.761
Warm	0.008	0.007

Table 9. Thermal comfort assessment for optimal scheduling of HP for optimal and reference operation

Weather data	**Cold**		**Intermediate**		**Warm**	
Case	Opt.	Ref.	Opt.	Ref.	Opt.	Ref.
Comfort category	Share of occupancy hours (%)					
I (best) (21-25ºC)	71	40	100	26	78	100
II (good) (20-25ºC)	100	98	100	100	100	100
III (acceptable) (18-25ºC)	100	100	100	100	100	100

4.2.3 Main findings and conclusions

- The optimization problem was solved with an average time of 7320 seconds for all three weather cases.
- The dynamic building model used allowed for an accurate representation of the building dynamics and heat transfer mechanisms that took place during the charging and discharging of the thermal mass.
- The optimized scheduling of the HP led to power peak shavings up to 27% and cost savings up to 42% for a 3-day horizon, where wind production was very high.
- Warm weather periods could not be considered for the optimal scheduling of the HP, since heating demand during that time was very low.

4.3 Utilizing thermal building mass for storage in district heating systems: combined building level simulations and system level optimization

Findings of this study have been submitted to Energy journal:

Franjo Dominkovic, D., Gianniou, P., Münster, M., Heller, A., Rode, C. (2018). Utilizing Thermal Building Mass for Storage in District Heating Systems: Combined Building Level Simulations and System Level Optimization. Submitted to Energy, November 2017.

Heat storage into the thermal mass of buildings is a cost-efficient way that can provide flexibility to the energy system as seen in the previous two studies. The utilization of building thermal mass on a district scale was the objective of this work that considered the whole district heating system of a Danish city that most buildings are connected to. The capture of building dynamics and transient heat transfer phenomena was of high importance for this goal.

4.3.1 Method

The holistic method that was applied is illustrated in Figure 23. First, a detailed building energy model was created in IDA ICE and dynamic simulations were run.

Second, the flexibility indicators for evaluation of thermal mass for storage were introduced that were mainly based on the previously presented ones (in Paper VI and VII) and different building heating scenarios were presented. Third, a linear continuous optimization model was introduced to represent the local energy system and can be found in the respective publication in the Appendix. A typical winter day was selected to conduct the simulations in the first modelling stage. This day was used as a pattern for all the cut-off events during the heating season (1st October to 30th April). A sensitivity analysis for the latter assumption was carried out and commented further. The optimization model included all the district heating energy plants supplying the district heating network of the city, electricity plants (wind and PVs) erected as a part of the energy transition plan of the city and transportation demand for fuels. As a short overview, the model optimized the energy system with an objective function set to minimize the total socio-economic costs of the energy system. The building thermal mass for storage, which was the focus of this study, was added to the model as a source of flexibility in the local energy system. The whole energy system was optimized during one year on hourly resolution.

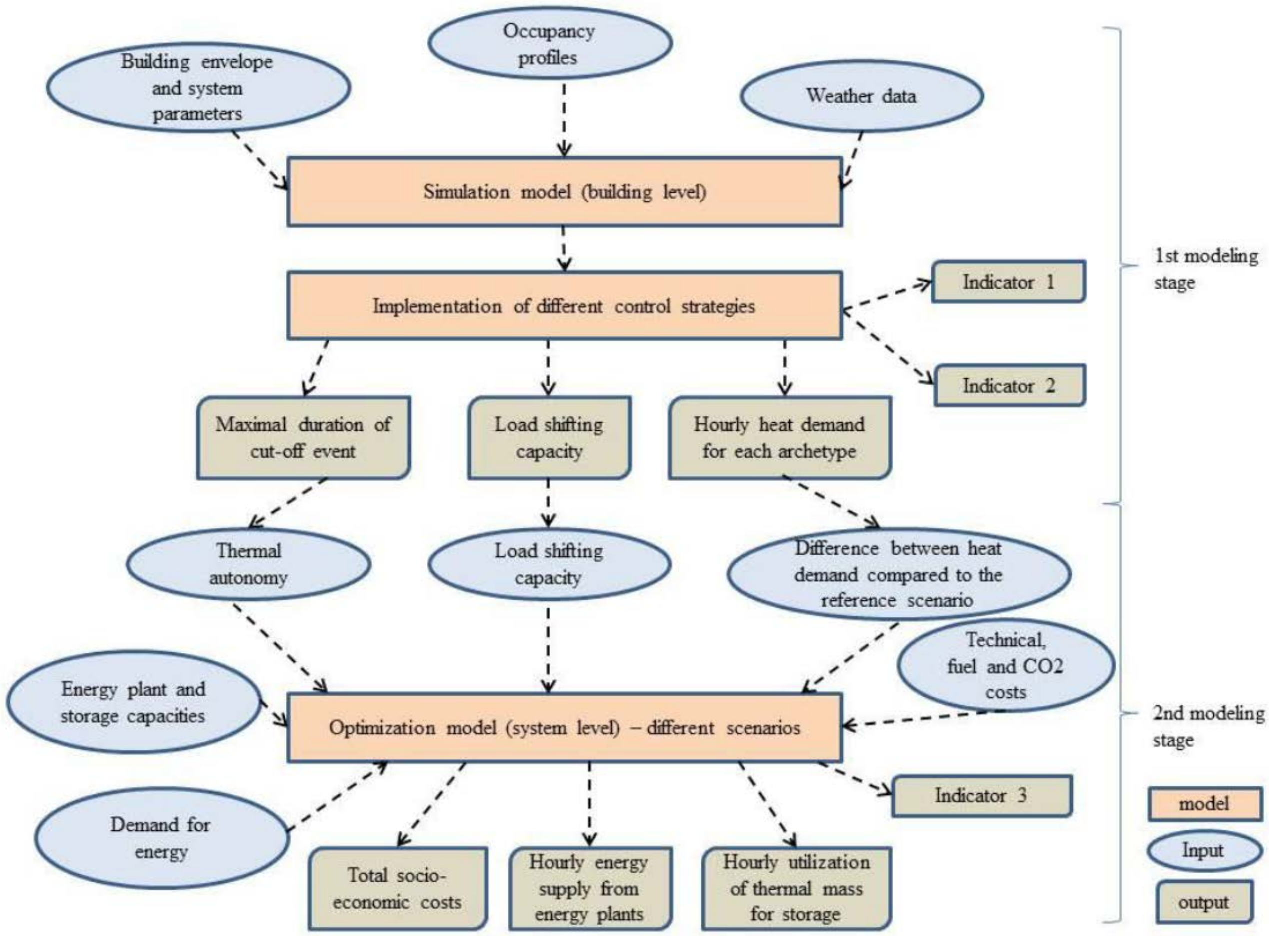


Figure 23. Flow diagram of the two-stage modelling process

An energy transition plan towards 2029 was also introduced, in which the municipality's goals to achieve carbon neutrality were included. In addition, an increased share of households connected to the district heating systems was assumed. The energy system optimization model used the output of building simulations in terms of the indicators, which were the thermal autonomy (equation 7), the difference in energy consumption compared to the reference case (equation 8) indicating the energy savings and a third indicator about the distribution of preheating and after cut-off heating demand as an input for running the optimization. This indicator represented the difference in the operational costs of the district heating system without including the costs of the other parts of the system (equation 9). The optimization was run using Matlab interface and Gurobi solver, one of the solvers with the fastest computation times for solving linear optimization problems [109].

$$Ind_1 = \min\{t|T(t) = 18^oC\} \tag{7}$$

$$Ind_2 = \frac{Q - Q_{ref}}{Q_{ref}} \tag{8}$$

$$Ind_3 = \frac{C_{operational,ref} - C_{operational}}{C_{operational,ref}} \tag{9}$$

Three different scenarios about the heating strategy were investigated and can be seen in Table 10: the first one included a complete cut-off of the heating system with no preheating or overheating strategy for a period that equals the thermal autonomy of each house (see equation 7). During the reference operation of the heating system, the setpoint was set to 21^{o}C. The second strategy considered 2 hours of preheating the house up to 24^{o}C followed by a complete cut-off of the heating system, while the third one considered 4 hours of preheating the houses up to 24^{o}C followed by a cut-off.

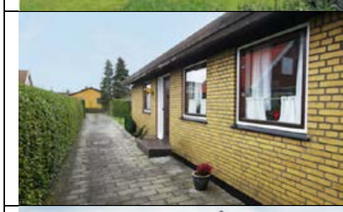
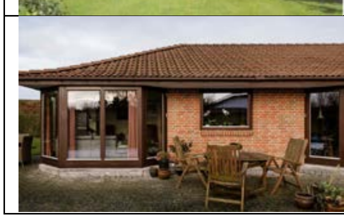
Table 10. DSM scenarios implemented to the building models

Scenario	Action	Preheating [h]	Notation
1	Heat cut-off	0	CO
2	Heat cut-off	2	CO_2hPH
3	Heat cut-off	4	CO_4hPH

The case study of this work was Sønderborg, Denmark. Around 53% of the area's heat demand is covered by the local district heating network [97]. The focus of the current study was on SFH, which represent the largest share of residential buildings in this area. The housing stock was treated with archetypes. One archetype building was modelled for each category and the results were aggregated.

The classification and characterization process can be seen in Table 11 along with some typical building characteristics of these archetypes. Based on available energy data that were used to calibrate the building models, six building archetypes were created representing the majority of the SFH in Sønderborg. The construction age of all the building models according to which the classification was made, their gross floor areas along with the area-weighted heat loss coefficient (U-value) of the total building envelope, the internal heat capacities, time constants, energy use intensity (EUI) and peak demand are presented in Table 11. These information represent the thermal resistance of the building envelope, the internal heat capacities, as well as the building energy performance. Moreover, the relative dominance of each building archetype is presented in Table 11 by giving the total floor area for each building archetype as a percentage of the total floor area for all dwellings in the Sønderborg area. The largest share of the local residential stock was represented by the newest archetype. The six building archetypes that were created represented 60% of the residential building stock and 55% of the district heating demand in Sønderborg. As expected, newer buildings had longer time constants.

Table 11. Building archetypes representing the examined Danish residential building stock

Building archetype	Constr. age	Floor area m^2	U-value* W/m^2K	Internal heat capacity e+07 J/K	Time constant h	EUI kWh/m^2 year	Peak demand kW	% of total floor area
	1930’s	85	0.72	1.44	20	225.7	7.5	5.9
	1950’s	87	0.47	2.34	43	148	4.4	6.8
	1960’s	140	0.54	2.33	29	143.6	7.3	7.8
	1960’s ref.	119	0.43	2.13	39	111.6	5.1	11.8
	1970’s	136	0.51	3.69	51	132.6	6.8	10.3
	1990’s	137	0.31	5.87	134	84.5	3.8	17.5

* Area-weighted U-value of total building envelope

4.3.2 Results

The results of the building energy analysis are presented in Table 12. According to those, the heat cut-off strategy resulted in energy savings (indicator 2) in all archetypes compared to the reference case, as expected. The effect of preheating control up to four hours was found to affect positively the thermal flexibility potential of buildings, but should be evaluated individually for each archetype. The experiment of the cut-off was implemented on a cold and grey day, so that the effect of extreme solar gains, was not considered, while transmission losses through the building envelope were increased. This choice, even though represented more than 80% of the Danish winter days, gave a pessimistic prediction to the results. The capacity of the heating system (i.e. hydronic radiators) and the duration of the cut-off were decisive for the peak loads that were created after the heat cut-off. The largest thermal autonomy was calculated for the newest archetype of the investigated sample and reached 6 hours. It should be noted that low-energy or nearly-zero energy buildings were not included in this study, which would result in significantly longer autonomy times.

Table 12. Results of heating strategies with regards to indicator 1 and 2 as implemented to the archetypes

	SCENARIO 1 (CO)		**SCENARIO 2 (CO_2hPH)**		**SCENARIO 3 (CO_4hPH)**	
Archetype	Ind_1 [h]	Ind_2 [%]	Ind_1 [h]	Ind_2 [%]	Ind_1 [h]	Ind_2 [%]
1930's	1	-0.3%	1	1.3%	2	4.2%
1950's	2	-1.9%	3	-1.5%	4	-0.7%
1960's	1	-1%	2	-2.1%	2	-1.4%
1960's ref.	2	-1.5%	3	-2.0%	4	-1.5%
1970's	1	-0.3%	2	-1.2%	2	0.4%
1990's	5	-14.6%	6	-18.1%	6	-15.2%

The total shifted heat demand during the year according to the optimization results, in terms of avoided heat demand during the cut-off events, for the 3 different strategies for the year 2015 and strategy two for the year 2029 (scenario CO_2hPH_2029) can be seen in Figure 24. In all the scenarios, the newer archetypes contributed more to the overall load shifting than the older ones. The oldest houses, represented by the 1930's archetype were utilized only in the cut-off (CO) scenario. All other archetypes contributed to the load shifting in all the scenarios. Archetypes 1960's ref, 1970's and 1980's accounted for 65%-70% of the total shifted load, in the first three scenarios. Moreover, the relatively old building archetype 1950's was often utilized, especially in scenarios CO_2hPH and CO_4hPH when it had larger thermal autonomy (3 and 4 hours compared to the 2 hours in the CO scenario). It was found feasible to utilise almost all of the identified load shifting potentials. The maximum possible capacity of the load shifting was 5.6% to 8.4% of the total DH demand of the city of Sønderborg in different scenarios. The real activation of the thermal mass for storage in different scenarios showed that the shifted load accounted for 5.5% to 7.7% of the total DH demand.

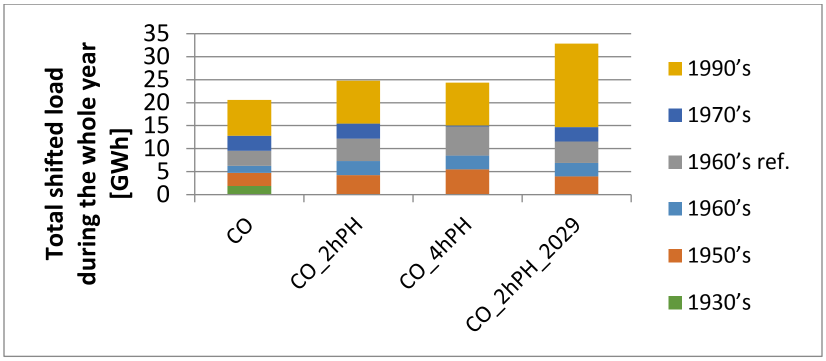


Figure 24. The total amount of shifted heat demand on the system scale in different scenarios for one whole heating season

Differences in operational costs of the district heating system, represented by Indicator 3, can be seen in Table 13. The economic savings according to the optimization findings of the district heating system of the city of Sønderborg were in the range of 0.7%-4.6% for the different scenarios. The largest saving potential occurred in the 2029 scenario. The main reason for much larger operational savings in 2029 was the larger capacity of solar thermal technology that was less curtailed when utilizing the thermal mass for storage. Moreover, more refurbished buildings meant that the larger share of buildings performed as the 1990's building archetype, having larger time constants and lower energy consumption (Indicator 2).

Table 13. Comparison of the operational costs of district heating in the city of Sønderborg in different scenarios (indicator 3).

	Reference (2015)*	CO	CO_2hPH	CO_4hPH	Reference (2029)*	CO_2hPH_2029
Operating costs of DH system [10^3€]	5,238	5,201	5,186	5,164	5,687	5,425
Savings compared to the reference case [%] (Indicator 3)		0.7%	1.0%	1.4%		4.6%

*Reference case costs were obtained by constraining the modelled system not to utilize the thermal mass for storage.

4.3.3 Main findings and conclusions

- The future scenario carried out for the year 2029 -that assumed carbon neutrality- had benefits of using the thermal mass for storage, which were greater in the future than in the current district heating system.
- The thermal autonomy was more important factor for triggering DSM events from the system point of view than the energy use intensity of different building archetypes.

- Thermal autonomy was mainly determined by the heat capacity of internal walls and heat losses from the external walls for old single-family houses during cold and cloudy weather conditions.
- Load shifting in relative terms was mostly utilized in the heat cut-off scenario, during 98% of the possible time.
- DSM events were usually triggered in the mornings, in the duration of 2-3 hours, and often in the evenings, especially during the winter time with higher overall DH demand.
- The characterization of building stock based on archetypes enabled the representation of the respective thermal flexibility potential.
- Smart controls should be invested in approximately 34% of the examined building stock, excluding buildings with very low thermal autonomy time.

5 Conclusion and Outlook

To implement residential energy demand policies, housing stock models are needed that can couple accuracy, reliability, simplicity, scalability and fast computation time. Aggregation of building energy demands can support decision-making policies in urban development projects. Two methods were tested with regards to aggregation of building energy demands on small district scale samples. It was found that the archetypal approach outperformed the micro-level building energy modelling method in the modelling and computation time. Uncertainty included in the input parameters affected both methods, with the archetypal approach being less dependent on it due to averaging phenomena across different buildings. Furthermore, the minimum information level that was needed to model reliably district-scale housing stocks consisted of basic typological information about the building, including construction characteristics and floor areas, as well as knowledge of the current refurbishment state of the building that highly influences the overall heat transfer coefficient of the building envelope.

The type of energy calculation methods that archetypal approaches are based on highly affected the accuracy of the results. Dynamic simulations resulted in more accurate predictions at aggregated level, with steady-state energy calculations having a higher deviation from measured energy data. However, it should be noted that simplifications required by urban scale energy models using steady-state conditions may increase results uncertainties to some extent, but the uncertainty involved in input parameters may induce larger uncertainties than uncertainties created by model simplifications. Dynamic simulations are particularly suitable for decision making for energy efficiency strategies and analyses where fine granular energy demand results are required.

The combination of physics-based methods with statistical methods can lead to reliable predictions making use of the scalability and flexibility that physical models contain. However, measured energy data at high resolution are needed that were not available at urban or district scale until recently. Based on an urban dataset, a simplified approach was proposed and demonstrated the estimation of equivalent temperature setpoints and thermal transmittance of the building envelope for a Danish residential building stock, which were validated with findings in literature. Setpoint temperature is key assumption of occupant behaviour that determines the heating demand in cold climates. The proposed method allowed to capture the heterogeneous behaviour among people that is reflected on temperature preferences. The estimation of the thermal transmittance of total building envelope can be used as an indication of energy refurbishment state of buildings, which is not regularly reported in building databases. The results can be transferrable across Scandinavia. They are applicable to urban building energy models, where uniform heating temperature setpoints for a whole heating season and U-values of the building envelope are important to define the building construction and thermal comfort. Similar methods combine lower complexity with higher generality and can be useful towards comparison among different buildings. Real measured energy data can be also useful for calibrating physical models. When coupled with administrative and land use records, they represent a critical resource for urban energy planning.

Due to the increasing deployment of smart meters in buildings, a large amount of granular data can be available. Data mining techniques and specifically, knowledge discovery in databases approaches have a large potential to identify hidden correlations between buildings and their influencing factors, as well as provide useful information about the consumer behaviour and patterns. A clustering-based analysis of district heating data of an urban-scale Danish housing stock was introduced and identified the representative load profiles and patterns. The number of representative groups of customers was highly dependent on the method employed, which was defined by the objective of the analysis. Nevertheless, the vast majority of the examined consumers was represented by fairly constant profiles. Two clustering methods were proposed: the first one aimed at determining heating consumption intensity groups, while the second one targeted the representative patterns of the load profiles. The former is suitable for energy benchmarking strategies and building classification schemes with regards to energy efficiency, while the latter is particularly useful for utilities and for developing targeted demand side management strategies (i.e. load shifting and peak shaving), where the time that the peak heating demand takes place is crucial.

A logistic regression analysis followed which determined that parameters, such as building age and floor area, were more correlated with heating demand than family size and age of occupants and thus, can be used in classification schemes to categorize detached houses with regards to energy use intensity. It was also concluded that consumption patterns for each individual customer were serially correlated, and groups with higher consumption intensity had lower variability in terms of the patterns. That means that a large share of consumers had regular and predictable consumer behaviours. Despite the homogeneous building sample that consisted of single-family houses, it is expected that residential district heating customers in Danish urban areas exhibit similar patterns, as the cultural habits are quite uniform in large Danish cities (e.g., work schedule, desired thermal comfort conditions). This research also highlighted the need for greater data transparency and data access from utility providers, where anonymization of data can greatly facilitate this process by addressing any privacy or confidentiality concerns. The resulting load profiles and patterns can support the development of characteristic consumer profiles that are attributed to archetypes and facilitate a scaled analysis of buildings at urban scale. In general, clustering analysis can handle the heterogeneity of buildings within housing stocks. Therefore, big energy data analytics combined with urban energy simulations set a new research direction towards urban zoning, as well as evaluating energy efficiency retrofitting strategies and smart energy management.

Considering a smart energy city context where the penetration of renewable sources is high, the balancing of the supply and demand sides is important to ensure security of supply. Buildings can play an important role to the flexible operation of the system, by integrating both active and passive strategies. The utilization of the thermal mass embedded in the building envelope was found to be a feasible solution due to its cost-effectiveness and applicability. The thermal flexibility that buildings can offer to the grid was quantified by two indicators: the first one reflected the thermal comfort degradation that occupants perceived after a specific heating control strategy, while the second one represented the stress on the heating system that was induced by control strategies. The former was an indication of thermal autonomy of the building, according to which a threshold was introduced as the lowest acceptable thermal comfort condition. It was found that thermal autonomy may vary significantly among residential buildings, with the newest ones providing the longest autonomy times. Considering a low-energy Danish apartment building, the energy flexibility indicators were greatly affected by the thermal transmittance of the building envelope and specifically, the external wall insulation. Furthermore, external factors such as solar gains and outdoor temperature also influenced highly the flexibility potential of the building.

The <u>flexibility</u> that buildings can offer to the energy system was investigated in two ways: i) by optimizing the operation of the heating system based on an electricity price signal for a single detached house using a dynamic model that could capture transient phenomena and ii) by solving an optimization problem for the whole energy system of a city, where the storage potential of the building thermal mass was added as a source of flexibility to the system. Results showed that the flexible operation of the grid-connected heat pump in the <u>single house</u> led to significant peak shaving and cost savings that were dependent on the weather (i.e. outdoor temperature) and on the wind penetration that affected the electricity prices. The electricity price signal reflected the lack or excess of renewables to the grid. The proposed method referred to future energy market designs, where the optimization of heating systems of individual buildings based on price signals would be possible.

As far as the <u>energy system optimization</u> was concerned, a third flexibility indicator was introduced, which characterized the economic viability of the scenarios. A future scenario for 2029 was also examined that included the municipal targets for carbon neutrality. It was found that the thermal autonomy was the main driver for triggering demand side management events from the system point of view instead of the energy use intensity of different building archetypes. In addition, the characterization and classification of the building stock into archetypes according to building age and refurbishment state represented the respective flexibility potential of buildings. The building stock consisted mainly of older houses, hence the heat capacity of internal walls and heat losses from the external walls determined the thermal autonomy of the buildings during typical Danish winter conditions. Preheating strategies had a less clear effect on the different archetypes. The future scenario carried out had economic benefits of utilizing the thermal mass for storage that were greater in the future than in the current district heating system due to the larger capacities of intermittent generation than can be successfully integrated. Finally, investing in smart controls that enable the implementation of demand side management techniques would be feasible for more than 30% of the examined housing stock, excluding buildings with very low thermal autonomy times.

To conclude, buildings can play an important role to future smart energy cities, ensuring security of supply and cost savings for end users and the system operator. Furthermore, application-specific model selection and development is essential and requires case-by-case consideration of all the aspects analysed in this thesis, including data availability, scale of analysis, accuracy of results and generality.

6 Future investigations and recommendations

Based on the findings of this work, the following issues are worthwhile to investigate and consider in the future:

- Clustering analysis can be used to develop representative consumer profiles that introduce stochasticity to housing stock models. Such profiles can be fed to urban building energy modelling tools and account for heterogeneity in consumer behaviour.
- Predictable and regular consumer behaviour is suitable to perform prediction analysis for specified groups of consumers. Data-driven algorithms can be employed to conduct this, such as Gaussian Mixture models or Artificial Neural Networks.
- The correlation of occupant-related information (e.g. number and age of occupants) with heating load profiles should be further investigated. Building classification based on occupant information should be based on the results of such analysis.
- The archetypal approach can be used for aggregation of building energy demands when a minimum information level for typological building characteristics is available.
- To better understand energy dynamics at urban scale, it would be useful to cover a diversity of building types, locations and climates.
- Uncertainty should be incorporated to housing stock models by probabilistic methods (i.e. Bayesian calibration, Monte Carlo simulations). When applied to urban scale, deterministic models (e.g. quasi-steady-state models) that allow the parameters to vary according to distributions derived by probabilistic methods can handle parameter uncertainty with reduced computation time. They can also be used for building archetype calibration.
- Data-driven models should be expanded towards multiple indexes to provide a comprehensive analysis of building performance including thermal comfort evaluation, instead of sole prediction of energy loads.
- The utilization of the thermal mass of buildings should be investigated in more complex systems, i.e. including additional storage mechanisms, and prioritization strategies should be developed.
- The increasing amount of macro-level energy data calls for secure management platforms developed by utilities, authorities and municipalities. Anonymization is key to handling privacy issues and enabling data distribution.

References

1. United Nations (2017). The Sustainable Development Goals Report 2017. [Online] Available at: https://unstats.un.org/sdgs/files/report/2017/TheSustainableDevelopmentGoalsReport2017.pdf [Accessed 17 January 2018].
2. Directorate-General for Environment, European Commission (2014). Resource efficiency in the building sector. Ecorys, Copenhagen Resource Institute. [Online] Available at: http://ec.europa.eu/environment/eussd/pdf/Resource%20efficiency%20in%20the%20building%20sector.pdf] [Accessed 20 September 2017].
3. M. Geidl, G. Koeppel, P. Favre-Perrod, B. Klöckl, G. Andersson, K. Fröhlich (2007). Energy hubs for the future. IEEE Power & Energy Magazine, 5(1): 24–30.
4. P. Meibom, K. B. Hilger, H. Madsen and D. Vinther (2013). Energy Comes Together in Denmark: The Key to a Future Fossil-Free Danish Power System. IEEE Power & Energy Magazine, 11 (5): 46-55. doi: 10.1109/MPE.2013.2268751 .
5. C.F. Calvillo, A. Sanchez-Miralles, J. Villar (2016). Energy management and planning in smart cities. Renewable and Sustainable Energy Reviews 55 (2016); 273–287.
6. Danish Government (2013). The Danish Climate Policy Plan - Towards a low carbon society. [Online] Available at: www.kebmin.dk [Accessed 15 June 2017].
7. Danish Ministry of Climate, E. a. B., (2013). Energy Policy Report 2013. [Online] Available at: http://www.ens.dk/sites/ens.dk/files/policy/danish-climate-energy-policy/dkenergypolicyreport2013_final.pdf [Accessed 20 July 2017].
8. European Union (2010). Directive 2010/31/EU of the European Parliament and of the Council on the Energy Performance of Buildings. [Online] Available at: http://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32010L0031&from=EN [Accessed 10 December 2017].
9. M. Meijer, F. Adriaens, O. Linden, W. Schik (2011). A next step for sustainable urban design in the Netherlands. Cities, Volume 28; 536-544.
10. J. O. Lewis, S. N. Hógáin, A. Borghi (2013). Cities of Tomorrow – Action Today. URBACT II Capitalisation. Building energy efficiency in European Cities. URBACT. [Online]. Available at: http://urbact.eu/sites/default/files/import/general_library/19765_Urbact_WS6_ENERGY_low_FINAL.pdf (Accessed 15 October 2017).
11. European Commission (2010). Covenant of Mayors: How to develop a sustainable energy action plan (SEAP)-Guidebook. Publication office of the European Union. URL: http://www.covenantofmayors.eu/index_en.html
12. C. Harrison, I.A. Donnelly (2011). A Theory of Smart Cities. Hull, UK, Proceedings of the 55th Annual Meeting of the ISSS.
13. European Commission (2014). Energy research challenges for Smart Cities. Energy Research Knowledge Centre. [Online] Available at: https://setis.ec.europa.eu/energy-research/sites/default/files/library/ERKC_PB_Smart_Cities.pdf [Accessed December 2, 2017].

14. R. Giffinger, et al. (2007). Smart cities - Ranking of European medium-sized cities. Final report. [Online] Available at: http://www.smart-cities.eu/download/smart_cities_final_report.pdf [Accessed 1 June 2014].
15. P. Neirotti, et al. (2014). Current trends in Smart City initiatives: Some stylised facts. Cities, Volume 38; 25-36.
16. H. Lund (2014). Renewable Energy Systems: A Smart Energy Systems Approach to the Choice and Modeling of 100% Renewable Solutions 2nd ed. Academic Press, Burlington, USA, (ISBN: 978-0-12-410423-5).
17. D.S. Chai, J.Z. Wen, J. Nathwani (2013). Simulation of cogeneration within the concept of smart energy networks. Energy Convers. Manag. 75; 453–465. http://dx.doi.org/10.1016/j.enconman.2013.06.045.
18. H. Lund, P. Alberg Østergaard, D. Connolly, B. Vad Mathiesen (2017). Smart energy and smart energy systems. Energy 137; 556-565.
19. F. Mosannenzadeh, A. Bisello, R. Vaccaro, V. D'Alonzo, G. Wayne Hunter, D. Vettorato (2017). Smart energy city development: A story told by urban planners. Cities 64; 54-65.
20. F. Mosannenzadeh, M.R. Di Nucci, D. Vettorato (2017). Identifying and prioritizing barriers to implementation of smart energy city projects in Europe: An empirical approach. Energy Policy 105; 191–201.
21. K. Halsnæs, K. B. Karlsson, T. H. Y. Føyn, B. Bach, R. Schütz, J. Schmidmayer (2011). Smart city initiatives in Denmark and Europe. In Risø energy report 10: Energy for smart cities in an urbanised world (pp. 17-22). Roskilde: Danmarks Tekniske Universitet, Risø Nationallaboratoriet for Bæredygtig Energi. (Denmark. Forskningscenter Risoe. Risoe-R; No. 1778(EN)).
22. EnergyLab Nordhavn (2015). [Online] Available at: URL: http://www.energylabnordhavn.dk/ [Accessed 13 December 2017].
23. H. Madsen (2013). CITIES - Centre for IT-Intelligent Energy Systems in Cities. Project Description.
24. International Energy Agency (2001). Annex 31 - Stock aggregation: methods for evaluation of the environmental performance. International Initiative for a Sustainable Built Environment (iiSBE).
25. L. Swan, V. Ugursal (2009). Modeling of end-use energy consumption in the residential sector: A review of modeling techniques. Renewable and Sustainable Energy Reviews 13; 1819-1835.
26. S.T. Moghadam, J. Toniolo, G. Mutani, P. Lombardi (2018). A GIS-statistical approach for assessing built environment energy use at urban scale. Sustainable Cities and Society 37; 70-84.
27. R. Buffat, A. Froemelt, N. Heeren, M. Raubal, S. Hellweg (2017). Big data GIS analysis for novel approaches in building stock modelling. Applied Energy 208; 277-290.
28. C. Koulamas, A.P. Kalogeras, R. Pacheco-Torres, J. Casillas, L. Ferrarini (2018). Suitability analysis of modeling and assessment approaches in energy efficiency in buildings. Energy and Buildings 158; 1662–1682.

29. M. Kavgic, A. Mavrogianni, D. Mumovic, A. Summerfield, Z. Stevanovic, M. Djurovic-Petrovic (2010). A review of bottom-up building stock models for energy consumption in the residential sector. Building and Environment 45; 1683-1697.
30. K. Sun, T. Hong (2017). A framework for quantifying the impact of occupant behaviour on energy savings of energy conservation measures. Energy and Buildings 146, pp. 383-396.
31. W. O'Brien, I. Gaetani, S. Carlucci, P-J. Hoes, J. L.M. Hensen (2017). On occupant-centric building performance metrics. Building and Environment 122; 373-385.
32. ISO13790, 2008. Energy performance of buildings - Calculation of energy use for space heating and cooling, CEN, European Committee for Standardization: Brussels, Belgium.
33. A. Heller (2002). Heat-load modelling for large systems. Applied Energy 72; 371-387.
34. L. Pedersen, J. Stang, R. Ulseth (2008). Load prediction method for heat and electricity demand in buildings for the purpose of planning for mixed energy distribution systems. Energy and Buildings, Volume 40; 1124-1134.
35. S. Matsumoto (2016). How do household characteristics affect appliance usage? Application of conditional demand analysis to Japanese household data. Energy Policy 94; 214-223.
36. S.I. Gallant (1995). Neural network learning and expert systems. 3rd ed. Cambridge, USA: The MIT Press.
37. R. Kramer, J. van Schijndel, H. Schellen, 2012. Simplified thermal and hygric building models: A literature review. Frontiers of Architectural Research 1; 318-325.
38. M.A. Rafe Biswas, M.D. Robinson, N. Fumo (2016). Prediction of residential building energy consumption: A neural network approach. Energy 117; 84-92.
39. M.T. Hagan, H.B. Demuth, M.H. Beale (1996). Neural network design. Boston: PWS Publishing Company.
40. R. Kumar, R.K. Aggarwal, J.D. Sharma (2013). Energy analysis of a building using artificial neural network: A review. Energy and Buildings 65; 352-358.
41. Y. Wei, X. Zhang, Y. Shi, L. Xia, S. Pan, J. Wu, M. Han, X. Zhao (2018). A review of data-driven approaches for prediction and classification of building energy consumption. Renewable and Sustainable Energy Reviews 82; 1027–1047.
42. N.R. Kristensen, H. Madsen, S.B. Jørgensen (2004). Parameter estimation in stochastic grey-box models. Automatica 40; 225-237.
43. P. Bacher, H. Madsen, H. Aalborg Nielsen, B. Perers (2013). Short-term heat load forecasting for single family houses. Energy and Buildings 65; 101-112.
44. M. Aydinalp-Koksal, V.I. Ugursal (2008). Comparison of neural network, conditional demand analysis, and engineering approaches for modeling end-use energy consumption in the residential sector. Applied Energy 85; 271-296.
45. J.H. Kämpf, D. Robinson (2007). A simplified thermal model to support analysis of urban resource flows. Energy and Buildings, 39(4); 445-453.
46. M. Lauster, et al., 2014. Low order thermal network models for dynamic simulations of buildings on city district scale. Building and Environment 73; pp. 223-231.

47. E.H. Borgstein, R. Lamberts, J.L.M. Hensen (2016). Evaluating energy performance in non-domestic buildings: A review. Energy and Buildings 128; 734–755.
48. ASHRAE (2013). ASHRAE Handbook: Fundamentals 2013.
49. E. Mata, A. Sasic Kalagasidis, F. Johnson (2014). Building-stock aggregation through archetype buildings: France, Germany, Spain and the UK. Building and Environment 81; 270-282.
50. S. Moffatt (2004). Stock aggregation e methods for evaluating the environmental performance of building stocks. In: Annex 31 energy related environmental impact of buildings. International Energy Agency (IEA).
51. M. Deru, K. Field, D. Studer, K. Benne, B. Griffith, P. Torcellini, B. Liu, M. Halverson, D. Winiarski, M. Rosenberg, M. Yazdanian, J, Huang, D. Crawley (2011). U.S. Department of Energy Commercial Reference Building Models of the National Building Stock. Technical report. National Renewable Energy Laboratory [Online] Available at: https://www.nrel.gov/docs/fy11osti/46861.pdf [Accessed 30 November 2017].
52. IEA EBC Annex 53 (2013). Final Report. Total energy use in buildings, November 2013.
53. C. Cerezo, J. Sokol, S. AlKhaled, C. Reinhart, A. Al-Mumin, A. Hajiah (2017). Comparison of four building archetype characterization methods in urban building energy modeling (UBEM): A residential case study in Kuwait City. Energy and Buildings 154; 321-334.
54. P. Gianniou, A. Heller, P.S. Nielsen, K. Negendahl, C. Rode (2015). Aggregation of building energy demands. In Proceedings Building Simulation 2015, Hyderabad, India.
55. I. Sartori, B. Jensen Wachenfeldt, A. Grete Hestnes (2009). Energy demand in the Norwegian building stock: Scenarios on potential reduction. Energy Policy 37; 1614-1627.
56. A.A. Famuyibo, A. Duffy, P. Strachan (2012). Developing archetypes for domestic dwellings—An Irish case study. Energy and Buildings 50; 150-157.
57. J. Sokol, C. Cerezo, C. Reinhart (2017). Validation of a Bayesian-based method for defining residential archetypes in urban building energy models. Energy and Buildings 134; 11-24.
58. M. Heine Kristensen, R. Choudhary, S. Petersen (2017). Bayesian calibration of building energy models: Comparison of predictive accuracy using metered utility data of different temporal resolution. In Proceedings of CISBAT 2017, Energy Procedia 122; 277-282.
59. A.T. Booth, R. Choudhary, D.J. Spiegelhalter (2012). Handling uncertainty in housing stock models. Building and Environment 48; 35-47.
60. J.N.B. Fonseca, M.J.N. Oliveira Panão (2017). Monte Carlo housing stock model to predict the energy performance indicators. Energy and Buildings 152; 503-515.
61. C.E. Kontokosta, C. Tull (2017). A data-driven predictive model of city-scale energy use in buildings. Applied Energy 197; 303-317.
62. C. Cerezo Davila, C.F. Reinhart, J.L. Bemis (2016). Modeling Boston: A workflow for the efficient generation and maintenance of urban building energy models from existing geospatial datasets. Energy 117; 237-250.

63. L. Frayssinet, l. Merlier, F. Kuznik, J.-L. Hubert, M. Milliez, J.-J. Roux (2018). Modeling the heating and cooling energy demand of urban buildings at city scale. Renewable and Sustainable Energy Reviews 81; 2318–2327.
64. Y. Chen, T. Hong, M.A. Piette (2017). Automatic generation and simulation of urban building energy models based on city datasets for city-scale building retrofit analysis. Applied Energy 205; 323-335.
65. C.F. Reinhart, C. Cerezo Davila (2016). Urban building energy modeling – A review of a nascent field. Building and Environment 97; 196-202.
66. G. Gröger, L. Plümer (2012). CityGML – Interoperable semantic 3D city models. ISPRS Journal of Photogrammetry and Remote Sensing 71; 12–33.
67. The European Parliament and the Council of European Union (2009). Directive 2009/72/EC concerning common rules for the internal market in electricity and repealing Directive 2003/54/EC. Official Journal of the European Union.
68. The European Parliament and the Council of European Union (2009). Directive 2009/73/EC concerning common rules for the internal market in natural gas and repealing Directive 2003/55/EC. Official Journal of the European Union.
69. European Commission (2014). Benchmarking smart metering deployment in the EU-27 with a focus on electricity. Report from the Commission, Brussels.
70. A.K. Pujari, K. Rajesh, D.S. Reddy (2001). Clustering Techniques in data Mining A survey. IETE Journal of Research 47(1-2); 19–28.
71. Z. Yu (2012). Mining Hidden Knowledge from Measured Data for Improving Building Energy Performance. PhD Thesis, Concordia University.
72. S. Darby (2010). Smart metering: what potential for householder engagement? Building Research & Information 38; 442-457.
73. P. Cabena, P. Hadjinian, R. Stadler, J. Verhees, A. Zanasi A (1998). Discovering data mining: From Concept to Implementation, Prentice-Hall, Inc.
74. U. Fayyad, G. Piatetsky-Shapiro, P. Smyth P (1996). From Data Mining to Knowledge Discovery in Databases. AI Magazine 17 (3).
75. S. D’Oca, S. Corgnati, T. Hong (2015). Data Mining of Occupant Behaviour in Office Buildings. In Proceedings of IBPC 2015, Energy Procedia 78; 585 – 590.
76. J.W. Han, M. Kamber, J. Pei (2012). Data Mining: Concepts and Techniques, 3rd ed., Morgan Kaufman Publishers, MA, USA.
77. C. Fan, F. Xiao, Z. Li, J. Wang (2018). Unsupervised data analytics in mining big building operational data for energy efficiency enhancement: A review. Energy and Buildings 158; 296-308.
78. Hsu D (2015). Comparison of integrated clustering methods for accurate and stable prediction of building energy consumption data. Applied Energy 160; 153–163.
79. S. Guha, R. Rastogi, K. Shim (1998). CURE: An Efficient Clustering Algorithm for Large Databases. ACM Sigmod Record 27(2); 73–84.

80. A.K. Mann, N. Kaur (2013). Review paper on clustering techniques, Global Journal of Computer Science and Technology Software and Data Engineering 13 (5).
81. P. K. Senyo, J. Effah, E. Addae (2016). Preliminary insight into cloud computing adoption in a developing country. Journal of Enterprise Information Management, 29(4); 400–422.
82. P.K. Senyo, E. Addae, R. Boateng (2018). Cloud computing research: A review of research themes, frameworks, methods and future research directions. International Journal of Information Management 38; 128–139.
83. Danish Energy Agency (2015). Danish energy and climate outlook. Published in December 2015 by the Danish Energy Agency, Copenhagen, Denmark.
84. Danish Energy Agency (2015). Regulation and planning of district heating in Denmark. [Online] Available at: https://ens.dk/sites/ens.dk/files/Globalcooperation/regulation_and_planning_of_district_heating_in_denmar520 k.pdf [Accessed 23 October 2017].
85. C.W. Gellings (1985). The concept of demand-side management for electric utilities. Proceedings of the IEEE, 73(10); 1468–1470.
86. IEA Annex 67 (2017). Energy Flexibility as a key asset in a smart building future. Position Paper of the IEA Energy in Building and Communities Program Annex 67 'Energy Flexible Buildings'.
87. Y. Zong, G.M. Boning, R.M. Santos, S. You, J. Hu, X. Han (2017). Challenges of implementing economic model predictive control strategy for buildings interacting with smart energy systems. Applied Thermal Engineering 114; 1476-1486.
88. T. Yan, R.Z. Wang, T.X. Li, L.W. Wang, I.T. Fred (2015). A review of promising candidate reactions for chemical heat storage. Renewable and Sustainable Energy Reviews 43; 13–31.
89. L. Navarro, A. de Gracia, S. Colclough, M. Browne, S.J. McCormack, P. Griffiths (2016). Thermal energy storage in building integrated thermal systems: A review. Part 1. active storage systems. Renew Energy 88; 526–547.
90. D. Picard, J. Drgona, M. Kvasnica, L. Helsen (2017). Impact of the controller model complexity on model predictivecontrol performance for buildings. Energy and Buildings 152; pp. 739–751.
91. TABULA Project Team (2012). Typology Approach for Building Stock Energy Assessment. Main Results of the TABULA project. Editors: T. Loga, N. Diefenbach, B. Stein. Institut Wohnen und Umwelt GmbH. http://episcope.eu/index.php?id=97
92. Danish Building Register, Bygnings- og Boligregistret (2018). [Online] https://bbr.dk/ [Accessed 5 January 2018].
93. M. Palonen, M. Hamdy, A. Hasan (2013). MOBO a new software for multi-objective building performance optimization. Proceedings of Building Simulation, Chambery, France, August 26-28.
94. K. Negendahl (2014). Parametric City Scale Energy Modeling - Perspectives on using Termite in city scaled models. Copenhagen, Denmark, iiESI European Workshop 2014, Technical University of Denmark, DOI:10.13140/RG.2.1.3624.5926.

95. Danish Meteorological Institute (2013). 2001-2010 Danish Design Reference Year. Technical Report 13-19. Danish Ministry of Climate, Energy and Buildings. Copenhagen, Denmark.
96. TABULA database (2012). [Online] Available at: http://webtool.building-typology.eu/#bm [Accessed 25 January 2018].
97. PlanEnergi (2014). Strategisk energiplan for Sønderborg - Sønderborgs energisystem i 2029.
98. Styrelsen for DataForsyning og Effektivisering. Kortforsyningen (2018). [Online] Available at: https://kortforsyningen.dk/ [Accessed 8 January 2018].
99. S. Aggerholm, K. Grau (2016). Bygningers energibehov. SBi-Anvisning 213. Danish Building Research Institute.
100. A.R. Day (1999). An investigation into the estimation and weather normalisation of energy consumption in buildings using degree-days. PhD Thesis, South Bank University, United Kingdom.
101. C.E. Kontokosta, C. Tull (2017). A data-driven predictive model of city-scale energy use in buildings. Applied Energy 197; 303-317.
102. S. Aggerholm (2013). Cost-optimal levels of minimum energy performance requirements in the Danish Building Regulations. SBi 2013:25. Danish Building Research Institute, Aalborg University.
103. K.B. Wittchen, L. Mortensen, S.B. Holøs, N. F. Björk, S. Vares, T. Malmqvist (2012). Building typologies in the Nordic countries, Identification of potential energy saving measures. Danish Building Research Institute, SBi2012:04.
104. E.K. Robert, W. Larry (1995). A reference Bayesian test for nested hypotheses and its relationship to the Schwarz criterion. Journal of the American Statistical Association 90(431); 928–934.
105. J. Yang, J. Leskovec (2011). Patterns of temporal variation in online media. In Proceedings of the 4th ACM international conference on Web search and data mining; 177-186.
106. C. Zhu, Z. Wang (2017). Entropy-based matrix learning machine for imbalanced data sets. Pattern Recognition Letters 88; 72–80.
107. S. Kiluk (2017). Diagnostic information system dynamics in the evaluation of machine learning algorithms for the supervision of energy efficiency of district heating-supplied buildings. Energy Conversion and Management 150; 904–913.
108. Association of Danish Energy Companies (2015). Electric Utility tariffs and electricity prices (Danish title: Dansk Energi. Elforsyningens tariffer & elpriser).
109. Gurobi (2018). [Online] Available at: http://www.gurobi.com/ [Accessed 15 October 2017].

Appendix – Published, accepted and submitted papers

Paper I. Gianniou, P., Liu, X., Heller, A., Sieverts Nielsen, P., Rode, C. (2018). Clustering-based Analysis for Residential District Heating Data. *Energy Conversion and Management 165*; 840–850.

Energy Conversion and Management 165 (2018) 840–850

Contents lists available at ScienceDirect

Energy Conversion and Management

journal homepage: www.elsevier.com/locate/enconman

Clustering-based analysis for residential district heating data

Panagiota Gianniou, Xiufeng Liu*, Alfred Heller, Per Sieverts Nielsen, Carsten Rode

Technical University of Denmark, Denmark

ARTICLE INFO

Keywords:
Clustering
Load pattern
Load profiling
Load transition
Variability

ABSTRACT

The wide use of smart meters enables collection of a large amount of fine-granular time series, which can be used to improve the understanding of consumption behavior and used for consumption optimization. This paper presents a clustering-based knowledge discovery in databases method to analyze residential heating consumption data and evaluate information included in national building databases. The proposed method uses the *K*-means algorithm to segment consumption groups based on consumption intensity and representative patterns and ranks the groups according to daily consumption. This paper also examines the correlation between energy intensity and the characteristics of buildings and occupants, load profiles of households, consumption behavior changes over time, and consumption variability. The results show that the majority of the customers can be represented by fairly constant load profiles. Calendar context has an impact not only on the patterns but also on the consumption intensity and user behaviors. The variability studies show that consumption patterns are serially correlated, the customers with high energy consumption have lower variability, and the consumption is more stable over time. These findings will be valuable for district heating utilities and energy planners to optimize their operations, design demand-side management strategies, and develop targeting energy-efficiency programs or policies.

1. Introduction

Information and communications technologies (ICTs) are revolutionizing today's energy management systems. The distinct characteristic is to use smart meters to monitor energy consumption. Smart meters are the digital devices that can collect energy consumption at a fine-granular time level, typically every 15 min [1]. Smart metering systems are being widely installed globally. European countries have set the goal of converting their legacy meters to smart meters by 2020 in line with the Third Energy Package in the Electricity Directive [2] and Gas Directive [3] issued by European Commission in 2006. It is expected that 72% of European consumers will own electricity smart meters by 2020 and 40% of them will have gas meters [4]. In Denmark, more than half of current electricity customers have smart meters [5]. In recent years, several large-scale smart meter installation projects have been carried out, including the project undertaken by Enel SpA in Italy covering 30 million customers between 2000 and 2005, the Linky pilot project in France involving 300,000 customers and the national Australian initiative in Victoria covering a total of 2.6 million electricity customers. Furthermore, smart meters have been installed by Chubu Electric Power in Japan for most of their high-voltage customers, office buildings, and individual households.

The main drivers behind the employment of smart metering in different countries include load management, peak or demand reduction, fraud reduction, accurate billing and water conservation [6]. Smart meters can also be utilized to develop more accurate prediction models and detailed analyses on the drivers of building energy consumption [7]. For example, the data can reveal potentially valuable information about buildings and end users that are useful to energy management. In particular, building-related data can unveil hidden correlations between energy use and its influencing factors in buildings. Smart meter data can also help developing and applying control strategies to improve building energy performance and efficiency [8]. In addition, building energy-related information can be provided to customers [9], which can help them use demand-response techniques to reduce energy consumption, improve energy efficiency, reduce carbon emissions and improve the use of renewable energy sources. There is also an expectation that smart metering or advanced metering infrastructure can contribute to demand and cost reduction and to the adoption of domestic low-and zero-carbon technologies [6]. Therefore, they can be utilized to decrease uncertainty related to building energy performance and provide detailed information on energy monitoring. Smart meter data analyses are therefore not only of value to the research community and customers but they also help utilities to improve their meter-to-cash processes. They enable them to use advanced tariff schemes and ultimately contribute to the value of metering

* Corresponding author.
E-mail address: xiuli@dtu.dk (X. Liu).

https://doi.org/10.1016/j.enconman.2018.03.015
Received 13 December 2017; Received in revised form 5 March 2018; Accepted 7 March 2018
0196-8904/

infrastructure [10,11].

Buildings account for nearly one-third of the world's energy consumption. Among others, half of the global energy consumption of buildings comes from space heating and cooling as well as hot water [12]. In heating-dominated climates, such as Scandinavia, the main source of building energy demand is space heating. The increasing availability of smart meter data makes it possible to gain insights into heating consumption of buildings to help with energy management, such as extracting hidden temporal patterns – knowledge which can not be captured at its detailed level without the use of smart meters. Despite the promising benefits, it is still difficult to obtain reliable and detailed heating data, largely due to commercial sensitivity and privacy issues. Moreover, heating sector management is typically more challenging than other energy sectors, such as electricity, due to the high variation of production and demand. It is therefore difficult to obtain reliable data on heat production, usage patterns, and production costs. In addition, the installation of heating meters is more challenging and costly than smart electricity meters [13]. This also means that data analyses in the heating sector receive much less attention.

In this paper, a clustering-based knowledge discovery in database approach is proposed for analyzing residential heating consumption data. The objective of this study is to provide the information to district heating utilities, which they can use for optimizing their operations and for better understanding their customers. At the same time, the analysis can be used by customers to understand their consumption profiles and behaviors to improve energy efficiency. In addition, this study can help utilities and decision makers to develop energy efficiency strategies and policies, as well as provide personalized energy services to specific customer segments. Due to the introduction of renewable energy sources and the electrification of the heating sector, the Danish energy sector is in a transition period, which means that the balancing of the power grid is challenging. Energy flexibility solutions can support this transition, including demand-side management techniques that balance utility and customer needs. Identification and mapping of consumption patterns enable district heating utilities to implement new operational strategies. This study will perform a statistical clustering analysis on heating consumption data of Danish dwellings connected to a local district heating network. The data consists of the load profiles of 8293 single-family households from Aarhus, Denmark. The study clusters the customers into different groups by the *K*-means clustering algorithm, ranks the groups according to their consumption intensity, and labels them using different alphabets. An exploratory analysis is conducted on the energy consumption and socio-technical data of the dwellings, which reveals the correlation between energy consumption and the characteristics of the buildings and occupants.

This paper makes the following contributions: First, this paper presents a clustering-based approach for district heating consumption data analysis, including data preparation, clustering, and the analysis based on the clustering results. Second, this paper uses the *K*-means algorithm to segment consumption intensity groups, studies the correlation between the consumption intensity and the characteristics of buildings and occupants, and analyzes the transition of consumption behaviors over time. Third, this paper identifies representative patterns by clustering normalized daily consumption patterns and studies the variability of consumption patterns using an entropy approach. Fourth, this paper leads to a number of interesting findings from the clustering-based analysis including non-intuitive results, such as the calendar impact on consumption, the serial correlation of consumption patterns, and high consumption with a lower variability.

The remainder of this paper is organized as follows. Section 2 gives a review of related work. Section 3 presents the methods applied. Section 4 presents the data and the results of the clustering-based analysis. Section 5 discusses the significant findings of this paper and the related issues. Section 6 concludes the paper and points to the future research direction.

2. Literature review

Clustering is one of the most commonly used techniques in data mining. It is used to discover groups and identify interesting distributions in the underlying data [14]. Specifically, a clustering problem is about partitioning a given data set into groups, classes or clusters such that the data points in the same cluster are more similar compared with the data points in other clusters [14]. The goals of cluster analysis are data reduction, hypothesis development and testing and prediction identification, based on groups [15]. The main steps of clustering include feature selection, choice of the clustering algorithm, validation and interpretation of results [16]. Clustering algorithms fall into the following categories based on how clusters are defined: partitional clustering, including *K*-means and *K*-medoids methods, and hierarchical clustering, which includes density-based and grid-based clustering [17]. Clustering has been applied to different fields such as biology and web mining. Clustering techniques are being applied to the analysis of smart meter data in buildings and district heating systems, because of the rapid introduction of smart meters. Knowledge about the characteristics of clusters of consumers with similar consumption patterns can facilitate the development of novel tariff schemes and improve network management [18].

Most of the studies on clustering analysis using smart meter data are found in the electricity sector where smart meters are common [19]. In [20], the clustering methods for electricity consumption pattern recognition were assessed, including hierarchical and *K*-means algorithms. It was found that the *K*-means algorithm is more useful for segmenting customer groups, and that information hidden in the consumption patterns has a great potential to be exploited towards load control, demand response actions, real-time pricing, etc. In [21], hourly electricity consumption data from 4500 smart meters, covering all categories of Danish customers were used to identify the load profiles and classify them. It was concluded that modeling of individual load profiles should be differentiated between different categories of customers, such as between industrial and residential uses, between weekday and weekend, and between summer and winter. In [18], an automatic classification method was proposed based on self-organizing maps, which identified a set of household properties that could be deduced from electricity consumption data, such as the size of household and the income of occupants. Two methods were introduced in [22] for modeling total energy consumption of buildings with regards to predictive accuracy and cluster stability: clusterwise regression and cluster validation methods to measure stability. Clusterwise regression gave very accurate predictions but relatively unstable clusters, while *K*-means gave more stable clusters but poor predictions in several clusters. Thus, clustering methods should be carefully selected considering the main objective, whether it is accuracy or stability. In [23], a Bayesian nonparametric clustering algorithm was used to distinguish specific features between electricity consumption patterns of 219 households in UK and Bulgaria. The main advantage of the method was that it was not necessary to pre-define the number of clusters before analyzing. However, the algorithm was a bit slower in data processing than other clustering techniques. It was found that features such as nationality, family size, and type of dwelling can be successfully assigned to the members of a specific cluster.

Fewer studies have been found on data analysis in the heating sector compared with the electricity sector. In [24], the heating data from 139 single-family houses in Denmark were investigated for identifying patterns in space heating profiles using the *K*-means algorithm. It was found that the heating load profiles varied with external load conditions (high, medium and low demand periods). Two main clusters of load profiles were identified for weekdays and weekend days, one of which was quite stable and the other had a higher variation throughout the day. The latter was characterized by two peaks, one in the morning and the other in the afternoon/evening. It was concluded that building characteristics like floor area, building year and type of space heating distribution system, the existence of children or teenagers and the postcode proved to explain the difference between the two patterns. In another study [25], typical daily heating profiles were identified in a three years data set for 19 high education buildings in Norway using a Partitioning Around Medoids clustering algorithm. The Pearson Correlation Coefficient was used to determine the dissimilarity measure to cluster the daily load profiles based on the similarity of the variability, rather than the magnitude similarity. The typical heating load profiles provided information on the peaks and troughs of the daily heating consumption, daily high heating consumption period and daily load variation. Two statistical analysis approaches were introduced in [26] to develop random domestic profiles based on 15-min interval time series data from 25 typical Danish households. The results showed that space heating consumption had larger variability during a year, compared with domestic hot water consumption, which was more sporadic and transient.

Many studies identify the shape of load profiles as the primary objective and determine stable representative clusters. These methods are commonly referred to as shape-based. A shape-based approach to classifying energy consumption profiles at the household level was presented in [27]. The method was based on a dynamic time warping method that reflected the effect of hidden patterns of regular end-user behaviors. This method resulted in a smaller number of clusters, higher clustering efficiency and a lower household variability. In another study [28], a k-shaped approach was introduced, relying on a scalable iterative refinement process that created homogeneous and well-separated clusters. In order to consider the shape of the time series, a normalized version of the cross-correlation measure was used as the distance metric for the clustering. It was concluded that the proposed shape-based approach outperformed all scalable and non-scalable approaches in terms of accuracy, it was domain-independent and highly efficient for time series analysis.

In summary, clustering algorithms have been widely used for customer segmentation, pattern recognition, and classification. Most of the literature uses electricity data, which are more readily available and accessible. There is much less research carried out for residential heating consumption data analysis. Load shape recognition is the major objective of the reviewed articles. In contrast, in this paper, the analysis is conducted based on the clustering results, including load profiling, consumption behavior change over time, and the variability of the heating consumption patterns.

3. Methods

This paper uses a three-stage approach to process and analyze district heating consumption data. Fig. 1 describes an overview of this approach, which includes data preparation, clustering and the analysis based on the clustering results. First, the heating consumption data are collected by the preparation module. The data usually contain some anomalous and missing consumption values that need to be cleaned. Data cleansing is an important step for ensuring data quality for the subsequent clustering-based data analysis. The data are pre-processed by filtering the single-family household data, removing the data with missing values, and adjusting the reading time stamp to align winter and summer time. The second module does the clustering on the load profiles. The K-means algorithm is employed for the clustering to segment different customer groups for further analysis in the third module. Two clustering analyses are applied. The first one is applied on the non-normalized data to identify the consumption intensity groups. The second clustering analysis is applied on the normalized data to identify the representative load patterns of all customers. It is expected that the first clustering analysis highlights the difference in the magnitude of heating consumption among customers. The second analysis points out the patterns of heating loads with regards to time. The optimal number of clusters is determined separately in both approaches, and different labels are assigned to the identified customer groups. The analysis of the clustering results is conducted in the third module, including customer segmentation according to the clustering results, logistic regression analysis to identify influencing factors on heating consumption, load profiling for individual customers, studying load transition over time, and consumption variability. The analysis algorithms are implemented using the in-database machine learning library, Apache MADlib [29], and run as database procedures in the open source database management system, PostgreSQL [30].

3.1. Clustering

The K-means clustering algorithm [31] is used in this analysis for clustering district heating consumption data. The goal of the clustering is to identify and segment the customers with similar load intensity and consumption patterns in the first place and secondly to specifically look at the consumption patterns on normalised data. Given a set of load patterns (daily load patterns based on hourly readings in this paper), the clustering algorithm classifies the load patterns into K groups or clusters. The patterns are similar within the same cluster, but dissimilar to the other clusters, according to a distance metric. K-means clustering is an iterative process with the aim of minimizing the intra-cluster inertia criterion defined by:

$$C(P,\mu) = \sum_{i=1}^{n} \sum_{X_i \in P_k} ||X_i - \mu_k||^2 \tag{1}$$

where $P = (P_1,P_2,\ldots,P_k)$ is the set of clusters, $\mu = (\mu_1,\mu_2,\ldots,\mu_k)$ is the set of cluster centers, and $||\cdot||$ is the L_2 norm associated to the distance metric.

In this paper, the clustering approach is used for identifying the consumption intensity and representative patterns of the customers. This is done as follows:

– *Consumption intensity:* This clustering is applied on the representative daily load profiles of all the customers without normalization, $(\widehat{X}_1,\widehat{X}_2,\ldots,\widehat{X}_n)$. The representative load profile of a customer $i,\widehat{X}_i$, is estimated by the mean of the hourly consumption:

$$\widehat{X}_i = \frac{1}{N} \sum_{n \in N} X(n) \tag{2}$$

where N is the number of days in a time series, and $X(n)$ is the n-th day's daily load profile, $X(n) \in \mathbf{R}^{24}$.

– *Representative patterns:* This clustering is applied on all the normalized daily load profiles of all customers. For a given customer i and the n-th day, the normalized load profile is defined as

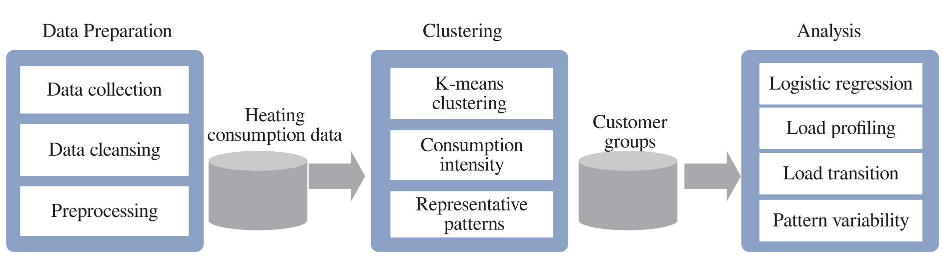


Fig. 1. Overview of the clustering-based analysis.

$$X_i^*(n) = \frac{X_i(n)}{S_i(n)} \tag{3}$$

where $X_i^*(n)$ and $X_i(n)$ denote the 24 h-period vectors of normalized and raw hourly consumption data, respectively. $S_i(n)$ is a scalar value representing the total consumption of the day n. The clustering on the normalized load profiles will only discover the patterns regardless of the consumption intensity.

3.1.1. Number of clusters K

The challenge of using the *K*-means algorithm is that it requires the user to provide the initial number of clusters K, and it is often difficult to determine the optimal number of clusters. There have been a number of approaches proposed, such as the random initialization for gaussian mixture models criterion in [32], using gap statistic to estimate the number of clusters [33], and an information measure for classification [34]. In this paper, the number of clusters for load intensity and consumption pattern are chosen according to the Bayesian Information Criterion (BIC) [35], which is based on the following considerations. In clustering, adding more clusters will decrease the variance (and increase the Bayesian likelihood). To avoid constantly adding centroids, the BIC works by penalizing the log-likelihood more when the complexity of the model (e.g. the number of parameters) increases [35]. Therefore, this approach can search for different values of k, score each clustering model according to the BIC value, and determine the optimal number of clusters. BIC is defined as

$$BIC(C|X) = \mathscr{L}(X|C) - \frac{p}{2}\log(n) \tag{4}$$

where $\mathscr{L}(X|C)$ is the log-likelihood of the data item of X belonging to a cluster C, and $p = k(d+1)$ is the number of parameters in the C with dimensionality of d and k cluster centroids.

3.1.2. Distance metrics

The clustering is based on the distance to decide which cluster an element belongs to. The Euclidean distance is one of the most commonly used metrics by clustering algorithms, which is defined as the following.

$$d(X,Y) = ||X-Y||_2 = \sqrt{\sum (X-Y)^2} \tag{5}$$

where X and Y are the two vectors with the same dimensionality. In the clustering of consumption intensity, they are the 24 h-vectors representing the hourly consumption values of the day. This paper uses an improved distance metric derived from the Euclidean distance to discover representative patterns. It is called *KSC-distance* [36], which is suggested by [37] to discover energy consumption patterns with a minor shift on the time scale. The KSC distance is defined as

$$d(X,Y) = \min_q ||X_q - Y||_2 \tag{6}$$

where X_q is a shifted version of X by the amount q that minimizes the instance. KSC-distance can remedy the “double-penalty” problem of the peaks that are only slightly different in timing when the peaks occur [37]. Fig. 2 shows the example where there are two vectors with the peak value of 1.0 kW h. If q is shifted to the left for one hour, i.e., −1, the distance d will be 0, otherwise $\sqrt{2}$. This means that the difference between the two patterns regarding a slight shifting along the time scale

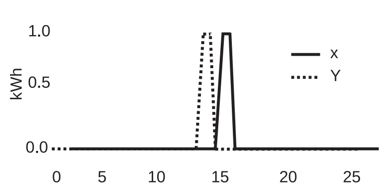


Fig. 2. Double-penalty problem.

can still be classified into the same cluster. This can solve minor pattern shifting problems. For example, a person getting up late with the morning peak at 8–9 am, instead of his/her usual morning peak at 7–8 am. The slight variance on the time scale should not lead to a wrong classification.

In this paper, the daily pattern is shifted between −1 and 1, i.e., the last hour of the previous day and the first hour of the next day. The distance is computed iteratively for the three shifts, and the minimal one is used for the clustering.

3.2. Correlation study based on the consumption intensity

The customer groups in terms of consumption intensity are identified by the clustering analysis without normalization. A correlation study is conducted to investigate the influencing factors to each of the consumption intensity groups based on the clustering results. The influencing factors include the characteristics of buildings and occupants. The purpose of this analysis is to determine the impact of building-related parameters on the heating consumption intensity, since these parameters are commonly used in national building databases to classify building stocks. Therefore, this study is to find out whether these parameters are representative to be used in classification schemes when the focus is on energy use intensity. Furthermore, occupant-related parameters are investigated to determine if they are suitable for characterizing heating consumption classification. The reason for selecting such parameters is that occupant behavior introduces large stochasticity to residential space heating demand [38,39]. This study is first done in an intuitive way simply using bar charts, followed by a logistic (binomial) regression analysis.

Logistic regression refers to a stochastic model in which the conditional mean of the dependent dichotomous variable (usually denoted $Y \in \{0,1\}$) is the logistic function of an affine function of the vector of independent variables (usually denoted $\boldsymbol{x}$). That is,

$$E[Y|\boldsymbol{x}] = \sigma(\boldsymbol{c}^T\boldsymbol{x}) \tag{7}$$

for some unknown vector of coefficients $\boldsymbol{c}$ and where $\sigma(x) = \frac{1}{1+\exp(-x)}$ is the logistic function. The logistic regression finds the vector of coefficients $\boldsymbol{c}$ that maximizes the likelihood of the observations. The relevance of studied characteristics is quantified by studying the significance of the coefficients.

3.3. Variability study based on the representative patterns

Based on the representative patterns discovered by the clustering with normalization, a variability study is conducted to investigate the consumption behavior change over time for households. In this study, each of the patterns is represented by an alphabet, and the daily consumption patterns of each customer will, therefore, be labelled as a sequence of the alphabets. Based on the alphabet sequence, the customer behavior changes over time will be quantified as the entropy in this study. Variability reflects the consumption habits of customers, and has a determining impact on operation cost for utilities. Utilities can rely on it to design energy efficiency programs for the relevant types of the users or neighborhoods that contribute to the variability. This study also provides the information for utilities to provision the supply of energy. The concept of entropy is hereby adopted to quantify variability. Entropy originally comes from the physics field and can be generally perceived as a measure of disorder. Entropy has been widely used in computer science and information theory mainly to characterize the uncertainty in the data [40,41]. In this case, a higher entropy value means that the user consumption is more variable, i.e., more unstable or less predictable. Otherwise, it is more predictable. There are typically two approaches for estimating the entropy: one is believing that the appearance of an alphabet α is uncorrelated, while the other is believing that the appearance is serially-correlated [37], i.e., the appearance of an alphabet for the next day depends on its previous days. For the first approach, the entropy can be calculated by

$$E^{uncor} = -\sum_{\alpha} p(\alpha)\log p(\alpha) \tag{8}$$

where $p(\alpha)$ is the posterior probability of an alphabet α. The probability can be estimated by

$$p(\alpha) = \frac{\#\{S(n) = \alpha\}}{N} \tag{9}$$

where the numerator is the count of an alphabet α in the sequence, and the denominator N is the length of the sequence, i.e., the number of days. The second approach estimates the entropy using the Lempel-Ziv compression algorithm [42]. The entropy is estimated as:

$$E^{cor} = \left(\frac{1}{N}\sum_{n} L_n\right)^{-1}\log(N) \tag{10}$$

where L_i is the length of the shortest sub-sequence at position n which doesn't previously appear from position 1 to n−1.

4. Analysis and results

This section will first describe the data, followed by a descriptive analysis of the data. Then, this section will do the clustering-based analysis to study consumption intensity, load transition, and variability.

4.1. Data

The district heating consumption data are from 8293 single-family households in Aarhus City, Denmark, with hourly time resolution. Single-family houses represent 44% of the total residential households in Denmark [43]. The readings were collected from different starting dates, with the earliest being March 02, 2009, but all end on November 29, 2015. These data are private and owned by the district heating supplier of the area, Aarhus Affaldvarme (AVA). AVA provided the authors with the data for research purposes after applying some anonymization techniques like removing personally identifiable information such as names, addresses and social security numbers to ensure the sources' privacy. The additional data include building information from the Danish National Building Register (BBR) such as the construction years, sites and areas. These data are publicly available at [44]. BBR is a nationwide register which includes the data of the majority of Danish buildings and households. It contains the information about 1.6 million properties, 3.8 million buildings and 2.7 million dwellings and commercial units [45]. It was originally established in 1977 by collecting the information from building owners via questionnaires. Since then, it has been updated by local authorities and by citizens [46]. Specific customer information was also provided to the authors, including family sizes, date of birth of residents, addresses and the dates of moving in and moving out.

4.1.1. Descriptive analysis of data

Initially, a descriptive analysis is conducted to visualize the time series of district heating consumption data. Fig. 3 shows three typical consumption time series (high, medium and low) in the year 2014. The figure indicates that most of the heating consumption occurs in the winter period between October and April of the following year, while there is very little heating consumption during the summer period May to September. This is in line with the typical assumption for the Danish heating season, starting on October 1st and ending on April 30th [47]. Fig. 4 represents the average hourly consumption of all households for weekdays and weekend days during January. The results show that the peak loads occur in the morning and evening on weekdays, while the morning peak is occurring later – approximately 3 h – in the weekends. Based on these results, it can be observed that there are different

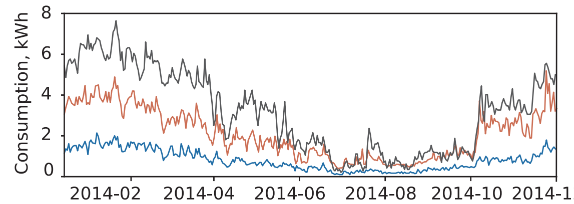


Fig. 3. Three typical examples of heating consumption time series.

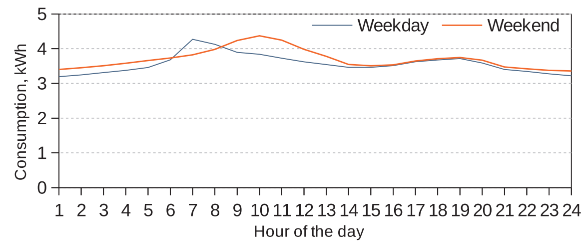


Fig. 4. Average daily heating consumption patterns.

consumption intensity levels. District heating data follow certain patterns and seasonality, which may be due to the weather conditions, building performance, living habits of customers and/or something else. In the following, this paper will conduct an exploratory analysis based on the clustering to study consumption intensity, load transition and variability.

4.2. Clustering for studying consumption intensity

In the following, an exploratory analysis will be conducted using the proposed clustering methods. Initially, the clustering is based on the daily load profiles without normalization. A daily load profile of day n is denoted as $X(n) = \langle x_0,...,x_h,...,x_{23}\rangle$, where x_h represents the consumption of the hour, h.

According to the proposed methods, the number of clusters is defined by using the BIC approach. Fig. 5 shows the evolution of BIC values with the increasing number of clusters. The BIC value decreases up to 5, $k = 5$ and increases progressively thereafter. Therefore the appropriate number of clusters is 5.

The clustering, therefore, segments the customers into five consumption groups. The silhouette score [1] [48] of the clustering is 0.641, indicating that the clustering achieves a satisfactory result in terms of intra-cluster cohesion and inter-cluster separation. The clusters are ordered by their daily average heating consumption, and labelled using the alphabet from A to E, which also represents the consumption levels ranging from the lowest to the highest. The purpose of this clustering is to identify the customers with a similar *consumption pattern and consumption intensity*. This information is valuable to utilities to target the customer for offering personalized services, e.g., giving energy-saving recommendations or heat emission system replacement. Fig. 6 shows the clustering results. In each cluster, the bold red line is the cluster centroid which is the representative group load profile, while the other lines are the mean daily load profiles of the customers in the cluster. Cluster A represents 24.2% of the households having the lowest heating consumption among all households. The cluster load profile is rather flat, having a slight peak between 5am and 8am. Cluster B represents the majority of households, which is 36.2% of the total examined stock. The heating consumption is a bit higher, while the cluster load profile is rather constant with a small variation. The daily heating consumption of the centre of cluster B does not exceed 2.5 kWh. Cluster C exhibits a similar pattern as B, with the heating consumption being a bit higher.

[1] Silhouette score is a measure for evaluating clustering results according to cohesion within a cluster, and the separation to other clusters. The score ranges from −1 to + 1, where a high value indicates a good matching [48].

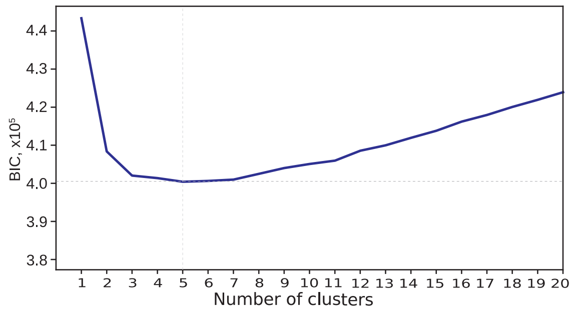


Fig. 5. Determining the optimum number of clusters.

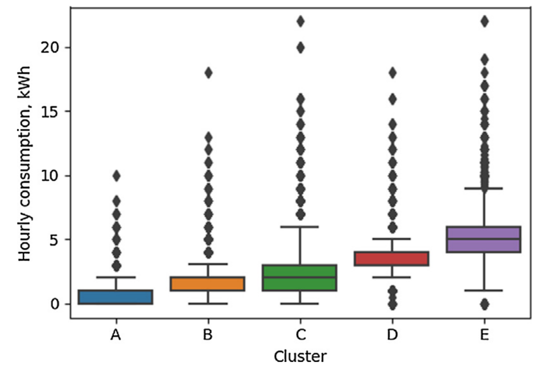


Fig. 7. Boxplot of hourly consumption distribution in each cluster.

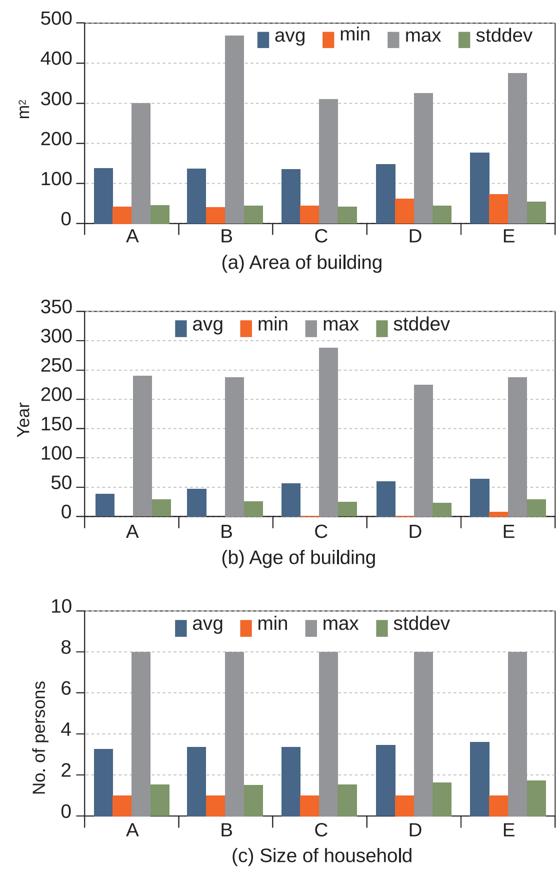


Fig. 8. Statistical indicator of each group.

The morning peak occurs between 6 am and 8 am, while the evening peak is very weak. Cluster *D* represents 11.2% of the households, with the daily heating consumption ranging from 2.7 to 4 kW h. Cluster *E* indicates the highest heating consumption among all households, representing only 2.1% of all dwellings. The morning and evening peaks are more pronounced, with the former occurring between 6am and 8am and the latter between 5 pm and 8 pm.

The hourly readings of the households are now explored in each cluster using the boxplot method. The boxplot method can unveil the descriptive statistics regarding dispersion and skewness, as well as outliers of the data [49]. Fig. 7 illustrates the results (the line in the box represents the median). In each cluster, there are some outliers that lay beyond the quartiles by 1.5 interquartile range. The distributions of the clusters *A*, *B* and *D* are more skew for their median lies at the quartile boxplot. The boxplots depict the distribution of the data in each cluster (the length of the box represents the variance, and the standard deviation range from 8.1 to 24.9 kW h). Cluster *A* has the lowest variability and consumption, while Cluster *E* has the highest variability and the highest consumption.

It is now further investigated if there exists any relevance between the energy consumption levels and building and occupant characteristics in an intuitive way. The analysis uses bar charts to determine the effect of floor area, the age of building and number of occupants on the heating load profile (see Fig. 8). Visually, the results have shown that buildings with large floor areas and old buildings have high heat demand on average. Looking at Fig. 8a, the average floor area is higher for cluster *E*, which represents the high heat demand profiles, as expected, since the heating consumption data are not normalised per floor area. Furthermore, it can be observed from Fig. 8b that the older a building is, the higher heat demand it has. Cluster *A*, which represents the lowest heat demand, includes buildings with the lowest average age of buildings. The effect of the number of occupants is not as clear among the different clusters (Fig. 8c).

The relevance is now quantified by the coefficient significance using logit regression analysis. The following features are selected: building age, floor area, the number of adults, the number of teenagers, and the number of children in a household. In this case, the independent variables are the house and occupant features. In logit regression, the dependent variable, y, only takes two possible values, 0 or 1. In this analysis, y corresponds to 1 or 0 if a sample is belonging to a cluster or not. For example, for the logit regression of cluster *A*, if a customer belongs to this cluster, then $y = 1$, otherwise, $y = 0$. The coefficients of the regression, along with the odds ratio for the coefficient, the Wald p-value and z are calculated and presented in the following. Table 1

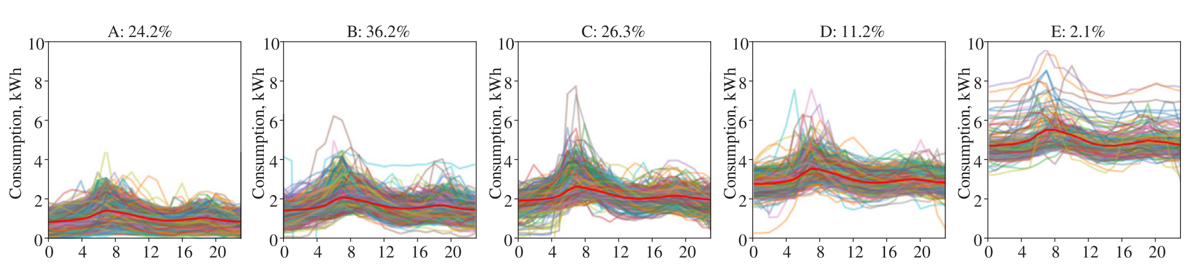


Fig. 6. Customer segmentation based on daily consumption patterns.

Table 1
Logit regression results for load profile segmentation.

Explanatory variables	A					B				
	Coef.	Odds	Std-err	z-value	p-value	Coef.	Odds	Std-err	z-value	p-value
Intercept	2.476	11.900	0.6931	3.573	0.00035**	0.665	1.945	0.536	1.240	0.214
Building area	−0.010	0.989	0.00272	−3.899	0.00009**	−0.004	0.996	0.0022	−1.727	0.084'
Building age	−0.038	0.962	0.0052	−7.318	0.000**	−0.013	0.987	0.0037	−3.369	0.0007**
No. of adults	0.055	1.057	0.0786	0.707	0.479	−0.049	0.952	0.0648	−0.757	0.448
No. of teenagers	−0.549	0.577	0.2191	−2.509	0.012*	0.063	1.065	0.1424	0.443	0.657
No. of children	−0.094	0.909	0.1282	−0.736	0.461	−0.048	0.952	0.1063	−0.454	0.649
	Df = 5 Log-likelihood = −90.4 Pseudo R^2 = 0.0526					Df = 5 Log-likelihood = −140.7 Pseudo R^2 = 0.075				

C					D					E				
Coef.	Odds	Std-err	z-value	p-value	Coef.	Odds	Std-err	z-value	p-value	Coef.	Odds	Std-err	z-value	p-value
−2.276	0.103	0.579	−3.930	8.45e−05**	−5.506	0.004	0.8423	−6.537	6.256e−11**	−12.919	2.449e−6	2.125	−6.078	1.211e−09**
0.002	1.002	0.00237	0.771	0.440	0.011	1.011	0.00327	3.414	0.0006**	0.036	1.037	0.007	4.724	2.307e−06**
0.016	1.016	0.00384	4.130	3.614e−05**	0.031	1.031	0.00530	5.898	3.664e−09**	0.047	1.048	0.0103	4.567	4.946e−06**
−0.066	0.936	0.0693	−0.959	0.337	0.121	1.129	0.0907	1.339	0.180	0.106	1.111	0.219	0.483	0.628
0.235	1.265	0.1465	1.605	0.108	0.000	1.000	0.2057	0.002	0.997	0.372	1.451	0.454	0.819	0.412
0.145	1.156	0.114	1.303	0.192	−0.058	0.943	0.1623	−0.361	0.717	−0.372	0.689	0.476	−0.782	0.433
Df = 5 Log-likelihood = −140.8 Pseudo R^2 = 0.102					Df = 5 Log-likelihood = −101.8 Pseudo R^2 = 0.214					Df = 5 Log-likelihood = −19.3 Pseudo R^2 = 0.384				

' Significance codes of p-value: 0.1.
* Significance codes of p-value: 0.05.
** Significance codes of p-value: 0.001.

presents the regression analysis results of the attributes of time-patterns for the five identified clusters. According to the Pseudo R^2, the models of the high energy consumption groups *D* and *E* have a better fit than the lower energy consumption groups (0.2–0.4 represents the good fit, which is equivalent to 0.7–0.9 in traditional R^2 [50]). According to the p-value, the explanatory variables, floor or building area and age of the building, are the most significant ones in all segments. The number of teenagers in the household appears to be positively associated with the low and relatively constant heat load represented by Cluster *A*. This may be attributed to teenagers spending less time at home compared with other age groups. In the high load segment, the number of teenagers changes 1.451 times the odds of a household to belong to that cluster. When looking at the medium load segment, the number of children changes 1.156 times the odds of a dwelling to belong to that cluster. The effect of the number of adults is not easy to interpret. Adult occupant behavior varies a lot and does not follow a steady pattern due to different thermal comfort preferences.

Fig. 9 shows the daily load profile of all customers after quantification of the clustering differentiated between weekdays (a) and weekends/holiday (b). All five segments show a similar trend for weekdays, but not for weekends. In particular, the magnitude of heating consumption differentiates the five clusters on weekdays. All of the five clusters show a remarkable morning peak. Furthermore, the segments of high and medium load profiles are characterized by a morning peak a bit earlier in the day (around 7 am), while in the two lower load profile segments the morning peak occurs approximately one hour later. In addition, cluster *E* which represents high load profiles has two soft evening peaks at 6 pm and 9 pm. Looking at the load profiling for weekends, the trends are less clear. Occupants do not follow a steady schedule during weekends and holiday. Cluster *D* which is the segment with the second highest heating consumption, has a striking morning peak at 9am. The behavior of the rest of the clusters is quite similar to each other and they show softer peaks during the day.

Fig. 10 shows how a typical customer's daily heating consumption profile changes cluster throughout the year (2014). In the customer segmentation process in Section 3.1 this customer is in cluster *E*. It means that this customer is classified into the highest heating consumption group. However, interestingly, the customer changes consumption pattern from day to day throughout the year, which means that this customer load profile fits better with another cluster the next day. The figure shows which cluster the customer is associated with at each day within the year. Although this customer on average is a high load profile consumer, the customer is fully associated with the low load profile cluster *A* during the warmer summer months, from May to September. During autumn in October and spring in April and half of May, the customer is associated with the medium heating consumption cluster *C*. This is in the two transition months between heating season and non-heating season where there is a medium heat demand. The customer is actually only associated with the high heating consumption cluster, *E*, during January and December, when ambient temperatures are low, probably sub-zero degrees Celsius. The customer is mostly associated with cluster *B* and *D* in February through April, as well as in November and December. The share of the time this customer is associated with the second lowest heat load cluster *B* during winter months is quite high and that indicates a well-insulated house that can possibly benefit significantly from solar gains. The relatively low heat loads can also be associated with specific calendar events, such as holidays, when occupants are not home. It should be noted that this house was built in 1929 and its floor area is 150 m^2. Therefore, it is an old house, which explains the high heating consumption during the winter months.

4.3. Clustering for studying load transition

This subsection investigates seasonal load transitions. Fig. 11 presents the evolution of consumer behavior from one cluster to another over the twelve months in the year of 2014. First, this figure shows the

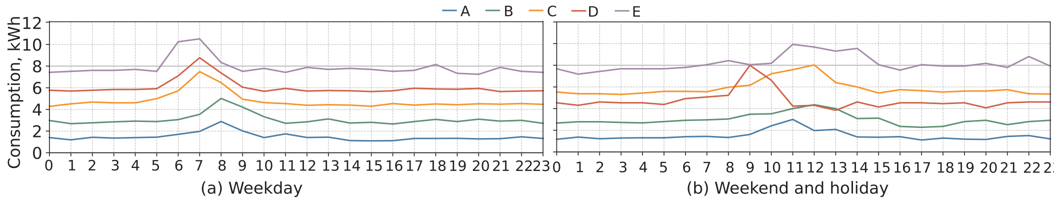


Fig. 9. Cluster daily load profiles into five groups for weekday and weekend/holiday.

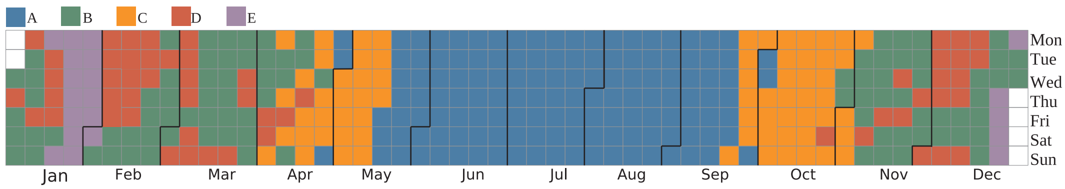


Fig. 10. Load profiling of a customer over the year of 2014.

clusters distribution of all households for each month. This can be explained by the behavior changes that are dependent on the ambient temperature and on calendar events. In particular, cluster *D*, which corresponds to the second highest heat load segment, represents the majority of the households during the winter months. It should be noted for comparison that cluster *D* represented 11.2% of the customers and cluster *E* 2.1% of the customers. The majority of the customers in January moved to clusters with lower heating consumption in February, namely cluster *B* and *C*. The proportions of the clusters, *D* and *E*, further decrease while moving towards the spring season. Cluster *C* is the dominating cluster in March, which is a transition month from the cold winter months to the non-heating season. During the summer season, the majority of the customers are represented by cluster *A*, which has the lowest heat load and a fairly constant daily pattern. Similar conclusions for the load transitions can be drawn looking at the autumn months.

Furthermore, the transition probability from one cluster to the others has been calculated to quantify the changes in heating consumption and consumer behavior (see Table 2). The probability in this table is the statistic data of the transition between daily patterns for all the customers in the year of 2014. The transition probability between months can also be calculated in a similar way. The table shows that the probability of remaining in the same state (cluster) is higher than transiting to another state. The transitions to the adjacent states are also detected, which depends on many factors, such as the changes of ambient temperature.

Table 2
Transition probability between clusters.

Clusters	*A*	*B*	*C*	*D*	*E*
A	0.9246	0.0624	0.0109	0.0019	0.0003
B	0.0990	0.7839	0.0327	0.0833	0.0012
C	0.1521	0.3037	0.3568	0.1716	0.0158
D	0.0046	0.1204	0.0294	0.7976	0.0479
E	0.0026	0.0065	0.0075	0.1816	0.8018

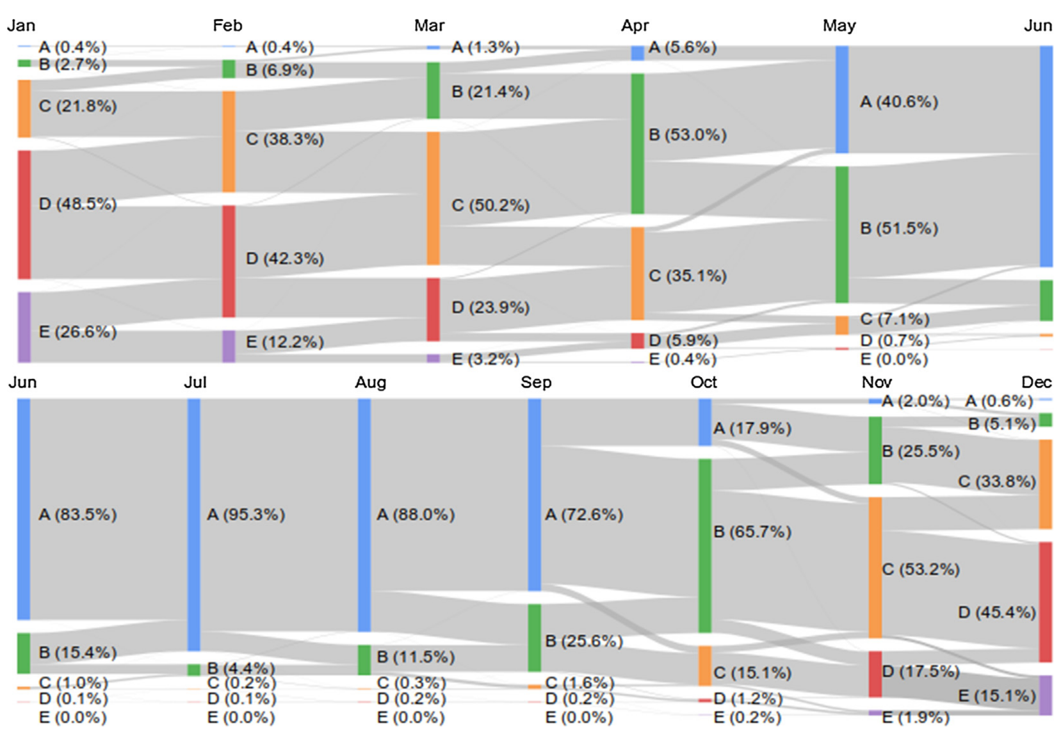


Fig. 11. Load transition over the months in the year of 2014.

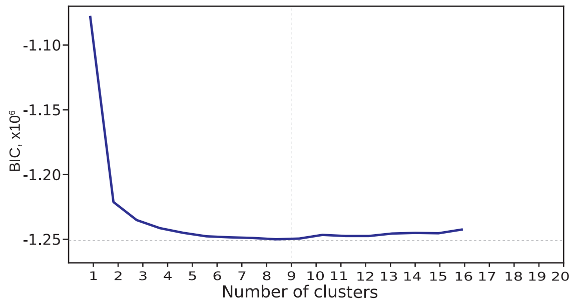


Fig. 12. Determining the optimum number of clusters for normalized consumption patterns.

4.4. Clustering for studying consumption pattern variability

To study the consumption variability regarding the daily patterns, this paper first applies clustering on the normalized patterns of all days of all the customers, then quantifies the variability of each customer based on the clustering results. The normalization is conducted according to the Eq. (3). In this clustering, the chosen number of clusters is nine with regards to the BIC value (see Fig. 12). The silhouette score of the resulting clusters is 0.396, which is a bit lower than the clustering on un-normalized data. The reason is that the KSC distance metric is used in the clustering, and the patterns within the same cluster are more cluttered. Fig. 13 shows the clusters ranked according to the percentage of the patterns in each cluster from the highest to the lowest. The clusters are labeled with the alphabets *A–I* (note that this labeling is different to the previous section which represents consumption intensity levels). As shown, cluster *A* represents the largest share of daily load patterns, 81.1%, where the patterns with morning and evening peaks are almost evenly distributed (as seen from the crowded lines of patterns). Since the KSC distance is used in the clustering, the minor shift of patterns are classified into the same cluster, and the centroid is flattened (represented by the bold red line). Cluster *B* has a pronounced morning peak around 9am (merged due to the use of KSC distance), and an evening lower peak around 7 pm. Cluster *D* also exhibits two peaks in the afternoon and evening respectively, while the remainders have one peak, which differentiates in appearance time during the day.

Based on the identified patterns, variability is quantified by the entropy metrics. The variability can be used to provision energy supply, optimize the design and operation of the future energy system by utilities. Depending on whether the patterns are serially correlated or not, the entropy is computed by the Eqs. (8) and (10), respectively. The distributions of the entropy are displayed in Fig. 14. The figure indicates that consumption at the individual level seems to be serially-

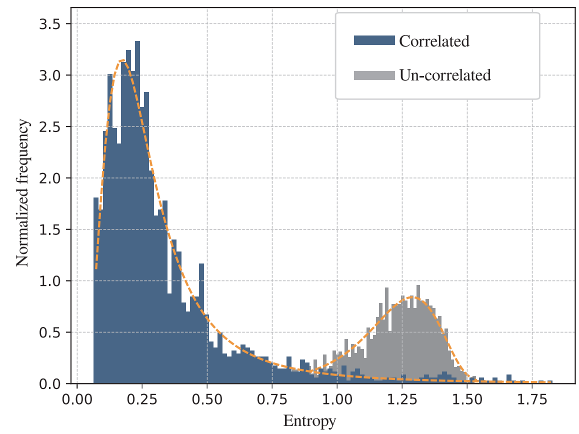


Fig. 14. Distribution of entropy.

correlated since in general $E^{cor} < E^{uncor}$, i.e., the consumption at the present depending on the past. The value of E^{cor} is much lower, which shows that the appearances of daily patterns are serially correlated. This can also be verified by the transition Table 2 which shows that the probability between two identical states is much higher. The entropy for the large portion of consumers is 0.25 for the correlated, and 1.25 for the uncorrelated.

Fig. 15 compares the average entropy for the customer groups identified by clustering approach in Section 3.1. The result shows that the lower consumption intensity groups have higher variability in terms of the entropy values, while the higher consumption intensity groups have lower variability. These findings can help utilities to identify the consumption of which customer or customer group are more predictable (or stable), and to identify the customers, for example, to participate in demand-response programs.

5. Discussion

One of the main objectives of this study has been to determine and map heating consumption patterns for district heating customers in Denmark. The analysis shows that heat load profiles of Danish residential customers living in single-family houses in Aarhus can be represented by five clusters with regards to load intensity. In this case, the daily consumption patterns are fairly constant with two weak peaks, one in the morning and one in the evening. Therefore, the space heating profile shows a pattern similar to the electricity load profile that has been identified in most of the electricity consumption pattern segmentation studies, i.e., with two distinct peaks. The morning peak is

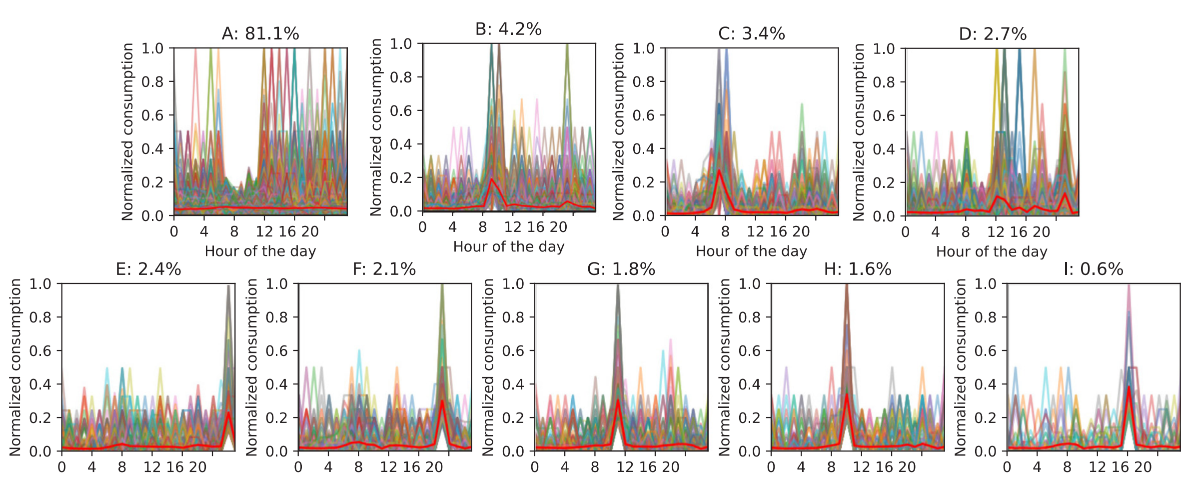


Fig. 13. Clusters of normalized daily consumption patterns.

Energy Conversion and Management 165 (2018) 840–850

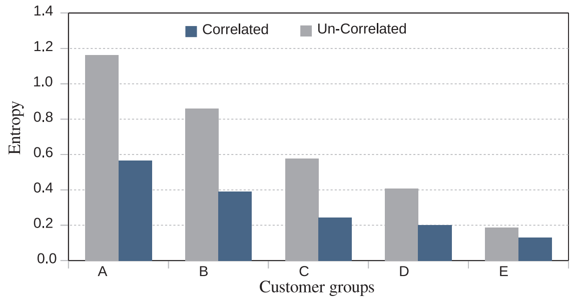


Fig. 15. The average entropy of each consumption intensity group.

more pronounced than the evening peak in the heating consumption profile, which is often the reverse for electricity. When the focus is placed on identifying specific consumption patterns, the customers are represented by nine clusters. The cluster that represents the vast majority of customers shows a fairly constant profile, while the remaining clusters are characterized by one or two peaks during the day. This finding is important as it indicates that the investigated district heating customers do not change their consumption during the day. Such information is valuable for district heating utilities to further optimize their network and apply advanced solutions (i.e. virtual storage on district heating network). Furthermore, the identification of the morning and/or evening peaks is important for implementing demand-side management strategies and for the possible utilization of thermal storage solutions considering the needs of the grid. During peak load periods, the grid is stressed and heating cut-off combined with activation of thermal energy storage mechanisms are solutions that could be examined in conjunction with the load profile and consumption patterns to balance the supply and demand sides.

The big data set that was used in the study increases the robustness of the method and the probability of finding statistically significant results. The sample of customers studied is very homogeneous, as the data set only includes single-family houses. Even though the results cannot be generalized for the rest of the country, it is expected that residential district heating customers in Danish urban areas show similar patterns, as the cultural habits are quite uniform in the bigger Danish cities (e.g., work schedule, desired thermal comfort conditions). If the study was to be expanded towards different building typologies, such as commercial buildings, the number of clusters to represent all customers would probably increase to accommodate more diverse heating patterns. Furthermore, the correlation of heat load patterns with building and occupant characteristics is an important step to determine if these factors can successfully reflect the heat load profiles, while partly explaining the dynamics of the space heating consumption. From a building physics perspective, more variables should be investigated, such as the time constant of the buildings, insulation level of the building envelope, and infiltration rate. In addition, it would be relevant to integrate a specific variable for holidays or calendar events in the model. The investigated building-related parameters (age of the building and floor area) were selected, because they have been most commonly used in national building classification schemes apart from the building types [51]. Therefore, the aim was to find out if they can accurately reflect residential heating load profiles. Results showed that they are strongly related to the load intensity of the clusters, hence they are suitable to be used for classification of Danish single-family houses (in urban areas) when the focus is placed on the energy use intensity. The family size and the age of occupants were also investigated to determine if they could substitute some of the already existing classification parameters to represent differences in occupant behavior. However, no strong correlation between the age of occupants with the consumption intensity could be drawn in the study. To draw conclusions on the effect of occupants' ages on the consumption patterns, the logit analysis can be repeated on the normalized data.

A different categorization of occupants with regards to age, such as children, adults and elderly or retired people could also be investigated, since pensioners are likely to have a different schedule than the rest of the occupants. The amount of investigated variables could also be expanded towards socio-economic characteristics, such as income per household and education level, in order to identify possible targeted groups for energy efficiency programs. Data on these variables are, however, more difficult to obtain because they are not available at urban-scale building databases.

The proposed techniques can easily be implemented on large data sets. The selected algorithm, *K*-means has been successfully used on large data sets because of its simplicity and linear time complexity [17]. This makes it computationally attractive. This study employs in-database analysis which has a better performance than using the traditional analysis environments such as *R* or *Matlab* as the data does not have to be read out of the database. The quantitative benchmarking studies can be found at [10]. The proposed approach can be a generic solution for other energy consumption data analysis, such as electricity, water or gas. The results can help utilities for better production-side management, such as developing new pricing policies, dimensioning new parts of the network and targeting specific customers to implement demand-side management solutions. Furthermore, the results can help customers to better understand their own consumption patterns and consumption behaviors to improve their own energy efficiency. However, smart meter data are difficult to acquire at a large scale, mainly due to privacy issues. The data anonymization applied before the analysis can greatly facilitate the process of data acquisition and publishing.

The results of the current study can also be used to characterize building performance behavior at urban scale. The proposed approach and the resulting load profiles can support a scaled analysis of buildings in large urban-scale groups. The findings are also useful for whole-building and district-scale simulations to calibrate the heating consumption profiles, optimize the design of large-scale heating systems and detect anomalous behaviors.

6. Conclusions and future work

Smart meters have increasingly been used for monitoring heating consumption. This paper has proposed a clustering-based knowledge discovery approach for understanding residential heating consumption data, including data preparation, clustering, and data analysis. The paper applied the *K*-means algorithm to cluster the daily load profiles from 8293 Danish single-family households in Aarhus. The results revealed that Danish district heating customers can be segmented into five clusters (according to the optimal number of clusters) with regards to their consumption intensity. The clusters were characterized by fairly constant load profiles with two weak peaks in the early morning and in the evening, respectively. The clusters were labeled with the alphabets, A–E, to represent heating consumption levels ranging from low to high. A discrimination between weekday and weekend/holiday profiles showed that weekend load profiles followed a similar pattern, but the morning peak was shifted a few hours later. To identify specific consumption patterns, a new clustering analysis was conducted on the normalized data which identified nine groups of customers, with the most dominant one showing a fairly constant profile, too. Thus, the vast majority of the examined district heating customers only made minor changes to their consumption during the day. The remaining clusters had one or two peaks in the morning or evening time, respectively. The paper also studied the correlation between heating consumption and characteristics of buildings and occupants, and used logit regression to quantify their relationship. The results indicated that building age, area and family size had a pronounced impact on the consumption, whereas the age of occupants was less pronounced. Therefore, it is appropriate to use the first two factors to categorize Danish housing stock, and particularly single-family households, in building classification schemes with regards to energy consumption intensity. In addition, the paper studied the load profile characteristics for a single customer using the clustering approach, and, based on the clustering results, illustrated the consumption transition probability over time and quantified the consumption variability using entropy methods. The results showed that the consumption patterns for each individual customer were serially correlated, and the higher consumption groups had lower variability in terms of the patterns.

For future work, clustering-based short-term energy demand forecasting will be investigated. In this study, it has been found that the majority of the customers have regular and predictable consumption behaviors. According to the transition probabilities, there is a high chance for a household to repeat the same consumption pattern the following few days. This indicates that clustering-based load forecasting is feasible. In addition, building samples could be expanded to study the consumption patterns by including more diverse building typologies. Finally, it is also interesting to use clustering-based approach to further reveal heating consumption characteristics of buildings and occupants.

Acknowledgements

This research was supported by the Danish research project CITIES (Centre for IT-Intelligent Energy Systems in cities) no DSF1305-00027B funded by Danish Strategic Research Council. The authors would like to thank Aarhus Affaldvarme, who enabled the distribution of the smart meter data to them.

References

[1] Liu X, Golab L, Golab W, Ilyas IF. Benchmarking smart meter data analytics. In: Proc of the 18th international conference on extending database technology; 2015. p. 385–96.
[2] The European Parliament and the Council of European Union. Directive 2009/72/EC concerning common rules for the internal market in electricity and repealing Directive 2003/54/EC. Off J Eur Union 2009.
[3] The European Parliament and the Council of European Union. Directive 2009/73/EC concerning common rules for the internal market in natural gas and repealing Directive 2003/55/EC. Off J Eur Union 2009.
[4] European Commission. Benchmarking smart metering deployment in the EU-27 with a focus on electricity. Report from the Commission, Brussels; 2014.
[5] European Commission. Countries reports – Denmark; 2014. Available at: < https://ec.europa.eu/energy/sites/ener/files/documents/2014_countryreports_denmark.pdf > [as of 2018-01-15].
[6] Darby S. Smart metering: what potential for householder engagement? Build Res Inform 2010;38:442–57.
[7] Kipping A, Trømborg E. Modeling and disaggregating hourly electricity consumption in Norwegian dwellings based on smart meter data. Energy Build 2016;118:350–69.
[8] Yu Z. Mining hidden knowledge from measured data for improving building energy performance PhD Thesis Concordia University; 2012.
[9] Beckel C, Sadamori L, Staake T, Santin S. Revealing household characteristics from smart meter data. Energy 2014;78:397–410.
[10] Liu X, Golab L, Golab W, Ilyas IF, Jin S. Smart meter data analytics: systems, algorithms and benchmarking. ACM Trans Database Syst (TODS) 2016;42(1).
[11] Liu L, Nielsen PS. An ICT-solution for smart meter data analytics. Energy 2016;115(3):1710–22.
[12] International Energy Agency. Technology Roadmap. Energy-efficient buildings: heating and cooling equipment; 2011. Available at < https://www.iea.org/publications/freepublications/publication/buildings_roadmap.pdf > [as of 2018-01-15].
[13] International Energy Agency. Heating without global warming; 2014. Available at < https://www.iea.org/publications/freepublications/publication/FeaturedInsight_HeatingWithoutGlobalWarming_FINAL.pdf > [as of 2018-01-15].
[14] Guha S, Rastogi R, Shim K. CURE: an efficient clustering algorithm for large databases. ACM Sigmod Rec 1998;27(2):73–84.
[15] Theodoridis S, Koutroubas K. Pattern recognition. Academic Press; 1999.
[16] Fayyad MU, Piatesky-Shapiro G, Smuth P, Uthurusamy R. Advances in knowledge discovery and data mining vol. 21. AAAI Press; 1996.
[17] Jain AK, Murty MN, Flynn PJ. Data clustering: a review. ACM Comput Surv 1999;31(3):264–323.
[18] Beckel C, Sadamori L, Santini S. Towards automatic classification of private households using electricity consumption data. In: Proceedings of BuildSys '12 of the fourth ACM workshop on embedded sensing systems for energy-efficiency in buildings; Toronto, Canada. p. 169–76.
[19] McLoughlin F, Duffy A, Conlon M. A clustering approach to domestic electricity load profile characterisation using smart metering data. Appl Energy 2015;141:190–9.
[20] Chicco G. Overview and performance assessment of the clustering methods for electrical load pattern grouping. Energy 2012;42:68–80.
[21] Andersen FM, Larsen HV, Boomsma TK. Long-term forecasting of hourly electricity load: Identification of consumption profiles and segmentation of customers. Energy Convers Manage 2013;68:244–52.
[22] Hsu D. Comparison of integrated clustering methods for accurate and stable prediction of building energy consumption data. Appl Energy 2015;160:153–63.
[23] Granell R, Axon CJ, Wallom DCH. Clustering disaggregated load profiles using a Dirichlet process mixture model. Energy Convers Manage 2015;92:507–16.
[24] Madeira C, do Carmo R, Christensen CH. Cluster analysis of residential heat load profiles and the role of technical and household characteristics. Energy Build 2016;125:171–80.
[25] Ma Z, Yan R, Nord N. A variation focused cluster analysis strategy to identify typical daily heating load profiles of higher education buildings. Energy 2017;134:90–102.
[26] Mendaza IDC, Pigazo A, Bak-Jense B, Chen Z. Generation of domestic hot water, space heating and driving pattern profiles for integration analysis of active loads in low voltage grids. In: Proc of innovative smart grid technologies Europe (ISGT Europe); 2013.
[27] Teeraratkul T, O'neill D, Lall S. Shape-based approach to household electric load curve clustering and prediction. IEEE Trans Smart Grid 2017;PP(99):1–13.
[28] Paparrizos J, Gravano L. k-Shape: efficient and accurate clustering of time series. ACM SIGMOD Rec 2016;45:69–76.
[29] Apache MADlib: big data machine learning in SQL. Available at < http://madlib.apache.org > [as of 2018-01-15].
[30] PostgreSQL. < https://www.postgresql.org > [as of 2018-01-15].
[31] Azimi R, Ghayekhloo M, Ghofrani M. A hybrid method based on a new clustering technique and multilayer perceptron neural networks for hourly solar radiation forecasting. Energy Convers Manage 2016;118:331–44.
[32] Figueiredo M, Jain AK. Unsupervised learning of finite mixture models. IEEE Trans Pattern Anal Mach Intell 2002;24(3):381–96.
[33] Tibshirani R, Walther G, Hastie T. Estimating the number of clusters in a data set via the gap statistic. J Roy Stat Soc B 2001:411–23.
[34] Wallace CS, Boulton DM. An information measure for classification. Comput J 1968;11:185–95.
[35] Robert EK, Larry W. A reference Bayesian test for nested hypotheses and its relationship to the Schwarz criterion. J Am Stat Assoc 1995;90(431):928–34.
[36] Yang J, Leskovec J. Patterns of temporal variation in online media. In: Proc of the 4th ACM international conference on Web search and data mining; 2011. p. 177–86.
[37] Albert A, Gebru T, Ku J, Kwac J, Leskovec J, Rajagopal R. Drivers of variability in energy consumption. In: ECML-PKDD DARE workshop on energy analytics; 2013.
[38] Sun K, Hong T. A framework for quantifying the impact of occupant behavior on energy savings of energy conservation measures. Energy Build 2017;146:383–96.
[39] O'Brien W, Gaetani I, Carlucci S, Hoes PJ, Hensen JLM. On occupant-centric building performance metrics. Build Environ 2017;122:373–85.
[40] Zhu C, Wang Z. Entropy-based matrix learning machine for imbalanced data sets. Pattern Recogn Lett 2017;88:72–80.
[41] Kiluk S. Diagnostic information system dynamics in the evaluation of machine learning algorithms for the supervision of energy efficiency of district heating-supplied buildings. Energy Convers Manage 2017;150:904–13.
[42] Schurmann T, Grassberger P. Entropy estimation of symbol sequences. Chaos 1996;6(3):414–27.
[43] Danmarks Statistik. Statistics Denmark. Available at < https://www.statistikbanken.dk/statbank5a/default.asp?w=2021 > [as of 2018-01-15].
[44] Bygnings- og Boligregistret (BBR). Danish Building Register. Available at < https://ois.dk > [as of 2018-01-15].
[45] Wittchen KB, Kragh J. Danish building typologies. Participation in the TABULA project, Hrsholm, Denmark: SBi, Danish Building Research Institute, Aalborg University; 2012.
[46] The Danish Ministry of Housing; 2006. Available at < http://boligejer.dk/ > [as of 2018-01-15].
[47] Stampe OB. Varme- og Klimateknik, Grundbog. In: Hansen HE, Kjerulf-Jensen P, editors. Danvak; 1992.
[48] Rousseeuw PJ. Silhouettes: a graphical aid to the interpretation and validation of cluster analysis. J Comput Appl Math 1987;20:53–65.
[49] Frigge M, Hoaglin DC, Iglewicz B. Some implementations of the boxplot. Am Stat 1989;43(1):50–4.
[50] Louviere JJ, Hensher DA, Swait JD. Stated choice methods: analysis and applications. Cambridge University Press; 2000.
[51] Wittchen KB, Mortensen L, Hols SB, Bjrk NF, Vares S, Malmqvist T. Building typologies in the Nordic countries, identification of potential energy saving measures. Danish Building Research Institute, Aalborg University; 2012 [SBi 2012:04].

Paper II: Gianniou, P., Reinhart, C., Hsu, D., Heller, A., Rode, C. (2018). Estimation of Temperature Setpoints and Heat Transfer Coefficients Among Residential Buildings in Denmark Based on Smart Meter Data. Submitted to *Building and Environment*, January 2018. Accepted.

Estimation of Temperature Setpoints and Heat Transfer Coefficients Among Residential Buildings in Denmark Based on Smart Meter Data

Panagiota Gianniou[a, *], Christoph Reinhart[b], David Hsu[c], Alfred Heller[a], Carsten Rode[a]

[a] Department of Civil Engineering, Technical University of Denmark, Kgs. Lyngby 2800, Denmark

[b] Department of Architecture, Massachusetts Institute of Technology, Cambridge MA 20139, USA

[c] Department of Urban Studies and Planning, Massachusetts Institute of Technology, Cambridge MA 20139, USA

* Corresponding author, e-mail: pgianniou@gmail.com

Abstract

Thermal comfort preferences of occupants and their interactions with building systems are top influential factors of residential space heating demand. Consequently, housing stock models are sensitive to assumptions made on heating temperatures. This study proposes a heat balance approach, inspired by the classical degree-day method, applied to an extensive urban dataset. The goal of this analysis is to determine heterogeneous characteristics, such as temperature setpoints of heating systems and thermal envelope characteristics from an overall population of residential buildings. Measured energy data are utilized for the purpose of the study from the city of Aarhus, Denmark, where the energy usage for heating of circa 14,000 households was monitored over time via smart meters. These data are combined with actual weather data as well as data extracted by a national building database. Using linear regression and heat balance models, temperature setpoints for the whole dataset are determined with a median and average of 19°C and 19.1°C, respectively. Furthermore, building related characteristics such as thermal and ventilation losses per building and overall heat transfer coefficients are extracted at urban scale. The reliability of the method over its complexity is discussed with regards to the big sample that has been applied to. In general, the overall performance of the approach is satisfactory achieving a coefficient of determination with an average of 0.8, and is found to be in line with previous findings, considering also the high uncertainty associated with building-related input parameters. The extracted setpoint distribution should be transferrable across Scandinavia.

Keywords: smart meter data, temperature setpoints, housing stock model, thermal comfort preferences, U-values, urban scale

1 Introduction

A number of modeling methodologies have been developed to obtain information on physical resource flows through the building stock [1]. These are mainly used to characterize and predict energy demand of residential building stocks and to estimate energy savings after energy retrofitting strategies. Housing stock models can thus play an important role in supporting energy policy-making. In order to be useful, they should be reliable, efficient and interpretable [2]. Housing stock models can be broadly classified into two categories: top-down and bottom-up approaches. Top-down models rely on historical energy data and cannot model in detail individual end-uses [3]. Bottom-up models consist of engineering-based and statistical models. Statistical methods usually include macroeconomic and socio-economic effects, enable the determination of end-use energy consumption and are easy to develop and be used [4]. However, they cannot model the impact of specific technologies implemented and are less flexible. Engineering-based housing stock models use actual building physics and overcome some of the limitations induced by statistical models [2]. However, the majority of them are developed at national scale to support policy making and disregard heterogeneity within a country. They are also usually time intensive and are fully dependent on input data, hence inducing a high degree of uncertainty. Therefore, there is a need to focus on regional housing stock models that handle heterogeneity.

According to the International Energy Agency in the Energy Buildings and Communities Program (IEA EBC) Annex 53: Total Energy Use in Buildings, the six driving factors of energy use in building stock are : i) climate, ii) building envelope, iii) building energy and services systems, iv) indoor design criteria, v) building operation and maintenance, and vi) occupant behavior. Even though significant progress has been made in quantifying these primary drivers, more emphasis on energy related occupant behavior in buildings is needed to develop reliable and standardized methods [5, 6]. Neglecting this aspect can lead to severe miscalculations and inaccurate conclusions about the energy performance of the building stock [7]. Occupants' interaction with building systems affects significantly the total energy use of buildings. The occupants' gratification with their thermal environment defines thermal comfort [8]. Therefore, the occupants' perception of comfort or satisfaction in the built environment drives them to perform various controls (e.g. on HVAC systems and window operations) [9]. The adjustment of thermostat setpoints and indoor thermal environment are the most influential factors of heating loads along with heated areas [10]. Some studies have even classified occupants as active, medium and passive users based on their heating setpoint preferences which impact the indoor thermal environment and energy consumption [11, 12]. Therefore, thermostat setpoints are crucial input parameters to building energy models due to their big influence on residential energy use [9]. Currently, the understanding of occupant behavior is still insufficient both in building design, operation and retrofit, leading to incorrect simplifications in modeling and analysis [5]. In the past, information about occupants' interactions with systems was based on sporadic visits to households and rough estimates of thermal preferences of occupants.

The increasing deployment of intelligent metering systems in buildings and district systems creates a vast amount of building energy use and occupant-related information. Following the Third Energy Package in the Electricity [13] and Gas Directive [14] issued by European Commission in 2006, European countries plan to convert part of their legacy meter stock to smart by 2020 with a focus on electricity. According to the projections, by 2020, it is expected that almost 72% of European consumers will have a smart meter for electricity and about 40% will have one for gas [15]. The enormous amount of information and data opens up endless opportunities for researchers and engineers to study building dynamics and performance at a large scale. In combination with weather data and cross-sectional data, they can be utilized to develop more accurate prediction models and detailed analyses on the drivers of building energy consumption [16]. Smart meter data can also help developing and applying control strategies to improve building energy performance and efficiency [17]. Therefore, they can be utilized to decrease uncertainty related to building energy performance and occupant behavior and provide detailed information on energy monitoring.

National building databases and registers can support housing stock energy analysis, by providing information about building typologies and construction characteristics. These databases are usually created with regards to building regulations and schemes. In some cases, information from building owners via questionnaires has also been collected. Building information can be updated by local authorities and by citizens. However, occupants' interventions on the building fabric (i.e. energy renovation measures) are not regularly reported to building databases. Therefore, there is a significant gap between the data that has been registered and the real energy performance of the building.

This study aims at utilizing a big urban dataset, consisting of smart meter data from more than 14,000 households in a Danish city, to estimate temperature setpoints and thermal transmittances on building level. In addition, actual weather data, as well as data collected from a national building register and a geographic information system (GIS) have been utilized. A heat balance approach is implemented to the measured energy data of one year applying linear regression analysis to extract parameters that represent the whole heating season and the total building envelope. This approach has been inspired by the degree-day theory and aims at providing a new useful tool for utilities and researchers to extract building and thermal comfort-related characteristics at urban scale based on smart meter data. The data used allows us to capture the full range of heterogeneous behavior among people, through their temperature preferences. The estimation of people's variation enables the development of customized solutions and messages for them. The estimated thermal transmittance of the building envelope indicates the refurbishment state of the building and thus, provides more accurate insights into the building stock. The generated results -in the form of distributions- can be used to improve urban building energy models for the Scandinavian housing stock.

The rest of the article is organized as follows. In section 2, related works and methods to predict room temperature setpoints are summarized. In Section 3, we present and apply the heat balance model to the smart meter dataset.

In Section 4, the dataset is presented and basic information about the examined housing stock is described. In Section 5, the results are compared and validated with previous findings and relevant literature. The applicability of the methods with regards to the considerations made and the data used is discussed in Section 6. Section 7 summarizes the research findings.

2 Background

To evaluate the potential impact of different energy retrofitting scenarios in urban areas, bottom-up urban building energy models (UBEM) have been introduced over the past years. UBEMs have the potential to become key planning tools for utilities, municipalities and urban planners [18]. A key input of UBEM models are building characteristics of a given building stock from thermal envelope properties to usage patterns including the number of occupants, equipment loads and schedules as well as thermostat settings. Some of those information may be derived from census data. However, there is generally a surprising lack of data available related to the thermal performance of buildings. A useful source of information can be derived from individual building energy audits [19]. In [19], the authors used the Monte Carlo method and created a physics-based housing stock model for energy performance prediction, where inputs were probability distributions based on an Energy Performance Certification national database. Another source are nationwide building databases which include information such as floor areas, construction materials, age of construction, etc. Nevertheless, these databases may have flaws or may not be updated frequently enough. Therefore, there is high uncertainty related to input parameters of UBEMs.

Several studies have been conducted on district or urban scale making use of statistical models and data mining techniques in order to extract hidden useful knowledge from building-related data, as well as to forecast energy consumption. The authors of [20] presented a data-driven approach to modeling end user consumption based on data from 6,500 buildings in Cambridge, Massachusetts, using linear regression analysis and Gaussian process regression. In [21], electric energy data of thousands of buildings were investigated to extract specific features based on socio-economic information. In [17], a data mining method was proposed to analyze building-related data in order to establish building energy demand predictive models and examine the influence of occupant behavior on energy consumption. Older studies had also made use of regression analysis based on billing data to determine household energy. A study by [22] used monthly energy billing data to decompose energy use to weather and non-weather dependent elements, as well as explain anomalies in energy use of some households.

Determining the internal temperatures or temperature setpoints has been of particular interest, especially in residential buildings in mostly heating dominated climate since the main source of building energy demand is driven by heating which in term directly depends on the temperature difference between inside and outside. Temperature setpoints and heating duration may differentiate across dwellings based on their preferred thermal comfort range, which affect the resulting internal temperatures. Nevertheless, many top-down urban scale models assume the same constant temperature setpoints for the whole building stock, while the rest calculate internal temperatures as a function of building envelope, occupancy and systems [23].

Most of the existing literature puts emphasis on predicting internal temperatures or measuring household room temperatures at district scale based on temperature recordings [24]. For example, the authors in [25] collected temperature and humidity data from 1,604 study dwellings in order to determine the effect of dwelling and household characteristics on indoor temperature variation. The median standardized daytime living room temperature was calculated to be 19.1ºC, while the night time bedroom temperature was found to be 17.1ºC. Temperatures were affected by the building envelope characteristics, thermal efficiency, number and age of occupants. The socio-economic status was not strongly related to them. In [26], temperature recordings from 821 English dwellings were collected and analyzed, which were monitored in different zones within the dwellings. The standardized internal temperatures were estimated by regressing the mean hourly indoor temperatures on outdoor temperature. The results showed that more efficient buildings have higher indoor temperatures at all outdoor conditions, as well as that households with children were the warmest. In another study [23], internal temperatures were predicted at high temporal resolution using panel methods based on data from 280 households. Temperature recordings were taken every 45 minutes for 6 months, while the model was generated using mean daily temperature data. It was concluded that thermostat settings play an important role in reducing total energy consumption. Moreover, centrally heated dwellings and detached homes had lower internal temperatures, while for each additional person living in a household the mean internal temperature increased by ~0.25ºC. In another study [27], internal temperature recordings were gathered by 54 households in China, which were found to be on average 13.5ºC for living rooms and 12.7ºC for bedrooms. These are far outside the ASHRAE steady state thermal comfort zone, highlighting the differences among climates in terms of construction and cultural behavior. Furthermore, the effect of the climate is pronounced on these findings, as this study referred to the 'hot summer-cold winter' climatic zone of China being characterized by relatively low ambient temperatures during winter season -which is quite limited- and very high temperatures during the rest of the year. However, these winter climate conditions are comparable to the ones in UK. The very low internal temperatures noticed in this study are a result of the fact that domestic heating is operated part-time-part-space in that region of China, meaning that only specific rooms are heated up instead of the whole house and only for a shorter time than a usual heating season duration.

Apart from the measurement of internal temperatures, some studies have tried to predict them. Most of them refer to smaller cases and make use of simple heat balance models. The authors in [28] used the Domestic Energy and Carbon Model [DECM] to estimate among others the annual internal temperatures in different types of English dwellings, which were on average 18.4ºC. The mean internal temperature was also found to be highly correlated with the CO_2 emissions of each dwelling and outweighed the effect of climate and building fabric construction. The authors of [29] proposed a method to explore future transformations in the UK housing stock based on the English House Condition Survey data. Besides energy demand prediction, they used the heat-balance method to infer the appropriate modeling temperatures for a base year in UK. In particular, the profiles for the internal temperatures varied from 14.1 to 18.7ºC for one of the two modeled zones, while these temperatures were 3ºC lower for the other zone. The study by [30] modeled 37 dwellings employing the BREDEM algorithms [31] and predicted the annual internal temperatures and the heating demand temperatures which equals the temperature setpoint in most cases. It should be noted that most of the afore-mentioned studies refer to the UK housing stock, while there is a lack of similar literature for Scandinavian or Danish building stock.

In cases where internal temperatures are not available, degree-days have been used as a tool to assess and analyze weather related energy consumption in buildings for over seven decades. The concept originates from agricultural research, where variation in outdoor air temperature is also important which is also the case for building energy use [32]. Degree-days can be defined as the summation of differences between the outdoor temperature and a building reference temperature (base) over a specified time period for both heating and cooling systems. The simplicity of the concept of heating degree-days (HDD) has led to a plethora of studies in the area of building energy use analysis. The influence of HDD and heating degree hour (HDH) on hourly electricity consumption of hundreds of Norwegian households was investigated and concluded that HDD achieve a slightly higher goodness of fit compared to HDH [33]. The authors in [34] proposed a city-scale degree-day method, according to which they extracted the average building heat loss rate and a city-scale base temperature for the area of Strasbourg in order to estimate the aggregate heating energy demand, while accounting for the urban heat island effect. In addition, residential heating energy requirements and fuel consumption for the city of Istanbul were estimated making use of HDD method and air temperature records, while studying different construction types [35]. A study by [36] utilized energy data to estimate the base temperature of a single building along with the heat loss coefficient and subsequently heating degree-days using Bayesian inference. The performance line method and energy signature method were presented in [37] to estimate the building's base temperature based on daily energy data and outdoor temperatures and to extract degree-days.

3 Method

The method applied in the current experiment takes an existing heat balance approach a step further. The temperature setpoints and total heat loss coefficients are estimated from smart meter data for thousands of buildings. The granularity of energy data is hourly and the sample of buildings covers a large share of a Danish urban building stock, as opposed to previous studies that have been found in literature and cover much smaller building samples. This allows to draw conclusions about the applicability and accuracy of the method applied.

3.1 Main considerations

To treat this urban-scale sample of buildings, the following considerations were made. Firstly, only the coldest days of the year were considered. In particular, only the data that correspond to days when average daily ambient temperature was lower than 15^{o}C were taken into account. In this way, a normal operation of the heating system would be ensured. Moreover, solar gains are reduced in Denmark especially during winter time. This results in transient phenomena being less dominating. Thus, steady-state conditions would be applicable. The investigated buildings are small enough that uniform air mixing was assumed. All households were treated as single-zone models in order to reduce complexity and computation times. In addition, the investigated housing stock is homogeneous with regards to building type (consisting only of single-family houses), allowing a somewhat accurate estimate of occupant density and construction standard. As no information about specified occupancy schedules was available either, internal heat gains were defined according to the Danish Building Research Institute guidelines [38] as 5 W/m^2 (corresponding to external floor areas), which is a sum for people loads and equipment loads for a residential building. This value was used as an average for all different spaces in households (kitchen, living room, bedrooms etc.), so that it is in line with the single-zone model approach. Furthermore, the 'equivalent' temperature setpoint, which represents the volume-weighted mean temperature across all conditioned and unconditioned spaces in every household, was assumed to be constant throughout the heating season that we investigated.

3.2 Calculation of equivalent thermostat heating temperatures and U-values

Inspired by the classical degree-day method and based on the steady-state heat balance for a room, the following formula (equation 1) can be derived. According to that, total energy loads for building space heating over a specified period are directly proportional to the building heat losses that vary with the change in the current indoor-outdoor air temperature gradient (i.e. natural ventilation and air infiltration, wall heat conduction). Similar equations to the following heat balance model, solved for Q_H, have been traditionally used to calculate the heat output of well controlled heating systems [37].

$$T_o = T_i - \frac{Q_H + Q_{SG} + Q_{IG}}{UA + c_p\,\rho\,n} \qquad (1)$$

where T_i is the equivalent internal temperature in ^oC, T_o is the ambient temperature in ^oC, Q_H is the heating demand in W, Q_{SG} stands for solar gains in W, Q_{IG} represents the internal heat gains from occupants and equipment in W/m^2, U is the overall heat transfer coefficient across the building envelope in W/m^{2}K, A is the total envelope area in m^2, n is the ventilation or infiltration rate in m^3/s, c_P is the specific heat capacity for air under constant pressure which equals to 1000 J/kg K and ρ is the density of air which equals to 1.2 kg/m^3. The heat loss factor due to the air change can be also simplified to equal $0.33\,N\,V$, where N is the ventilation rate in air changes per hour (ACH) and V is the volume of the building in m^3.

Due to the extensive data quantity both in terms of number of monitored buildings and data granularity, complicated approaches would be less suitable. So, the methodology applied in this case was simple linear regression analysis. The ordinary least squares (OLS) method was used and run for each household's data. According to [39], the OLS method is the most efficient for urban scale models having the lowest deviation from measured data.

The four main statistical assumptions regarding OLS method were tested to prove whether they were satisfied on building level: the regression is linear in parameters, the sampling of observations is random, the conditional mean is zero and there is no multi-collinearity.

Thus, in the current model, the ambient temperature T_o was modeled as an affine function of the total energy load, so that $T_o = T_i - \frac{load}{losses}$, where $load = \ Q_H + Q_{IG} + Q_{SG}$ and $losses = UA + 0.33\ N\ V$. The internal temperature or temperature setpoint and the losses are expressed by the intercept and the slope of the linear model, respectively. The losses represent the envelope and ventilation losses. The energy load quantity will always be positive. Q_H represents the space heating consumption of the households which was given on hourly resolution. However, in order to account for latent thermal mass effects in the buildings, it was decided to aggregate the hourly energy data on daily basis and report the aggregate value of energy consumption (heating). Thus, internal temperatures in equation 1 refer to the mean of internal temperatures recorded, representing the relative heating profile of a dwelling. The internal gains were also aggregated on a daily resolution based on the indicated hourly values.

In the afore-described model, the heat losses are mainly attributed to conduction through the building envelope and to ventilation losses. For that reason, we are focusing only on the winter time or heating season when the ambient temperatures are low; furthermore, the solar gains are significantly lower. Thus, as mentioned before, the regression model was restricted to periods when the mean ambient temperature was less than 15ºC. Information concerning building characteristics and occupancy were retrieved by the national building database and by national building regulations, where no detailed information was available. Actual weather data on hourly resolution were acquired for the area of Aarhus covering the examined period (2014-2015) and included ambient temperature, solar radiation and relative humidity.

The calculation of solar gains was based on the actual weather data available for the area consisting of direct and diffuse radiation. An exemplary building was made in DIVA software [40], which is a daylighting and energy modeling plug-in for Rhinoceros, according to a Radiance calculation, where time-integrated solar irradiance on each surface was calculated as presented in [41, 42]. Radiance is a backward raytracer that was originally developed at Lawrence Berkeley National Laboratory [43]. An average value of integrated solar irradiance for all four surfaces of an exemplary building was taken, which was then inserted into the model. According to the following equation (2), the solar gains are calculated as the product of overall solar heat gain coefficient (*SHGC*), projected area of fenestration A_f in m^2 and incident total irradiance E_t in W/m^2.

$$Q_{SG} = SHGC_t\ A_f\ E_t \qquad (2)$$

To estimate the equivalent overall heat transfer coefficient of the building envelope for each household, the losses factor in equation 1 had to be determined. The geometry of the houses (i.e. envelope areas) was available through an additional dataset consisting of GIS data at city scale. Therefore, the ventilation and infiltration rate had to be determined taking into account the type of mechanical systems installed in each house, possible insulation state of building envelope and air leakages.

4 Data

4.1 Typical dwelling characteristics in Denmark

Approximately 45% of Danish dwellings are detached houses [44], although their share is significantly less in larger cities (e.g. Copenhagen or Aarhus). A large share of heating (namely 63%) in private Danish houses is provided by district heating, which covers both space heating and domestic hot water demand [45]. Hydronic systems are the most common heat emission systems found in residential buildings. Mechanical cooling and ventilation is mostly available in commercial buildings. The national average energy consumption for Danish single-family houses constructed in the period before 1980 is 151 kWh/m^2 per year and 102 kWh/m^2 per year for houses constructed between 1980 and 2000 [46]. That gives an indication of the energy efficiency of the investigated building stock. Moreover, according to the Danish building regulations, there should be a sufficient amount of thermal mass and insulation layer in newer buildings, which results in more airtight building envelopes. More information and typical building construction examples of Danish single-family houses can be found in [47].

4.2 Description of smart meter data

Our approach is evaluated on basis of smart meter data collected initially from 15,063 households in Aarhus, Denmark at a 60-minute granularity of one full year. The data is provided by Aarhus AffaldVarme (AVA), which supplies district heating to the inhabitants of Aarhus. The data ranges from August 2014 to August 2015. The data contains the heating consumption of residential AVA customers. The households are all single-family houses (SFH). The periods that the data cover range depending on when each of the smart meters was installed. There is a slightly smaller number of smart meters that have recordings for the whole investigated period. Thus, energy data from 14,182 households are utilized in the experiment, which correspond to the whole examined period. The data has been cleansed and aggregated at a daily time scale and the results of this analysis are presented in the results' section.

4.3 Building characterization

The dataset covers a big spectrum of single family houses in Aarhus, ranging from buildings constructed in 1800 to 2015. The floor areas range from 25m^2 to 504m^2, with a mean of 134m^2. The number of floors varies between 1 and 3, with the majority of the buildings being single-floor houses. These information are presented in Figure 1.

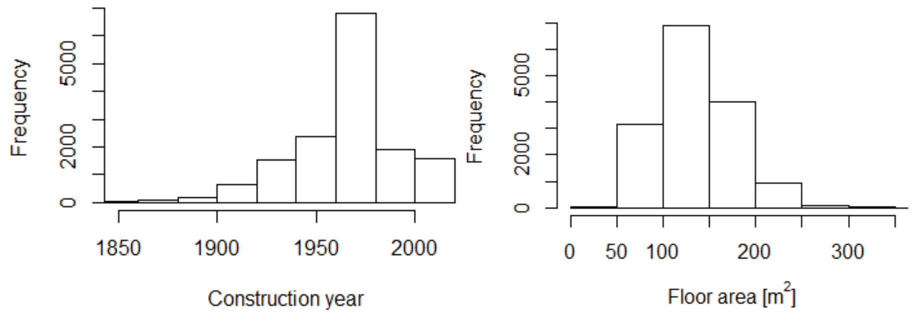


Figure 1. Building information of dataset

5 Results

The afore-mentioned method was applied to smart meter data of 14,182 households over one full year and the results are presented hereafter.

5.1 Variables' distributions

Figure 2 presents the main results of this study, which are the temperature setpoint distribution and the distribution of the total envelope and ventilation loss factor for all examined households, as estimated according to equation 1. As it can be seen, the mean equivalent temperature setpoint that was calculated for the whole sample of buildings was 19.1°C, while the median value was 19°C. The standard deviation was 1.54°C. This temperature is a bit lower than the expected indoor temperature value, which would be close to 20°C according to estimations of the Danish Building Research Institute [48] included in Danish Standards [49]. It can be attributed to two reasons. First, every household is modelled as a single-zone, which means that any unconditioned spaces are also included and represented with this specific equivalent parameter. Second, this figure refers to equivalent temperature setpoint, which is inferred over a whole season. So, unoccupied periods and hours with night-setback are also included in this estimation. The internal heat gains that were assumed in the model proved to have a decisive impact on the temperature results, so higher internal gains than the ones defined in the standard would lead to higher temperatures. Nevertheless, the uncertainty associated with the internal and solar gains did not result in high inaccuracy of the model. Daily internal heat gains ranged from 5 kWh to 110 kWh among the different households. The temperature setpoint distribution seems to approach normality, while the building envelope and ventilation losses graph approximates a log-normal distribution curve. The mean building envelope and ventilation losses were calculated to be 0.1 kW/K with a standard deviation of 0.03 kW/K. Additional conclusions can be drawn on how close to expectation the current used model has performed and the goodness of the accuracy that can be achieved at such a large scale. Despite the steady-state heat balance model that was applied on the building sample and the considerations that were made about input parameters, the results indicate a reasonable range of temperature setpoints and loss factor across the dwellings. In order to interpret the thermal environment and thermal comfort that occupants perceive based on the estimated equivalent internal temperature, the

Predicted Mean Vote (PMV) and Predicted Percentage Dissatisfied (PPD) indices proposed by Fanger [50] were used. The average PMV was calculated to be -0.6, which corresponds to PPD equal to 12.9%. Based on these estimations, the thermal environment falls into Category C out of the three desired thermal environment categories as described in [51].

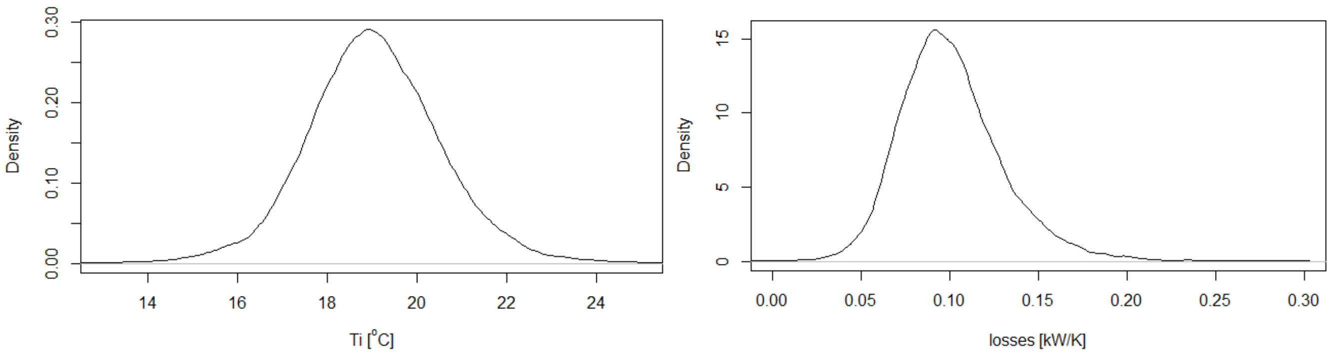


Figure 2. Distribution of temperature setpoint (left) and total building envelope and ventilation losses (right)

To assess the goodness of the fit for the simple linear regression models, the coefficient of determination (R^2) was utilized to estimate the variance of the predictable variable from the independent variable. The results are presented in Figure 3. It has to be noted that a relatively high R^2 is achieved with respect to the nature of the experiment, which regards end users energy use. The mean and median value of the coefficient of determination were found to be 0.8 and 0.83, respectively, with a standard deviation of 0.11. That confirms the good fit of the linear regression models to the smart meter data, as well as a quite consistent model, considering the large amount of data that has been used.

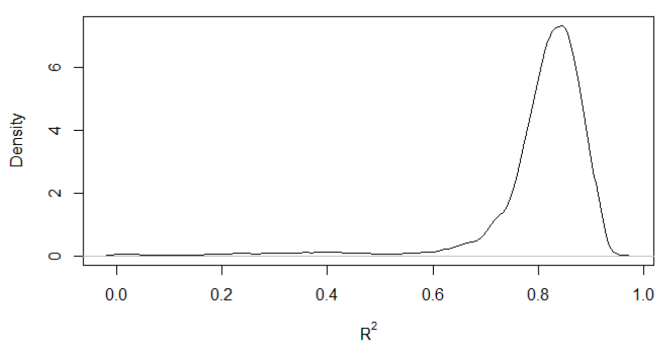


Figure 3. Distribution and boxplot for R^2 for all houses' fits

In Figure 4, the residual standard deviation distribution is illustrated, which can indicate the variability of predictions in the regression. This shows the deviation of the errors and not the errors of the regression themselves. It can be observed that the deviation of the residuals has a mean of 1.53, which is satisfactory considering also the big building sample.

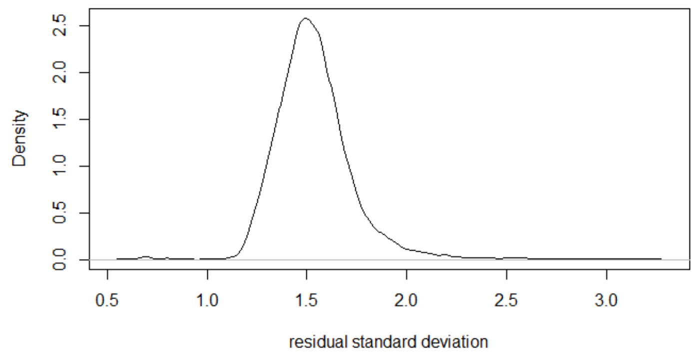


Figure 4. Residual standard deviation distribution for all houses

5.2 Exemplary households

The regression models for two randomly selected households are presented hereafter. The intercept of the models represents the estimated temperature setpoint, while the slope stands for the coefficient of envelope and ventilation losses. The coefficient of determination for these two cases were R^2=0.84 and R^2=0.8, respectively. The estimated temperature setpoints for these two models were 19.39°C and 18.91°C, respectively. Both models are quite similar, with the household presented to the right having a bit more scattered energy use data and higher number of outliers.

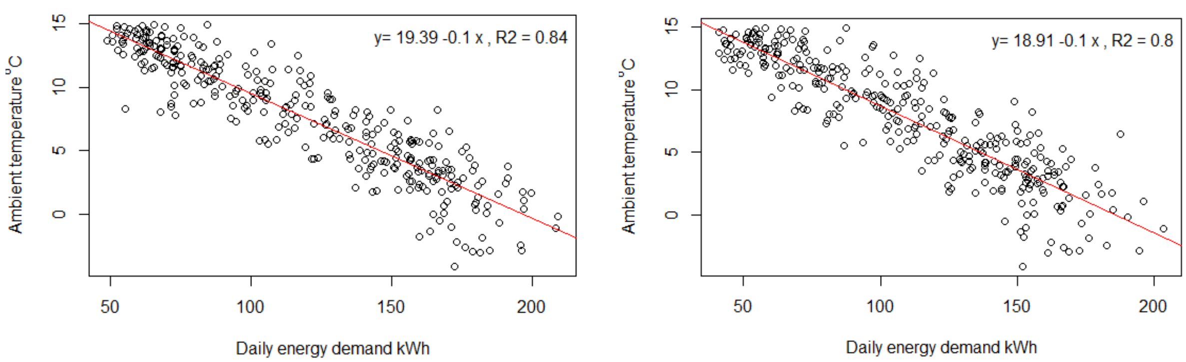


Figure 5. Regression fit for two random households

5.3 Regression diagnostics

The main assumptions of OLS were tested for numerous randomly selected models and the following conclusions were drawn. First, normality was tested by checking the probability plot of the standardized residuals against the values expected under normality, concluding that the normality assumption is satisfied. Second, the residuals did not have non-linear patterns, confirming the linearity condition. Third, the constant variance assumption was met, since the residuals were spread equally and randomly along the ranges of predictors. The errors of the regression models had zero mean as required by the Gauss-Markov theorem. The independence assumption was also satisfied. Lastly, the outliers that were influential to the regression results were determined. That means that these may be influential to the regression analysis and the results will change if we exclude these cases.

5.4 Estimation of equivalent total U-value of building envelope

After the total loss factor for each household was calculated, the equivalent overall heat transfer coefficient, U-value, for the building envelope was calculated. This would provide additional valuable information for the insulation state of the buildings and indicate any possible energy refurbishments. Thus, the ventilation and infiltration rate on building level had to be determined. According to [52], the infiltration rates for Danish single-family houses vary between 0.1 and 1 ACH. These were estimated for each household in accordance with available information about the age of the building and construction characteristics extracted from the national building register. Specifically, an algorithm was used according to which the infiltration rate took values in *[0.1, 1]* based on the construction year of the building and a binary variable that indicated if the house had undergone any energy renovation that has been registered in the national building database. These two factors would collectively indicate any thermal bridges on the building envelope affecting the infiltration rate. The estimated infiltration rate for each household was assumed to be constant throughout the course of the day. The mean value of natural ventilation rate for the investigated building stock was estimated to be 0.5 ACH. Rough information for the geometry of the houses and the envelope areas on building level was available through the Danish Building Register (BBR) [53]. BBR is a nationwide register including data for the majority of Danish buildings and households. Nowadays, it contains information about 1.6 million properties, 3.8 million buildings and 2.7 million dwellings and commercial units [54]. It was originally set up in 1977 by collecting information from building owners via questionnaires. Since then, it has been updated by local authorities and by citizens [55]. Data contained in BBR -provided for every registered house- are categorized into areas, building constructions and installations. Information about areas can be summarized to the following: total building/residential/commercial area, built-up area, number of storeys, attic and total basement area. The information extracted for the current analysis was the building footprint area and the number of floors for each building. This information was then coupled and validated with an open-source GIS dataset for the city of Aarhus [56]. Thus, the total envelope area for each building was extracted, as well as the total building volume.

The results of this analysis are presented in Figure 6. The distribution approximates a log-normal curve. The mean value of the equivalent U-values for the examined residential building stock is calculated to be 0.58 W/m^{2}K with a standard deviation of 0.22 W/m^{2}K. The median value is 0.54 W/m^{2}K. This finding gives insight into the state of the building envelopes across Danish SFH, as well as their current energy refurbishment state. According to the Danish Building Research Institute [57], the average area-weighted U-values of Danish single family houses -as calculated with regards to the national building regulations- mainly vary from 0.3 to 0.65 W/m^{2}K depending on the construction age of the building. It should be noted that the majority of the investigated single-family houses were constructed in the 1960's, which have an average area-weighted U-value of 0.52 W/m^{2}K according to [57]. The overall heat transfer coefficients of the buildings do not always comply with the ones defined in the national building databases or introduced by the building regulations. Therefore, the calculated U-values can determine the level of energy refurbishment that might have been implemented to the buildings. Such updated information, although very important for energy calculations, is missing from the majority of the Danish building register [58]. The results from this study can also validate existing values in building databases. In addition, since this equivalent U-value summarizes the heat transfer coefficient for the whole building structure, no specific conclusions can be drawn for the particular building components. However, they can be assessed in combination with the rest of building characteristics and reveal valuable information about the building energy performance of the stock.

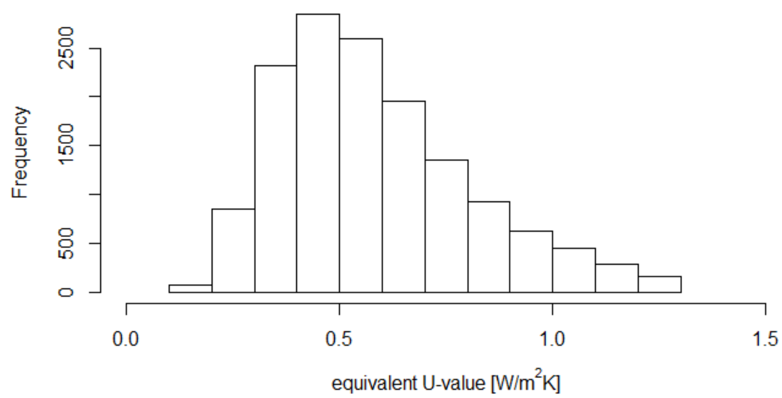


Figure 6. Distribution of equivalent U-values for the building stock

6 Discussion

The proposed method aims at investigating the applicability of a simplified heat balance approach on smart meter data to derive temperature setpoints and thermal transmittances at urban scale. Our analysis focused on fitting the overall heating energy distribution across the investigated building stock. The lack of information about the share of heated and unheated space resulted in the use of an equivalent temperature setpoint, which represented the volumetric mean temperature indoors. That means that there may be rooms that are warmer (i.e. occupied spaces) and rooms that are cooler in a building. However, the scope of our analysis was to treat the whole building as an entity. In Danish building regulations and directives, a dimensioning internal temperature of 20ºC is commonly used for the majority of Danish houses. However, it is also mentioned that in older poorly insulated houses, the room temperature can be even lower than 20ºC or some rooms are unheated to reduce the heating costs [48]. The thermal environment evaluation was conducted assuming that the living spaces (i.e. whole house) were heated. However, if unheated spaces were to be evaluated, the desired thermal environment follows a different classification. According to steady-state conditions of the model, changes in temperature sepoints were not taken into account. The use of a steady-state model resulted in extreme conditions and transient states not being taken into account. If a transient model was to be used instead, the temperature setpoint would no longer be represented by a single value for the whole heating season but replaced possibly with daily or even hourly ones. Moreover, the assumptions on constant infiltration rates throughout the day and over the heating season and constant internal gains throughout the day and among the different households would have to be adjusted. In particular, varying heat flow rates from occupants and appliances could be used based on hourly schedules and differentiating between weekdays and weekends. Nevertheless, the current study aimed at producing results that could be utilized subsequently by urban building energy models or housing stock models, which in their majority require a single heating temperature setpoint per building or zone. Thus, it would be outside the scope of this work to calculate dynamic temperature setpoints. In addition, this would increase the computational complexity enormously and subsequently the computation time for urban-scale or district-scale building stocks.

The biggest uncertainty on the total energy load factor comes from the internal gains and solar gains factor. Internal gains were assumed to differentiate according to floor area based on Danish standards. As a result, the residents of the investigated area were assumed to have the same occupancy patterns. Internal gains depend on occupancy behavior and schedules, as well as cultural patterns that can be estimated but not be predicted accurately at a large scale. Thus, a certain amount of assumptions based on recommendations by standards had to be taken, which resulted in a higher uncertainty in the housing stock model predictions. Solar gains had comparatively smaller impact on the energy load and the assumption about an average orientation did not seem to affect significantly the results. This should be considered in combination with the local climate, which in this case is characterized by relatively low solar radiation during the heating season and low ambient temperatures.

Despite the considerations that were made regarding the envelope properties, the performance of the proposed approach was satisfactory compared to existing knowledge and statistical values included in literature. The analysis was run on a private scientific cloud (SciCloud) using a web-based interface to interact with the data [59]. SciCloud consists of 18 physical servers with 80 cores and 564 GB of memory, as well as 4.2 TB of node storage, plus 1.2 TB of network storage. The environment used for the data analysis was *R*. The computation time for the whole housing stock analysis did not exceed two hours, due to the low complexity of the heat balance models and the OLS method. Therefore, the proposed model is expected to balance predictive accuracy and parsimony. If more complicated heat transfer phenomena were to be included in the model, the computation time would increase significantly. The big amount of historical measured energy data led to statistically significant results and strengthened our methodology. On the other hand, such large-scale energy data are mostly available to utilities and less to the research community due to privacy issues. Anonymization techniques can be applied and facilitate significantly the data acquisition and publishing process.

Measurements of internal temperatures for a smaller sample of the examined building stock would validate this methodology. The installation of smart sensors in newly constructed buildings measuring internal temperatures in rooms would allow this. However, the share of newly constructed buildings remains quite low in the city of Aarhus and the employment of smart sensors in the existing building stock would be a relatively long and costly procedure. Moreover, the application of the proposed method on a much smaller sample, which more detailed building and occupant-related information (i.e. infiltration rates, occupancy schedules) would be available for, would be a next step to continue this study. In this direction, occupant data could be used to approximate internal heat gains more accurately. The current sample is a relatively homogeneous building stock, including only single-family houses in the same urban area. Their thermal envelope characteristics varied a lot, though, in terms of building envelope insulation, construction materials and geometry. Thus, a wide range of building construction types was covered. Our method proved to work reliably on the examined stock with regards to the number of assumptions that were made. This resulted in the estimated parameters including high second-order uncertainty. To cope with this uncertainty, probabilistic methods would be required which would be less attractive for this case of thousands of buildings. Despite the uncertainty included in the input variables and estimated parameters, heterogeneous preferences of occupants regarding thermal comfort were determined successfully. Nevertheless, further work needs to be carried out to investigate how it can be expanded upon more diverse building stocks. Finally, the Danish climate is characterized by decreased solar gains and cold winters, which make the transient phenomena being less dominating. If this work was to be reproduced to different climates with increased solar gains, the quasi-steady-state conditions would be less valid. Also, equation *(1)* would have to be adjusted accordingly so that a utilization factor for solar heat gains (e.g. function of time constant) is included that accounts for the dynamic heat flows within the building.

7 Conclusion

This study has utilized an urban dataset of more than 14,000 households in Aarhus, Denmark to derive temperature setpoints and overall heat transfer coefficients at house level. This dataset comprised of measured daily heating energy data, actual weather data, basic building typological data and geometry information extracted from a building register and GIS data. A heat balance model –inspired by the degree-day theory- was proposed and applied in combination with linear regression analysis. The results showed that a good fit was achieved overall in the majority of the examined households. The results provided distributions of i) equivalent temperature setpoints, ii) a factor for building envelope and ventilation losses, as well as iii) equivalent U-values for the building envelopes. The average equivalent temperature setpoint was calculated to be 19.1^{o}C across the investigated Danish dwellings, considering both heated and unheated spaces in the buildings. This value represented the mean volumetric temperature indoors. The mean overall heat transfer coefficient for the total building envelope of all houses was estimated to be 0.58 W/m^{2}K. The mean value of the coefficient of determination (R^2) was 0.8, indicating a good fit of the linear regression models. It was found that Danish homes differed in heating setpoint temperatures and envelope insulation state. The energy data was proven to be highly correlated with the weather data (consisting of ambient temperature and solar radiation), as expected. Statistically significant results have been reported due to the big sample size and the consistent and granular energy measurements. Therefore, the proposed steady-state approach is applicable and recommended for urban-scale building samples when a uniform setpoint temperature is adequate to be extracted for the whole heating season. Furthermore, overall heat transfer coefficients for the whole building envelope can be used to determine any possible energy retrofit measures that have been applied to the envelope. Moreover, this method enabled the capturing of the full range of heterogeneous behavior among people, as reflected on their temperature preferences. These findings are important to characterize the thermal comfort preferences of occupants and their interactions with the building systems, which are top influential factors of residential heating energy loads.

The interest of using this method goes beyond the results that are presented here. The study provided insights that will help direct future research in identifying ways to estimate temperature setpoints, assess indoor thermal comfort and consequently improve urban building energy models. Thus, customized messages and solutions for occupants can be developed. Furthermore, it provided better understanding of the Danish building stock and its occupants, which is crucial in building energy demand calculations. Thus, the given methodology can be applied to large scale smart meter datasets to acquire building-related and thermal comfort-related characteristics. In addition, it can provide valuable information to utilities to further optimize their network and apply advanced technologies (i.e. virtual storage on district heating network) based on the temperature setpoint distribution of the district. The proposed simplified approach opens up new possibilities of building performance analysis at urban scale. Next step would be to couple this energy dataset with internal temperature recordings, which would enable the validation of the estimated variables at urban scale. In that way, the current findings could be utilized to challenge assumptions used in Scandinavian housing stock models regarding heating patterns and mostly, temperature setpoints throughout the heating period.

8 Acknowledgements

This work was undertaken as a part of the CITIES (Centre for IT-Intelligent Energy Systems in cities) project number DSF1305-00027B funded by the Danish Strategic Research Council. Special thanks should be given to Aarhus AffaldVarme, who enabled the distribution of the smart meter data to the author.

9 References

1. International Energy Agency (2001). Annex 31 - Stock aggregation: methods for evaluation of the environmental performance. International Initiative for a Sustainable Built Environment (iiSBE).
2. A.T. Booth, R. Choudhary, D.J. Spiegelhalter (2012). Handling uncertainty in housing stock models. Building and Environment 48, pp. 35-47.
3. L. Swan, V. Ugursal (2009). Modeling of end-use energy consumption in the residential sector: A review of modeling techniques. Renewable and Sustainable Energy Reviews 13, pp. 1819-1835.

4. M. Kavgic, A. Mavrogianni, D. Mumovic, A. Summerfield, Z. Stevanovic, M. Djurovic-Petrovic (2010). A review of bottom-up building stock models for energy consumption in the residential sector. Building and Environment 45, pp. 1683-1697.
5. T. Hong, S.C. Taylor-Lange, S. D'Oca, D. Yang, S. P. Corgnati (2016). Advances in research and applications of energy-related occupant behavior in buildings. Energy and Buildings 116, 694–702.
6. K. Sun, T. Hong (2017). A framework for quantifying the impact of occupant behavior on energy savings of energy conservation measures. Energy and Buildings 146, pp. 383-396.
7. W. O'Brien, I. Gaetani, S. Carlucci, P-J. Hoes, J. L.M. Hensen (2017). On occupant-centric building performance metrics. Building and Environment 122, pp. 373-385.
8. ANSI/ASHRAE, ANSI/ASHRAE Standard 55: Thermal Environmental Conditions for Human Occupancy, ANSI/ASHRAE, 2004.
9. Z. O'Neill, F. Niu (2017). Uncertainty and sensitivity analysis of spatio-temporal occupant behaviors on residential building energy usage utilizing Karhunen-Loeve expansion. Building and Environment 115, 157-172.
10. M. Bonte, F. Thellier, B. Lartigue (2014). Impact of occupant's actions on energy building performance and thermal sensation. Energy Build 76, 219–227.
11. V. Fabi, R.V. Andersen, S.P. Corgnati (2013). Influence of occupant's heating set-point preferences on indoor environmental quality and heating demand in residential buildings. HVAC&R Research 19 (5), 635–645.
12. S. D'Oca, V. Fabi, R.K. Andersen, S.P. Corgnati (2014). Effect of thermostat and window opening occupant behavior models on energy use in homes. Building Simulation 7, 683–694.
13. The European Parliament and the Council of European Union (2009). Directive 2009/72/EC concerning common rules for the internal market in electricity and repealing Directive 2003/54/EC. Official Journal of the European Union.
14. The European Parliament and the Council of European Union (2009). Directive 2009/73/EC concerning common rules for the internal market in natural gas and repealing Directive 2003/55/EC. Official Journal of the European Union.
15. European Commission (2014). Benchmarking smart metering deployment in the EU-27 with a focus on electricity. Report from the Commission, Brussels.
16. A. Kipping, E. Trømborg (2016). Modeling and disaggregating hourly electricity consumption in Norwegian dwellings based on smart meter data. Energy and Buildings 118, pp. 350-369.
17. Z. Yu (2012). Mining Hidden Knowledge from Measured Data for Improving Building Energy Performance. PhD Thesis, Concordia University.
18. C.F. Reinhart, C. Cerezo Davila (2016). Urban building energy modeling – A review of a nascent field. Building and Environment 97, pp. 196-202.

19. J.N.B. Fonseca, M.J.N. Oliveira Panão (2017). Monte Carlo housing stock model to predict the energy performance indicators. Energy and Buildings 152, pp. 503-515.
20. J. Zico Kolter, J. Ferreira (2011). A Large-scale Study on Predicting and Contextualizing Building Energy Usage. In: Proceedings of the Twenty-Fifth AAAI Conference on Artificial Intelligence, San Francisco, USA.
21. C. Becker, L. Sadamori, T. Staake, S. Santin (2014). Revealing Household Characteristics from Smart Meter Data. Energy, Vol 78, pp. 397-410.
22. E. Hirst, R. Goeltz, D. White (1986). Determination of household energy using 'fingerprints' from energy billing data. Energy Research 10, pp. 393–405.
23. S. Kelly, M. Shipworth, D. Shipworth, M. Gentry, A. Wright, M. Pollitt, D. Crawford-Brown, K. Lomas (2013). Predicting the diversity of internal temperatures from the English residential sector using panel methods. Applied Energy 102, pp. 601-621.
24. K. Vadodaria, D.L. Loveday, V. Haines (2014). Measured winter and spring-time indoor temperatures in UK homes over the period 1969-2010: A review and synthesis. Energy Policy 64, pp. 252-262.
25. T. Oreszczyn, S.H. Hong, I. Ridley, P. Wilkinson, The Warm Front Study Group (2006). Determinants of winter indoor temperatures in low income households in England. Energy and Buildings 38, pp. 245-252.
26. I.G. Hamilton, A. O'Sullivan, G. Huebner, T. Oreszczyn, D. Shipworth, A. Summerfield (2017). Old and cold? Findings on the determinants of indoor temperatures in English dwellings during cold conditions. Energy and Buildings 141, pp. 142-157.
27. B. Lin, Z. Wang, Y. Liu, Y. Zhu, Q. Quyang (2016). Investigation of winter indoor thermal environment and heating demand of urban residential buildings in China's hot summer – Cold winter climate region. Building and Environment 101, pp. 9-18.
28. V. Cheng, K. Steemers (2011). Modelling domestic energy consumption at district scale: A tool to support national and local energy policies. Environmental Modelling & Software 26, pp. 1186-1198.
29. S. Natarajan, G. J. Levermore (2007). Predicting future UK housing stock and carbon emissions. Energy Policy 35, pp. 5719-5727.
30. S. K. Firth, K. J. Lomas, A. J. Wright (2010). Targeting household energy-efficiency measures using sensitivity analysis. Building Research & Information 38, pp. 25-41.
31. B. R. Anderson, P. F. Chapman, N. G. Cutland, C. M. Dickson, G. Henderson, J. H. Henderson, P. J. Iles, L. Kosmina, L. D. Shorrock (2002). BREDEM-12. Model description, 2001 update. BRE, Department for Environment, Food and Rural Affairs.
32. The Chartered Institution of Building Services Engineers (CIBSE) (2006). Degree-days: theory and application. TM41: 2006, London.
33. A. Kipping, E. Trømborg (2016). Modeling and disaggregating hourly electricity consumption in Norwegian dwellings based on smart meter data. Energy and Buildings 118, pp. 350-369.

34. M. Kohler, N. Blond, A. Clappier (2016). A city scale degree-day method to assess building space heating energy demands in Strasbourg Eurometropolis (France). Applied Energy 184, pp. 40-54.
35. A. Durmayaz, M. Kadioglu, Z. Sen (2000). An application of the degree-hours method to estimate the residential heating energy requirement and fuel consumption in Istanbul. Energy 25, pp. 1245-1256.
36. D. Lindelöf (2017). Bayesian estimation of a building's base temperature for the calculation of heating degree-days. Energy and Buildings 134, pp.154-161.
37. A.R. Day, I. Knight, G. Dunn, R. Gaddas (2003). Improved methods for evaluating base temperature for use in building energy performance lines. Building Services Engineering Research and Technology 24, pp. 221-228.
38. S. Aggerholm, K. Grau (2016). Bygningers energibehov. SBi-Anvisning 213. Danish Building Research Institute.
39. C.E. Kontokosta, C. Tull (2017). A data-driven predictive model of city-scale energy use in buildings. Applied Energy 197, pp. 303-317.
40. DIVA for Rhino plug-in. [Online] Available at: http://www.solemma.net/Diva.html (accessed September 5, 2017).
41. C. F. Reinhart, J. Wienold (2011). The daylighting dashboard. A simulation-based design analysis for daylit spaces. Building and Environment 46, pp. 386-396.
42. C. F. Reinhart, O. Walkenhorst (2001). Validation of dynamic RADIANCE-based daylight simulations for a test office with external blinds. Energy and Buildings 33, pp. 683-697.
43. G. Ward, R. Shakespeare (1998). Rendering with RADIANCE. The Art and Science of Lighting Visualization, Morgan Kaufmann Publishers.
44. Statistics Denmark. [Online] Available at: http://www.dst.dk/en/Statistik/emner/levevilkaar/boligforhold/boliger (accessed May 15, 2017).
45. Danish Energy Agency (2015). Regulation and planning of district heating in Denmark. [Online] Available at:https://ens.dk/sites/ens.dk/files/Globalcooperation/regulation_and_planning_of_district_heating_in_denmark.pdf
46. K. Wittchen, J. Kragh (2016). Danish building typologies and building stock analyses. Danish Building Research Institute, Aalborg University, SBi 2016:18.
47. European Union (2014). Inclusion of New Buildings in Residential Building Typologies. Steps Towards NZEBs Exemplified for Different European Countries. EPISCOPE Synthesis Report No. 1. [Online]. Available at: http://episcope.eu/fileadmin/episcope/public/docs/reports/EPISCOPE_SR1_NewBuildingsInTypologies.pdf [Accessed October 5, 2017).
48. S. Aggerholm (2013). Cost-optimal levels of minimum energy performance requirements in the Danish Building Regulations. SBi 2013:25. Danish Building Research Institute, Aalborg University.
49. DS418 : 2011. Calculation of heat loss from buildings. Danish Standard.

50. P.O. Fanger (1970). Thermal Comfort-Analysis and Applications in Environmental Engineering, in: C.D.T. Press. (Ed.).
51. ISO 7730 (2005). Ergonomics of the thermal environment - Analytical determination and interpretation of thermal comfort using calculation of the PMV and PPD indices and local thermal comfort criteria. International Standard.
52. P. Gianniou, A. Heller, P. Sieverts Nielsen, C. Rode (2016). Identification of Parameters Affecting the Variability of Energy Use in Residential Buildings. Proceedings of CLIMA Conference, Aalborg, Denmark.
53. Bygnings- og Boligregistret (BBR). Danish Building Register. [Online] Available at: https://ois.dk/ (accessed May 20, 2017).
54. K. B. Wittchen, J. Kragh (2012). Danish building typologies. Participation in the TABULA project, Hørsholm, Denmark: SBi, Danish Building Research Institute, Aalborg University.
55. The Danish Ministry of Housing, U. a. R. A., 2006. Boligejer.dk. [Online] Available at: http://boligejer.dk/ [Accessed July 1, 2016].
56. Styrelsen for Dataforsyning og Effektivisering. Kortforsyningen. [Online] Available at: https://kortforsyningen.dk/ (accessed July 15, 2017).
57. K.B. Wittchen, L. Mortensen, S.B. Holøs, N. F. Björk, S. Vares, T. Malmqvist (2012). Building typologies in the Nordic countries, Identification of potential energy saving measures. Danish Building Research Institute, SBi2012:04.
58. P. Gianniou, A. Heller, P.S. Nielsen, K. Negendahl, C. Rode (2015). Aggregation of building energy demands for city-scale models. In Proceedings of Building Simulation 2015, Hyderabad, India.
59. X. Liu, P.S. Nielsen, A. Heller, P. Gianniou (2017). SciCloud: A Scientific Cloud and Management Platform for Smart City Data. In 2017 28th International Workshop on Database and Expert Systems Applications (DEXA), pp. 27-31. IEEE. DOI: 10.1109/DEXA.2017.22

Paper III: Franjo Dominkovic, D., Gianniou, P., Münster, M., Heller, A., Rode, C. (2018). Utilizing thermal building mass for storage in district heating systems: combined building level simulations and system level optimization. *Energy 153*; 949-966.

Energy 153 (2018) 949–966

Contents lists available at ScienceDirect

Energy

journal homepage: www.elsevier.com/locate/energy

Utilizing thermal building mass for storage in district heating systems: Combined building level simulations and system level optimization

D.F. Dominković [a,*], P. Gianniou [b], M. Münster [c], A. Heller [d], C. Rode [b]

[a] Department of Energy Conversion and Storage, Technical University of Denmark (DTU), Frederiksborgvej 399, Roskilde, Denmark
[b] Department of Civil Engineering, Technical University of Denmark (DTU), Niels Koppels Allé, Kgs, Lyngby, Denmark
[c] Department of Management Engineering, Technical University of Denmark (DTU), Produktionstorvet, Kgs, Lyngby, Denmark
[d] Department of Civil Engineering, Technical University of Denmark (DTU), Brovej, Kgs, Lyngby, Denmark

ARTICLE INFO

Article history:
Received 1 December 2017
Received in revised form
3 March 2018
Accepted 16 April 2018
Available online 18 April 2018

Keywords:
Thermal energy storage in buildings
Thermal mass
District heating
Energy supply optimization
Thermal autonomy
Heat flexibility

ABSTRACT

Higher shares of intermittent renewable energy in energy systems have raised the issue of the need for different energy storage solutions. The utilization of existing thermal building mass for storage is a cost-efficient solution. In order to investigate its potential, a detailed building simulation model was coupled with a linear optimization model of the energy system. Different building archetypes were modelled in detail, and their potential preheating and subsequent heat supply cut-off periods were assessed. Energy system optimization focused on the impact of thermal mass for storage on the energy supply of district heating. Results showed that longer preheating time increased the possible duration of cut-off events. System optimization showed that the thermal mass for storage was used as intra-day storage. Flexible load accounted for 5.5%–7.7% of the total district heating demand. Furthermore, thermal mass for storage enabled more solar thermal heating energy to be effectively utilized in the system. One of the sensitivity analyses showed that the large-scale pit thermal energy storage and thermal mass for storage are complimentary. The cut-off duration potential, which did not compromise thermal comfort, was longer in the newer, better insulated buildings, reaching 6 h among different building archetypes.

1. Introduction

District heating systems produce heat centrally and distribute it to the end consumers via transmission and distribution pipes. Heat storage can be used when a mismatch between the timing of production and demand for heat occurs. Other solutions include peak boilers that can quickly be dispatched. When comparing heat storage and peak boilers, the former usually has large capital costs and low operating costs, while the latter usually have larger operational costs and lower capital costs.

All buildings, which are connected to district heating systems, have certain thermal capacities for storing heat inside the structure of the buildings. Contrary to the usual heating storage types such as hot water tanks or water pits, the capital costs for the utilization of thermal mass for storage is close to zero, as the building structure does not have to be modified additionally. Thus, utilizing thermal mass for storage could be an efficient solution for load shifting and/or peak shaving in district heating grids. The objective of this paper is to analyse the potential of utilizing thermal mass for storage in district heating systems in order to reduce the operational costs of the district heating systems by optimally shifting the district heat load. District heating systems are dominated by the peak demand during a few morning hours [1], which results in higher operational costs of the district heating systems.

One of the main findings from a recent review of the district heating and cooling systems has indicated that district energy systems are more efficient than individual heating and cooling systems, based on many projects reviewed across the world [2]. Thermal energy storage was one of the emphasized technologies that has a potential to further increase efficiency into current district energy systems [2]. It has been anticipated that district heating should play an important role in future renewable energy systems [3]. Moreover, authors have concluded that the future smart thermal grids will involve more energy efficient buildings, as well as integration with electricity and gas grids [3]. For the case of future energy system of Denmark, the share of 55–57% of district heating

* Corresponding author.
E-mail addresses: dodo@dtu.dk (D.F. Dominković), pagian@byg.dtu.dk (P. Gianniou), maem@dtu.dk (M. Münster), alfh@byg.dtu.dk (A. Heller), car@byg.dtu.dk (C. Rode).

https://doi.org/10.1016/j.energy.2018.04.093
0360-5442/

in the total heat demand could be cost-effective from the energy system point of view, although significant heat savings in the building sector have been anticipated [4]. Another recent international review of district heating and cooling systems claimed that those systems have strong potentials to be feasible supply options in a future world [5]. An author has reported that large national district heating research projects are being supported in Denmark, Germany, Sweden and China [5]. It was further concluded that heat recovery and heat based on renewable energy sources is larger in the European Union than in the rest of the world [5]. Nevertheless, the future energy system will need to balance out the potential energy savings in the building sector with the renewable energy supply in a cost-effective way. One study showed that the energy demand of buildings could be cost effectively reduced by 12–17% by the year 2015 [6]. The same study has shown that larger savings are to be expected in individual heating areas than in district heating areas.

Thermal energy storage has proven to be a technology that can be beneficial towards the energy efficiency of a building by contributing to an increased share of renewable energy and/or reduction in energy demand or peak loads for both heating and cooling [7]. Thermal building mass for storage could serve as a supplement to already existing storage solutions, such as hot water tanks. The reason for the latter is the low capital costs in thermal building mass storage type, as no physical alterations to the buildings are needed. Many different thermal storage options have been researched and some are already implemented on a large scale. Thermal storage can be realized in several different ways. They can be central (closer to the supply side of the system) or decentral (close to the consumer side of the system). Another division considers the thermodynamic nature of the way heat is stored, i.e. whether it is latent, sensible or thermochemical storage. Seasonal thermal energy storage has been reviewed in Ref. [8] and it was concluded that although it is a promising technology, its cost does not make it applicable to all projects, even less for single family houses. Furthermore, a review of promising candidates for chemical heat storage has been reviewed in Ref. [9], highlighting its significant potential due to the high thermal storage density, but also its low efficiency, special consideration of safety and large initial investment that is required. Thermal energy storages using phase change materials (PCMs) have been reviewed in Ref. [10]. One of the main aspects of PCMs is their low thermal conductivity (usually between 0.2 and 0.7 W/mK), thus requiring the use of complex heat exchanger geometries to obtain required heat transfer rates from latent heat storage containers. Regarding thermal storage building integrated systems, one study reviews it extensively [11]. The authors have concluded that active storage systems in the building envelope could be used when constructing new buildings. The integration of active thermal storage in buildings should be planned during a design phase in order to overcome the problems of availability of space for installations. In the same study, it has been claimed that both commercial and public buildings have huge potential on implementing thermal energy storage in double skin façade as well as in ventilation systems.

The utilization of short term heat storage in the sensible thermal mass of the buildings has been investigated in a number of studies during the last years. The zero investment cost that is required for the utilization of the thermal mass along with the capacity that is available in the majority of buildings in northern climates makes it a promising storage solution. One study of combined thermal energy storage and buildings has also dealt with a potential of using thermal mass of buildings for sensible heat storage [7]. It was concluded that the thermal energy storage can result in increased energy efficiency in buildings, reduced emissions, increased efficiency of HVAC equipment and reduced peak loads in system [7]. It was further argued that it is important always to fulfil specific demands and conditions that differ from building to building [7]. In a Danish study [12], two residential buildings with different states of insulation and air tightness were examined in terms of heat storage and heat conservation. The findings showed that the potential of the thermal mass depends on many factors (level of insulation, heat emission system etc.) and varies significantly over the season. The poorly insulated building could offer short thermal autonomy or heat flexibility meaning the time where the building can perform without activating a heating system, while the energy efficient passive house had a much higher time constant. This means that large amounts of heat could be shifted for shorter periods of time in poorly insulated buildings. On the contrary, a complete switch-off of the heating system could be achieved in the passive house for more than 24 h without violating the thermal comfort of the occupants.

Demand side management (DSM) can be defined as a modified consumer energy demand through various methods. Usually, the balancing of intermittent generation and load shifting from peak demand hours to off-peak demand hours are the most important targets of the DSM. A study by Ref. [13] investigated the potential of structural thermal mass of a single family dwelling for demand-side management (DSM) equipped with an air-to-water heat pump coupled with low temperature heat emission system, as well as a photovoltaic system in South-eastern Europe. The findings showed that the structural storage capacity has strong potential for shifting peak electricity loads for heating to off-peak hours. The DSM potential was found to be higher for massive buildings than for light-weight buildings.

Furthermore, the interaction between the heating system and the available thermal mass is significant. The authors in Ref. [14] have addressed that even after very short overheating periods, the heating demand for the following hours can be reduced significantly (up to 20%) utilizing the thermal storage capacity of the examined building, which included a hydraulic radiator-based heating system. The main limiting factors to the discharging rate were the slow temperature increase within the thermal mass and the heat conduction into the deeper wall layers. Moreover, the influence of the ambient temperature to the storage performance of the thermal mass has been highlighted. The authors conclude that good DSM can be achieved with shorter overheating periods at cold weather conditions. In addition, a Swedish pilot study [15] investigated the storage potential of the thermal inertia of five multi-family residential buildings connected with a district heating (DH) system. Results showed that heavy-weight buildings, with a structural core of concrete, can tolerate large variations in heat deliveries while still maintaining an acceptable indoor climate. Thus, the control can be applied in many buildings in DH systems at a relatively low cost. The study also demonstrated that degree hours instead of a fixed time constant can be a more accurate metric to represent variations in indoor temperature caused by the utilization of the thermal mass of the buildings. Although many examples of the simulated uses of thermal mass for storage have been reviewed on a building scale, there is a lack of cases calculating the potential of thermal mass for storage on a system scale.

A few papers dealt with the analysis of the thermal mass for storage potential on a system scale. Authors in Ref. [16] have presented a framework for planning cost effective operation of HVAC systems utilizing multi-building thermal mass. They have used a business-economic optimization approach and optimized thermal mass for storage of commercial buildings [16]. However, in their approach, thermal mass for storage was used to impact only the power sector while their time frame was one day. A simulation platform and different control strategies for utilizing the thermal mass for storage has been presented in Ref. [17]. The authors have concluded that single buildings only have marginal influence on the energy system flexibility and that extension of the models to the entire city is needed [17]. They have also assessed the potential of utilizing the thermal mass for storage to increase the flexibility of the power sector. A detailed dynamic and grey box model has been developed in Ref. [18]. They have shown that between 3% and 14% of the load can be shifted by utilizing the thermal mass for storage [18]. However, they have also considered the impact of flexibility solely on the power sector and not on the district heating sector [18]. Furthermore, the buildings they have modelled were well insulated and no behaviour of older buildings has been presented. The primary energy supply in the power sector was dominated by gas (40%), followed by nuclear (35%), waste and renewables (10%) and hydro (9%) [18].

To continue on the latter, one paper has proposed a resistance-capacitance representation of building thermodynamics in order to incorporate thermal mass for storage into an integrated planning model [19]. They have used it for an Irish case study and their focus was on partially decoupling of heat and electricity demand, without including district heating systems into the model [19]. The authors concluded that by utilizing building thermal inertia, electrified residential heat costs can be reduced to the cost of benchmark technology, which were gas boilers in the Irish case [19]. The energy supply in the power sector was dominated by wind (40%), coal, peat, gas, oil and hydro. Other authors used Balmorel model with the thermal building model add-on in order to represent thermal mass for storage [20]. The model has not taken transient behaviour of heat transfer into account, while the average annual building energy consumption was taken from the literature without detailed modelling [20]. They have concluded that by utilizing thermal mass for storage, peak load could be shaved and that heat pumps operation could be prioritized for hours with low marginal generation costs [20].

So far, most of the research papers dealt with detailed calculations of thermal mass for storage on a building scale for specific building archetypes. The majority of the papers that modelled thermal inertia on a system scale have not captured dynamics of the heat transfer and transient behaviour well enough during the hours after the DSM event using reduced-order models. Moreover, almost no literature has been found, which focuses on the impact of utilizing thermal mass for storage on district heating supply in general; the presented papers have rather focused on the potential benefits that the thermal mass for storage could have on the power sector. Thus, the aim of this paper is to analyse building performance of different archetypes during the DSM events and to analyse the impact of DSM events on district heating systems. Moreover, the main objectives of the paper is to give an overview of the energy flexibility potential that an urban residential building stock can give to the energy grid and to project this potential into the future. In order to have a sufficiently detailed model, a two level modelling approach is proposed. First, detailed simulation of existing building archetypes shall be carried out in order to estimate their thermal autonomy or heat flexibility potential, obtain detailed heat demand pattern after the DSM events and possible additional peaks in heat demand of different building archetypes. Second, simulation output data shall be used as input for linear optimization model that optimizes the whole energy system and analyses the potential of thermal mass for storage on a system scale. In this way, the realistic building energy performance is captured in great detail while the analysis of the impact of DSM events on the energy system on a district scale brings clarity about the total potential of smart control in district heating systems.

The outline of the paper is structured as follows: the building simulation and system optimization models are described in the Methods section. The city of Sønderborg was chosen for the case study and specifics of its energy system and the representative buildings are described in the Case study section. The potential of the thermal mass for storage is presented in the Results section. The results are put into perspective and compared with the findings from the other studies in the Discussion. Finally, the main points are summarized in the Conclusions section.

2. Methods

The following section presents the methods that were applied to model the investigated building stock, calculate the indicators and conduct the energy system optimization analysis. First, a detailed building energy model is presented. Second, the system indicators for evaluation of thermal mass for storage are introduced and different building heating scenarios presented. Third, an energy system optimization model is described. The energy system optimization model used the output of building simulations in terms of thermal autonomy, a difference in energy consumption compared to the reference case and the distribution of preheating and after cut-off heating demand as an input for running the optimization. Different steps used in the model are presented in Fig. 1.

2.1. Building energy modelling

The current analysis was conducted by use of building models. The building model and energy simulations were run in IDA ICE Version 4.7 [21]. The model behind the simulation tool is a detailed physical representation of the transient heat transfer phenomena taking place in a building. The model describing the external walls is a finite differences model of a multi-layer component. The buildings were simulated as single-zone models. The models were run with the Danish Design Reference Year (DRY) weather file [22]. These weather conditions are characterized by very cold winter temperatures. The monthly degree-days for the given climate and the annual outdoor temperature distribution according to [23] are presented in Appendix B. Based on literature, the utilization of thermal mass during times of very low ambient temperatures can have a good potential for DSM strategies. A cold and grey winter day was selected out of the DRY file to run the following experiment, during which the solar gains were very low and the average ambient temperature was -3 °C. Thus, the heat losses were increased for the specific examined case. The global radiation for an exemplary week in January is presented in Appendix B. However, the ambient temperature on the selected day did not fall below −12 °C, which is the dimensioning temperature for heating systems in Denmark. Deterministic occupancy profiles were modelled representing typical house living schedules, where there was no occupancy during 8am-3pm on weekdays. The heat supply system was DH and the heat emission system were hydronic radiators. No mechanical cooling or ventilation was assumed being installed in the buildings. The infiltration rate was determined based on wind-driven air flow. Internal walls and furniture mass were modelled taking into account the floor area that they covered, their material, thickness and the convective heat transfer coefficient.

The segmentation and characterization of the building stock was done according to [24]. Six building archetypes were created to represent different types of buildings according to their construction age and energy refurbishment level. These can be seen in Table 2. The properties of the building envelope and systems were created according to the TABULA [25], EPISCOPE project [26] and national building regulations. The building models were calibrated according to [27] based on measured hourly energy data acquired from 54 households in the city of Sønderborg.

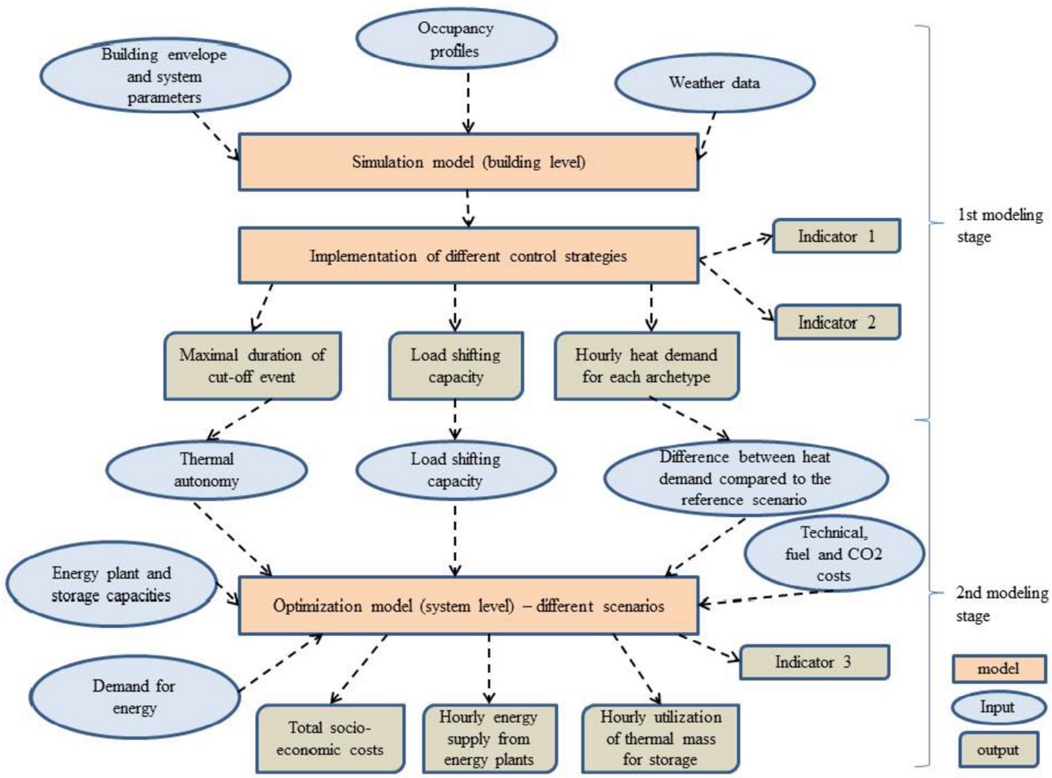


Fig. 1. The flow diagram of the two-stage modelling process.

2.2. Definition of thermal autonomy and adopted indicators

The thermal autonomy of the building was defined as an indicator of heat flexibility standing for the duration of the thermal comfort period [28], during which the operative temperature inside the building does not fall below 18 °C (1). This threshold was selected according to the Danish standard [29] corresponding to the third acceptable thermal comfort category, to reflect the degradation of indoor comfort in a straight-forward way. The thermal autonomy [h] was calculated on the basis of a heat supply cut-off experiment in the building models and was defined as the first indicator in our study.

$$Ind_1 = \min\{t|T(t) = 18^oC\} \tag{1}$$

The second indicator was defined as the difference in space heating demand of the building caused by the heating strategy implemented as opposed to the 'reference' heating demand (2). It was calculated on the basis of the thermal autonomy potential of each house. Indicator 2 [%] was calculated on the 24-h time period, starting from the hour 432 in the modelled year.

$$Ind_2 = \frac{Q - Q_{ref}}{Q_{ref}} \tag{2}$$

The exact time period was chosen as in none of the cases, the DSM event resulted in heat demand differences of more than 2% after the 24-h period.

2.3. Overview of different strategies of DSM events

To be able to determine the potential for thermal autonomy for all houses, heating control strategies were implemented. First, three different scenarios were created, which can be seen in Table 1. The first scenario included a complete cut-off of the heating system with no preheating or overheating strategy for a period that equals the thermal autonomy of each house, according to equation (1), so that operative temperature would not decrease below 18 °C. In order to make our results more robust, this strategy was implemented on the afore-mentioned cold weather conditions, so that an unfavourable case is represented. The heating set point of the reference operation of the heating system was set to 21 °C in all scenarios based on the findings from calibrating the building models.

Then, the effect of preheating on the utilization of the thermal mass was investigated. The second scenario considered 2 h of preheating the houses up to 24 °C, followed by a complete cut-off of the heat supply on the same cold day as the first scenario, during the same time. The duration of the cut-off was found such that it equalled the result of indicator 1 after the preheating was applied, so that the temperature would not decrease below 18 °C. According to the third scenario, 4 h of preheating strategy up to 24 °C was applied, followed by a heating cut-off on the same day, as previously explained. The duration of the heat cut-off was again adjusted so that it matched the new autonomy potential after the 4 h of preheating. The three different strategies were implemented in the IDA ICE models through the option of variable controller heating set

Table 1
DSM scenarios implemented to the models.

Scenario	Action	Preheating [h]	Notation
1	Heat cut-off	0	CO
2	Heat cut-off	2	CO_2hPH
3	Heat cut-off	4	CO_4hPH

points ranging from 21 °C in the reference case to 24 °C in the preheating phase and to 18 °C during the heat cut-off phase, which were continuous throughout the entire simulation year.

2.4. Energy system optimization

A linear continuous optimization model was used in order to represent an energy system. The optimization was run using Matlab interface and Gurobi solver, one of the solvers with the fastest computation times for solving linear optimization problems [30]. After the building simulations had been run, inputs from the building simulations such as autonomy time, the maximum capacity of avoided energy consumption during a cut-off event for each archetype, increased demand before and after the cut-off event and their corresponding distributions were incorporated in the optimization model. The simulated day from the first modelling stage was used as a pattern for all the cut-off events during the heating season (1st October to 30th April). A sensitivity analysis for the latter assumption was carried out and commented on in the Discussion section. Between 1st May and 30th September load shifting utilization was not allowed in the model as the space heating demand during that period was low. The optimization model includes the power, heat, gas and transport sectors. The optimization model used was developed, presented and validated in Ref. [31]. As a short overview, the model optimizes the energy system with an objective function set to minimize the total socio-economic costs of the energy system (3). It optimizes the whole energy system during one year on hourly resolution.

$$\begin{aligned}\min Z = &\sum_{i=1}^{n}(fix\_O\&M_i + lev_{inv\ i})x_i + \sum_{j=1}^{m}\left(var\_O\&M_j + \frac{fuel_j}{\eta_j}\right.\\ &\left. + CO2_j \cdot CO2_{inten}\right)x_j + \sum_{k=1}^{p}(el\_imp\_exp_k + gas\_imp\_exp_k\\ &+ dies\_imp_k + petr\_imp_k)x_k\end{aligned} \tag{3}$$

Energy plants were modelled using a black-box approach. The total socio-economic costs of the system incorporated levelized investment costs over the lifetime of the energy plants, fixed and variable operating and maintenance costs, fuel costs and costs of CO_2 emissions. Other taxes were not included. A discussion about the reasons for including the cost of CO_2 emissions and excluding the cost of other taxes in the total socio-economic costs can be found in Ref. [32]. The main reasoning for the latter is that the CO_2 emissions cost represents the internalized negative externality in terms of climate change costs imposed to the society. A detailed explanation of all the terms used in (3), as well as corresponding units, can be found in the Nomenclature, located at the end of this paper. As the goal of the optimization is to minimize the total socio-economic costs, different storage solutions will only be utilized if the corresponding load shifting can bring savings in the overall energy system costs. Based on the decision of the modeller, the model can optimize investments, too.

Furthermore, the model included all the district heating energy plants supplying the district heating network of the city, electricity plants (wind and PVs) erected as a part of the energy transition plan of the city and transportation demand for fuels. Electricity could be exported/imported over the system boundary (constrained by the transmission capacity), as well as fossil fuels for meeting the transport demand. Large (sensible) heat storage is another flexibility option that existed within the system boundary. Finally, gas CHP plant did not have a constraint on the minimum load level, which made it available to behave as a flexible generation technology. The building thermal mass for storage, which is the focus of this paper, was the last source of flexibility in the local energy system.

Congestion of the district heating grid and electricity grids was not modelled. Moreover, the model does not include dynamic modelling of the district heating grid due to the complexity of the size of the current model.

For this paper, the model was significantly expanded in order to include the potential of thermal mass for storage of different building archetypes. For each modelled building archetype, a set of variables for cut-off event, increased after cut-off heating demand and increased preheating were introduced. In order to meet the increased heat demand after the cut-off event, as well as before the cut-off event when the heating strategy included preheating, inequality constraints (4) and (5) were introduced.

$$\begin{aligned}&cutoff_{t,x} \cdot C_{x,y} \le preheating_{t,x,y}\\ &\forall t \in \{1,\dots,8760\}, \forall x \in \{1,\dots,X\}, \forall y \in \{1,\dots,Y\}\end{aligned} \tag{4}$$

Where $cutoff_{t,x}$ represents the avoided heat demand due to the cut-off event [MWh] for each building archetype x in every hour t of the year. $C_{x,y}$ represents coefficients of hourly difference in district heat demand in each preheating hour y (compared to the reference heating demand pattern), obtained as a part of calculation for indicator 2 of the first step of the model. Finally, $preheating_{t,x,y}$ represents the increased heat demand [MWh] prior to the anticipated DSM event.

$$\begin{aligned}&cutoff_{t,x} \cdot C_{x,z} \le afterheating_{t,x,z}\\ &\forall t \in \{1,\dots,8760\}, \forall x \in \{1,\dots,X\}, \forall z \in \{1,\dots,Z\}\end{aligned} \tag{5}$$

Where $afterheating_{t,x,z}$ represents increased heat demand after the cut-off event [MWh] in each after cut-off heating demand hour z.

The $C_{x,z}$ and $C_{x,y}$ coefficients were used to realistically capture the dynamics of the heating demand in hours before and after the cut-off events took place.

The sum of preheating hours y, the duration of cut-off event (equal to the value of Indicator 1) and after cut-off heating demand hours z were set to 24 h in order to match the indicator 2 from the first part of the model (6).

$$y + Ind_1 + z = 24 \tag{6}$$

In order to further integrate the first and the second part of the model, only one *maximum capacity* cut-off event was allowed in any 24-h period (7). The latter constraint was introduced in order to acknowledge the calculation method used for the estimation of indicator 2. However, the system could choose to have more than one cut-off event during any 24-h period, as long as the sum of the avoided heat demands during all the cut-off events did not exceed the maximum capacity of a single cut-off event.

$$cutoff_{t,x} + cutoff_{(t+1),x} + \dots + cutoff_{(t+23),x} \le cutoff_{max,x} \tag{7}$$

Where $cutoff_{max,x}$ represents maximum avoided heat demand during the cut-off event for each building archetype x, calculated in the first part of the model. Inequality constraint (7) allows more than one cut-off event during the 24-h period as long as the sum of them does not exceed maximum possible cut-off demand.

Finally, after the optimization part of the model had been run, the economic indicator (Indicator 3) was calculated (8).

$$Ind_3 = \frac{C_{operational,ref} - C_{operational}}{C_{operational,ref}} \tag{8}$$

Indicator 3 represents the difference in the operational costs of the district heating system, without the costs of the other parts of the energy system. Operational costs of the district heating system did not include investment costs of the energy plants. They were calculated using the equation (3) for the plants operating in the district heating network, without including the first and the third summation terms. The economic indicator encompassed operational costs of the district heating system only, as the district heating sector was the focus of research carried out in this paper. For the purpose of calculating the operational costs of DH systems, the income from cogeneration units selling electricity to the grid was included as revenue in the operational costs of the district heating systems, reducing the overall operational costs of the district heating system, while in the same time all the fuel costs of running the cogeneration units were included as expenditure.

3. Case study

The city of Sønderborg was chosen as the current case study. It is located in the south of Denmark with a population of 27,500 inhabitants as reported in 2011. It includes several types of energy supply plants and the whole municipality has started a transition towards net zero carbon until 2029. The highest share of heating demand is attributed to residential demand, accounting for 69% of the total heating demand (Fig. 2). In Denmark, housing accounts for 64% of the total heat demand. As Sønderborg can be considered representative in terms of heat demand, and ambitious in terms of integrating variable renewable energy sources, it was decided to focus on the residential building stock of Sønderborg to investigate the potential for energy flexibility provided by the thermal mass included in the building envelope and internal walls. Around 53% of the area's heat demand is covered by the local district heating network [33].

3.1. Characterization of Sønderborg' building stock

The focus of the current study is on single-family houses (SFH), which represent the largest share of residential buildings in this area. The building stock of Sønderborg was represented by archetypes as afore-mentioned. The categorization follows the general guidelines of the TABULA database. According to the TABULA project, in general there are ten typical Danish building archetypes corresponding to SFH [25]. One archetype belongs to each of the ten proposed age bands, which reflect a shift in building tradition and the introduction of building energy codes. The same age bands were used in our case to extrapolate the results of simulated building archetypes to the system level. Based on the available energy data that we had and used to calibrate the building models, six building archetypes were created representing the majority of the SFH in Sønderborg. The models of the archetypes are presented in Table 2. The construction age of all the building models, their gross floor areas along with the average heat loss coefficient (U-value) of the total building envelope and internal heat capacities are presented. The time constant of the buildings is also presented in Table 2. It is calculated as the ratio of the total heat capacity of the building and the total heat loss coefficient including transmission, ventilation and infiltration losses. The time constant gives an indication of the response pace of the building to different stimulations such as change in heating or outdoor temperature, hence being very relevant to the current study. It is evident that newer buildings have longer time constants. The total energy use intensity and peak energy demand over a full simulated year, which are also presented in Table 2, give an indication of the energy efficiency of the investigated building stock. The energy demand results were calculated after the models were calibrated according to the measured energy data and with the system properties and internal conditions that were described for the reference case. There are two archetypes corresponding to SFH built in the 1960's, since one of the two (1960's ref.) had undergone more extensive energy refurbishments based on the calibration findings. The U-values give an indication of the insulation state of the building envelope relating to how airtight each building is. The lower the U-value, the more airtight the building is and fewer thermal bridges it has. The refurbishment implemented in each building model/archetype -also affecting the heat loss coefficient of the building envelope- was decided based on the results of the calibration with measured energy data as mentioned earlier. The majority of the houses were made of heavy-weight insulated brick walls with an average thickness of 340 mm. The internal walls in the majority of the investigated houses consisted of aerated concrete of 75 mm thickness and were designed based on typical floor plans of Danish SFH. The idea of using archetypes to characterize the building stock is that one archetype represents a building category having uniform characteristics in terms of building construction and systems that are regulated by building codes. Thus, the energy performance of buildings belonging in the same category will be quite similar making one archetype representing each category sufficient. More information about this approach can be found in Ref. [24]. The six building archetypes that were created represented 60% of the residential building stock and 55% of the district heating demand in Sønderborg. Most of the building stock not represented in Table 2 (26.4% of the building stock and 36.9% of district heating demand), were buildings built before 1930. The latter building stock has similar characteristics to the first archetype in terms of energy use intensity, dominated by low time constants.

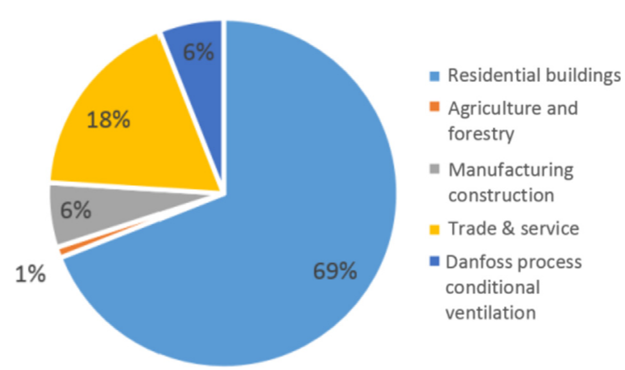


Fig. 2. Allocation of heat demand on the building categories in Sønderborg 2007 [29].

3.2. Overview of Sønderborg district heating system

In 2015, the total district heat supplied to the customers in the city was 288.95 GWh (not including 23% of distribution losses), based on information provided by the Sønderborg Fjernvarme, which is the operator of the district heating system. The Sønderborg DH operation data, such as volume flows, supply and return temperatures, as well as the circulation pumps electricity demand, can be found in great detail in Appendix B. According to the energy transition plan towards 2029, an increased share of households connected to the district heating systems has been anticipated [33]. The current district heating plants that supply the city of Sønderborg with heat are listed in Table 3.

All household owners in Denmark are obliged to report certain information about their real estate, including the year of construction and last refurbishment of the building, energy consumption, heat energy source, type of the heating system, etc. in

Table 2
Building archetypes representing the examined Danish residential building stock.

Building archetype	Construction age	Floor area [m^2]	Average U-value of total building envelope [W/m^2K]	Internal heat capacity e+07 [J/K]	Time constant [h]	Energy use intensity [kWh/(m^2year)]	Peak demand [kW]	Share of buildings represented [m^2/m^2] [%]
	1930's	85	0.72	1.44	20	225.7	7.5	5.9
	1950's	87	0.47	2.34	43	148	4.4	6.8
	1960's	140	0.54	2.33	29	143.6	7.3	7.8
	1960's ref.	119	0.43	2.13	39	111.6	5.1	11.8
	1970's	136	0.51	3.69	51	132.6	6.8	10.3
	1990's	137	0.31	5.87	134	84.5	3.8	17.5

*Pictures taken from Refs. [34–36].

Table 3
The current district heat supply plants [31].

Energy plant type	Electrical power [MW_e]	Heat generation capacity [MW_{th}]
Waste CHP	4.5	20
Gas CHP	40	53
Gas boilers		100
Solar heating (centralized)		5.2
Bio-oil		5.4
Geothermal + biomass driven absorption heat pump		12.5

the so called BBR Database [37]. For the purpose of this paper, we have obtained the BBR data for the city of Sønderborg through the ProjectZero company, the managing company of the net zero carbon energy transition of Sønderborg. Data has been cleansed removing unreasonable values. The high uncertainty in some of these variables used as inputs to the building models has been balanced by the calibration process as earlier mentioned, according to which measured energy data were used to calibrate the inputs of the building models so that the predicted and actual energy demand match.

The energy system of Sønderborg should undergo significant changes until the year 2029 and different renewable energy plants are anticipated [33,38]. The whole municipality of Sønderborg will try to achieve carbon neutrality by 2029. Important contribution towards this goal should be made by the district heating sector. Thus, in order to assess the effects of the potential utilization of the thermal mass for storage in the future energy system, an additional scenario (the 4th one, designated CO_2hPH_2029) was modelled. An overview of the plants, which should be installed until the year 2029 can be seen in Table 4. Only zero-carbon energy technologies are anticipated for the year 2029. Furthermore, it was complicated to extract the capacity of solar thermal anticipated in the DH network of the City of Sønderborg, as the transition plans are made for the whole municipality of Sønderborg. In order to cope with the latter, it was assumed that the maximum solar thermal generation would correspond to the 15% of the yearly DH demand in 2029. Effective solar thermal generation supplied to the district heating could be lower if there would not be enough demand during the peak generation hours and there would be a lack of heat storage capacity in the system.

It should be noted from Table 4 that the capacity of biomass boilers in the year 2029 was not fixed before the optimization was run. Opposite to the current portfolio of the DH supply plants, the future investments in the plants can be projected taking into account different technologies or flexibility options, such as utilizing the thermal mass for storage. In order to account for the potential savings of the capital costs, by reducing the installed capacity in the DH system, biomass boilers as a peak technology was also optimized in the CO_2hPH_2029 scenario. Thus, the resulting capacity of biomass boilers was the minimum capacity still being able to satisfy the DH demand in all hours during the year.

Table 5
Average wholesale electricity prices (based on Nordpool el-spot [41]) and CO2 emission prices used in the scenarios.

	2015	2029[a]
Average electricity price (€/MWh)	22.89 [41]	59.64 and [42]
CO2 emission price (€/t)	7.6 [42]	13.6 [42]

[a] Real future prices, i.e. the future prices were adjusted for inflation (2016 prices).

Furthermore, it is expected that the increased share of households will connect to the district heating grid. The DH demand should increase by 13% compared to the 2015 share, taking into account significant energy savings, which are anticipated in the building sector [33]. Furthermore, energy retrofit of buildings will increase the share of the buildings with longer thermal autonomy, i.e. more buildings will have better air-sealed and thermally insulated envelope. Historically, the energy retrofit rate of the buildings in Denmark has been 1% per year [39,40]. We adopted the same rate until the year 2029. Moreover, we assumed that the retrofitted buildings behave in the same way as the 1990's archetype, as those were the newest modelled buildings. Finally, for the 2029 scenario, the second strategy (with two hours of the preheating time) was used. The latter also means that the indicators 1 and 2 are equal in CO_2hPH and CO_2hPH_2029 scenarios. CO_2 emission costs, as well as the average electricity prices can be seen in Table 5.

One should also note that the assumed volatility of the prices was greater in the year 2029, as more intermittent renewable energy sources was assumed to be installed in the power grid. The spread between the highest and the lowest electricity price was 192.3 EUR/MWh in 2015, while the assumed spread was 302.6 €/MWh in 2029.

4. Results

4.1. Results of the building energy modelling

As mentioned in the Methodology, one representative winter day had to be identified and selected as the date when the experiments in the simulation part of the model would be carried out. These results were then reproduced according to the optimization results within the heating season of a year in the second stage of the model, making the system level results of the optimization part of the model robust. So, the heat cut-off was applied on January 19, which was found to be one average cold day of the reference year with the mean daily ambient temperature being -3 °C and very low solar gains in order to isolate their effect on heat flexibility.

Table 4
Energy plants capacities anticipated in 2029.

	Installed capacity 2015 (MW_{th})	Installed capacity 2029 (MW_{th})
Gas boilers	100	0
Gas CHP	53	0
Waste CHP	20	20
Geothermal + biomass driven absorption heat pump	12.5	12.5
Biomass boilers (including bio-oil)	5.4	[a]
Large scale heat pumps	0	12 (heating capacity)
Solar heating	5.2	40
Heat storage	4000 MWh	4000 MWh
Wind turbines	14.6	180
Photovoltaics	14.8	60

[a] Optimized by the model (a minimum capacity needed to satisfy the heating demand in all hours throughout the year).

Table 6
Results of indicator 1 (thermal autonomy) and 2 (difference in heating demand) 2.

Archetype	SCENARIO 1 (CO)		SCENARIO 2 (CO_2hPH)		SCENARIO 3 (CO_4hPH)	
	Ind_1 [h]	Ind_2 [%]	Ind_1 [h]	Ind_2 [%]	Ind_1 [h]	Ind_2 [%]
1930's	1	−0.3%	1	1.3%	2	4.2%
1950's	2	−1.9%	3	−1.5%	4	−0.7%
1960's	1	−1%	2	−2.1%	2	−1.4%
1960's ref.	2	−1.5%	3	−2.0%	4	−1.5%
1970's	1	−0.3%	2	−1.2%	2	0.4%
1990's	5	−14.6%	6	−18.1%	6	−15.2%

The effect of the preheating strategies on the energy use for heating can be seen in Table 6. These results should be interpreted with regard to the information presented inTable 2. The buildings with the lowest average U-values have the highest thermal autonomy potential (Ind_1), indicating that the effect of heat losses in the examined buildings is dominating. It should be also noted that even though the 1970's archetype has a longer time constant than the 1960's refurbished archetype, its better insulated envelope (thus lower U-value) results in a longer thermal autonomy potential. This validates the previous finding that heat losses have a more significant effect than the heat capacity of the building or its thermal mass for the specific investigated buildings under the cold and grey weather conditions. It should be noted that the internal wall mass for all archetypes was modelled in a very similar way, having similar construction characteristics (materials and wall thickness), as well as the share of the volume of internal walls to the floor area was assumed to be almost the same for all the different models. It is also evident that as the duration of preheating increases, the results of thermal autonomy (Ind1) increase in most archetypes. The effect of preheating is not pronounced on the newest archetype, which represents SFH built in the 1990's. This is attributed mainly to the fact that the building has airtight envelope leading to low heat losses, so its thermal autonomy potential is already high (6 h) and is slightly improved by preheating, which is not evident due to rounding. According to [28], the effect of transmission losses exceeds the effect of thermal mass when it comes to thermal autonomy potential of a building and the peak load that is created after overheating strategies. Since the weather conditions on the day of the experiment were cold, the effect of the transmission losses on the specific indicators is outlined. Therefore, it is confirmed that buildings with airtight building envelopes and very low overall heat transfer coefficients would perform better in general with regards to heating demand. Looking at indicator 2, we observe that the cut-off of heating led to savings in space heating demand in all archetypes compared to the reference case when heating was always on, as expected. The heat losses were decreased due to the lower average internal temperature during the day of the cut-off. The highest decrease was observed for the newest archetype, which allowed the longest duration of cut-off. Indicator 2 could be positive or negative for the two preheating scenarios, as the duration of the preheating and of the subsequent cut-off may or may not lead to savings of space heating energy use. In the majority of the houses, even in the case of four hours of preheating up to 24 °C, the heating demand was still lower than the reference case except for the 1930's archetype and the 1970's archetype. Thus, it can be concluded that longer preheating times do not have the same effect in all buildings.

Operative temperatures and heat consumption of the 1990's archetype and the 1960's archetype during the 24-h period around the simulated cut-off event can be seen in Fig. 3 and Fig. 4, respectively. The latter represents the archetype with short time constant and a low thermal autonomy (1–2 h in different scenarios), while the former presents the archetype with the longest time constant and longer thermal autonomy (5–6 h). For the implemented DSM strategies, we can see that operative temperature inside the 1990's and 1960's archetypes varies between 18 °C and 22 °C. Furthermore, the preheating scenarios CO_2hPH and CO_4hPH lead to higher initial internal temperatures for both archetypes. The thermal comfort remains to be acceptable in all cases according to [29]. When comparing the reference case and the cases with the occurring DSM events, one can note new peaks in heat demand. These peaks occur both during preheating and after heat cut-off phase. We can observe that the effect of preheating strategies is minimized on the archetype of 1990's leading to similar thermal autonomy results. The magnitude of the peak load that is created after the cut-off is similar for all three scenarios and it is defined by the capacity of the heating system. For the same archetype, even though the preheating set point is set to 24 °C, this cannot be achieved within the given time, with the maximum operative temperature inside the house being 21.8 °C for the CO_4hPH scenario. Despite the 4-h preheating, the thermal autonomy is not much prolonged as mentioned earlier due to the long-time constant and the low heat loss coefficient of the building envelope of the 1990's archetype. We can see that the operative temperature stabilizes in the three scenarios at the end of the day reaching almost the same value of the reference case where no cut-off was applied. Looking at the heat power graph of Fig. 3, it can be observed that the peak load that is created after the heat cut-off is almost equal for the three scenarios, independent of the preheating strategy as already mentioned. Similarly, 1960's archetype has a slightly increased thermal autonomy after preheating scenarios CO_2hPH and CO_4hPH as seen in Fig. 4. The magnitude of the peak load seems again to be independent of the preheating strategy and is very much defined by the maximum capacity of the heating system that is almost fully utilized in these cases of a sudden internal temperature drop. It should be noted that the heating system

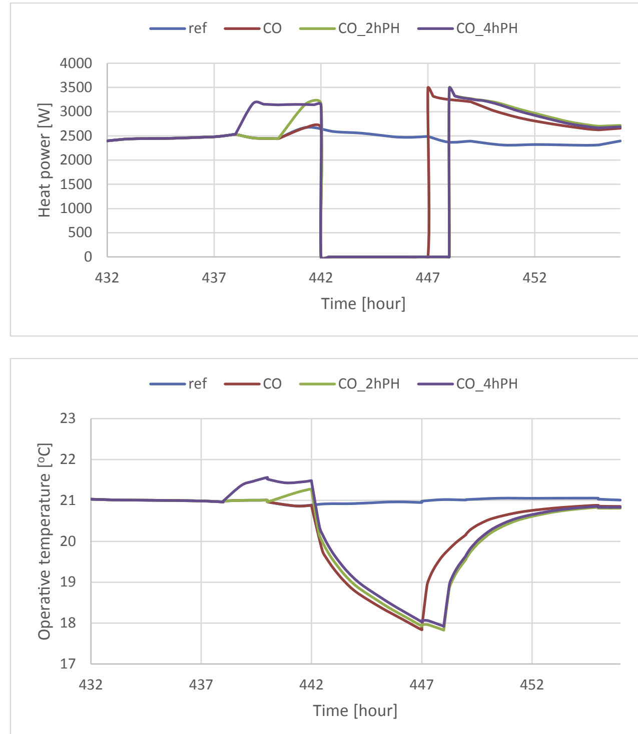


Fig. 3. Operative temperatures (up) and heat demand (down) of the 1990's archetype in different scenarios.

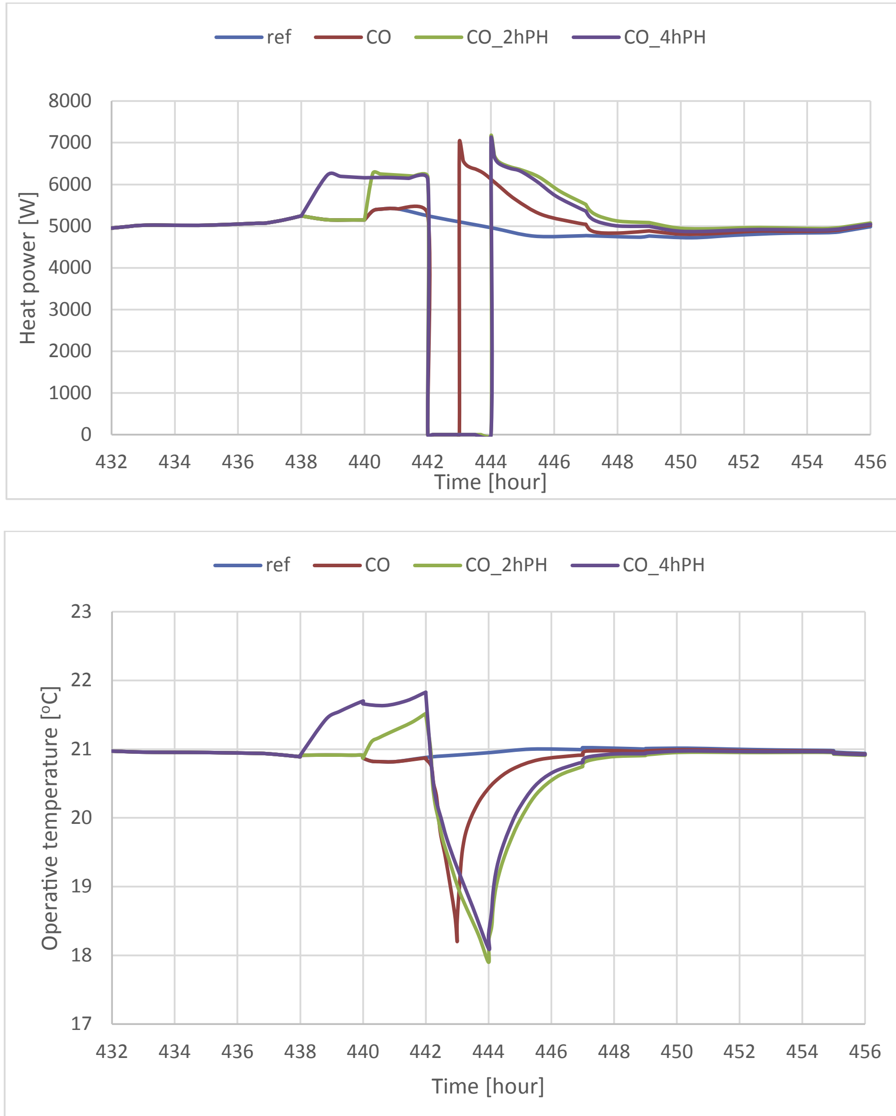


Fig. 4. Operative temperatures (up) and heat demand (down) of the 1960's archetype in different scenarios.

is dimensioned in the construction phase of the building. So, when the building undergoes energy refurbishments, the heat demand may be reduced but the capacity of the heating system remains the same as before. That is the reason that the older archetype has higher peak loads after the cut-off compared to the newer archetype. The operative temperature of the 1960's archetype in the CO scenario does not seem to reach 18 °C due to the setup of the controller that was set to take average values of very short time steps to eliminate temperature changes within a few minutes.

4.2. Results of the energy system optimization

The total shifted heat demand during the year, in terms of avoided heat demand during the cut-off events, for the 3 different strategies for the year 2015 and strategy two for the year 2029 (scenario CO_2hPH_2029) can be seen in Fig. 5.

In all the scenarios, the newer archetypes contributed more to the overall load shifting than the older ones. The oldest houses, represented by the 1930's archetype were utilized only in the CO scenario. All other archetypes contributed to the load shifting in all the scenarios. Archetypes 1960's ref, 1970's and 1980's accounted for 65%–70% of the total shifted load, in the first three scenarios. Moreover, the relatively old building archetype 1950's was often utilized, especially in scenarios CO_2hPH and CO_4hPH when it had larger autonomy (3 and 4 h compared to the 2 h in the CO scenario). It was found feasible to utilize almost all of the identified load shifting potential. The maximum possible capacity of the load shifting was 5.6%–8.4% of the total DH demand of the city of Sønderborg in different scenarios. The real activation of the thermal mass for storage in different scenarios showed that the shifted load accounted for 5.5%–7.7% of the total DH demand in the different scenarios.

Differences in operational costs of the district heating system, represented by Indicator 3, can be seen in Table 7.

The largest saving potential occurred in the 2029 scenario. The energy system of the 2029 is anticipated to have more capital intensive, but efficient technologies, such as large scale heat pumps and centralized solar thermal systems (Table 4). The main reason for much larger operational savings in 2029 was the larger capacity of solar thermal technology that was less curtailed when utilizing

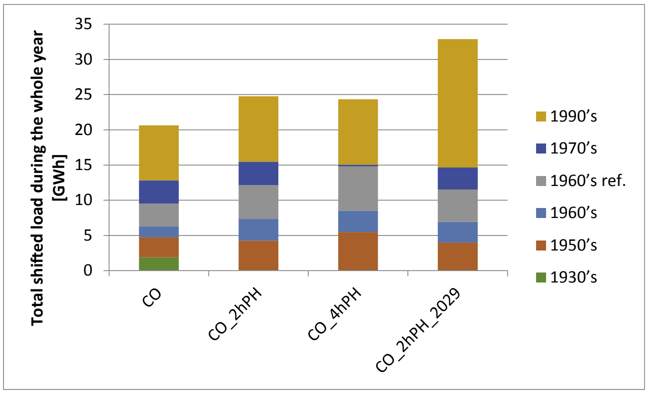


Fig. 5. The total amount of shifted heat demand on the system scale in different scenarios for one whole heating season.

the thermal mass for storage. Moreover, more refurbished buildings meant that the larger share of buildings behaved as the 1990's building archetype, having larger time constants and lower energy consumption (Indicator 2).

The resulting capacities of the peak biomass boilers for the year 2029 were 52.2 MW in the reference (2029) case and 37.7 MW in the CO_2hPH_2029 scenario. The latter means that taking into account the thermal mass for storage and corresponding reduction in heat demand (indicator 2) when planning future investments, reduced the needed peak capacity. The reduced peak boiler capacity of 14.5 MW represents 14.6% of the peak DH demand of the city of Sønderborg throughout the year. Although it was not the focus of our study, reduced capacity of the installed plants reduced the total capital costs of the whole energy system by 2% in 2029, equal to 770,000 EUR.

Comparing the CO_2hPH_2029 scenario and the reference scenario for the year 2029, the optimization results showed that large scale solar-thermal was more utilized. Gas boilers and large scale heat pumps had significantly lower heat generation. Although the capacity of solar thermal DH was larger in the year 2029 than in the 2015, the overall capacity of it was still relatively low. However, the results showed that it could be possible that in systems with larger shares of solar thermal DH, the utilization of thermal mass for storage would be even more attractive economically.

5. Discussion

One of the main goals of this paper was to focus on the impact of utilizing building thermal mass for storage on district heating supply and its operating expenses. Our analysis showed that operational savings can be much larger in the future, with larger capacities of the intermittent sources installed in the DH system, larger average electricity prices and higher electricity price fluctuations in the energy system. Load shifting in relative terms was mostly utilized in the CO scenario, during the 98% of the possible time. Although in the CO_2hPH_2029 load shifting possibility was utilized the most in the absolute numbers (Fig. 5), in relative terms its utilization was lower than in the CO scenario, i.e. it utilized 96% of the total thermal mass for storage capacity throughout the year. DSM events were usually triggered in mornings, in the duration of 2–3 h, and often in the evenings, especially during the winter time with higher overall DH demand. Heating savings related to the activation of thermal mass for storage were assessed using the indicator 2. Only two archetypes did not have heating savings in CO_2hPH, CO_4hPH and CO_2hPH_2029 scenarios. All other building archetypes had heating savings, resulting in a lower overall DH demand. One should note that the latter means that certain share of load shifting occurred due to the energy savings itself. In order to check the share of savings that comes from reduced average temperature when activating the thermal mass for storage, simple sensitivity analyses were carried out for CO_2hPH and CO_2hPH_2029 scenarios. To carry out the sensitivity analyses, the reference heat demand was reduced in order to be equal to the total yearly heat demand in the CO_2hPH and CO_2hPH_2029 scenarios. The sensitivity analyses showed that the operational economic savings in the CO_2hPH scenario reduced from 1.0% to 0.17%. In the CO_2hPH_2029 scenario, the operational economic savings reduced from 4.6% to 3.1%. However, one should note that the actual implementation of the proposed preheating scenarios and heat supply cut-off to the residential building stock requires a central management system, the same one that allows the activation of thermal mass for storage, which would impose a non-negligible capital cost to the system.

All scenarios showed that despite most of the energy plants reduced their generation (Table 8) solar DH increased its useful output in all the scenarios. The latter could be especially observed in the carbon neutral scenario carried out for the year 2029, when 3.8 GWh more solar thermal heat generation was effectively

Table 7
Comparison of the operational costs of district heating in the city of Sønderborg in different scenarios (indicator 3).

	Reference (2015)[a]	CO	CO_2hPH	CO_4hPH	Reference (2029)[a]	CO_2hPH_2029
Operating costs of DH system [10^3€]	5238	5201	5186	5164	5687	5425
Savings compared to the reference case [%] (Indicator 3)		0.7%	1.0%	1.4%		4.6%

[a] Reference case costs were obtained by constraining the modelled system not to utilize the thermal mass for storage.

Table 8
Difference in generation of district heating plants in the city of Sønderborg compared to the reference case (the first three scenarios compared with the reference case for the year 2015, the last scenario compared with the reference case for the year 2029).

	CO [MWh]	CO_2hPH [MWh]	CO_4hPH [MWh]	CO_2hPH_2029 [MWh]
Solar heating	111	235	247	3762
Geothermal + biomass driven absorption heat pump	−1	−44	−52	−734
Biomass boilers	−45	−74	−68	−6755
Waste CHP (heat generation)	0	0	0	30
Gas CHP (heat generation)	98	−5	−36	0
Gas boilers	−1361	−1852	−1460	0
Heat pumps	0	0	0	169

*The full results of the generation of different energy plants in the reference cases can be seen in Appendix A.

utilized compared to the reference scenario for the year 2029. Moreover, the thermal autonomy was more important factor for triggering DSM events from the system point of view than the energy use intensity of different building archetypes.

There were four main flexibility sources in the modelled energy system: flexible generation of gas driven plants, import/export of electricity over the system boundaries, heat (pit thermal energy) storage and the building thermal mass for storage. In order to check the operation and mutual influence of pit thermal energy storage and the activation of the thermal mass for storage, another sensitivity analysis was carried out for the CO_2hPH_2029 scenario. In the sensitivity analysis, except the biomass boilers capacity which was not constrained, the capacity of pit thermal energy storage was not constrained either. The analysis showed that the optimal capacity of pit thermal energy storage, that minimizes the total socio-economic costs of the energy system, was 158,000 m^3, behaving as a seasonal storage. The significantly increased capacity of the thermal storage further reduced the need for additional 10 MW of biomass boilers due to the increased utilization of both solar thermal and the thermal mass for storage. The total load shifted by activating the thermal mass for storage in the sensitivity analysis was 34.1 GWh, compared to the 32.9 GWh in the CO_2hPH_2029 scenario. The latter shows that the two storage flexibility options are mutually complementary. The thermal mass for storage behaved as intra-day storage, shaving daily peaks in demand, while the pit thermal energy storage behaved more as seasonal storage, shifting lots of excess solar thermal generation during the summer time to the winter time.

In the first three scenarios, the majority of savings came from less utilization of gas boilers. In the CO_2hPH_2029 scenario, the majority of energy savings came from less utilization of biomass boilers. The latter also shows that implementation of thermal mass for storage could lead to savings in consumed biomass, allowing larger amounts of sustainable biomass to be utilized for the transition of the heavy-weight transport sector. It was shown that the transition of the heavy-weight part of the transport sector to the renewable one will be especially energy demanding [43].

Compared to the other studies that assessed the potential of thermal mass for storage and were presented in the literature review, our study simulated real existing buildings from our case study. The simulations were detailed, taking transient behaviour into account as opposed to steady-state or quasi-steady-state models. The latter has allowed us to capture the peaks in heat demand just before (when preheating was applied) and after the cut-off events, a finding that would be hard to obtain using less detailed models. Moreover, our study focused on the DH supply system in a holistic way; not just on its electrified part or on integration of intermittent electricity sources by utilizing the thermal mass for storage. On the other hand, because of the coupling of detailed simulation model with the holistic energy supply model, we had to use a two-level approach and not the integrated model. Furthermore, the DSM strategies were implemented on cold weather conditions with very low solar gains and mean daily temperature of -3 °C, which represented a cold and grey winter day, which is quite common in Danish heating season so that our system level results become more robust. Two other cases were also modelled to estimate the flexibility potential on a more favourable case and on a least favourable one with regards to the thermal autonomy potential of buildings. These were i) an equally cold but entirely sunny day, with an average ambient temperature of -3 °C and ii) an equally grey day but colder with an average ambient temperature of -6 °C. The experiment was run for the newest archetype representing SFH built in the 1990's and showed that the effect of solar gains is significant. In particular, for the clear day the thermal autonomy potential of the newest archetype increased by 7 h in the cut-off and preheating scenarios due to the highly increased solar gains that coincided with the time of the cut-off. When a colder grey day was investigated, the thermal autonomy potential of the same archetype decreased by 2 h due to the increased thermal losses. Thus, the flexibility results are subject to significant changes based on the weather conditions. However, it should be noted that the extremely sunny day that was examined represents only 9% of the days in a typical Danish heating season (October–April), while the share of very cold days, when the average daily temperature falls below -3 °C, is 10%. In addition, the examined 1990's archetype in this sensitivity analysis was characterized by a very long time constant and low overall heat loss coefficient, thus benefiting significantly from high solar gains. These results would be different for the older houses. Therefore, it is assumed that our initial choice of day to run the experiments represents the building performance during a Danish heating season quite well. The effect of warmer days on heat savings [kWh] could be further investigated. Overall, it is expected that much warmer days would lead to lower heat savings due to decreased heating demand. . Testing extremely cold ambient temperatures falling below −12 °C was beyond the scope of this analysis, as this is the dimensioning temperature of heating systems according to the Danish standards, below which the heating systems would not be able to operate sufficiently to achieve acceptable thermal comfort.

Out of other studies, one study showed that between 3% and 14% of the load can be shifted by utilizing the thermal mass for storage [18]. Our findings showed that the theoretical potential for load shifting is between 5.6% and 8.4% and the economic potential between 5.5% and 7.7%, according to the different scenarios. Thus, our results were similar with a smaller variation in different scenarios. A grey-box model applied for in Ref. [19] showed very low total system cost savings when utilizing the thermal mass for storage. However, if significant electrification of the heating sector would be pursued, thermal mass for storage would lead to more significant savings, up to 15% [19]. Although our results are not directly comparable, the total system costs in our case were also only marginally lower compared to the reference case. Finally, the linear optimization model developed in Ref. [20] showed that the thermal mass for storage was mostly utilized to shift the morning peaks and to a some extent the late afternoon peaks. The same behaviour occurred in our case, especially in the colder periods. They have further claimed that smart controls should be invested in approximately 34% of the buildings. Our findings showed that the system would economically benefit, in terms of operational costs of the district heating system, by equipping up to 98% of the modelled buildings with smart controls, which equals to 59% of the total housing stock in the city of Sønderborg. Only the buildings with very low autonomy time should not spend money on implementing smart controls. It is yet unclear what the price of implementing smart controls in district heating systems on this wide scale would be. However, based on the CO_2hPH scenario and assumed lifetime of smart controls of 15 years, the maximum investment in smart controls of 261 EUR per household would be economically feasible. Larger investments in smart controls would not be recuperated by operational savings in the district heating system.

Regarding the building energy simulations, IDA ICE uses a variable time step solver to capture the dynamics of the system. However, the cut-off duration was rounded up on hourly intervals to reduce the complexity of the optimization problem and thus, the computation time. If a simpler optimization problem was to be solved, the implemented cut-off times would be converted to 30-min intervals or less. Thus, the effect of preheating strategies on the heat flexibility potential of the archetypes would be more evident since rounding would be avoided. In addition, it should be noted that despite the preheating set point of 24 °C, the operative temperature inside the buildings could not exceed 22 °C even in scenario 3 (CO_4hPH), which is due to the high inertia of the internal walls and mass, as well as the relatively short preheating duration. The lowest operative temperature of 18 °C was selected to determine the duration of the heating cut-off, so that acceptable thermal comfort was ensured according to ASHRAE and Danish standards and also provide some flexibility to the system. One can argue that this temperature might be too low when occupants would be present or too high if no occupancy is assumed during the cut-off times. Therefore, further investigations are proposed that will study a wider range of lowest temperature threshold and its effect on heat flexibility. Another proposal would be to introduce an additional indicator that will represent the deterioration in the thermal comfort from occupant's side. That would give the human perspective to the effect of DSM approaches. Moreover, some tests were run where internal wall mass was neglected from the building models which resulted in a much lower inertia of the building. Consequently, the thermal autonomy potential was decreased in the majority of archetypes since the total heat capacity decreased. Thus, internal walls proved to be decisive for the inertia of the building, which in combination with the U-value of the building envelope determine the heat flexibility indicators. Previous work on Danish low-energy apartments [44] indicated that the thermal autonomy was mainly determined by the heat capacity of internal walls and heat losses from the external walls. Our work validated this finding for older single-family houses, too, which consist of less well-insulated building envelope and larger external envelope areas. If extensive refurbishments were to be applied to the building envelope, the effect of the thermal mass would be more pronounced.

The current case only investigated the residential building stock and specifically single-family houses. Thus, the building sample was very homogeneous. It should be pointed out that the investigated building stock did not include low-energy or nearly-zero energy buildings, which would have even longer thermal autonomy potential due to their very well-insulated building envelope. Furthermore, no apartment blocks or multi-family buildings were modelled in the current study, which would have lower thermal losses due to decreased external envelope area, leading to potentially higher heat flexibility. If different building typologies had been included, the thermal autonomy results could have been different, resulting in a different triggering of DSM events on the system scale. Also, if commercial buildings are to be included in the analysis, the internal gains and the occupant schedules should be adjusted accordingly. In that case, high internal gains from employees and equipment could potentially create overheating problems especially in newer office buildings, which is a challenge that has to be addressed during the design phase. Hence, the duration and set-points of the preheating strategy would have to be adjusted accordingly, so that they do not lead to very high and uncomfortable internal temperatures. It should be pointed out that this analysis focused only on the heating demand and the heat flexibility that could be provided to the district energy system due to the cold climate. No mechanical cooling is installed in the majority of residential buildings in Denmark. However, in warmer climates, district cooling is one of the promising solutions for increasing energy efficiency of system [45]. Thus, if a warmer climate had been studied, the thermal flexibility from a cooling perspective would have been relevant, too, as a demand side management technique.

There are several refinements, which could be assessed in the future. First, the simulation strategy we adopted was to return to the operative temperature as quickly as possible after the DSM event. This was facilitated by the cold ambient temperature and the relatively short overheating periods, which led to fast discharging of the thermal mass.

However, one could try to assess the performance of buildings if much slower temperature increase would be adopted. However, the latter needs to be balanced taking into account the expected thermal comfort, too.

Second, a future scenario with much more excess heat, solar thermal, heat pumps, geothermal and other renewable (and possibly intermittent) supply sources in the district heating grid could be analysed. Having more options and capacities of low operational cost technologies in the system could increase the activation of thermal mass for storage in order to avoid the utilization of high operating cost technologies.

Third, as the literature review showed that nearly zero energy buildings have very long autonomy times, one could try to assess the impact of utilizing the thermal mass for storage in a system with a very large share of nearly zero energy buildings. The latter system could represent newly built neighbourhoods connected to DH systems. According to [44], the thermal autonomy times of newly-built low-energy apartments in Denmark could exceed 15 h with a 4 h-preheating strategy, similar to our CO_4hPH.

Fourth, the system optimization model currently does not have the possibility of dynamic modelling of the operation of the DH grid. Hence, one of the future research pathways could be soft-linking of the current system optimization model with some of the dynamic models of the DH grid operation. The latter could make results robust for a wide range of operational cases.

Implementing smart controllers in district heating grid on very large scale is a significant feat. Thus, one potential implementation strategy would be to implement it when building refurbishments take place. That would reduce the installation costs and implement the controllers in the buildings that would have longer thermal autonomy times, after being refurbished.

6. Conclusions

This study investigated the potential of thermal building mass for storage in district heating systems. It was conducted as a two-stage analysis, where building performance simulations were run followed by a system optimization analysis. The building stock of Sønderborg was characterized by six archetypes that represented the 60% of single-family houses in the area. The results were evaluated on basis of two flexibility indicators: thermal autonomy potential that was defined during a heat supply cut-off while the internal temperature did not fall below 18 °C and the savings in heating demand compared to the reference case. A third indicator was used to evaluate the economics of the system. Three different strategies were investigated: (i) a heating cut-off for a certain number of hours, and preheating of (ii) two or (iii) four hours followed by a heating cut-off. The heating cut-off resulted in energy savings in all archetypes compared to the reference case, as expected. The effect of preheating control up to four hours was found to affect positively the heat flexibility potential of buildings, but should be evaluated individually for each archetype. The experiment of the cut-off was implemented on a cold and grey day, so that the effect of additional gains, such as solar gains, was not considered, while transmission losses through the building envelope were increased. This choice gave a pessimistic prediction to our results. The peak loads that were created after the heating cut-off were mostly determined by the capacity of the heating system (i.e. hydronic radiators) and the duration of the cut-off, which determined the internal temperature. It was concluded that the highest potential for utilization of building thermal mass is provided by houses built after the 1980s, which have well-insulated building envelope and thus, have lower transmission losses. Furthermore, the thermal autonomy potential is better described by the total heat loss coefficient of the building envelope and less by the time constant during these cold and cloudy weather conditions, since heat losses were found to be more dominating in the specific buildings than the embedded thermal mass.

Operational savings in the DH system occurred in all the cases when thermal mass for storage was utilized. The economic savings in operational costs of the district heating system of the city of Sønderborg were in the range of 0.7%–4.6%, not taking the cost of smart controls into account. It would be feasible to invest up to 261 EUR per household in the installation of smart controls. The scenario carried out for the year 2029 showed that the benefits of using the thermal mass for storage are much greater in the future than in the current district heating system, due to the larger capacities of intermittent generation that can be successfully integrated in the district heating supply. Moreover, load shifting that is made possible with activation of thermal mass for storage allows for larger load factors of capital intensive-low operating cost technologies, such as central heat pumps.

One of the sensitivity analyses showed that the large-scale pit thermal energy storage is complimentary to the thermal mass for storage, as the former is mainly used for seasonal shifting of load and the latter one for intra-day shifting of load.

All the scenarios showed that the thermal mass for storage allowed more solar thermal district heating to be effectively utilized. The most significant energy savings originated from the less utilized central gas boilers, as well as biomass boilers in the CO_2hPH_2029 scenario.

Acknowledgments

This work was undertaken as a part of the CITIES (Centre for IT-Intelligent Energy Systems in cities) project no DSF1305-00027B funded by Danish Strategic Research Council. Cooperation with the 4DH project funded by the Innovation Foundation and the ProjectZero, including their representative Nicolas Bernhardi, on obtaining the data for calculations is greatly appreciated. Nevertheless, we are thankful to Sønderborg Fjernvarme, the operator of the district heating system, for providing us a data about the district heating system.

Nomenclature

$CO2\_inten_j$	CO_2 intensity of a certain technology or energy within the system boundaries, ton/MWh
$CO2_j$	Costs of CO_2 emissions, €/ton
$dies\_imp_k$	Price of import of diesel in a specific hour, €/MWh
$el\_imp\_exp_k$	Price of import or export of electricity in a specific hour, €/MWh
$fix\_O\&M_i$	Fixed operating and maintenance costs of energy plants, €/MW
$fuel_j$	Fuel cost of specific energy type, €/MWh$_{fuel}$
$gas\_imp\_exp_k$	Price of import or export of gas in a specific hour, €/MWh
lev_inv_i	Levelized cost of investment over the energy plant lifetime, €/MW
$petr\_imp_k$	Price of import of gasoline in a specific hour, €/MWh
x_i	Capacity variables of energy plants, transmission grid and gas grid, MW
x_j	Generation capacities of energy plants (8760 variables for each energy plant, representing the generation in each hour during the one year), MWh
x_k	Import or export across the system boundaries of different types of energy (8760 variables per one type of energy, representing the flow in each hour during the one year), MWh
η_j	Efficiency of technology, MWh$_{energy}$/MWh$_{fuel}$
T	Operative temperature [^oC]
Q	Heating demand of building after the applied strategy [Wh]
Q_{ref}	Heating demand of building under normal operation of heating system
CO	Heat supply cut-off strategy
CO_2hPH	Heat supply cut-off strategy with a 2 h preheating
CO_4hPH	Heat supply cut-off strategy with a 4 h preheating

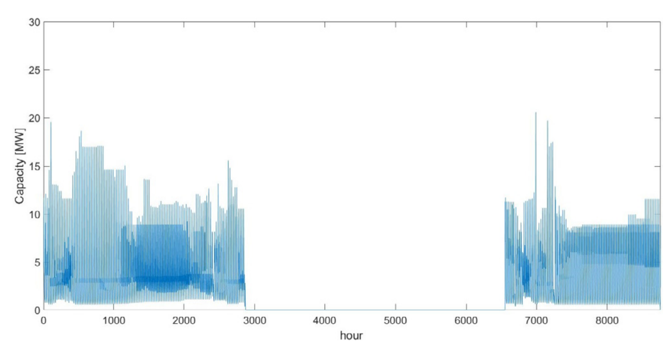
Appendix A

Generation of different plants in reference cases for the year 2015 and 2029.

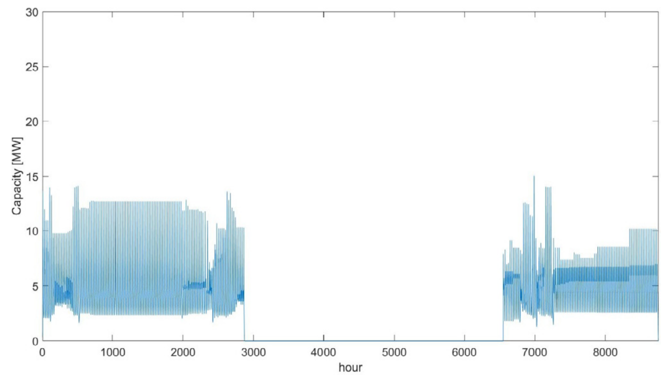
	Reference (2015) [MWh]	Reference (2029) [MWh]
Solar heating	3362	30,683
geothermal + biomass ab. Heat pump	77,930	60,108
biomass boilers	27,034	95,593
waste CHP heat	170,346	174,637
gas CHP heat	17,833	0
gas boilers	80,184	0
heat pumps	0	65,503
Wind turbine	37,488	454,875
PV	13,150	54,101
el grid import	334,300	20,030
el grid export	329	242,455
waste CHP ele	38,366	39,333
gas CHP ele	20,265	0
gas import	602,207	169,274
gas export	0	0
gasoline import	253,169	253,169
diesel import	253,169	253,169

Activation of cut-off events on hourly resolution in different scenarios:

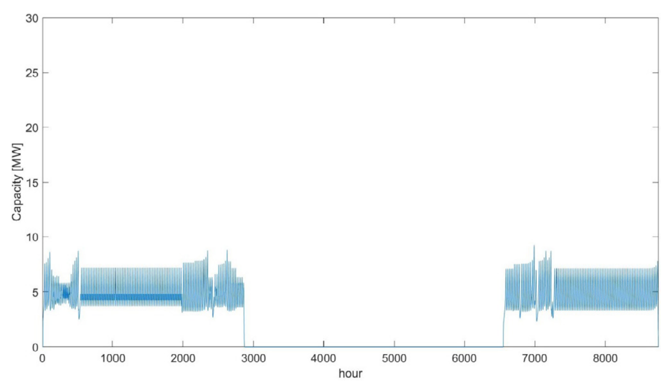
CO scenario:



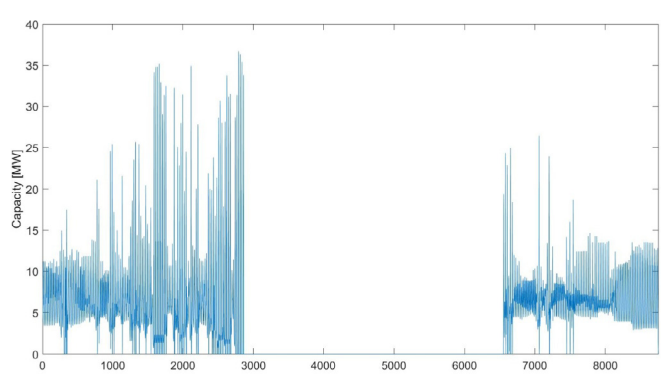
CO_2hPH scenario:



CO_4hPH scenario:



CO_2hPH_2029 scenario:

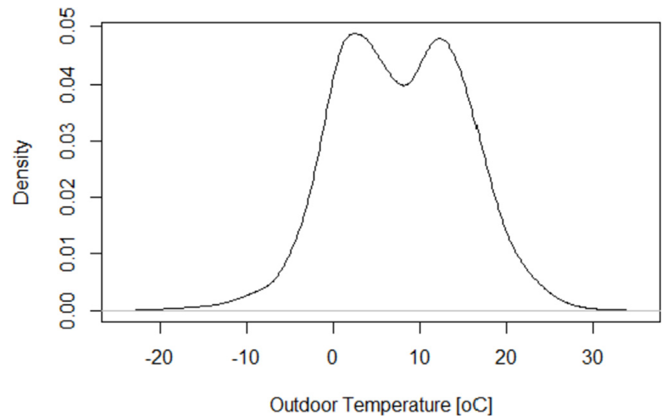


Appendix B

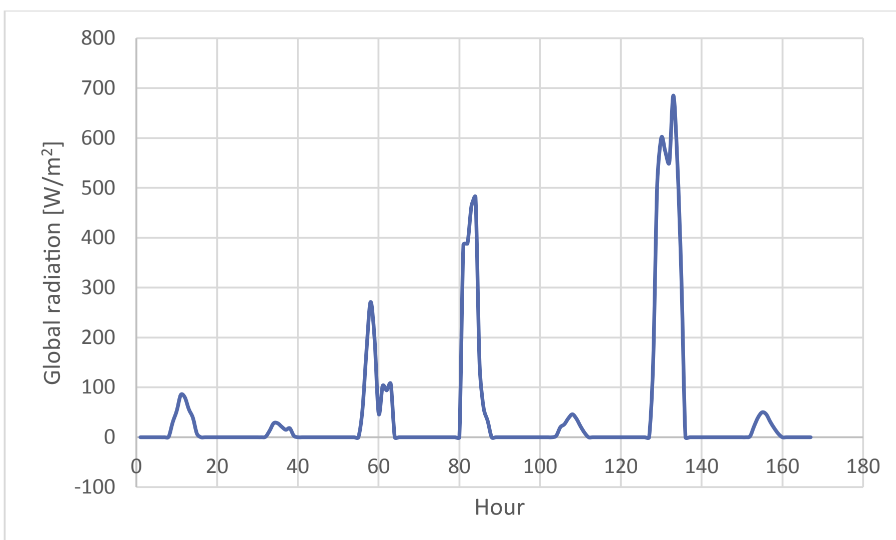
Monthly degree-day values for a year according to the given Danish climate file:

Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
536	533	463	355	177	95	60	50	126	237	324	469

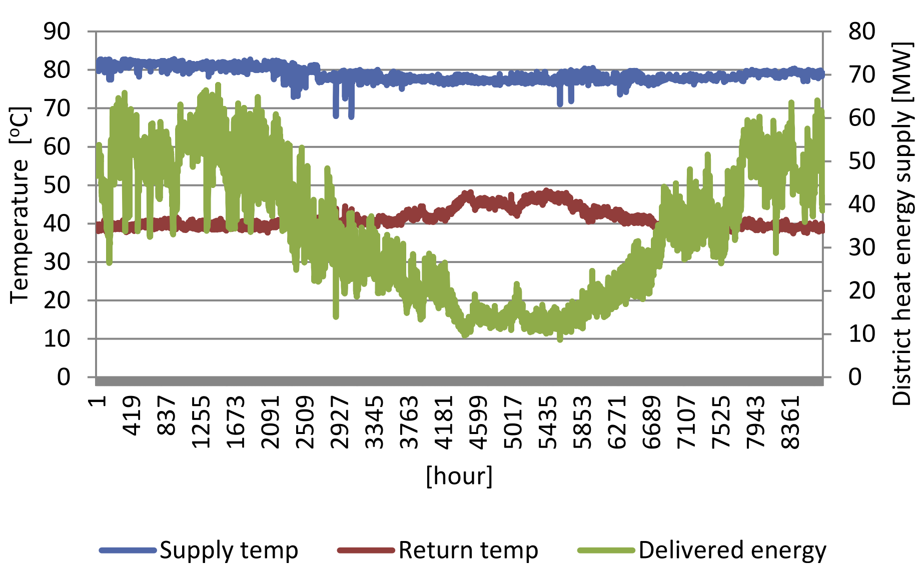
Annual outdoor temperature distribution:



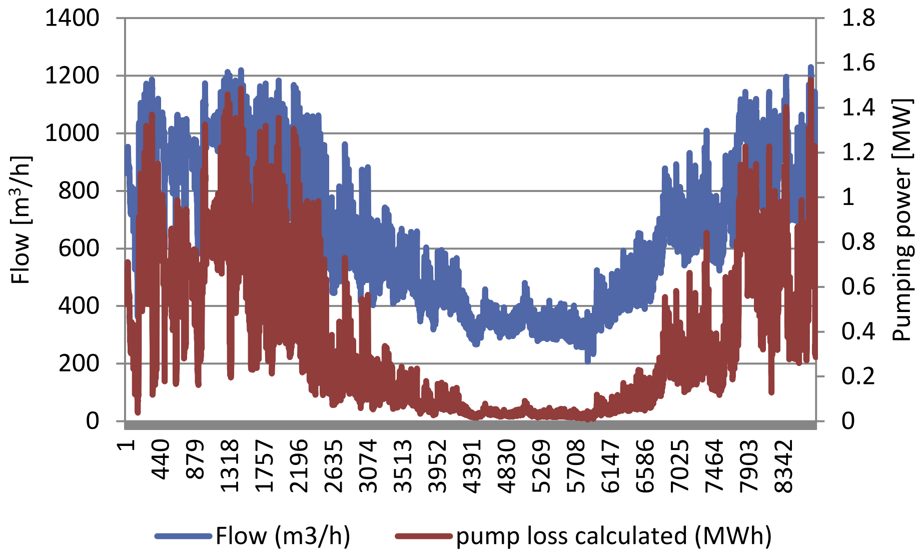
Global radiation amount for an exemplary week in January:



Hourly supply and return temperatures, as well as total delivered district heating energy to the end users, obtained from Sønderborg Fjernvarme (the operator of the DH grid):



Hourly volume flow (obtained from Sønderborg Fjernvarme) and estimated hourly pumping demand in the DH grid (total yearly pumping electricity demand is equal to the 1% of the final DH energy demand, in line with the pumping energy demand reported in Ref. [46]):



References

[1] Noussan M, Jarre M, Poggio A. Real operation data analysis on district heating load patterns. Energy 2017;129:70–8. https://doi.org/10.1016/j.energy.2017.04.079.

[2] Lake A, Rezaie B, Beyerlein S. Review of district heating and cooling systems for a sustainable future. Renew Sustain Energy Rev 2017;67:417–25. https://doi.org/10.1016/j.rser.2016.09.061.

[3] Lund H, Werner S, Wiltshire R, Svendsen S, Thorsen JE, Hvelplund F, et al. 4th Generation District Heating (4GDH). Integrating smart thermal grids into future sustainable energy systems. Energy 2014;68:1–11. https://doi.org/10.1016/j.energy.2014.02.089.

[4] Münster M, Morthorst PE, Larsen HV, Bregnbæk L, Werling J, Lindboe HH, et al. The role of district heating in the future Danish energy system. Energy 2012;48:47–55. https://doi.org/10.1016/j.energy.2012.06.011.

[5] Werner S. International review of district heating and cooling. Energy 2017: 1–15. https://doi.org/10.1016/j.energy.2017.04.045.

[6] Zvingilaite E, Balyk O. Heat savings in buildings in a 100% renewable heat and power system in Denmark with different shares of district heating. Energy Build 2014;82:173–86. https://doi.org/10.1016/j.enbuild.2014.06.046.

[7] Heier J, Bales C, Martin V. Combining thermal energy storage with buildings - a review. Renew Sustain Energy Rev 2015;42:1305–25. https://doi.org/10.1016/j.rser.2014.11.031.

[8] Hesaraki A, Holmberg S, Haghighat F. Seasonal thermal energy storage with heat pumps and low temperatures in building projects—a comparative review. Renew Sustain Energy Rev 2015;43:1199–213. https://doi.org/10.1016/j.rser.2014.12.002.

[9] Yan T, Wang RZ, Li TX, Wang LW, Fred IT. A review of promising candidate reactions for chemical heat storage. Renew Sustain Energy Rev 2015;43: 13–31. https://doi.org/10.1016/j.rser.2014.11.015.

[10] Pereira da Cunha J, Eames P. Thermal energy storage for low and medium temperature applications using phase change materials - a review. Appl Energy 2016;177:227–38. https://doi.org/10.1016/j.apenergy.2016.05.097.

[11] Navarro L, de Gracia A, Colclough S, Browne M, McCormack SJ, Griffiths P, et al. Thermal energy storage in building integrated thermal systems: a review. Part 1. active storage systems. Renew Energy 2016;88:526–47. https://doi.org/10.1016/j.renene.2015.11.040.

[12] Le Dréau J, Heiselberg P. Energy flexibility of residential buildings using short term heat storage in the thermal mass. Energy 2016;111:991–1002. https://doi.org/10.1016/j.energy.2016.05.076.

[13] Reynders G, Nuytten T, Saelens D. Potential of structural thermal mass for demand-side management in dwellings. Build Environ 2013;64:187–99. https://doi.org/10.1016/j.buildenv.2013.03.010.

[14] Wolisz H, Harb H, Matthes P, Streblow R, Müller D. Dynamic simulation of thermal capacity and charging/discharging performance for sensible heat storage in building wall mass. Inst Energy Effic Build Indoor Clin n.d.: 2716–23.

[15] Kensby J, Trüschel A, Dalenbäck J-O. Potential of residential buildings as thermal energy storage in district heating systems – results from a pilot test. Appl Energy 2015;137:773–81. https://doi.org/10.1016/j.apenergy.2014.07.026.

[16] Pavlak GS, Henze GP, Cushing VJ. Evaluating synergistic effect of optimally controlling commercial building thermal mass portfolios. Energy 2015;84: 161–76. https://doi.org/10.1016/j.energy.2015.02.073.

[17] Müller D, Monti A, Stinner S, Schlösser T, Schütz T, Matthes P, et al. Demand side management for city districts. Build Environ 2015;91:283–93. https://doi.org/10.1016/j.buildenv.2015.03.026.

[18] Masy G, Georges E, Verhelst C, Lemort V, André P. Smart grid energy flexible buildings through the use of heat pumps and building thermal mass as energy storage in the Belgian context. Sci Technol Built Environ 2015;21:800–11. https://doi.org/10.1080/23744731.2015.1035590.

[19] Heinen S, Turner W, Cradden L, McDermott F, O'Malley M. Electrification of residential space heating considering coincidental weather events and building thermal inertia: a system-wide planning analysis. Energy 2017;127: 136–54. https://doi.org/10.1016/j.energy.2017.03.102.

[20] Hedegaard K, Balyk O. Energy system investment model incorporating heat pumps with thermal storage in buildings and buffer tanks. Energy 2013;63: 356–65. https://doi.org/10.1016/j.energy.2013.09.061.

[21] Equa. IDA Indoor climate and energy. EQUA Simulation AB; 2017.

[22] Kern-hansen C. Technical report 13-19 2001 – 2010 Danish design reference year - reference climate dataset for technical dimensioning in building, construction and other sectors Peter Grunnet Wang Kim Bjarne Wittchen Colophon. 2013.

[23] Coskun C, Ertürk M, Oktay Z, Hepbasli A. A new approach to determine the outdoor temperature distributions for building energy calculations. Energy Convers Manag 2014;78:165–72. https://doi.org/10.1016/j.enconman.2013.10.052.

[24] Gianniou P, Heller A, Nielsen PS, Negendahl K, Rode C. Aggregation of building energy demands for city-scale models. In: BS 2015-14th int IBPSA conf; 2015. p. 704–11.

[25] TABULA Project Team. TABULA-Typology Approach for Building Stock Energy Assessment. Main Results of the TABULA project n.d.:43. www.building-topology.eu (accessed July 30, 2017).

[26] Wittchen K, Kragh J. Danish building stock typologies and building stock analyses. Participation in the EPISCOPE project. 2016.

[27] ASHRAE. Measurement of energy and demand savings. ASHARE Guidel 14-2002. 2002. 8400:1–165.

[28] Sarran L, Foteinaki K, Gianniou P, Rode C. Impact of building design parameters on thermal energy flexibility in a low-energy building. In: Proc. BS2017 15th conf. Int. Build. Perform. Assoc. San Fr. USA; 2017.

[29] DS/EN 15251. Indoor environmental input parameters for design and assessment of energy performance of buildings addressing indoor air quality,

thermal environment, lighting and acoustics. Dansk standard; 2007.
[30] http://www.gurobi.com/. Gurobi 7.5 2017.
[31] Dominković DF, Bačeković I, Sveinbjörnsson D, Pedersen AS, Krajačić G. On the way towards smart energy supply in cities: the impact of interconnecting geographically distributed district heating grids on the energy system. Energy 2017:20. https://doi.org/10.1016/j.energy.2017.02.162.
[32] Dominković DF, Bačeković I, Ćosić B, Krajačić G, Pukšec T, Duić N, et al. Zero carbon energy system of south East Europe in 2050. Appl Energy 2016;184: 1517–28. https://doi.org/10.1016/j.apenergy.2016.03.046.
[33] PlanEnergi. Strategisk energiplan for Sønderborg - Sønderborgs energisystem i 2029. 2014.
[34] SparEnergi.dk. https://sparenergi.dk/n.d. https://sparenergi.dk/(accessed November 28, 2017).
[35] Aggerholm S. Cost-optimal levels of minimum energy performance requirements in the Danish Building Regulations. 2013.
[36] Wittchen KB, Mortensen L, Holøs SB, Björk NF, Vares S, Malmqvist T. Building typologies in the Nordic countries : identification of potential energy saving measures. 2012.
[37] Database. BBR database n.d.
[38] COWI. Sønderborg kommune: Strategisk energiplan, kortlægning og basisfremskrivning. 2015.
[39] Danish Government. Strategy for energy renovation of buildings. The route to energy-efficient buildings in tomorrow's Denmark. 2014. p. 54.
[40] Wittchen KB, Kragh J. Energy savings in the Danish building stock until 2050. In: 10th Nord. Symp. Build. Phys. 15-19 June 2014 Lund; 2014. p. 1277–84.
[41] Nordpool. Nordpool Elspot prices n.d. http://www.nordpoolspot.com/Market-data1/Elspot/Area-Prices (accessed May 27, 2016).
[42] Energinetdk. Energinet.dk's analysis assumptions 2014-2035. 2014. Update September 2014.
[43] Dominković DF, Bačeković I, Pedersen AS, Krajačić G. The future of transportation in sustainable energy systems : opportunities and barriers in a clean energy transition. Renew Sustain Energy Rev 2017:16. https://doi.org/10.1016/j.rser.2017.06.117.
[44] Foteinaki K, Heller A, Rode C. Modeling energy flexibility of low energy buildings utilizing thermal mass. In: Proc. 9th Int. Conf. indoor air Qual. Vent. & energy Conserv. Build.; 2016.
[45] Dominković DF, Bin Abdul Rashid KA, Romagnoli A, Pedersen AS, Leong KC, Krajačić G, et al. Potential of district cooling in hot and humid climates. Appl Energy 2017;208:49–61. https://doi.org/10.1016/j.apenergy.2017.09.052.
[46] Frederiksen S, Werner S. District heating and cooling. Studentlitteratur; 2013.

Paper IV: Gianniou, P., Heller, A., Nielsen, P. S., Negendahl, K., & Rode, C. (2015). Aggregation of building energy demands for city-scale models. In Proceedings of *Building Simulation 2015,* Hyderabad, India.

AGGREGATION OF BUILDING ENERGY DEMANDS FOR CITY-SCALE MODELS

Panagiota Gianniou[1], Alfred Heller[1], Per Sieverts Nielsen[2], Kristoffer Negendahl[1], Carsten Rode[1]
[1]Department of Civil Engineering, DTU, Lyngby, Denmark
[2]Department of Management Engineering, DTU, Roskilde, Denmark

ABSTRACT

Smart Cities initiatives and its focus on city level energy policy management has emphasised the need for development of aggregated energy demand models. This study models aggregated energy demand in 16 single-family houses, which are investigated on energy performance. The energy performance is modelled with Termite which uses Danish Be10 for energy performance calculations according to EN ISO 13790. Two methods of aggregating the energy demand were examined: the first method was based on modelling the individual energy performances, while the second method used building typologies and archetypes. The results highlight that the latter represents quite well the respective buildings, but deviates from the actual measured heat consumption in the buildings. However, the modelled annual aggregated heat demand was found to be very close to the measured consumption. Extensive discussion on the challenges and uncertainties of the suggested city scale energy model is presented.

INTRODUCTION

Aggregation is an important step towards estimating building energy demands, which is extensively discussed by many studies nowadays. It starts with the analysis of individual buildings' energy demands and continues with an upwards aggregation, which can be used to evaluate the performance of the whole building stock. This can be expressed as a bottom-up approach to determine the energy demand in the building stock. The observation of energy use trends and profiles for total aggregate stock also facilitates prediction of energy demand of buildings in the future (IEA, 2001).

The scale of building energy aggregation can range from small neighborhoods for single projects to national building stocks for residential, commercial or industrial sectors. Either way, the aggregation method and the structure of the data processing are the same.

Aggregating building energy demands has been studied for years especially in connection to central energy grids, electricity and district heating (Heller, 2000), (Heller, 2002). This scientific interest for aggregation has increased the last years, as the concepts of sustainable energy and smart cities have become popular topics for city administrations. This is also due to its potential to support decision-making in many urban development projects. More specifically, its contribution to decision-making can be summarized in two ways (IEA, 2001):

• It gives designers a better overview of the impact individual buildings have on the aggregated stock and adapt their design accordingly.

• Planners and policy-makers can have a holistic database including energy use data, which directly affects energy resources use, power system stability, as well as environmental emissions.

The aggregation of building energy demands can be classified with two different methods: i) the energy estimates of individual buildings can be added up to calculate the total energy use of the building stock, ii) reference buildings and building categories can be used, which are representative for the whole stock and weighting factors can be used proportionally for every category (Choudhary, 2011), (Matsuoka et al., 2013). Both methods are valid for all types of building energy demand, including space heat demand and domestic hot water (DHW) demand. They are presented analytically below.

METHODOLOGY

This paper investigates two methods to aggregate building energy demands. Method 1 uses the simplest form of aggregation, namely a sum of the measured energy consumption of the stock. Method 2 is based on an aggregation of building typologies. These typologies are generated from various information level indicators such as building age, net floor area and location. The indicators are used as input to a simulation tool, which in turn generates an estimate of the energy consumption of each individual building. The paper investigates how these two methods are applied on a small scale. This includes a qualitative investigation of selected information level indicators.

Method 1 - Aggregation of individual buildings' energy estimates

Aggregation of individual building energy demands is calculated in a plethora of studies. However, very few studies have described analytically the methodology of the aggregation. The simplest and most common way to aggregate energy demands is to add them up. Thus, the total energy consumption of the aggregated stock Y will be a sum of the individual buildings' energy demands, as presented in the following equation.

$$Y = \sum_{i=1}^{n} X(i) \qquad (1)$$

where

Y = total energy demand of the examined building stock [kWh]

n = number of individual buildings

X = energy demand per building [kWh].

It is assumed that each building's load profile is independent and normally distributed for simplification reasons.

The majority of scientific studies that deal with aggregate energy data in a variety of applications make use of this method. The advantage of it is the simplified methodology and the accuracy of the total results. On the other hand, detailed data, measurements and a load of energy simulations are required to calculate the energy demand of every building. This data-intensive approach proves to be very time-consuming and expensive. For this reason, it is mainly used at local scale, when a small neighborhood or a district is examined and access to such data is possible. However, if aggregation aims at a national level, then estimating the energy use for every building becomes even more difficult.

Method 2 - Aggregation based on reference buildings

To overcome the difficulties that arise from aggregating every individual building's energy demand, reference buildings have been introduced as a concept to stock aggregation. These represent fully the features of buildings included in every category. The process of aggregation is in this way greatly simplified, since the number of reference buildings is only a small part of the total stock. What is more, they are not so much affected by poor data quality. Databases of reference buildings enable the addition of new buildings or even existing ones, when their data become available and can be registered to them (IEA, 2001).

According to the methodology of using reference buildings as an aggregation method, the aggregation of building energy demand is made in two steps: i) multiplying the results of each building type with either the total number of buildings included or their total floor area and ii) summing the sub-totals to calculate the aggregate energy demand. These steps are indicated by the equation (2) below.

$$Y = \sum_{j=1}^{N} X(j)\ B(j) \qquad (2)$$

Where

j = building type

N = total number of building types describing the stock

B = total number of buildings included in every type.

In the present study, every archetype's energy demand is estimated per floor area - also known as Energy Use Intensity (EUI) [kWh/m^2]. So, the above equation is changed as follows:

$$Y = \sum_{j=1}^{N} EUI(j)\ A(j) \qquad (3)$$

where

EUI = energy demand per floor area [kWh/m^2] for each building type

A = total floor area [m^2] of all buildings included in the respective type.

These aggregating methods were applied in the present study to a real case consisting of 16 single-family houses located in Sønderborg, Denmark. The energy performance in the examined houses was modelled by Termite, a newly-developed parametric modelling tool, which uses Rhinoceros design interface, Grasshopper parametric options and Be10 for energy performance calculations according to EN ISO 13790 (Negendahl, 2014).

Information levels

One of the most significant challenges of bottom-up city simulation is data requirement. For this reason, four different information levels were investigated regarding the amount of data available for the specific examined case in Sønderborg. These data levels are summarized below:

A. Simple typological data. These data regard the type of the examined house according to building typologies. In the current study, TABULA database was used, as mentioned before.

B. Information collected by online public databases. In particular, the construction year, the floor area, a general overview of the building's design and the construction materials were collected from the Danish Building Register (BBR).

C. Information acquired by Google Maps, Street View. More specifically, these were used to acquire basic design information, such as the houses' ground plan and the houses' orientation, as well as the type and placement of windows and doors.

D. Information acquired by personal visits to the examined houses, on site measurements and distributed questionnaires to the occupants. In particular, two sources of information are included in this information level for the present study: i) the number of occupants and ii) whether the houses had undergone any refurbishments that affected their energy performance. The energy refurbishment state is indicated by the U-values of building materials and windows, which in case of renovation are significantly lowered according to TABULA database. Personal visits to the examined houses may also determine if any additional systems (solar PV panels) have been installed.

Information levels A, B, C and D were used to model energy performance in the 16 single-family houses in Sønderborg. Six different scenarios were investigated, which are illustrated in Table 1. The scenarios represent every possible combination of the available information (A, B, C and D) that can be created. To facilitate the easy understanding and analysis of the six created scenarios, Table 2 presents a brief description of the information included in every scenario.

Table 1 Presentation of the six created scenarios based on the four different information levels

Inform. Level	S1	S2	S3	S4	S5	S6
A	X	X	X	X	X	X
B	X	X	X	X	X	X
C		X		X		X
Di				X	X	X
Dii			X		X	X

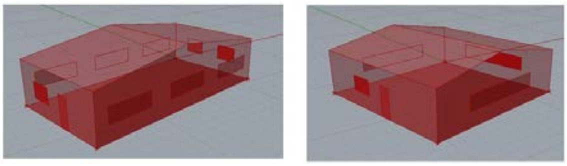


Figure 1 Design of House 1 for Scenario 1 (left) and 6 (right)

Table 2 Description of specific information contained in every Scenario

Scenario No.	Description of information included
S1	BBR
S2	BBR & Google Maps/Street View
S3	BBR & Refurbishment
S4	BBR & Google Maps/Street View & Occup. Loads
S5	BBR & People Loads & Refurbishments
S6	BBR & Google Maps/StreetView & Occup. Loads & Refurbishments

The first scenario was initially constructed using only the basic information, included in information levels A and B, which are the type of the building, the age of construction, the floor area and the typical U-values of the houses according to TABULA Webtool. The design of houses was created based on generic means, aiming at an abstract representation of the actual design. Thus, the floor plan for all 16 houses was assumed to be a square area, having equal sides. The glazing area was set to be the same for all houses, and the exact orientation was not defined in Termite.

The second scenario was made by adding information level C, thus the exact design of the houses. The dimensions of the houses and windows, as well as their placement were taken into account at this stage. This design information was collected by Google Maps and Street View.

The third, fourth and fifth scenarios of the houses' model were made according to information level D: questionnaires distributed to occupants and on site observations. These scenarios' models included the actual occupancy loads based on the real number of occupants, while they also contained information about any possible energy refurbishments applied to the houses. The existence of energy refurbishments to the building envelope affects significantly the thermal characteristics of the building materials and windows (U-values, g-values). The information level

D was divided into these three scenarios (3, 4, 5) so as to separate the impact of the individual parameters. More specifically, Scenario 3 only takes the information about any energy refurbishments in the houses into account, neglecting the actual people loads and the real dimensions or design of the houses. In this scenario, a standard occupancy load of 1.5 W/m^2 for all 16 house models is considered. Scenario 4 makes use of the real design of the house, as well as the real occupancy loads, but without any information about energy renovations. Scenario 5 considers both the actual occupancy loads and the energy refurbishment state, but without including the real design of the houses. When the refurbishment state of each house was known, the adapted U-values to the renovated building envelope were collected by TABULA database.

Finally, a sixth scenario (Scenario 6) was investigated, which combined all the existing available information (information levels A, B, C, D), thus creating a realistic model of the 16 single-family houses. More specifically, the houses were modelled using their real design - both dimensions and orientation- and considering the real occupancy loads and the energy refurbishment states.

Figure 1 illustrates indicatively the design of House 1 created in Grasshopper for Scenarios 1 and 6, respectively. The building's floor area has remained constant in all cases, but the dimensions of the building envelope as well as the dimensions and placement of the windows differentiate between these two cases.

To calculate the aggregate energy demand, the first aggregating method was applied. Thus, the aggregation was made by summing up the individual buildings' energy demands for the different case-scenarios (Equation 1). This data-intensive method was possible to implement in the specific case, because only a small sample of buildings in the Sønderborg area was examined. Since the computation time of Termite remains very small even for a very large number of buildings, the main difficulty faced in city models is the building data collection and analysis.

In cases where the number of investigated buildings increases a lot and the examined area extends to a district or city level, the second methodology as presented previously is proposed. Thus, the aggregation of building energy demand is based on archetypes of reference buildings, using Equation 3. TABULA database and Webtool were used to create the archetypes, as already described. The 16 examined houses were all of the same type (single-family house), thus, the differentiation among them was based on their construction age. All Danish building stocks are categorized into nine time periods according to Wittchen & Kragh, 2012; the examined 16 houses correspond to five of them. Based on these time periods, five building types were created covering all 16 examined houses. One example building was created for each type. The characteristics of the example buildings are presented in Table 3. The example buildings share similar technical characteristics to the buildings that they represent, such as building envelope, U-values, HVAC systems etc. The floor area, U-values and g-values were acquired by TABULA Webtool for every building type.

Table 3 Characteristics of the example buildings

Building type	Construction period	No. of incl. buildings	Total floor area [m^2]
A	1931-1950	2	238
B	1951-1960	2	180
C	1961-1972	10	1,530
D	1973-1978	1	117
E	1979-1998	1	122

Model setup

First, a building model was created for house number 0 in Grasshopper. Danish building databases, as well as the TABULA WebTool were used as information sources to collect data. The geometry of the house included the building envelope (external walls, floor, roof), as well as the glazing and external doors. To create the remaining building models (Houses 1-15) a number of lists were made in Grasshopper, which contained the input data. The main idea behind that was the parametric modelling; thus, by introducing a main controller which would be able to move from one house model to the other, all 16 house models would be simulated with just one model setup by simultaneously changing all input parameters. So, the only differentiation among the various house models was the input parameters, which are altered from house to house. These inputs' lists covered every building component and characteristic that had to be defined. Regarding geometry and design, lists of 16 different inputs were made for the dimensions of the houses, the windows' sizes and placement, as well as orientation. All lists were connected with the main controller that determined which house would be modelled every time. Therefore, the main controller could take values from 0 up to 15, corresponding to the house model number. The automatic movement of the main controller from one building model to the other via a timer enabled the modelling of an unlimited number of buildings varying from hundreds to millions. Of course, the larger number of buildings, the more complicated the model setup would be.

RESULTS

According to the collected weather data for the Sønderborg area, the lowest outdoor temperature of the examined period was -12°C, which coincided with the dimensioning outdoor temperature defined by Termite. Thus, simulation results and measurements were possible to be compared with each other.

According to the measurements of monthly heat consumption in the 16 examined single-family houses, the aggregate heat consumption was calculated based on the aggregation methodology of individual buildings.

To be able to evaluate the results of the 6 different scenarios of information levels, the results were compared with the ones acquired by the real measurements. In Table 4, the aggregated heat demand is presented as calculated by Termite for every scenario including all 16 houses, as well as the average value of the deviation between the specific scenario's monthly heat demand results and the monthly measured heat consumption. In addition, the deviation of the calculated annual heat demand for each scenario from the aggregate measured one is illustrated in Table 4.

Table 4 Results of scenarios - deviation compared to the real measurements of heat demand

Scenario	Total heat demand [kWh]	Average deviation of monthly demands	Deviation of yearly demand
1: BBR	421,295	45%	40%
2: BBR &GoogleMaps/Str. View	448,439	54%	49%
3: BBR & Refurb.	252,700	18%	16%
4: BBR & Google Maps/Str. View & Occup. Loads	445,742	53%	48%
5: BBR &Occup. Loads&Refurb.	250,658	19%	17%
6: BBR & Google Maps/Str. View & Occup. Loads & Refurb.	266,045	14%	12%

The results presented in Table 4 help the prioritization of the available building information towards an accurate energy simulation. It is observed that the least accurate results, hence the highest deviation from the measured values, are the results of Scenarios 1, 2 and 4, which lack the information about the current energy refurbishment state of the houses. When this information is included in the modeling (Scenario 3, 5, 6), the heat demand results improve significantly and become much more realistic compared to the real measurements. In particular, just the knowledge of the energy refurbishment state of the house improves the deviation of yearly demand by 24% compared to the simplest model including just BBR information. In the scenario where BBR data, realistic design acquired by Google Maps/Street View and occupancy loads are known, the addition of information about energy refurbishment improves the deviation of annual calculated heat demand from the measured one by 36%. Scenario 6, which combines all the available information, has the lowest deviation (12%) from the aggregate measured heat consumption, as it was expected.

The knowledge of the real design of houses and windows, as obtained by Google Maps and Street View, does not seem to affect the calculated heat demand results a lot compared to the scenarios that do not include it. This is due to the fact that the dimensions of windows, as well as the total glazing area of every house, were calculated based on generic means. Thus, for some houses the generic glazing area was lower than the actual one, resulting in lower heat loss through the transparent components calculated by Termite. Therefore, heat demand was lower and more close to the measured one. When the knowledge of real occupancy loads was added, the calculated heat demand results did not improve much. This may be attributed to the fact that the default occupancy load of 1.5 W/m^2 used in Scenarios 1, 2 and 3 is not that different compared to the real occupancy loads based on the exact number of occupants in every house. However, this result would have been different if houses with more occupants were examined.

Figure 2 presents the results of the six modelling scenarios of the examined houses in Sønderborg in terms of heat demand. The results are aggregated for all 16 houses and are presented throughout one year period. All models follow the same pattern, having the highest heat demand during winter months and the lowest during summer, as expected. The graph supports the conclusions based on Table 4 above. It is clear that knowing whether a house has undergone any energy refurbishments improves the calculated heat demand results significantly. Scenario 6, which contains all available building information that can be gathered, is most closely aligned with the real measurements' curve compared to the rest scenarios. The highest deviation between Scenario 6 and measurements is observed between May and September. The explanation is that Be10 does not consider any heating in summer months. However, according to the real measurements, there is still some demand for space heating during the summer period.

The least realistic approaches are the ones of Scenarios 1 (BBR data), 2 (BBR & GoogleMaps/Str. View) and 4 (BBR & GoogleMaps/Str. View & Occup.Loads), as already mentioned. These scenarios exclude the information of refurbishments of the investigated houses. Their highest deviation from the measured consumption values is observed in winter period. This was expected since the models created based on these scenarios contain much higher U-values of the building envelope and the windows than the actual ones. Thus, heat loss is calculated to be much higher and the heat demand is therefore significantly increased compared to the real measurement.

Figure 2 also indicates that knowing the exact occupancy loads (Scenario 4, 5) does not necessarily lead to better results, as already indicated with the sensitivity analysis. Thus, the curves of these scenarios differentiate too little compared to the ones that do not include these information. The same goes also for the knowledge of the real design of the windows and the house. Even though glazing area and its placement in terms of orientation does affect the Be10 results a lot, the way they were generated in the present Scenarios (1, 3, 5) was based on a generic pattern for all houses, thus cannot outline such difference.

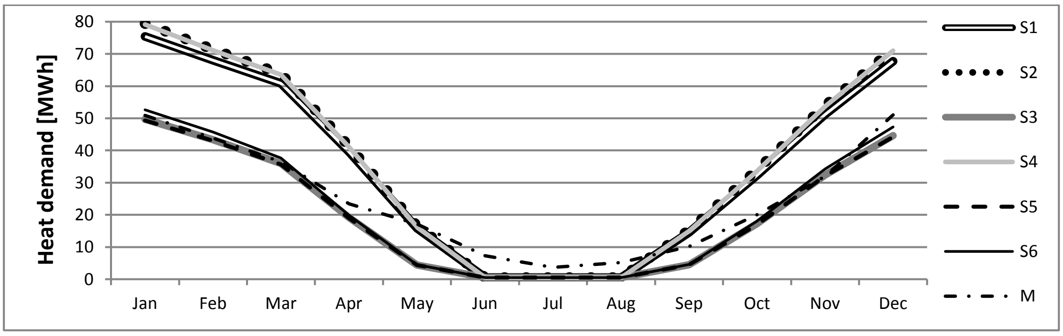


Figure 2 Aggregate heat demand results of the six different scenarios compared to real measurements conducted in the 16 single-family houses

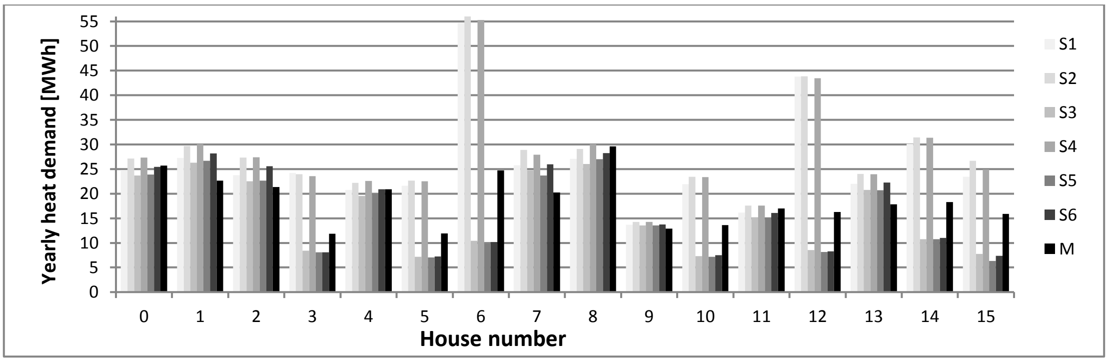


Figure 3 Yearly heat demand per house for all six examined scenarios compared with real measurements.

Figure 3 illustrates the annual heat demand results of the 16 examined houses individually for all six modelling scenarios. In addition, the consumption measurements are included in Figure 3 based on the examined year. First of all, the houses that have undergone renovations can easily be identified through the large deviations among the six models. In particular, Houses 3, 5, 6, 10, 12, 14 and 15 have undergone some kind of energy renovation as indicated by the large deviations between Scenarios 1 (BBR data), 2 (BBR & GoogleMaps/Str. View) or 4 (BBR & GoogleMaps/Str. View & Occup.Loads) and scenarios 3 (BBR & Refurb.), 5 (BBR & Occup. Loads & Refurb.) or 6 (BBR & Google Maps/Str. View & Occup. Loads & Refurb.). A similar refurbishment state was noticed for all these renovated houses. However, it was observed that benefits from energy renovations in houses were highly depended on their construction age. These benefits increased significantly for the oldest of the examined houses (Houses 6, 12) constructed in the 1930s and 1940s. In particular, House 6 and 12 reduced their heat demand significantly by 81% and 80%, respectively, between Scenario 1 (BBR data) and Scenario 3 (BBR and refurbishment data). Therefore, some usual renovation measures in old houses can lead to impressive reductions in the energy demand. Since these houses were constructed in the thirties and forties, their U-values according to TABULA database were the worst ones in terms of energy performance. Similarly, the renovated houses constructed in the 1960s (House 2, 14) had a decrease in energy demand of 65% and 64% respectively. In general, the newer a building is, the more advanced renovation measures are required to achieve high energy savings.

Moreover, it is observed that results according to Scenario 6, which represents the highest information level and the most realistic results, do not match the real measurements' results when looking into the individual house's energy performance. In particular, the highest deviations are noticed in House 6 and 15. It is also worth mentioning that the newly-built House 9 presents the lowest deviation among all models compared to the measurements. This indicates that new buildings must fulfill specific standards and requirements regarding their construction process and thus, their thermal characteristics are determined by national regulations. So, the actual U-values of their building envelope comply with the ones presented in official databases. These databases are not as representative when it came to older buildings, which did not comply with specific technical rules. Furthermore, Figure 3 validates what Figure 2 also illustrated: Scenarios 1, 2 and 4, which did not take the corrected U-values based on the energy refurbishment state into account, are far from realistic. However, the existence of actual design information and people loads slightly differentiate them from each other.

Aggregation based on reference buildings

Afterwards, the aggregate heat demands based on Equation 3 were calculated. The heat demands per floor area, thus EUI, for every building type are presented in Table 5. It has to be reminded that the presented energy demand includes heat demand for space heating, as well as DHW demand of the houses.

Table 5 Energy results of the example buildings per building type

Building type	**EUI [kWh/m^2]**	**Total floor area [m^2]**	**Energy demand [kWh]**
A	82	238	19,516
B	91	180	16,380
C	158	1,530	241,740
D	118	117	13,806
E	104	122	12,688

Adding up the above-presented estimations results in an aggregate heat demand of 304,130 kWh. Compared to the real aggregate consumption based on the measurements, an impressively small deviation of 1% is observed (Table 6). This outlines that even if the example buildings failed to represent the individual measured heat demands, they resulted in a very accurate result concerning the aggregate heat demand of the 16 single-family houses. However, the small range of the selected sample of the 16 houses does not allow general conclusions on the accuracy of the followed method.

Table 6 Comparison of aggregate heat demand results between measurements and archetypes

Calculation method	Aggregate heat demand [kWh]	Deviation
Measurements	300,858	1%
Archetypes-Example buildings	303,582	

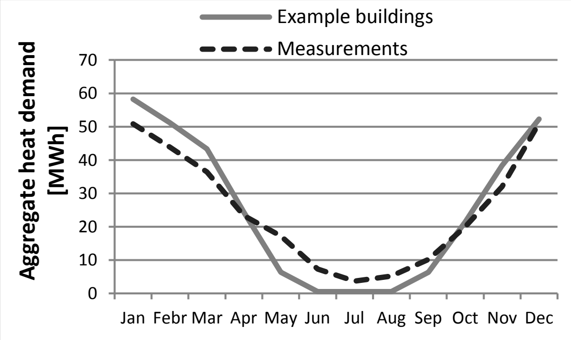


Figure 4 Aggregate heat demand results based on example buildings and measured values

Finally, Figure 5 illustrates the aggregate heat demand results on a monthly basis that were calculated based on the example buildings and the respective measured consumption. It is observed that overall, the aggregate heat demand of example buildings follows a similar trend to the real consumption in the 16 houses, as measured. The highest deviation between these two is noticed during the first three months of the year, as well as from May to August. During the summer period, this difference is attributed to the fact that Be10 does not consider any demand for space heating, independently of the ambient temperature. Even though the deviation between the annual heat demand results of example buildings and measurements is infinitesimal (Table 6), it is observed that their monthly demands do not coincide throughout the year (Figure 5).

CONCLUSION

According to the **first aggregation method**, data availability was found to be crucial. Thus, six scenarios based on different information levels were constructed and examined. It was found that the information level which is necessary to conduct building energy simulations includes BBR data. However, to be able to define the real U-values of the examined houses, it has to be known whether they have undergone any refurbishments that affect their energy performance. In particular, just adding the knowledge of the energy refurbishment state of each house to the basic scenario was found to improve the model's results by 24%. In addition, it was observed that the knowledge of the real occupancy loads from people did not affect the accuracy of the model's results very much. Furthermore, it was observed that benefits of energy renovations depend highly on the building's age. In particular, a similar energy refurbishment led to higher energy savings for the oldest buildings than for the newest ones. Thus, standard renovation measures can have impressive reductions of energy demand in older buildings.

The model that was found to represent the real heat consumption of the examined houses best was the one that included the highest information level -consisting of BBR data, realistic geometry, energy refurbishment state and occupancy loads- as expected.

According to the **second aggregation method**, the examined houses were classified into five building types based on their construction age, as indicated by TABULA. One example building was defined and modelled for each building type. It was found that the specific example buildings represented the individual houses' performance as modelled based on the highest information level quite well, but did not represent the real heat consumption as measured very well. The error increased significantly when the respective houses had undergone energy renovations, but the example building representing them did not include improved U-values. Thus, a different classification strategy may have been more effective, which would differentiate between renovated and not renovated buildings. In any case, the importance of knowing the exact energy refurbishment state of every investigated house was outlined in these results, as well. Moreover, on an annual basis, the calculated aggregate heat demand was found to be almost identical to the real consumption as measured. Furthermore, when the aggregate heat demand was studied on a monthly basis, many differences were noticed between the calculated heat demand of the archetypes and the measured one.

The results of the modelled heat demand demonstrated a deviation from the measured heat consumption. The sample of the building stock examined in the current study was very limited, covering only the category of single-family houses and a restricted number of building subcategories. Thus, it would be inappropriate to draw generalized conclusions. Extensive research is proposed focusing on a larger building population case to evaluate the observed trends and patterns. However, the deviation between modelled and measured demands may have been quite smaller, if a different energy simulation tool was used or integrated in Termite. In particular, dynamic energy simulation tools, such as IDA-ICE or EnergyPlus, may give the opportunity to model the houses with higher accuracy in terms of internal gains, climate data etc. The introduction of occupancy profiles can also lead to a better representation of the actual building energy demand. Furthermore, results will be in hourly values which enable the investigation of further parameters' effect. However, it increases data requirement both at the modelling level and data to define the reference of real measurements (based on hourly heat consumption data).

ACKNOWLEDGEMENT

This study is part of the Danish research project CITIES (Centre for IT-Intelligent Energy Systems in cities).

REFERENCES

Choudhary, R., 2011. A probabilistic model for assessing energy consumption of the non-domestic building stock. Sydney, Australia, Proceedings of 12th Conference of IBPSA: Building Simulation.

Heller, A., 2000. Demand Modelling for Central Heating Systems. Department of Buildings and Energy, Technical University of Denmark, Lyngby, Denmark. Report, no. R-040

Heller, A., 2002. Heat-load modelling for large systems. Applied Energy, Volume 72, pp. 371-387.

IEA, International Energy Agency, 2001. Annex 31 - Stock aggregation: methods for evaluation of the environmental performance, Ottawa, Canada: CMHC.

Matsuoka, A., Yamaguchi, Y., Suzuki, Y. & Shimoda, Y., 2013. Urban scale modelling of energy demand of retail facilities. Chambéry, France, Proceedings of 13th Conference of IBPSA: Building Simulation.

Negendahl, K., 2014. Parametric City Scale Energy Modeling - Perspectives on using Termite in city scaled models. Copenhagen, Denmark, iiESI European Workshop 2014, Technical University of Denmark, DOI:10.13140/RG.2.1.3624.5926

Wittchen, K. B. & Kragh, J., 2012. Danish building typologies. Participation in the TABULA project, Hørsholm, Denmark: SBi, Danish Building Research Institute, Aalborg University.

Paper V: Gianniou, P., Heller, A., & Rode, C. (2015). Building energy demand aggregation and simulation tools: a Danish case study. In Proceedings of *CISBAT 2015*, Lausanne, Switzerland.

BUILDING ENERGY DEMAND AGGREGATION AND SIMULATION TOOLS – A DANISH CASE STUDY

P. Gianniou; A. Heller; C. Rode

Department of Civil Engineering, Technical University of Denmark, Kgs. Lyngby

ABSTRACT

Nowadays, the minimization of energy consumption and the optimization of efficiency of the overall energy grid have been in the agenda of most national and international energy policies. At the same time, urbanization has put cities under the microscope towards achieving cost-effective energy savings due to their compact and highly dense form. Thus, accurate estimation of energy demand of cities is of high importance to policy-makers and energy planners. This calls for automated methods that can be easily expandable to higher levels of aggregation, ranging from clusters of buildings to neighbourhoods and cities. Buildings occupy a key place in the development of smart cities as they represent an important potential to integrate smart energy solutions. Building energy consumption affects significantly the performance of the entire energy network. Therefore, a realistic estimation of the aggregated building energy use will not only ensure security of supply but also enhance the stabilization of national energy balances.

In this study, the aggregation of building energy demand was investigated for a real case in Sønderborg, Denmark. Sixteen single-family houses -mainly built in the 1960s- were examined, all connected to the regional district heating network. The aggregation of building energy demands was carried out according to typologies, being represented by archetype buildings. These houses were modelled with dynamic energy simulation software and with a simplified simulation tool, which is based on monthly quasi-steady state calculations, using a visual parametric programming language (Grasshopper) coupled with a 3D design interface (Rhinoceros). The estimated heat demand of the examined houses from both simulation tools is compared to actual measured data of heat consumption. An assessment of the two different types of tools follows, which will indicate the suitability of each tool depending on the desired accuracy of results and on the purpose of analysis.

Keywords: Building energy, Heat demand, Archetypes, Energy Simulation Tools

INTRODUCTION

Aggregation of building energy demands is key to depicting national building stocks, while supporting urban development decisions and Smart Cities development. In particular, it is considered to be one of the most efficient methods for analysing stock performance due to its bottom-up approach [1]. The building sector within the European Union (EU) accounts for 35%-40% of the final total energy consumption and 25%-40% of the associated CO_2 emissions [2]. Thus, it can play a pivotal role on Smart Cities development. Especially the residential sector has received extended political attention and so, better statistics and knowledge exist for it, which facilitates its analysis. According to [3], building stocks can be represented by sample buildings and archetypes. Their difference lays in the fact that the former require knowledge about the actual building parameters and measurements, while the latter are statistical composites of features of a specific building type [2]. The methodology of archetype buildings can cope with the lack of physical description of buildings which is very common at a large scale.

Building energy simulation software is widely used for load modelling and building energy estimations. Simulation methods are categorized upon their energy balance calculation to dynamic and quasi-steady-state methods. According to dynamic methods, heat balance is calculated with short time-steps (e.g. hourly) considering the thermal inertia of the building. In particular, their main characteristic is that an instantaneous surplus of heat during the heating period results in an increase of the indoor temperature above the set-point, while it removes the surplus heat by extra transmission, ventilation and accumulation, if no mechanical cooling exists [4]. Furthermore, changes in room temperature highly depend on the heat stored in or released from the mass of the building. On the contrary, quasi-steady-state methods calculate heat balance for a long time (e.g. monthly/seasonal) considering dynamic effects through correlation factors. Regarding heating, a utilization factor is introduced, which considers that only a part of the internal and solar heat gains can reduce the heat demand, while the rest of the heat gains result in an increase of the room temperature above the set-point. In this study, both dynamic and quasi-steady-state energy simulation tools were used. The objective is to investigate their suitability and accuracy at an aggregated level of building energy demands, while using archetypes.

METHOD

Sixteen single-family houses located in Sønderborg, Denmark were investigated upon their heat demand. All of them were connected to the local District Heating (DH) network. No mechanical ventilation or auxiliary heating sources were installed in the houses. These were classified into five building types based on their construction age, following changes in building traditions and energy requirements according to the Danish Building Regulations [5]. Some characteristics of the examined buildings are illustrated in Table 1.

Building type	Construction Period	No. of incl. buildings	Total floor area [m^2]
A	1931-1950	2	238
B	1951-1960	2	180
C	1961-1972	10	1,530
D	1973-1978	1	117
E	1979-1998	1	122

Table 1: Characteristics of archetype buildings

One archetype building was created to represent each type. The characterization of the archetype buildings was based on information gathered by national building databases, TABULA Webtool and on statistical values applied by the Danish Building Research Institute on a national evaluation of the energy demand for the Danish building stock in total. This included information about the median floor and glazing area, U-values of the thermal envelope and glazing, internal gains, as well as infiltration. In particular, the window area of single-family houses is 16% compared to the floor area for those built in the 1930s [6], 22% for those built in the 1960s [6] and 15% for those built in the 1980s and 1990s [7]. Some houses had undergone usual energy refurbishments mainly affecting the U-values of the thermal envelope and glazing, which was taken into account in the building models.

The examined sixteen single-family houses were modelled initially with IDA-ICE. This required extensive information about the building parameters. Multi-zone models were created for all five archetypes according to the standardized parameters. Their space heat and domestic hot water (DHW) demand was simulated for one whole year. Afterwards, a simplified simulation tool was implemented, Termite, which required a much smaller amount of inputs in the models. Termite is a newly-developed parametric modelling tool, which uses Rhinoceros design interface, Grasshopper parametric options and Danish building performance simulation engine Be10 for energy performance calculations according to EN ISO 13790 [8]. Be10 calculates the energy needs of buildings in accordance with the energy requirements of the Danish Building Regulations BR10 and DS418, based on steady state monthly calculations [9]. For the calculation of energy consumption it uses only a whole building approach (single-zone building models). Be10 includes both static and non-static parameters in the energy performance calculation [10], but it does not enable the change of the input weather data, which are based on the Danish Design Reference Year (DRY) [11]. During the energy performance calculation with Termite, input lists were created to enable the fast modelling of all sixteen houses while using only one model setup. Analytical description of the model setup in Termite can be found in [12]. The results of both tools were compared to hourly heat consumption measured data obtained from the examined houses.

RESULTS

Figure 1 and Figure 2 present monthly heat demand (DH and DHW) as calculated by IDA-ICE and Termite for the five building types, being represented by the respective archetype buildings. These are compared to the measured data of heat consumption, which were averaged for each building type.

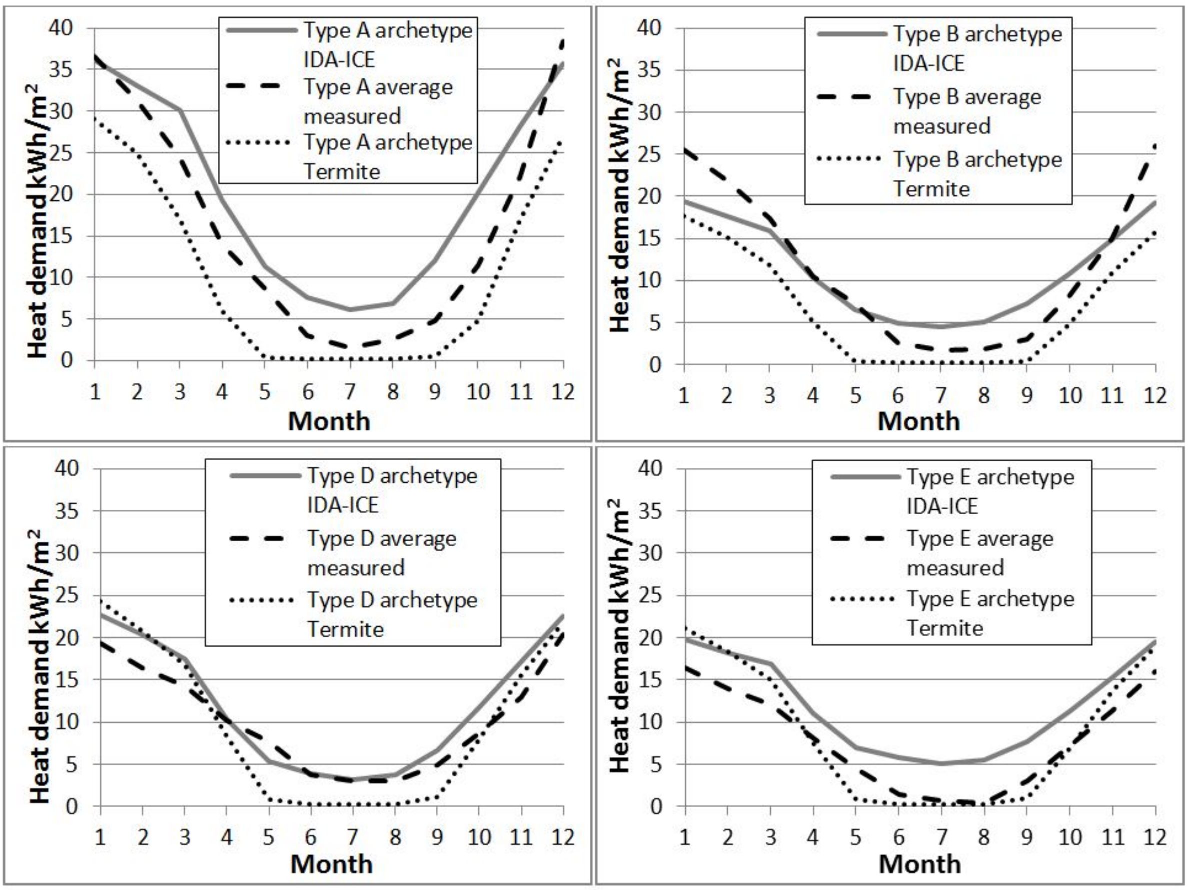


Figure 1: IDA-ICE and Termite monthly results of heat demand for building Type A, B, D and E compared to average measured data

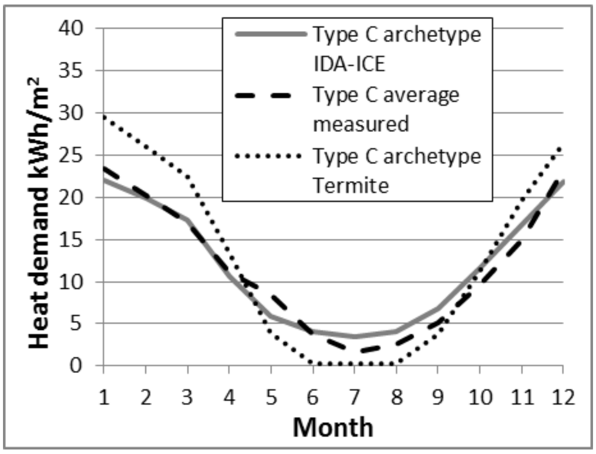


Figure 2: IDA-ICE monthly results of heat demand for building Type C (most representative) compared to average measured data

The calculated heat demands presented in Figure 1 and Figure 2 generally follow similar trends with the measured consumption for all types. The heat demand of the houses reduces as the construction age decreases, as expected. Type C, presented in Figure 2, is considered to be the most representative since the highest number of the examined houses are classified to this type. The rest building types consist only of one or two houses. The deviation between IDA-ICE heat demand and measured heat consumption is quite low for Type C. On the other hand, Termite heat demand differentiates more from the measured heat consumption during both winter and summer period for this type. This may be attributed to the fact that Be10, which is the energy simulation core of Termite, uses DRY climate data that are characterized by very low temperatures in winter and very high temperatures in summer. On the contrary, IDA-ICE makes use of real weather data. The estimated DHW demand by both simulation tools was not estimated accurately, which is due to no information about occupant behaviour.

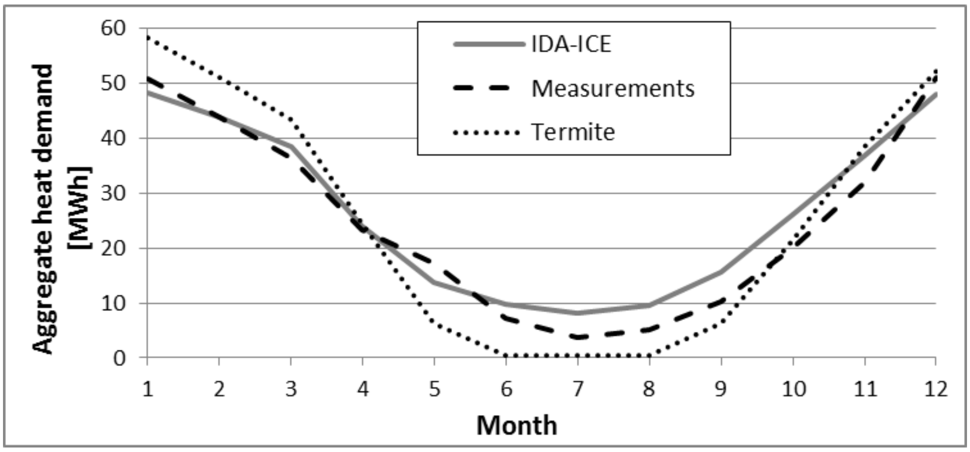


Figure 3: Monthly aggregated heat demand results from IDA-ICE and Termite compared to measured data

Afterwards, the monthly Energy Use Intensity (EUI) [kWh/m^2] estimated by IDA-ICE and Termite for each archetype building was multiplied by the total floor area of all houses classified in the specific type. This aggregated heat demand is illustrated in Figure 3 as calculated by IDA-ICE and Termite for a whole year and is compared to the measured total heat consumption for the specific year. It is observed that IDA-ICE led to slightly more accurate estimation of the heat demand at an aggregated level compared to Termite. Especially during the winter period, IDA-ICE seems to result in more realistic calculations than Termite. However, it still does not match fully the measured aggregated heat consumption.

DISCUSSION

The lack of information about the building physical properties as well as the occupants' behaviour is a common problem met at large-scale analyses. This was overcome to some extent through the use of archetypes. Even though the sample of the examined houses was small, the methodology described in the present study aims at city-scale projects. For that reason, the values of many building parameters were assumed according to statistical data and national building databases. However, information about the occupants of the houses was not known which would be useful to draw realistically their profiles. The uncertainty in describing physically the examined houses affected the accuracy of results of both IDA-ICE and Termite. Furthermore, the monthly time-step was selected to enable the comparison between these two energy simulation software. Thus, the ability of IDA-ICE to produce hourly results was not exploited. If a yearly time-step had been selected instead, the results would be quite different, but no detailed analysis could be made. Type C was the most representative one, as it included the highest number of houses. The rest of the building types including only one or two houses could not lead to very representative archetypes.

CONCLUSION

The reported results indicated that the estimation of heat demand of clusters of building is a demanding procedure, where uncertainty plays an important role. The latter affected both simple building models, which require fewer inputs and more advanced models, where the number of parameters increases and therefore, the demand for detailed insight into the building construction and user behaviour increases. At aggregated level, it was found that IDA-ICE archetype models represented slightly better the heat demand of the sixteen single-family houses than Termite. Even though the energy calculations conducted by IDA-ICE were more advanced than Termite, the uncertainty in several model parameters counterbalanced its accuracy. However, its estimation of heat demand still remained the closest to the actual measured heat consumption.

Moreover, it was expected that the application of standardized parameter values would lead to rather exact results which was not the case for the current very limited sample of buildings. This could be attributed to the fact that the current buildings were not representative of Danish buildings, or it could be due to the fact that the variability is a dominating factor. The next step of the present work would be to expand the sample of examined buildings, which likely will lead to a more representative case.

The suitability of simulation tools depends highly on the purpose of analysis. If the purpose of this work was to study energy demand shifting in residential buildings, then dynamic simulation software would be the only way to go, since high resolution results would be needed. As the scale moves from building level to district or city level, simplified energy simulation tools may be seen as more preferable. Nevertheless, the methodology of building typologies and archetypes enables the use of advanced dynamic simulation tools, which in any another case would have vast and thus, restrictive computation times. Lastly, the combination of quasi-steady-state and dynamic energy simulation methods could be the solution towards fitting measured consumption data.

ACKNOWLEDGEMENTS

This study has been a part of the Danish strategic research project CITIES (Centre for IT-Intelligent Energy Systems in cities).

REFERENCES

1. International Energy Agency: Annex 31 - Stock aggregation: methods for evaluation of the environmental performance. Ottawa, Canada: CMHC, 2001.
2. Mata, É., Sasic Kalagasidis, A., Johnsson, F.: Building-stock aggregation through archetype buildings : France, Germany, Spain and the UK. Building and Environment, Vol 81, pp. 270-282, 2014.
3. Swan, L. G. & Ugursal, V. I.: Modeling of end-use energy consumption in the residential sector: A review of modeling techniques. Renewable and Sustainable Energy Reviews, 13(8), pp. 1819-1835.
4. ISO13790: Energy performance of buildings - Calculation of energy use for space heating and cooling. CEN, European Committee for Standardization, Brussels, 2008.
5. Wittchen, K. B., Kragh, J.: Danish building typologies. Participation in the TABULA project, SBi, Danish Building Research Institute, Aalborg University, Hørsholm, 2012.
6. Aggerholm, S.: Cost-optimal levels of minimum energy performance requirements in the Danish Building regulations. Danish Building Research Institute, SBI 2013:25, 2013.
7. The Danish Ministry of Housing: Building Regulations 1982, BR82. The National Building Agency, Copenhagen, 1983.
8. Negendahl, K.: Parametric City Scale Energy Modeling - Perspectives on using Termite in city scaled models. iiESI European Workshop 2014, DTU, Copenhagen, 2014.
9. Danish Building Research Institute: SBi anvisning 213 - Bygningers energibehov.
10. Christensen, J. E., Schiønning, P., Dethlefsen, E.: Comparison of simplified and advanced building simulation tool with measured data. Proceedings of BS2013:13th Conference of International Building Performance Simulation Association, Chambéry, 2013.
11. Danish Ministry of Climate, Energy and Building: 2001-2010 Danish Design Reference Year. Technical Report 13-19. Danish Meteorological Institute, Copenhagen 2013.
12. Gianniou, P., Heller, A., Sieverts Nielsen, P., Negendahl, K., Rode, C.: Aggregation of building energy demands for city-scale models, 2015 (submitted).

Paper VI: Gianniou, P., Foteinaki, K., Heller, A., & Rode, C. (2017). Intelligent Scheduling of a Grid-Connected Heat Pump in a Danish Detached House. In Proceedings of *Building Simulation 2017*, San Francisco, United States.

Intelligent Scheduling of a Grid-Connected Heat Pump in a Danish Detached House

Panagiota Gianniou[1], Kyriaki Foteinaki[1], Alfred Heller[1], Carsten Rode[1]
[1]Department of Civil Engineering, DTU, Kgs. Lyngby, Denmark
Contact: pagian@byg.dtu.dk

Abstract

This study proposes a methodology for intelligent scheduling of a heat pump installed in a refurbished grid-connected detached house in Denmark. This scheduling is conducted through the coupling of a dynamic building simulation tool with an optimization tool. The optimization of the operation of the system is based on a price-signal considering a three-day period for different weather cases. The results show that the optimal scheduling of the system is successful in terms of reducing the peak load during times when electricity prices are high, thus achieving cost savings as well as maintaining good thermal comfort conditions. The proposed methodology bridges dynamic building modelling with optimization of real-time operation of HVAC systems offering a detailed model for building physics, especially regarding thermal mass and a stochastic price-based control.

Introduction

As the share of renewable energy sources (RES) in power generation is constantly increasing in many countries, imbalances arise between the supply and demand side. Specifically for Denmark, renewable energy is expected to cover about 80-85% of electricity consumption and up to 65% of district heating in 2020 (Danish Energy Agency, 2015). One of the main goals is that the entire energy supply is covered by RES by 2050. That calls for demand-side management (DSM) approaches that can facilitate the operation of the smart grid while enabling controllability of the electric loads.

The potential of buildings' participation in DSM approaches has been investigated, implementing strategies to reduce power consumption during peak periods (peak-shaving) and/or to shift the power consumption from peak periods to off-peak periods (load-shifting). There is a number of studies indicating the importance of the structural thermal mass of buildings, for example Reynders et al. (2013), while some others implemented DSM by adjusting the HVAC systems of the building, for example Arteconi et al. (2013) and Masy et al. (2015). There are also a few studies that tried to reschedule the operation of shiftable plug loads, such as dishwashers, washing machines and tumble dryers, as in D'hulst et al. (2015).

The two main types of control are direct control and indirect or price-based control. Schibuola et al. (2015) concluded that HVAC systems will not undergo direct control in future smart grids due to discomfort issues that this might bring to the occupants. Thus, price response strategies will be the main focus of this study. Particularly in countries like Denmark, where power generation achieved by RES accounted for 25% of the adjusted gross energy consumption, which describes total observed energy consumption adjusted for fluctuations in climate with respect to a normal weather year (Danish Energy Agency, 2015), this leads to very low electricity spot prices or even negative ones. Thus, price-based control can achieve peak-shaving that is much needed by the grid to ensure decrease or displacement of peak loads.

Computational methods of design optimization have proven to be advantageous in several building studies according to Evins (2013). Optimization algorithms are mainly categorized into heuristic and meta-heuristic when it comes to building co-simulation. Heuristic algorithms include direct search such as pattern search and linear or non-linear programming. Meta-heuristic algorithms consist of evolutionary algorithms such as genetic algorithms (GA) or particle swarm optimization (PSO). So far, optimization in building modelling has been used with regards to building envelope, systems, energy generation and holistic approaches considering many aspects in building operation (Evins, 2013). Extensive studies have focused on optimizing the building envelope and dimensioning of the systems mostly during the design phase. The optimization of the control of the HVAC systems has been implemented in some studies. Particularly Zhou et al. (2003) investigated the minimization of electricity cost of varying cooling set-points for different algorithms and concluded on a most suitable one. Miara et al. (2014) investigated how heat pumps provided with heat storage and floor heating system may take advantage of real-time pricing. The authors determined that the most important parameter to ensure this was the water heat storage tank. However, the flexibility that thermal mass, which is embedded into the building structure, can offer along with demand response management of the heating systems, have not been thoroughly discussed. Especially when using reduced order models or grey-box models, many optimization techniques and algorithms can ensure their feasibility. Their application has not been extensively highlighted though, in cases of flexible operation of HVAC systems in residential buildings. Particularly in the case of white-box building models, which describe building physics and heat transfer mechanisms in full detail, little effort has been made to use them as basis for optimal scheduling of the HVAC systems based on electricity pricing. This is mainly due to their complexity, which impedes their coupling with advanced optimization tools. However, the potential that the thermal mass of the building gives with regards to energy flexibility is crucial to such studies and should be modelled with every detail possible.

The current study aims at developing a methodology for optimally scheduling the systems of a grid-connected single-family house equipped with an air-to-water heat pump and low temperature radiators based on a price response strategy. This strategy will indicate the potential of the systems for optimization, which will define a generic methodology while enabling the utilization of the building's thermal mass that can be applied in every residential building with similar HVAC systems.

Model description

The current analysis was conducted by use of a building model. This corresponds to a typical Danish single-family house (SFH) built between 1961-1972 (Figure 1), which is the most common type of SFH in Denmark according to Danmarks Statistik (2016). The properties of the building envelope and systems have been created according to TABULA database (2016). The building area is 153m^2, while the glazing covers an area of 34m^2. Its building envelope has been extensively renovated and the average U-value of the building is 0.27 W/m^{2}K. The material layers that consist the building components can be seen in Table 1. The house contains an air-to-water heat pump (HP) of 13 kW with COP of 3.5. The heat emission system installed is a hydronic one with low-temperature water radiators. The controller of the hydronic system is a thermostat set at 21ºC with 2ºC deadband based on the operative temperature.

Table 1: Material layers

Building component	Layer material	Thickness (m)	U-value (W/m^{2}K)
External walls	Brick Light insulation Brick	0.05 0.25 0.05	0.13
Roof	Light insulation Wood Gypsum	0.34 0.03 0.013	0.10
Floor	Wood Light insulation Gypsum	0.02 0.28 0.02	0.12

Deterministic profiles were created for the internal gains. The occupants living in the house were assumed to be two according to data from Danish Statistics (Danmarks Statistik, 2016) following a typical house living profile (Figure 2) and with sedentary activity level (met-value=1.2) according to DS/EN 15251 (2007). Their clothing was variable ranging from 1 during the winter season to 0.5 for the summer season. The schedules of equipment and lighting can also be found in Figure 2. Internal blinds were drawn in cases of excessive solar radiation. It was also assumed that windows started to open when indoor temperature reaches 25ºC and opened fully at 27ºC. There was no mechanical ventilation or cooling installed in the building. The infiltration rate was 0.2 ACH representing a quite air-tight refurbished envelope. The location was set to Copenhagen, Denmark. The main façade of the house was assumed to be oriented towards the south. The weather file used was the Danish Design Reference Year (DRY) (DMI, 2013).

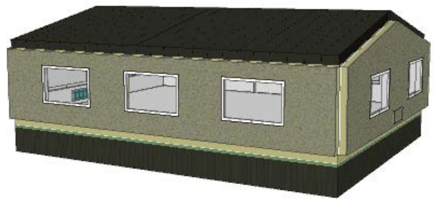


Figure 1: Design of examined SFH

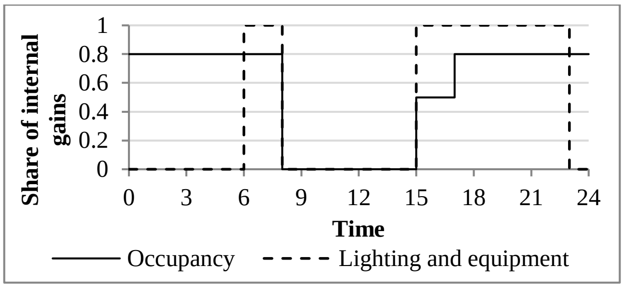


Figure 2: Daily schedule for internal gains

Since emphasis will be placed on the utilization of the thermal mass of the building in the present study, the heat transfer mechanisms taking place inside the building structure are important. The wall model used is a finite difference model of multi-layered components. Each material layer is discretized into four nodes. Thus, it can provide accuracy when alternative models, such as RC-networks, fail to. Furthermore, nonlinear effects in the thermal dynamics of buildings, which are usually oversimplified in RC-networks (Thavlov and Bindner, 2015), (Zong et al., 2017), are described in detail in the current model.

Methodology

In this paper, a price-based control strategy for the management of the heat pump was implemented that enabled the exploitation of dynamic electricity prices and thus management of the energy load according to the grid's imbalances. In this way, the interaction of the detached house with the grid was improved. The scheduling of the HP comprised two parts: the control of the system, and its optimization.

The scheduling of the heat pump was carried out considering an advanced controller for the radiators. This can be seen in Figure 3. The base heating setpoint was set to 20ºC. An additional controller, which contained an algorithm to smooth its output as an approximation to a P-controller so that no events (abrupt turn-ons and offs) were created, was added in the loop to allow for the system's flexibility when the electricity prices were low. Therefore, the operation of the HP was forced to increase during low-tariff hours. The hourly tariff concept was assumed to reflect the lack of or excess of RES energy in the grid and thus represent the stress on the grid (Dar et al., 2014). The low electricity prices were defined as the ones that were lower than the average electricity price of the specific month. This means that when the electricity price was low, a positive deadband of 3ºC was added to the base setpoint, resulting in 23ºC upper threshold, which allowed the system to take advantage of the low prices. In this way, the increase in the heating setpoint would enable heat to be stored into the thermal mass during low price periods, and be released back to the room when prices were higher. The electricity price signal was imported into the model as a source file connected to the P-controller. The proportional band of the P-controller was set to -0.5 as this referred to heating mode. The capability of the system to keep the indoor environment within comfort conditions was a significant advantage of the selected control strategy. According to EN/DS 15251 (2007), Category II of acceptable thermal comfort ranges from 20-25ºC.

The optimization was defined as minimization of the operating cost of the HP. This cost depended on the variable electricity prices and on the consumption of the HP. The cost-optimal control of the heat pump on a 3-day horizon would ensure cost savings in the electricity bill along with the desired peak shavings in the heat load of the building. Also, the 3-day horizon was selected so that any phenomena of cumulative heat storage into the thermal mass can be observed. Due to increased computation time, it was not selected to investigate a longer period. It has to be clarified that no electricity was assumed to be sold back to the grid or modelled in this case for simplification reasons.

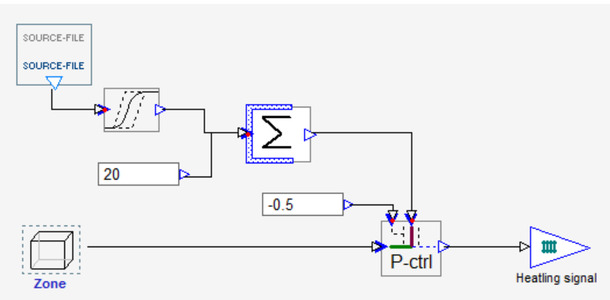


Figure 3: Control of heating system based on price signal

The optimization problem of the intelligent scheduling of the heat pump according to the price signal was formulated as following:

$$Minimize\ C_{HP} \tag{1}$$

where C_{HP} is the operating cost of the HP, calculated as following:

$$C_{HP}=p_{el} \cdot W \tag{2}$$

where p_{el} is the total electricity price and W is the energy consumption of the HP.

The optimization of the system operation was conducted through the boiler schedule which was characterized by ten variables per day, b_i, corresponding to the schedule of the part load operation of the HP. This means that a different schedule of the HP operation was optimized for each day.

$$b_i \in [0,1]\ for\ i = 1 \ldots 30 \tag{3}$$

The optimization was conducted with the use of the open-source software MOBO. MOBO is a generic freeware designed to handle single-objective and multi-objective optimization problems with continuous or discrete variables (Rosli et al., 2016). It provides the possibility to be coupled to building performance simulation tools, while selecting different algorithms to perform the calculations. Furthermore, MOBO has the feature of running multiple simulations in parallel, thus reducing the overall computation time with a factor equal to the number of threads available in the computer. In this case, different optimization algorithms were selected, and they were compared upon the optimal solution and the number of runs. In particular, a deterministic algorithm (Hooke-Jeeves), two genetic algorithms (NSGA-II, Omni-optimizer) and a hybrid one (GA and Hooke-Jeeves) were tested. As defined in Wetter and Wright (2004), GA are algorithms that operate on a finite set of points, called a population. The different populations are called generations. They are derived on the principles of natural selection and incorporate operators for fitness assignment, selection of points for recombination, recombination of points, and mutation of a point. The simple GA iterates either until a user-specified number of generations is exceeded, or until all iterates of the current generation have the same cost function value (Wetter and Wright, 2004). Initially, the optimization algorithms were implemented having the same solver settings. These were 8 populations, 20 generations, 0.05 mutation probability and 0.9 cross-over probability. The meaning of the cross-over probability is that when this is not met, the individuals do not continue to the new population as explained in Evins et al. (2010). The crossover process continues until the new population is full. The reason for using a small population size was to reduce computation time. The deterministic algorithm that was tested does not enable parallel computing, thus being significantly slower. The settings of the deterministic algorithm were ρ=0.05 (convergence parameter) and ε=0.01 (minimum criterion to stop the search).

Furthermore, different solver settings were applied to determine their effect on achieving the optimal solution. Based on the cost-optimal solution and the number of simulations carried out until convergence was reached, one optimization algorithm was selected and tested upon its accuracy for various numbers of generations. Then, the solver settings that reached the biggest cost reduction were selected and implemented into the building model. The schedule for the boiler/HP operation was thus defined according to the optimal solution. These results were compared to the ones of the reference case with normal operation of the HP, as previously described in the model description.

Simulation

The simulation was run in IDA ICE Version 4.7 (EQUA, 2013) using MOBO. The building was simulated as a single-zone model. The reference model was initially run with the DRY weather file. Three climate cases were selected out of the entire simulation period, the coldest 3-day period of the heating season (January) with a mean ambient temperature of -4^{o}C, the warmest 3-day period of the heating season (September) with a mean ambient temperature of 11^{o}C and an intermediate one (April) when the average ambient temperature was 3^{o}C.

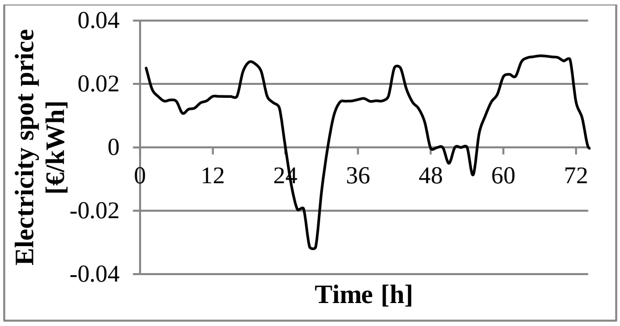


Figure 4: Electricity spot prices over 3 days

The prices for electricity were set in the model through a stochastic profile reflecting real electricity prices in 2015 according to the Nordic electricity market Nord Pool for East Denmark. During this period, wind power generation accounted for 74% of the total electricity production, which explains the low electricity prices that comprise the price signal that was coupled to the IDA ICE model. These prices reflect the total electricity prices consisting of variable el-spot prices (commercial), which account for 32% of the total price in Denmark, while the remaining 68% are fixed taxes for local network, grid and system tariffs, public service obligation tariff and further subscriptions to electrical companies based on the Association of Danish Energy Companies (2015). The total electricity prices were used in the model so that they correspond to the ones that electricity customers have to pay. The same price-signal was used for all three cases of weather data to ensure comparability, so it has not been correlated to the climate data. Figure 4 presents the electricity spot prices during this 3-day period. The negative prices indicate the surplus of wind power generation during these days.

Result analysis and discussion

As mentioned before, different optimization algorithms were tested in MOBO in order to determine which one to select based on the biggest cost reduction they could achieve. The selected genetic and deterministic algorithms were implemented and tested for the three different weather data having the same solver settings, as described in the Methodology section. The implemented algorithms are presented in Table 2. A detailed explanation of all the algorithm parameters is beyond the scope of this paper and we refer the readers to Wetter and Wright (2003). The cost-optimal solution refers to the minimum cost (€) that each optimization algorithm achieved which means the total operating cost of the HP needed to heat up the building for the examined 3 days.

Due to the increased number of variables used in the optimization problem, some discontinuities could be expected in the cost function which would make optimization rather difficult to be achieved (Wetter, 2004). For this reason, different optimization algorithms were tested, some of which (i.e. Hooke-Jeeves) are less likely to converge to a discontinuity far from the optimal solution (Wetter, 2011). Moreover, the total cost reduction observed might be limited for the above-described reasons and is not assessed per se but in comparison with the rest of the results. Furthermore, some uncertainty in the results has to be considered due to randomness of the optimization. The hybrid algorithm, GA and Hooke-Jeeves, proved to be the best one among the rest, even though the differences were very small, calculating the biggest cost reduction in all three weather cases. It has to be noticed that the hybrid algorithm led to a higher number of simulations, due to which it converged to a lower cost-optimal solution. Thus, it was selected to run some further analysis on the impact that the number of generations has on achieving the optimal solution.

Table 2: Results of different optimization algorithms

Optimization algorithm	Type of algorithm	Cost-optimal solution [€]	Simulations until convergence (output)	Weather data
NSGA-II	Genetic (evolutionary)	1.802	144	Cold
		0.787	152	Intermediate
		0.008	152	Warm
Hooke-Jeeves	Deterministic (pattern-search)	1.775	221	Cold
		0.778	166	Intermediate
		0.008	60	Warm
GA and Hooke-Jeeves	Hybrid	1.774	311	Cold
		0.778	363	Intermediate
		0.007	311	Warm
Omni-optimizer	Genetic (evolutionary)	1.806	152	Cold
		0.780	152	Intermediate
		0.007	152	Warm

Three numbers of generations were applied to the hybrid algorithm being comprised of GA and Hooke-Jeeves, and are presented in Table 3 along with the number of simulations and the solution they resulted in. It can be observed that the higher the number of generations and the more simulations conducted, the bigger cost reduction was achieved, as it was expected. Therefore, the optimal solution found with 100 generations for each weather case was used for the following analysis. However, it has to be pointed out that even though a scenario of 100 generations leads to approximately twice as many simulations as the one of 50 generations, the cost reduction is not significantly bigger. Specifically, this was 2% for the cold and intermediate weather data and 1% for the warm case, (Table 3). Therefore, it was decided not to investigate an even bigger number of generations.

Table 3: Impact of generations on optimal solution

GA and Hooke-Jeeves	**Cost-optimal solution [€] (Number of simulations)**		
Generations	Cold	Intermediate	Warm
25	1.7741 (311)	0.7778 (363)	0.0075 (311)
50	1.7727 (512)	0.7771 (568)	0.0074 (800)
100	1.7418 (913)	0.7612 (909)	0.0073 (1491)

Figure 5 shows the power consumption for the reference and optimized case for each of the weather cases. As it was expected, the pattern is different for the different climate cases examined. For the cold weather, when using the thermostat as in the reference case, the HP turned on and off in very small intervals, in order to cover the high demand, which results in high power peaks. On the other hand, for the optimal case, the power consumption presented a considerably more smooth pattern, which was fully in line with the price signal. This is also attributed to the smoothing effect that the controller applied in the optimal scenario has. There were small parts of the load that were shifted towards a different timing than in the reference case, but this schedule mainly achieved considerable peak shaving. It is clear that the optimization was conducted successfully, as during times with low electricity prices, the HP was forced to increase its part load and vice versa. For the intermediate climate case, it can be seen that the power consumption pattern is similar for the reference and optimal cases. The time when the power consumption peaks is almost identical, but still there was peak shaving achieved with the optimal case and the operation pattern of the HP is more smooth. This means that continuous on/offs are avoided, which would lead to a decrease in HP's lifetime and inefficient operation. For the warm climate case, it is obvious that the energy demand is amost zero, so the difference between the two cases is marginal. This minimum heat load left no room for the optimization to take place.

Table 4. Total energy consumption results

Weather data	**Total energy consumption (kWh)**	
	Reference case	**Optimal case**
Cold	191	180
Intermediate	92	104
Warm	8.80E-05	8.70E-05

Regarding the total energy consumption, the results were different for the different climate cases. These are presented in Table 4. For the cold climate, there was an energy consumption reduction of 6% with the optimal case, while for the intermediate climate there was a 13% increase for the optimal case. This indicates that when implementing a DSM strategy, i.e. peak shaving, this might result in an overall over-consumption. However, the goal of this strategy is to use the energy produced from RES, which is, in this study, reflected in the low prices. So, the focus is not to achieve energy minimization, but to "force" the operation of the HP to follow the production pattern of renewable energy in the energy system, while maintaining the thermal comfort and achieving cost savings for the occupants. As previously mentioned, the heat demand for the warm climate is minor, so there is only 1% decrease in the total energy consumption after the optimal case was applied.

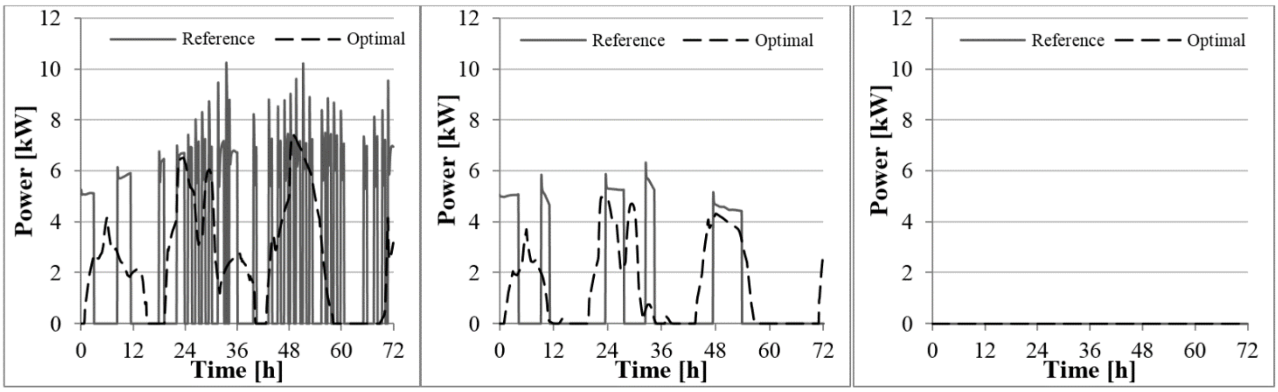


Figure 5: Power results for the optimized scheduling of the HP and the reference operation for cold weather data (left), intermediate (middle) and warm (right)

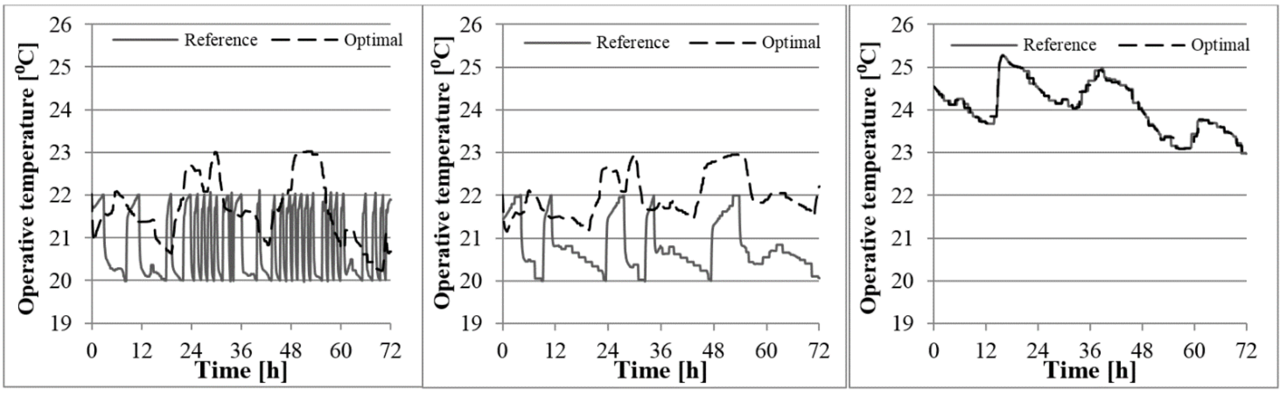


Figure 6: Operative temperature results for the optimized scheduling of the HP and the reference operation for cold weather data (left), intermediate (middle) and warm (right)

As far as cost reduction is concerned, the results varied, too, with regards to the different climate cases and can be seen in Table 5. In the cold weather case, optimally scheduling the HP resulted in 42% decrease in the total operating cost of the HP for the examined 3-day period. In the intermediate case, the cost savings were 22% compared to the reference operation of the HP, while in the warm weather case these were significantly lower, approximately 13%, as it was expected.

Table 5. Operating cost results of the HP for 3 days

Weather data	Total operating cost (€)	
	Reference case	Optimal case
Cold	2.998	1.742
Intermediate	0.979	0.761
Warm	0.008	0.007

Regarding peak shavings, which is the main goal of this demand side management approach, the results showed a clear decrease in peak consumption at the optimal case both for the cold and the intermediate weather. Taking the highest peak in power for each case, there was a reduction of 27% (2.75 kW) for the cold weather and 21% (1.32 kW) for the intermediate weather (Figure 5). This was achieved due to the smoother operation pattern of the heat pump that shifted parts of the load in time, which led to reduced peak demand. The magnitude of the potential peak decrease would always depend on the price signal, but the main outcome is that the implemented control is an effective peak shaving strategy.

Figure 6 presents the operative temperature for the reference and optimized cases for each of the weather cases. It can be observed that there is a trend for the optimal cases to have higher operative temperature than the respective of the reference cases at all climate cases. This can be explained by the control of the heating system, which was able to increase the setpoint up to 23°C when the electricity prices were low. It has to be pointed out that the electricity prices according to which the optimal scheduling of the HP was conducted in the three weather cases were found to be low compared to the average monthly price, which resulted in utilization of the increased heating setpoint most of the time. If a different threshold was used instead of the monthly average electricity price, the price signal might not be characterized as low most of the time and the results would be somewhat different as the decision to add the deadband to the base setpoint would alter. The lowest threshold of the thermal setpoint (considering the given deadband) was set to be equal for the reference and the optimal cases (20°C), so that they both retain similar thermal comfort conditions. Furthermore, there is a clear correlation between the power consumption pattern and the operative temperature pattern. This explains why the temperature of the reference case of the cold climate had frequent fluctuations, whereas the temperature of the optimal case of the cold climate and both cases of the intermediate climate had a smooth pattern. Due to the cold external temperatures that reached -18°C, the HP was forced to switch on and off very frequently creating high peaks in order to maintain the desired indoor temperature. In the intermediate weather case, the significantly higher ambient temperatures allowed a more flexible operation of the HP. In the warm climate case, there was almost no difference observed in the operative temperatures between the optimal and the reference case since the HP operated for a very short time. In all cases, the operative temperature stayed within the acceptable limits of EN/DS 15251 ensuring that with the applied control and cost optimization, the thermal comfort for the occupants was not compromised.

This can be verified by the amount of occupancy hours into the different comfort categories, which were in all cases within the acceptable categories according to EN/DS 15251, as it can be seen in Table 6. In particular, Table 6 indicates the cumulative share of occupancy hours that fell into the three comfort categories according to EN/DS 15251. Category IV is not presented since no occupancy hours belonged to that. The optimized scheduling of the HP increased the amount of occupancy hours belonging to comfort Category I for the cold and intermediate weather case, while slightly decreased these for the warm weather case. Overall, the optimized operation of the HP maintained the good thermal comfort corresponding to the design conditions of the heating system in all three weather cases.

Table 6: Thermal comfort assessment for optimal scheduling of HP

Weather data	**Cold**		**Intermediate**		**Warm**	
Case	Opt.	Ref.	Opt.	Ref.	Opt.	Ref.
Comfort category	Share of occupancy hours (%)					
I (best) (21-25ºC)	71	40	100	26	78	100
II (good) (20-25ºC)	100	98	100	100	100	100
III (acceptable) (18-25ºC)	100	100	100	100	100	100

Summing up, the optimization problem has been solved using MOBO coupled with IDA ICE. The simulations were performed on a computer with Intel i7 4-core processor clocking at 2.1 GHz and 8 GB of RAM. The average total time of solving was 7250 seconds for the cold and warm weather case and 7460 seconds for the intermediate one. In this specific case, the optimization problem was solved for a three-day period.

It has to be pointed out that the goal of this study was the proposal of a methodology to schedule the HP operation intelligently using building performance simulation tools according to the grid's imbalances that are reflected on the price signal and not the investigation of optimization techniques. In addition to this, the complicated building model and the large number of continuous variables used in the optimization problem increased significantly the computation time. On the other hand, the advanced building model that was used modelled accurately the thermal building physics mainly of the thermal mass and the complicated heat transfer mechanisms that take place during the charging or discharging of the thermal mass. These were critical to estimate the flexible operation of the HP.

Due to increased computation time, it was chosen to investigate a three-day design period for each weather case instead of a whole- year simulation that would be more representative. The optimization horizon was not restricted to one day, since the authors wanted to investigate any heat storage mechanisms into the thermal mass of the building, which are cumulative over time. Furthermore, the proposed methodology refers to future energy market designs, where the optimization of individual buildings' heating or cooling systems based on price signals will be possible. In this case, day-ahead prices will be available, so the optimization problem will be solved only for the following day, which will result in shorter computation time. However, attention should be paid to the initial and final conditions of the thermal model, so that they are consistent with the ones of the previous and following day.

Computationally expensive simulations also led to tight solver settings (small number of populations and low mutation probability), which increased the chances of not converging to a stationary minimum point of the cost function.

The positive 3ºC deadband provided to the base heating setpoint that allowed for the system flexibility was selected such that the thermal comfort of occupants was not compromised with the optimized scheduling of the heating system. If a looser comfort threshold was to be achieved, this deadband could be further expanded so that lower operative temperatures than 20ºC should be allowed. Furthermore, different price signals could have been applied to the case so that an estimate of the range of peak and power savings could be made and also a realistic correlation to the weather data would be possible. Lastly, the use of historical weather data for the examined period would have resulted in much more clear optimization results and would have avoided uncertainty. Hence, they ought to replace the DRY weather file as a next step to this study.

Conclusion

In this study, the cost optimization of the control of a HP in a detached Danish house was investigated. A dynamic building performance simulation tool was coupled to an optimization software. A hybrid genetic algorithm was selected as the most suitable one to achieve the biggest cost reduction of the operation of the HP. The optimal scheduling of the HP that was achieved based on a price signal for three weather data was compared to the reference case, which was normal operation of the HP system. The comparison was made with regards to power peak shavings, cost savings, operative temperature and thermal comfort. The results showed that the optimization was conducted successfully, as the price-based control managed to reduce the peak loads during high price times and increase the energy load during the low price times. This resulted in a smoother pattern for the operation of the HP for the cold and intermediate weather data. In the warm weather case, the very low heating demand left no room for optimization. Furthermore, the optimal scheduling of the HP maintained good thermal comfort of the occupants with regards to the operative temperature. The operative temperatures during the optimal scheduling of the system were maintained in a good comfort category according to EN/DS 15251. Even though the detailed building model that was used increased the number of variables in the optimization problem and the computation time, it modelled the heat dynamics into the thermal mass sufficiently well, which allowed for the flexible operation of the HP.

Finally, the three selected weather cases represented different climatic conditions during a heating season in Denmark. The cold and the intermediate weather cases proved to be representative of extreme and mild winter weather conditions and enabled the optimization of the HP. The warm weather data could not be considered for the optimal scheduling of the HP, since the heat demand was very low. Therefore, should a holistic overview of the intelligent scheduling of a HP in a similar building and climate be given, a combination of the two aforementioned weather data has to be considered.

Acknowledgements

This study is part of the Danish research project CITIES (Centre for IT-Intelligent Energy Systems in cities) and EnergyLab Nordhavn. The authors would like to thank Mika Vuolle, Erkki Karjalainen and Ala Hasan for their useful comments on MOBO and Rune Korsholm Andersen for his help regarding the implementation in IDA ICE.

References

Arteconi, A., Hewitt, N. J., Polonara, F. (2013). Domestic demand-side management (DSM): Role of heat pumps and thermal energy storage (TES) systems. *Applied Thermal Engineering*, *51*(1–2), 155–165.

Association of Danish Energy Companies (2015). Electric Utility tariffs and electricity prices (Danish title: Dansk Energi. Elforsyningens tariffer & elpriser)

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

Danish Energy Agency (2015). Danish energy and climate outlook. Published in December 2015 by the Danish Energy Agency, Copenhagen, Denmark.

Danmarks Statistik (accessed 2nd December 2016). https://www.statbank.dk/statbank5a/default.asp?w=1920

Dar, U. I., Sartori, I., Georges, L., Novakovic, V. (2014). Advanced control of heat pumps for improved flexibility of Net-ZEB towards the grid. *Energy and Buildings*, *69*, 74–84.

DMI, Danish Meteorological Institute (2013). Danish Ministry of Climate, Energy and Building: 2001-2010 Danish Design Reference Year. Technical Report 13-19.

EN/DS 15251 (2007). Indoor environmental input parameters for design and assessment of energy performance of buildings addressing indoor air quality, thermal environment, lighting and acoustics.

Evins, R. (2013). A review of computational optimisation methods applied to sustainable building design. Renewable and Sustainable Energy Reviews, 22, 230–245.

Evins, R., Pointer, P.,Vaidyanathan, R. (2010). Configuration of a genetic algorithm for multi-objective optimisation of solar gain to buildings. In: Proceedings of the genetic and evolutionary computation (GECCO). NewYork, NY, USA, 1327–1328.

EQUA (2013). IDA Indoor Climate and Energy. User Manual. Version 4.5. http://www.equaonline.com/iceuser/pdf/ICE45eng.pdf .

Masy, G., Georges, E., Verhelst, C., Lemort, V., Andre, P. (2015). Smart grid energy flexible buildings through the use of heat pumps and building thermal mass as energy storage in the Belgian context. *Science and Technology for the Built Environment*, *4731*(August), 800–811.

Miara, M., Günther, D., Leitner, Z. L., Wapler, J. (2014). Simulation of an Air-to-Water Heat Pump System to Evaluate the Impact of Demand-Side-Management Measures on Efficiency and Load-Shifting Potential. Energy Technology, 2(1), 90–99.

Reynders, G., Nuytten, T., Saelens, D. (2013). Potential of structural thermal mass for demand-side management in dwellings. *Building and Environment*, *64*, 187–199.

Rosli, E. M.A.M., Sopian, K., Mat, S., Sulaiman, M. Y. (2016). Renewable Energy in the Service of Mankind Vol II. *Selected Topics from the World Renewable Energy Congress WREC 2014*, *I*, 583–590.

Schibuola, L., Scarpa, M., Tambani, C. (2015). Demand response management by means of heat pumps controlled via real time pricing. Energy and Buildings, 90, 15–28.

TABULA WebTool (accessed 30th November 2016). http://webtool.building-typology.eu/#bm .

Thavlov, A., Bindner, H. (2015). A Heat Dynamic Model for Intelligent Heating of Buildings. International Journal of Green Energy, 12, 240–247.

Wetter, M. and Wright, J. (2003). Comparison of a generalized pattern search and a genetic algorithm optimization method. In Augenbroe, G. and Hensen, J., editors, Proc. of the 8-th IBPSA Conference, volume III, pages 1401-1408, Eindhoven, NL.

Wetter, M. and Wright, J. (2004). A comparison of deterministic and probabilistic optimization algorithms for nonsmooth simulation-based optimization. Building and Environment, 39, 989-999.

Wetter, M. (2011). GenOpt, generic optimization program, user manual, version 3.1.0. Lawrence Berkeley National Laboratory, Berkeley, CA, USA.

Zhou, G., Ihm, P., Krarti, M., Liu, S., Henze, G. (2003). Integration of an internal optimization module within EnergyPlus. In: Proceedings of the building simulation, Eindhoven, Netherlands.

Zong, H., Bönning, G. M., Santos, R. M., You, S., Hu, J., Han, X. (2017). Challenges of implementing economic model predictive control strategy for buildings interacting with smart energy systems. Applied Thermal Engineering, 114, 1476–1486.

Paper VII: Sarran, L., Foteinaki, K., Gianniou, P., & Rode, C. (2017). Impact of Building Design Parameters on Thermal Energy Flexibility in a Low-Energy Building. In Proceedings of *Building Simulation 2017*, San Francisco, United States.

Impact of Building Design Parameters on Thermal Energy Flexibility in a Low-Energy Building

Lucile Sarran[1], Kyriaki Foteinaki[1], Panagiota Gianniou[1], Carsten Rode[1]
[1]International Centre for Indoor Environment and Energy, Department of Civil Engineering, Technical University of Denmark, Kgs.Lyngby, Denmark

Abstract

This work focuses on demand-side management potential for the heating grid in residential buildings. The possibility to increase the flexibility provided to the heat network through specific building design is investigated. The role of different parts of the building structure on thermal flexibility is assessed through a parameter variation on a building model. Different building designs are subjected to heat cut-offs, and flexibility is evaluated with respect to comfort preservation and heating power peak creation.

Under the conditions of this study, the thermal transmittance of the envelope appears to have the largest impact on thermal flexibility. The importance of window design, namely the size, U-value and orientation, is underlined due to its critical influence on solar gains and heat losses. It is eventually observed that thermal mass has a secondary influence on the evaluated indicators; its variation only affects thermal flexibility if the thermal resistance of the envelope is sufficient.

Introduction

The share of renewable energy sources should represent at least 55% of the European gross energy consumption in 2050 (European Commission, 2011). Their production being inflexible and highly dependent on weather conditions, a high penetration of renewable sources might create a risk for energy security of supply. One of the solutions to this challenge is demand-side management as defined by Gellings (1985), which aims at adapting energy consumption to a fluctuating production rather than the opposite. This idea is particularly studied in the building sector, where implementation of demand-side management would permit to reduce the energy footprint of buildings, while providing flexibility to the energy grid as a whole (Kolokotsa, 2015).

EnergyLab Nordhavn is a full-scale research project investigating the possibility of intelligent energy operation across the neighbourhood of Nordhavn, Copenhagen. As a part of that project, the present study aims at understanding how energy efficient buildings can adapt perturbations in a city's heat and power grids. Focus is set here on thermal energy flexibility, specifically defined in this work as the capacity of a building to provide a good indoor comfort in spite of changes in delivered heating power. The understanding of this potential can then be used by the grid operator to optimise heating schedules. Thermal demand-side management is moreover seen as a tool to ensure electrical grid stability, via a coupling of the heat and power grids (EnergyLab Nordhavn, 2015; Müller et al., 2015).

An interest presented by low-energy buildings in the flexibility issue is their ability to retain heat for a long period of time, therefore acting as a storage medium in the heating grid. However, even though buildings' thermal flexibility potential is being investigated, few of the studies specifically evaluate the influence of design features on a building's flexible behaviour. The impact of different wall properties on their ability to store energy is well documented in literature (Asan, 2006; Borresen, 1973; Ma & Wang, 2012; Moffiet et al., 2014; Orosa et al., 2012; Wang et al., 2014) and their importance for load shifting was also specifically studied by Reynders (2015). Yet, it is valuable to extend the impact assessment to other building components than its sole thermal mass: all parameters having an influence on indoor comfort variations with time (through heat gains, losses and storage) are of interest when assessing a building's resilience to heating perturbations. Therefore, window parameters must also be investigated. Some researchers pay attention to the role of windows and solar gains (Orosa et al., 2012; Reynders, 2015; Wang et al., 2014), however more with respect to thermal buffering than flexibility. Similarly, research has been carried out on the role of window orientation in energy savings, but literature about its contribution to flexibility potential is rare (Reynders, 2015; Zhu et al., 2009).

Moreover, the achieved flexibility is often quantified in terms of financial savings (Masy et al., 2015; Reynders, 2015), shifted heating energy (De Coninck & Helsen, 2016) or capacity (Oldewurtel, 2013). The implications of heating control strategies on the occupants' comfort are generally not quantified.

This study is the first step of a larger project aiming at assessing the ability of low-energy buildings to ensure an active role in a city's heat and power grids. The goal of this preliminary study is to determine how the design parameters of a given low-energy building can influence its capacity to adapt simple perturbations on the heating grid in which it is integrated. Focus was set on passive flexibility strategies, in particular heat storage in the building structure. A parameter variation was carried out on a simulation model of a dwelling, investigating the role of the different building components in the flexibility potential of the building regarding its heat load. This paper does not consider the role of building services systems, e.g. space heating and ventilation, on flexibility since this is the theme of a parallel work that will be published separately.

This work has two main outcomes. First, a set of two indicators was built, quantifying theoretical heat flexibility in a building in terms of indoor comfort and heating demand. Second, this work provides an understanding of the heat storage processes in a low-energy building, permitting to identify the issues related to heating power perturbations and the extent to which building design can respond to it.

Methodology

Investigated building

The impact of design features on thermal flexibility was evaluated on a model of an apartment created according to the geometry, materials and systems of an existing apartment in Copenhagen, Denmark. The modelled apartment is located in the northern district of Nordhavn, in a nearly-zero-energy building currently under construction. The chosen building is representative of the current and future constructions in Denmark, which are bound to low energy consumption due to the strict Danish Building Regulations. It was chosen to focus the study on a single apartment in order to be able to get a thorough understanding of the obtained results. This choice allows getting deeper into the heat transfer mechanisms occurring at a smaller scale and to give a straightforward explanation of the findings.

Parameter variation

As the output of a literature study, six design parameters were chosen to be investigated further. Several values were chosen for each of them, covering a realistic range of possibilities as found in the literature (Gianniou et al., 2016; Reynders, 2015) and which reflects the construction characteristics of Danish building stock. Even though the ranges of values investigated for the different parameters are heterogeneous, they can all be interpreted as the complete range from the worst to the best-performing building components currently available – or in the case of the glazing-to-wall ratio and building orientation, as the whole set of values observed in Danish buildings. The corresponding difference in flexibility potential between extreme design cases thus represents the overall potential improvement that can be triggered by a realistic change of the considered parameter.

Table 1 presents the chosen parameters, the value of these parameters in the investigated apartment, and the variation range for each of them.

Different versions of the original apartment model were created, each giving a different value to a single one of the investigated parameters. The goal here was to perform a local sensitivity analysis in order to isolate the impact of each of the parameters on thermal flexibility.

Table 1: Investigated parameters, value in the base apartment model and variation range

Parameter	**Original value**	**Studied range**
External wall concrete thickness (cm)	15	[2 ; 30]
External wall insulation thickness (cm)	27.5	[0 ; 32]
Floor concrete slab thickness (cm)	6	[3 ; 25]
Glazing-to-external-wall ratio[1] (-)	0.42	[0.1 ; 0.55]
Window U-value (W/m²K)	0.81	[0.75 ; 2.5]
Orientation of the main facade[2]	S	N; E; S; W

Flexibility assessment and indicators

The flexibility assessment performed on the different models included two elements: a protocol, and a set of indicators quantifying the apartment´s flexibility potential. The overall protocol was the following: a cut-off in the heating schedule was performed, and the consequences of this perturbation were measured with focus on two aspects: the occupants' comfort and the heating power profile. This led to the two flexibility indicators developed.

The first test, focusing on occupants' comfort, consisted of cutting heat off in the apartment at 7 AM on Monday, January 19th and observing the decrease in operative temperature in the living room. The climate files and schedules used in the simulation are detailed in the "Model description" section. The living room was chosen as the target of the study since it is assumed to be the space where occupants are the most present during daytime and therefore where thermal comfort is valued the most. Moreover, its large glazed surface makes it likely to suffer from large temperature swings. This heating control strategy corresponds to a peak shaving operation under its most extreme form: heat is completely cut off in the residential buildings to shave the morning heating peak. This scenario was chosen for its

[1] The glazing-to-external-wall ratio was modified by changing the size of the windows in the different rooms simultaneously and by the same factor.

[2] The orientation was changed by rotating the entire building model.

simplicity: indeed, implementing a preheating period of a specific duration or reducing heat supply by a certain percentage would create a bias on the results, while this bias is reduced when using a fundamental signal such as a total cut-off. The corresponding indicator was the time (measured from the cut-off) after which the operative temperature dropped below the lower bound of the occupants' acceptability range. In this work, this minimum acceptable temperature was set to 20°C, corresponding to Category II of EN/DS 15251. Figure 1 gives a simplified graphical representation of the indicator calculation.

In practice, given the large fluctuation of indoor temperature, temperature can drop for some minutes below 20°C and rise again, which is not a real threat to indoor comfort. Therefore, a tolerance factor was introduced, which made the indicator practically calculated as the duration of the time interval, starting from the cut-off, during which operative temperature in the living room has been above 20°C for 90% of the time since the cut-off.

This indicator was introduced to reflect in simple terms the degradation of indoor comfort in the living space. It is a tangible measure of occupants' perception. Following the variation of this indicator when modifying the apartment's design parameters permits to understand the theoretical ability of these investigated parameters to preserve indoor comfort, in case of a change in heating schedule ordered by the thermal grid operator. The simplicity of the protocol permits to obtain a fundamental comparison of the parameters' role – more complex control strategies will be implemented in a further study.

The second test consisted of cutting heat off in the apartment between 7 AM and 4 PM on Monday, January 19th, which corresponds to a heat prioritization strategy: dwellings are heated up in the evening and during the night, and offices during working hours. It was chosen to measure the maximum heating power level reached in the following two days (between the 19th and the 21st of January) in the simulation with heating cut-off and to compare it with the maximum power level in case of

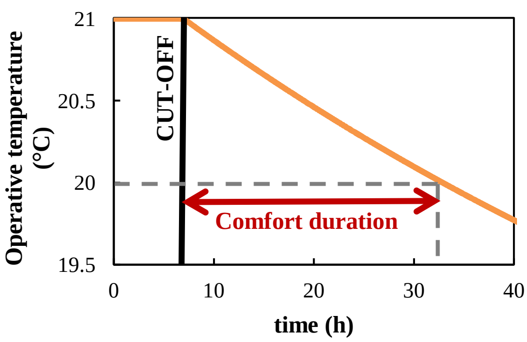


Figure 1: Duration of the comfort period

normal operation (with a setpoint at 21°C) during the same period. The heating power is defined as the heat input in the floor heating system for the whole apartment. This output constitutes the second flexibility indicator (1). Q_{max} and $Q_{max,ref}$ are represented in Figure 2.

$$Ind_2 = \frac{Q_{max} - Q_{max,\ ref}}{Q_{max,\ ref}} \qquad (1)$$

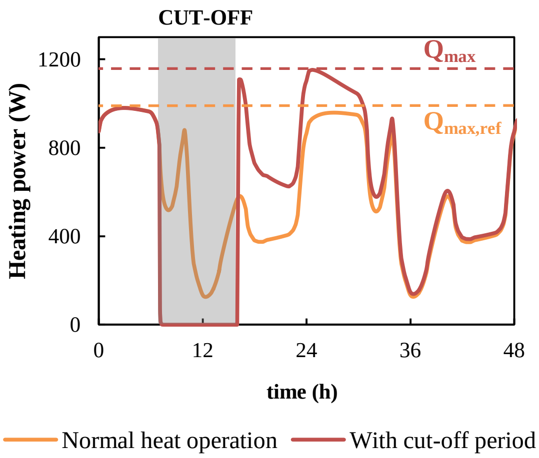


Figure 2: Peak power variation

The goal of this indicator is to assess the pressure set on the heating system when performing a heating control operation. Indeed, in order to satisfy indoor comfort requirements, there is a risk that the system reacts by a large increase in heating power when the cut-off period is over, which can in some cases go against the goals of performing a load-shifting strategy. Calculating this indicator under different design solutions permits to identify the magnitude of this potential problem; if the problem can be predicted; and if it is particularly affected by specific design parameters. In a future work, different controls strategies will be applied to mitigate this problem, in particular a ramp for the temperature setback rather than the simple on/off control tested here. In this indicator, the peak levels that are compared are not necessarily occurring at the same time: they are the overall maxima over the considered days. This approach gives information about the extra capacity that would be needed to accommodate the new peaks.

The two flexibility indicators were calculated for each of the models, and their variation was related to the change in the parameters' value by a graphical representation. The impact of the different parameters on the chosen indicators could then be appreciated.

Model description

This study was based on a set of numerical models created on the building performance simulation software IDA ICE 4.7.

Figure 3: Building overview (source: Alectia)

Geometry

The building that was investigated is a 5-floor low-energy residential building located on the waterfront, with the water-facing facade oriented 10° from South. One apartment of this building was modelled in this study. It is a 95m², 3-room apartment located on the 4th floor, with the main facade facing the waterfront.

The geometry of the IDA ICE model is shown in Figure 4. The rest of the building volume was included in the model, but only with regards to shading calculations: it was assumed that the adjacent apartments are similarly heated spaces, so there is no heat exchange with the rest of the building.

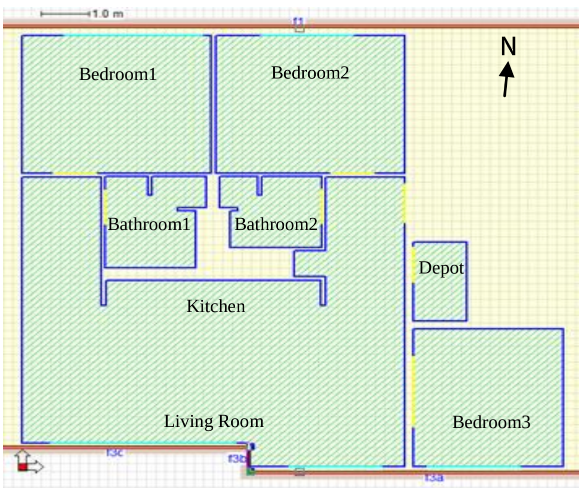


Figure 4: IDA ICE apartment model - floor plan

Building materials

The external walls of the building included a thick layer of concrete on the inside and a brick external façade, with mineral wool insulation in-between. Two sorts of internal walls were used: bearing walls made of a single concrete layer, and non-bearing walls made of aerated concrete, used to separate the bedrooms and in the kitchen corner. The floor of the bedrooms, living room and depot room was covered by oak planks lying on a concrete screed. The windows consisted of three glass panes filled with argon, with an aluminium and wood frame. The total U-value of the apartment is equal to 0.49W/m²K. Table 2 details the different construction layers.

Systems

All the apartments were equipped with a floor heating system connected to the district heating network, with supply and return temperatures of respectively 40 and 35°C. The heating setpoint was chosen to be 21°C and no cooling system was included. The domestic hot water circuit was not modelled in this study due to its little influence on heat flexibility in the absence of individual hot water storage system for each apartment. The mechanical ventilation system was balanced CAV (constant air volume) with 80% heat recovery. Fresh air was supplied in the bedrooms and the living room with a setpoint of 17°C and the return air was exhausted from the kitchen and the bathrooms.

Outdoor conditions and internal gains

In order to isolate the results from the influence of fluctuating outdoor conditions, the base outdoor temperature was set constant to -5°C, representing a cold winter day, with 90% relative humidity. Due to the difficulty to predict the wind-driven infiltration in the context of a semi-exposed building, a constant value for infiltration of 0.1 L/s/m^2 floor area was used. Some solar gains were defined based on an average radiation in a short winter day. This pattern was repeated every day of the simulation period.

A standard pattern for the heat flux from occupants and electrical appliances was defined. The day was divided in four periods, and different uses were defined for weekdays and weekends in the different rooms. It was considered that four people live in the apartment, which is common for this specific building: Bedroom 1 was considered double. The four occupants were always present in the apartment from 5 PM to 9 AM, and all day during the weekends. None of them was present from 9 AM to 5 PM during workdays. Equipment units were positioned in every room apart from the depot room and the bathrooms.

A daylight-related control was designed for lighting

Table 2: Composition and U-value of building parts

Building part	**Layers (from outside/from bottom)**	**Total U-value (W/m²K)**
External wall	• Brick (10.8 cm) • Mineral wool (27.5 cm) • Concrete (15 cm)	0.13
Bearing internal wall	Concrete (20 cm)	3.43
Non-bearing internal wall	Aerated concrete (10 cm)	1.13
Floor	• Hollow concrete (22 cm) • Termotec insul. (14.6 cm) • Concrete screed (6 cm) • Oak planks (3 cm)	0.27
Window	Triple-pane Argon-filled	0.81

as advised in BR15 (The Danish Building Regulation, 2015). Shading devices were included and supposed to be manually activated by the occupants in case of excessive lighting conditions.

Simulation parameters

A dynamic simulation was chosen: the program made an initial guess and reached convergence while simulating from the 1st of January to the 18th of January, then started the simulation from the 19th of January, day of the cut-off. The maximum simulation time step was 12 minutes, but it was automatically reduced when more accuracy was needed.

Results

The graphical representation of the obtained flexibility results is shown in Figure 5. The two indicators (duration of the comfort period in the left graphs and peak power variation in the right graphs) are plotted against the different design parameters, which are gathered in three groups: material layers, window features and main facade orientation. In each of the graphs, the points representing the original building are darker and circled in black.

Material layers

As seen in Figure 5a and Figure 5b, the thickness of the concrete layer in the external walls shows impact on none of the two indicators under the considered conditions. Between a 2-cm and a 30-cm concrete layer, the duration of the comfort period in the living room varies between 72 and 74.5 hours, which is almost negligible in comparison to the impact observed for other parameters. Identically, the peak power in a period perturbed by a heat cut-off gets approximately 19% higher than under normal operation in the same period, independently of the concrete thickness in the external walls. This result questions the role of heat storage in the external walls in response to a heat cut-off: in this particular case, the increase in thermal storage capacity does not affect the apartment's energy flexibility potential.

On the contrary, the duration of the comfort period after cut-off in the living room shows a large dependence on the thickness of the insulation layer in the external walls. While a non-insulated external wall permits to retain heat in the living space for 21 hours before the comfort limit is reached, adding 10 centimeters of insulation brings this figure to 69 hours. Above 10 centimeters of insulation, the corresponding improvement in comfort conditions is relatively small, namely reaching 32 centimeters of insulation results in an increase in the comfort period duration from 69 to 81 hours. This result shows that a minimum thermal resistance is necessary to allow heat retention: if the insulation is too poor, energy is lost too rapidly towards the ambient to be stored in the building structure (external walls). The insulation thickness of the external walls also has a significant impact on the heat consumption peak: a heat cut-off in a non-insulated apartment leads to a rapid temperature decrease, and therefore to an important heating peak when heat is turned on again at 4 PM: the peak gets 60% higher than the maximum heating power in normal operation (with no cut-off). The peak power variation due to a cut-off gets less important when the insulation layer gets thicker: for an apartment with more than 18 centimeters insulation, the peak due to a cut-off is around 20% higher than the peak power in normal operation. Both tests show that even though a minimum level of insulation is required, the apartment's flexibility potential is not significantly improved for an insulation thickness higher than 20 centimeters, which is a valuable piece of information when drawing design guidelines.

Varying the thickness of the concrete slab that embeds the floor heating pipes from 3 to 12 centimeters leads to a change in the comfort period from 72 to 93 hours (Figure 5a), while increasing the slab thickness from 12 to 25 centimeters increases the comfort period duration by 3 hours only, from 93 to 96 hours. This result may be due to the fact that the pipes are located only one centimeter below the surface of the concrete slab in all of the cases, which limits the penetration of heat in the lower layers of the concrete. The study of the impact of the depth of the pipes into the concrete was not within the scope of this work. There is nevertheless room for improvement in the investigated building: increasing the concrete slab thickness from 6 cm to 12 cm could save up to 20 hours of heating in the living room. The analysis of the peak power variation due to a cut-off does not show a clear dependence on the slab thickness (Figure 5b). The results nevertheless show an optimal behaviour when the slab thickness approximates 12 centimeters.

Window features

Both the window U-value and the glazing-to-external-wall ratio show a large impact on the calculated flexibility indicators. As seen on Figure 5c, when varying the window U-value from 0.75 to 2.5 W/m²K, the duration of the comfort period is approximately divided by 4 (from 75 to 19 hours). This result demonstrates that the heat loss from windows has a large responsibility in the heat retention performance of the considered building, in this particular case of a large window surface (glazing-to-floor ratio of 0.42) and relatively cold outdoor conditions. It is also found that the higher the U-value of the windows, the higher the peak in heating power observed after a cut-off period (Figure 5d). This is due to the larger heat loss during the cut-off period when window U-value is increased, leading to a need for higher heating power when heat supply is re-established. The window size has a very clear influence on the flexibility results. The larger the share of glazed surface in the external walls, the shorter time can heat be retained indoors, as seen in Figure 5c. It has to be noticed that the scale of Figure

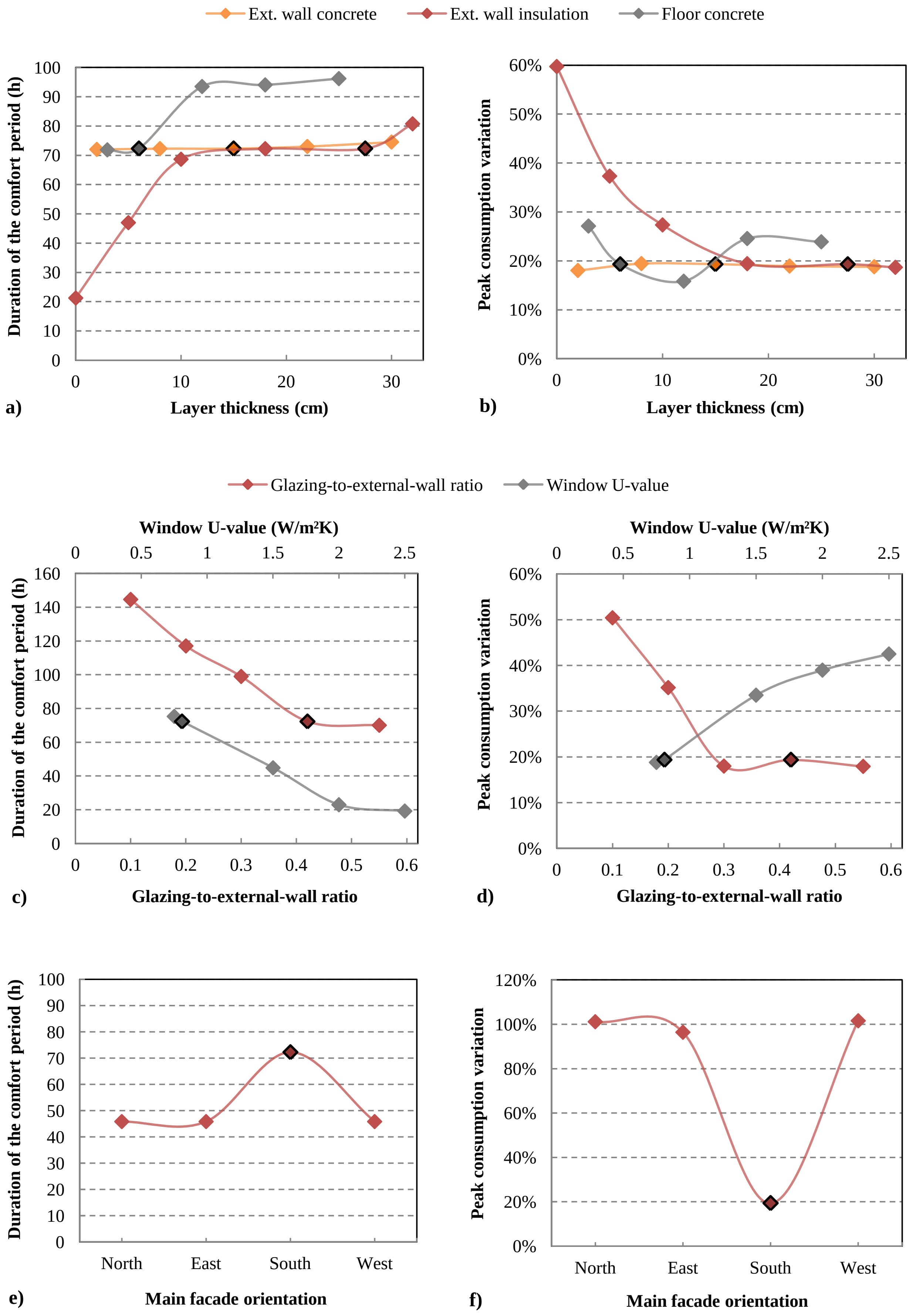


*Figure 5: Parameter variation results: duration of the comfort period (**a**,**c**,**e**) and peak consumption variation (**b**,**d**,**f**) in function of the thickness of different wall and floor layers (**a**,**b**), window properties (**c**,**d**) and orientation (**e**,**f**)*

5c has been extended, since for a glazing-to-wall ratio lower than 30%, good comfort conditions can be kept in the living room for more than 100 hours.

This result also confirms the predominance of thermal resistance in the flexibility issue: indeed, windows are both a source of heat loss and of solar gains, but the losses seem to have a much larger impact on flexibility in this apartment since reducing the window size has such a positive effect on heat retention. However, a further analysis has to be carried out investigating whether this effect is due to the greater wall volume when window size is reduced, increasing the apartment's thermal mass. The analysis of the heating peak gives different results (Figure 5d). For glazing-to-wall ratios below 30%, the solar intake is low and does not contribute much to heating up the space. Consequently, the impact of cutting heat off is larger and the change more brutal, leading to a relatively high peak when heat is turned on at 4 PM (up to 50% higher than normal). For glazing-to-wall ratios higher than 30%, the height of the peak stabilizes between 18 and 19% above its level in absence of a cut-off. Solar gains constitute an immediate source of heat able to balance heat losses, hence their impact on the heating power required after a heating control operation. For large window sizes, the positive impact of solar gains is balanced by the increased heat loss from the extra surface, keeping the peak power surplus constant for glazing-to-wall ratios of 30 to 50%. The Danish BR15 building regulation sets a minimum of 15% for the glazing-to-floor ratio, which corresponds in the present apartment to a glazing-to-wall ratio of 24%. Going above this value shows to be beneficial from the point of view of peak power minimization, but leads to a shorter comfort period in case of a cut-off. However, daylight being an important component of indoor comfort, it is not advisable to introduce a lower threshold for window size in a new version of the building regulation.

Main facade orientation

As seen on Figure 5e and Figure 5f, all orientations but the South give similar flexibility results in the studied apartment for both indicators: the comfort period lasts around 45 hours, and a several-hour long heat cut-off triggers a heating power peak twice the height of the maximum power in normal operation. Orienting the facade 10° to the South, as done in the existing building, permits to extend the comfort period up to 72 hours, which is a favourable scenario when considering energy flexibility. This result permits to qualify the interpretation of the findings from the window size investigation. Indeed, even though solar gains show to have, in the present case, a smaller impact than heat losses through windows on thermal comfort preservation after a heating control operation, their absence significantly degrades the performance described by the indicator.

A South orientation permits to greatly limit the heating peak, making it only 20% higher than the maximum power in normal operation. This result confirms the finding from the glazing-to-wall ratio investigation: solar gains have a prevailing moderating impact on the heating power increase following a cut-off.

Holistic analysis

In order to get a better understanding of the findings, all the investigated cases were gathered. The two flexibility indicators were plotted over both the UA-value of the apartment and its total effective internal heat capacity, giving a point for each of the different design cases. The latter was calculated including the thermal mass of external and internal walls, floors and ceilings, calculated accordingly to DS/EN ISO 13790 (2008). The results can be seen in Figure 6.

The thermal resistance of the envelope appears as the primary factor able to guarantee a satisfactory response to a control operation in terms of indoor comfort. As shown in Figure 6a, the duration of the comfort period shows a clear dependency on the apartment's total UA-value, with little dispersion. As highlighted in the study, the insulation level of the external walls and the U-value of the windows are critical parameters in the preservation of thermal comfort during a heating control operation such as, in the most extreme case, a complete heat cut-off. Reducing the window size also permits to preserve comfort for a longer time, even though this results in a lower solar energy intake. These three parameters impact the heat loss rate through the envelope, thus influencing the duration of the comfort period.

The conducted parameter analysis shows that the envelope's thermal performance has a decisive impact on the possible extreme heat power peaks following a heat cut-off. Indeed, a poor thermal performance leads to high heat losses during the cut-off period, thus the need to quickly increase heating power to satisfy indoor comfort requirements. However, Figure 6b makes it clear that even though an increase in UA-value globally leads to an increase in peak surplus, another factor has a much higher influence on this parameter: the facade orientation, responsible for a doubling of the peak height when changed from the South to any other direction. Solar gains show to strongly attenuate the risk of creating heat consumption peaks after a cut-off period.

Through the analysis of different orientations and windows sizes, it is made clear that solar gains participate in preserving indoor comfort after a heating perturbation, but that in the current case, the amount of heat lost through windows is higher than the amount gained from solar radiation, which limits the relevance of increasing the window size. In the studied apartment, heat accumulation in the thermal mass (specifically external walls and floor) plays a limited role, as highlighted by the study and confirmed by Figure 6c. Increasing the heat storage capacity in the external walls triggers no change in

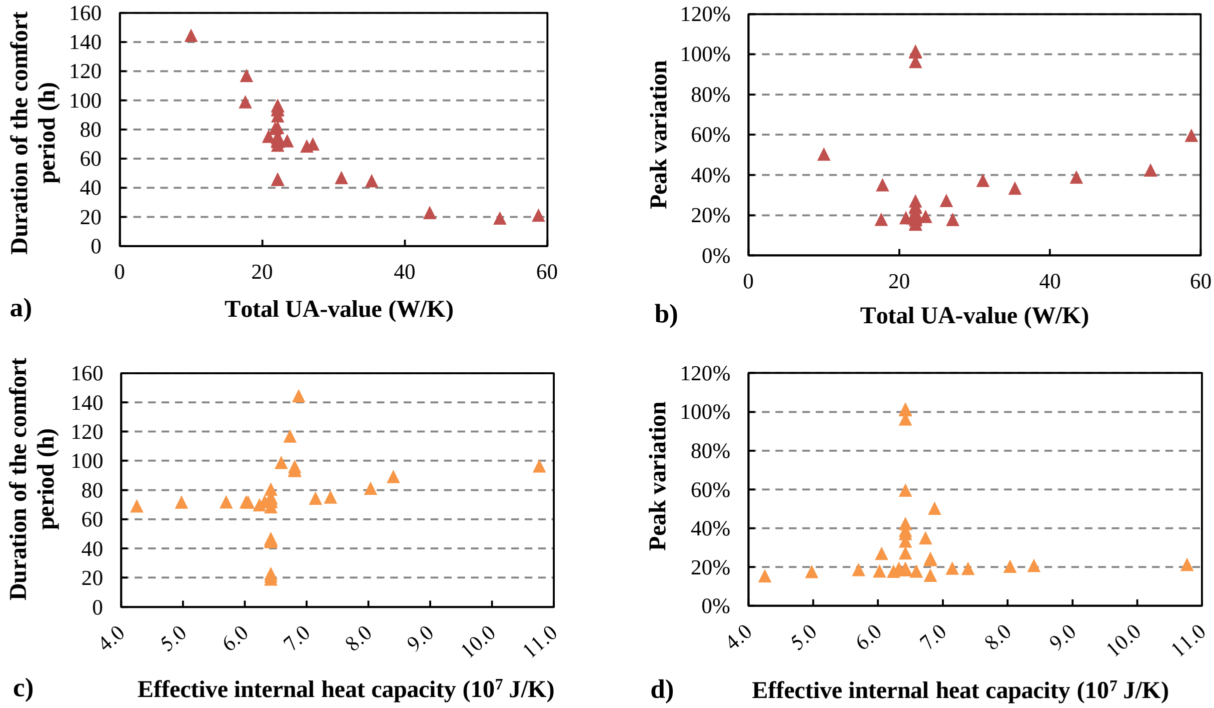


*Figure 6: Flexibility indicators as function of total UA-value (**a**,**b**) and effective internal heat capacity (**c**,**d**)*

the duration of the comfort period, while increasing the available storage volume in the floor slab that contains the floor heating pipes helps counteract heat losses in the space up to a certain volume (Figure 5a). Figure 6c shows a globally flat profile with a high dispersion, demonstrating that internal heat capacity is not a decisive parameter in this case. This result could find an explanation in the large window surface resulting in a limited external wall volume. In order to confirm this hypothesis, the impact on the indicators of a simultaneous change in the window area and in the heat storage capacity should be investigated. Identically, the building's thermal mass shows no direct impact on heat power peaks following a cut-off (Figure 6d). As an example, the different design cases that keep an internal heat capacity of around $6.5{\times}10^7$ show a large variability in peak power variation, which indicates no correlation between the indicator and the internal heat capacity. The value of the indicator is rather sensitive to other parameters such as the window orientation or the insulation thickness, which variations have no influence on the internal thermal capacity.

Discussion

This study is based on simulations that aim at understanding and comparing the influence of different design parameters on heat flexibility. Assumptions and simplifications were used in order to be able to isolate the influence of the investigated parameters from other factors. Thus, the results cannot be used as such to establish a heating control strategy for example, since they have been obtained in a context that could deviate from other cases. The purpose of the obtained flexibility results is rather to show the evolution of the flexibility potential when varying a parameter and to observe the different impact that they would have. The results have to be read in relative terms, the trend being more interesting and reliable than the level itself. Among the main simplifications are the constant outdoor temperature and infiltration, the periodic heat gains and the scheduled occupants' behaviour. An analysis of the sensitivity of the present flexibility results to these assumptions is available in a complementary publication by Zilio et al. (2017).

Moreover, this paper describes a local sensitivity analysis, which has the advantage of isolating every parameter's impact while requiring a limited number of simulation runs. The influence of one of the parameters on the results given by the other, through a multi-parameter analysis, is under investigation for some particular parameter combinations and the results are to be published in a coming article. This complementary analysis is expected to permit to make the whole analysis applicable to a wider range of buildings.

Another assumption having an impact on the results is the choice of the cut-off time. In particular, the choice of the cut-off period in the heat power investigation is likely to give a bias to the role of solar gains in the peak power height, since the cut-off period includes the whole solar radiation period. The heating control strategies themselves include a bias in the obtained results: as an example, Wolisz at al. (2013) show that in the case where a pre-heating period is performed before the heat cut-off, the building's thermal mass plays a more important role.

However, even though the investigation mode itself induces some bias in the result, the difference in result presented by the two studies gives some valuable information. While the first indicator is an index of building performance from the occupants' point of view, the second indicator can rather be used when focusing on grid stability. Moreover, both measurements give information about the ability of the apartment to retain heat indoors, but under different angles. Measuring the duration of the comfort period after a cut-off answers the following question: how long can the building provide an acceptable temperature indoors without heating? Focusing on the heating peak after a temporary cut-off rather answers this question: after 9 hours without heating, how far is the temperature in the living room from acceptable conditions (namely the heating setpoint)? This difference in time span explains in particular why increasing the window size has such a positive impact on reducing the heating peak while its effect on the comfort period is overall negative: solar gains help increasing indoor temperature on the short term but thermal losses are predominant on a longer term.

Eventually, it would have been of great interest to be able to couple this study, which is based on a simulation model of the apartment, to actual measurements in the investigated building. The edifice still being under construction, it was impossible. However, measurements are scheduled in the context of the EnergyLab Nordhavn project and this aspect will be the topic of a further study. The influence of the occupants and of the outdoor conditions will in particular be made more clear and the evaluation of indoor comfort more detailed by distributing questionnaires to the occupants.

Conclusion

The present study deals with design of buildings as a tool to improve their ability to retain heat and thus adapt changes in heating schedule, the end goal being to provide energy flexibility to the heating network. In order to understand which of the building design parameters influence its flexibility potential and in which way each of them contributes to it, two investigations were performed on a number of design solutions for the building model. In the first investigation, a complete heat cut-off was performed at 7AM, and the time during which the operative temperature in the living room remained above 20°C was measured, constituting the first flexibility indicator. In the second investigation, heating was cut off between 7AM and 4PM and the heating power peak around that period was measured. Its relative difference with the peak level under normal heating operation constituted the second flexibility indicator. By comparing the results under different design versions, the influence of design parameters on heat flexibility in the present building was assessed.

It was found that the design parameters with the most influence on the temperature drop in the inhabited space after a heat cut-off are those impacting the heat losses through the building envelope. The insulation of the external walls is the parameter showing the largest impact on flexibility. Improving the U-value of the windows can multiply the duration of the comfort period by up to a factor of 4. Heat retention time strongly decreases when the glazing-to-external-wall ratio increases, showing that the losses due to a larger glazed surface impact more the indoor temperature than the increased solar gains when the window gets larger. The orientation of the apartment shows a moderated but clear impact on heat retention: any other orientation than South significantly reduces the time in the comfort range after a cut-off.

As to the thickness of the concrete layers of the different building components, its impact is less important than expected. The layer which thickness has the most significant influence on flexibility is the floor slab, since it embeds the heating pipes, but this is only true up to a certain thickness. The thickness of the concrete layer of the external walls shows a negligible influence under the conditions of this study.

A temporary heat cut-off performed during the day is followed by a heat power consumption peak that is influenced by the building design. The relative height of this peak with respect to the peak power in normal conditions is not affected at all by thermal storage in the building structure, but is greatly influenced by the thermal resistance of the envelope, namely by the insulation thickness and U-value. However, according to the present calculation methods, the main influence on the peak power seems to come from the presence of solar gains: orienting the main facade in any other direction than South greatly increases the heating power peak. Similarly, while smaller windows guarantee a better heat retention, they also lead to an important increase of the peak power.

This work has permitted to understand the mechanisms that lead to transmission and retention of heat indoors, which give the district heating operator the possibility to apply restrictions on heat supply when best for the system as a whole. The importance of an optimal building design is now obvious, and the role of the main building components has been clarified.

The present work is a preliminary theoretical comparison of the role of different building components in heat perturbations adaptation. This work also aims at giving an estimate of the energy flexibility potential that these newly-constructed apartment blocks can offer to the grid and which can be further utilized by the future heat supply of the area. Some multi-parameter investigations as well as analyses with different building shapes and indoor space organisation will complete this work, and help forming a set of guidelines for an optimal building design with focus on heat flexibility. An impact assessment of outdoor conditions and internal gains on the present flexibility results is developed by Zilio et al. (2017). The following steps of this work include yearly simulations using realistic weather data permitting to assess the building's response to both summer and winter conditions. Moreover, field measurements are being collected on the studied building and will be used to support the present simulation work. Finally, a study of different heating schedules is planned, including price-dependent scenarios.

Acknowledgements

This study is part of the project EnergyLab Nordhavn - New Urban Energy Infrastructures and the Danish research project CITIES (Centre for IT-Intelligent Energy Systems in cities).

References

Asan, H. (2006). Numerical computation of time lags and decrement factors for different building materials. *Building and Environment*, *41*(5), 615–620.

Borresen, B. A. (1973). Heat storage in walls. *Build Serv Eng*, *41*, 17 – 18.

De Coninck, R., & Helsen, L. (2016). Quantification of flexibility in buildings by cost curves - Methodology and application. *Applied Energy*, *162*, 653–665.

DS/EN ISO 13790 - Energy performance of buildings – Calculation of energy use for space heating and cooling. (2008).

EnergyLab Nordhavn. (2015). EnergyLab Nordhavn - public project description. Retrieved January 11, 2016, from http://energylabnordhavn.dk

European Commission. (2011). Energy Roadmap 2050 (Vol. 53, pp. 1689–1699).

Gellings, C. W. (1985). The concept of demand-side management for electric utilities. *Proceedings of the IEEE*, *73*(10), 1468–1470.

Gianniou, P., Heller, A., Nielsen, P. S., & Rode, C. (2016). Identification of Parameters Affecting the Variability of Energy Use in Residential Buildings. In *Proceedings of CLIMA Conference, Aalborg*.

Kolokotsa, D. (2015). The role of smart grids in the building sector. *Energy and Buildings*, *116*, 703–708.

Le Dréau, J., & Heiselberg, P. (2016). Energy flexibility of residential buildings using short term heat storage in the thermal mass, *111*, 991–1002.

Ma, P., & Wang, L. S. (2012). Effective heat capacity of interior planar thermal mass (iPTM) subject to periodic heating and cooling. *Energy and Buildings*, *47*, 44–52.

Masy, G., Georges, E., Verhelst, C., & Lemort, V. (2015). Smart grid energy flexible buildings through the use of heat pumps and building thermal mass as energy storage in the Belgian context. *Science and Technology for the Built Environment*, *4731*(August), 800–811.

Moffiet, T., Alterman, D., Hands, S., Colyvas, K., Page, A., & Moghtaderi, B. (2014). A statistical study on the combined effects of wall thermal mass and thermal resistance on internal air temperatures. *Journal of Building Physics*, *38*(5), 419–443.

Müller, D., Monti, A., Stinner, S., Schlösser, T., Schütz, T., Matthes, P., … Streblow, R. (2015). Demand side management for city districts. *Building and Environment*, *91*, 283–293.

Oldewurtel, F., Sturzenegger, D., Andersson, G., Morari, M., & Smith, R. S. (2013). Towards a standardized building assessment for demand response. *Proceedings of the IEEE Conference on Decision and Control*, 7083–7088.

Orosa, J. A., & Oliveira, A. C. (2012). A field study on building inertia and its effects on indoor thermal environment. *Renewable Energy*, *37*(1), 89–96.

Reynders, G. (2015). *Quantifying the impact of building design on the potential of structural storage for active demand response in residential buildings*. KU Leuven.

The Danish Building Regulation. (2015). Executive Order amending the Executive Order on the Publication of the Danish Building Regulations 2010 (BR10).

Wang, L.-S., Ma, P., Hu, E., Giza-Sisson, D., Mueller, G., & Guo, N. (2014). A study of building envelope and thermal mass requirements for achieving thermal autonomy in an office building. *Energy and Buildings*, *78*, 79–88.

Wolisz, H., Harb, H., Matthes, P., Streblow, R., & Müller, D. (2013). Dynamic simulation of thermal capacity and charging / discharging performance for sensible heat storage in building wall mass. In *13th Conference of International Bulding Performance Simulation Association* (pp. 2716–2723).

Zhu, L., Hurt, R., Correia, D., & Boehm, R. (2009). Detailed energy saving performance analyses on thermal mass walls demonstrated in a zero energy house. *Energy and Buildings*, *41*(3), 303–310.

Zilio, E., Foteinaki, K., Gianniou, P., & Rode, C. (2017). Impact of Weather and Occupancy on Energy Flexibility Potential of a Low-energy Building.

Paper VIII: Zilio, E., Foteinaki, K., Gianniou, P., & Rode, C. (2017). Impact of Weather and Occupancy on Energy Flexibility Potential of a Low-energy Building. In Proceedings of *Building Simulation 2017*, San Francisco, United States.

Impact of Weather and Occupancy on Energy Flexibility Potential of a Low-energy Building

Emanuele Zilio[1], Kyriaki Foteinaki[1], Panagiota Gianniou[1], Carsten Rode[1]
[1]International Centre for Indoor Environment and Energy, Department of Civil Engineering, Technical University of Denmark, Kgs. Lyngby, Denmark

Abstract

The introduction of renewable energy sources in the energy market leads to instability of the energy system itself; therefore, new solutions to increase its flexibility will become more common in the coming years. In this context the implementation of energy flexibility in buildings is evaluated, using heat storage in the building mass. This study focuses on the influence of weather conditions and internal gains on the energy flexibility potential of a nearly-zero-energy building in Denmark. A specific six hours heating program is used to reach the scope. The main findings showed that the direct solar radiation and the outdoor temperature appeared to have the larger impact on the thermal flexibility of the building. Specifically, the energy flexibility potential of the examined apartment can ensure its thermal autonomy up to 200 h in a typical sunny winter day.

Introduction

The high share of Renewable Energy Sources (RES) expected in the energy market for the coming years will lead to lower stability of the energy system itself due to the high dependence on weather conditions (e.g. solar radiation and wind). Consequently, there will be problems with matching the production side with the demand side. To overcome this issue, new solutions have to be implemented to increase the energy flexibility of the energy system (Lund and Lindgren et al., 2015) (Rotger-Griful and Hylsberg Jacobsen, 2015). Few decades ago, Gelling coined the term "Demand-Side Management" (DSM) to indicate all actions that can be taken to influence the users' consumption profile in order to adapt it to the energy production (Gellings and Parmenter, 1985). Such solutions (e.g. peak shaving or load shifting) are becoming more relevant in the energy management field. For instance, energy storage devices can be implemented at final-user level, in this way, when there is surplus of energy from renewable production, it can be stored and used in a second moment when there is a lack of it (Lund and Lindgren et al., 2015).

The integration of RES in the energy system also leads to a more decentralized structure, since these sources are distributed in the territory. For this reason, a "smart city" model has to be used to allow communication between energy producers, loads and components (Aduda et al., 2016) (Lund and Lindgren et al., 2015). The aim of the new energy system is to improve the flexibility on the demand side, since, due to the high maintenance and operational costs, it would be more expensive to act on the supplier side (Aduda et al., 2016) (X. Xue et al., 2015) (Labeodan et al., 2015). In the building sector, a certain amount of storage capacity can be installed while ensuring good indoor comfort for the occupants (Zheng et al., 2015). Referring to the Danish energy market, its correlation between energy price and amount of renewables production ensures that flexible consumers can benefit from lower electricity prices (Neupane et al., 2015).

As the previous studies presented by Marin et al. (2016) and Ling et al. (2006) state, the thermal mass of buildings can be used to store energy. For example, the thermal inertia of the floor heating system allows the use of intermittent heating strategies, achieving load shifting. On the other hand, the envelopes of refurbished or new buildings lead to lower heating demand and higher influence of internal gains on the heating consumption. Firląg et al. (2013) presents a first assessment of the influence of the internal gains on the energy consumption. In new buildings, heat gains can cover up to 60% of the heat losses and therefore considerably affect the energy consumption. They can affect the indoor comfort and with a dedicated heating control strategy, it is possible to take advantage of them, with the consequent reduction in energy consumption, up to 30%. In Lazos et al. (2014) it is described how heat gains prediction from occupancy and weather conditions could be used to reduce the energy consumption in a range 15-30%, highlighting the importance of heat gains prediction in the energy management of a building. Sikula et al. (2012) found that the highest impact is given by solar gains that can affect the energy consumption around 20%. In Chen et al. (2012) it is explained that the energy consumption of a building could be forecasted based on weather forecast. Le Dréau and Heiselberg (2016) showed that a single-family passive house can be totally autonomous in terms of heating demand up to 48 h, while ensuring the operative temperature was always kept above 20 °C. In the same way, Ingvarson and Werner (2008) investigated the energy storage capacity in the building mass. The results show that the heating system can be switched-off in newly built or renovated dwellings, while ensuring acceptable indoor temperature for more than 24 h.

The aim of this paper is to investigate energy flexibility in a newly constructed building through energy storage in building mass during the heating period. The building investigated follows the Danish building regulation (BR10) for a low energy building in class 2015. The energy supply for heating, ventilation, cooling and domestic hot water per m^2 of heated floor area does not exceed 30.1 kWh/m^2/year. An important step of the evaluation is the definition of a dedicated flexibility indicator, evaluated with regards to different weather conditions and internal gains. Furthermore, two heating strategies are implemented: the first one is used to define the flexibility indicator and the second one is used to assess the effect of achieving energy storage on a daily basis with a specific heating program. The relevance of this study is based on the type of building chosen for the investigation. Due to its design characteristics, it ensures a unique contribution in terms of energy flexibility.

Methodology

Model description

The reference building is located in Copenhagen, Denmark, and is newly constructed. It consists of 85 apartments of different sizes, divided in five storeys. In this case study, a reference apartment is chosen, which is 95 m^2, located at the 4^{th} floor and facing the waterfront on the south-east façade. The apartment has three bedrooms, a kitchen-living room, two bathrooms and a storage room. Figure 1 gives an overview of the apartment.

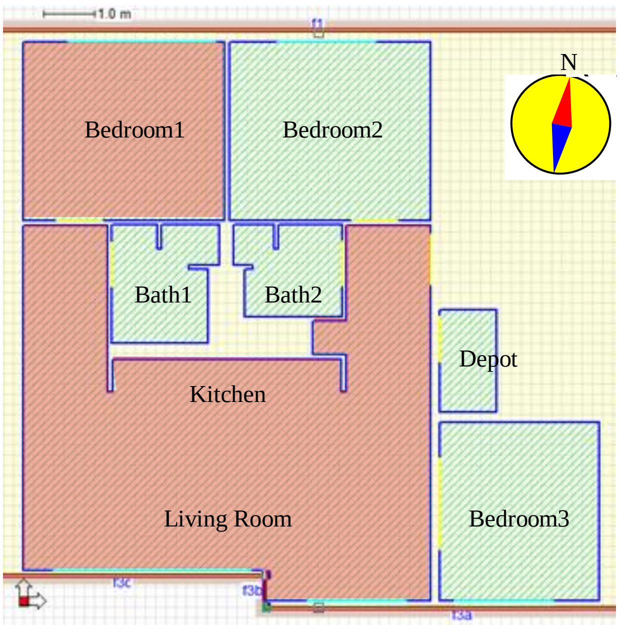


Figure 1: Floor plan of the apartment investigated

The rooms used for the analysis in this study are the Kitchen-Living Room (45 m^2) and bedroom 1 (14 m^2), highlighted in red in Figure 1. The building itself has a concrete-structure, where both internal and external bearing walls are reinforced concrete walls. Information about the composition of the building's components can be found in Table 1. The external walls have a U-value of 0.12 W/m^2K, while for the internal walls is 3.43 W/m^2K. Triple glazing windows are used with an overall U-value equal to 0.8 W/m^2K. The glazing-to-wall ratio is 0.42.

Table 1: Material layers

	Layer material	Thickness (m)
External walls	Concrete (inside)	0.150
	Mineral wool	0.275
	Air gap	0.025
	Brick (outside)	0.108
Internal floor	Oak planks (top)	0.030
	Concrete	0.060
	Termotec insulation	0.146
	Hollow concrete (bottom)	0.220
Bathroom floor	Concrete (top)	0.100
	Air	0.126
	Hollow concrete	0.220
	Lightweight concrete (bottom)	0.060
Bearing internal wall	Concrete	0.200
Non-bearing internal wall	Aerated concrete	0.100
Bathroom wall	Lightweight concrete	0.080

The building is connected to the local district heating network with supply water temperature at 60 °C and return at 40 °C. The space heating (SH) system is a hydronic floor-heating system with a heating design power of 16 W/m^2 (excluding the storage room) and a supply water temperature of 40 °C. The pipes are embedded in the 10 cm concrete slab (1 cm depth from the top), which is covered by a 3 cm wooden covering. The mechanical ventilation is a balanced constant air volume (CAV) system with heat recovery of 80% efficiency. The air is supplied in bedrooms and living room and is exhausted from kitchen (20 l/s) and bathrooms (15 l/s each). The supply air temperature is set at 17 °C. No mechanical cooling is implemented in the building. The external infiltrations are considered constant and are set to 0.1 l/s m^2.

The domestic hot water system (DHW) is not considered in the simulations. In the building, the system is equipped with a water tank, which provides hot water to the entire building. Since this study focused on one apartment, a scaling procedure of the water tank would have led to not precise results. Moreover, the DHW system is decoupled from the SH system, therefore this assumption does not influence the results.

The study is performed using the simulation software IDA ICE and a multi-zone model of the apartment is created with it. The simulations are run for one month, focusing on the heating season, with specific settings for weather conditions, heating schedule and occupancy characteristics as described in the following paragraphs.

Weather file

To assess the influence of the weather conditions on the flexibility potential, it is decided to vary only two weather parameters (outdoor temperature and solar radiation). For each case, three specific weather files are created based on realistic conditions of Denmark's heating season. The Test Reference Year (TRY) file representing the weather conditions in Vanløse, suburb of Copenhagen, is used as reference file to create the new ones. The outdoor temperature pattern during the heating period is analysed and three representative constant temperatures are selected. The temperatures chosen are 5 °C, 0 °C and -12 °C, respectively the highest, the average and the lowest temperature in January (coldest month in Denmark). In order to limit the influence of temperature variations, the three selected temperatures are kept constant during the simulation period.

The second parameter considered is the direct solar radiation. Similar to the temperature profile, three typical conditions are selected from the TRY file: large solar radiation, narrow solar radiation and no direct solar radiation, based on the hours that direct solar radiation is available. The large solar radiation considers 8 h of direct solar radiation, with a peak of 500 W/m^2 in the middle of the day and 0 W/m^2 at the beginning and the end of it; the narrow solar radiation considers 4 h of direct solar radiation with the same pattern. Regarding the diffuse radiation, it is assumed to have a peak of 50 W/m^2 in the middle of the day, in accordance with the direct solar radiation. These three cases are considered to occur on daily basis throughout the simulation period. The relative humidity is kept constant at 80% in all cases and absence of wind is considered.

Internal gains

In addition to solar radiation, the influence of the occupancy is investigated, simulating with two, three and four occupants. Table 2 gives a detailed description of the schedule applied to the occupants

Table 2: Schedule for four occupants working

Room	Time	No. People [MET] Weekdays	No. People [MET] Weekend
Living Room-Kitchen	7:00 - 9:00	2.1 [1]	2.1 [1]
	9:00 - 17:00	0 [0]	1.2 [1]
	17:00 -23:00	2.1 [1]	2.1 [1]
	23:00 - 7:00	0 [0]	0 [0]
Bedroom 1	7:00 - 17:00	0 [0]	0.4 [1]
	17:00 - 23:00	0 [0]	0.4 [1]
	23:00 - 7:00	2 [0.7]	2 [0.7]
Bedroom 2/3	7:00 - 17:00	0 [0]	0.2 [1]
	17:00 - 23:00	0 [0]	0.2 [1]
	23:00 - 7:00	1 [0.7]	1 [0.7]
Bathrooms	7:00 - 9:00	0.5 [1.2]	0.5 [1.2]
	9:00 - 17:00	0 [1.2]	0.2 [1.2]
	17:00 -23:00	0.5 [1.2]	0.5 [1.2]
	23:00 - 7:00	0 [0]	0 [0]

The occupancy schedule applied to the occupants evaluates two different configurations. The first one considers working people that are not at home during weekdays and the second one considers retired people that spend more time at home throughout the week. For the schedule of retired people, the weekend pattern is applied for all the weekdays.

Heating program

The aim of the heating program is to allow the possibility of load shifting. Two cases are created in order to achieve two different aims during the investigations:

- Cut-off program
- Daily heating program

The first one is used to assess the flexibility potential in terms of the number of hours of acceptable indoor comfort that can be maintained in the apartment once the heating is switched-off. It is related to the first flexibility indicator explained in the next paragraph, and it is used to evaluate the flexibility potential using different internal gains. This heating program is called "cut-off program" since it considers the complete switch-off of the SH system. More specifically, it considers a continuous control of the heating from the 1st of January until the 14th of January, with the room set point at 20 °C. Afterwards, on the 15th of January, from 0:00 until 6:00, the heating set point in the rooms is risen to higher values (21 °C, 22 °C, 23 °C and 24 °C), in order to enable heat storage in the building mass. Afterwards, the heating is switched-off and the temperature drop in the apartment is analysed.

The second program implemented considers a daily pattern. It is used to determine whether a daily heating program could provide flexibility by storing heat into the building mass. The heating pattern used is the same as the one used for the cut-off program, but in this case after the 15th of January, the SH system is set to work every day from 0:00 to 6:00. As previously, different set points are tested during the six hours heating period (21 °C, 22 °C, 23 °C and 24 °C).

Flexibility indicator

In a residential apartment, it is necessary to guarantee the occupants' comfort. In this study, the thermal comfort, and in particular the operative temperature is used as indicator to evaluate the flexibility potential of the building. The requirements provided by the European Standard DS/EN 15251 (Dansk Standard, 2007) for a building in comfort category II are followed, that require minimum operative temperature at 20 °C during the heating season.

The first flexibility indicator represents the number of hours that the operative temperature is kept above the lowest limit recommended by the standard (20 °C), once the heating is switched-off according to the cut-off program. Figure 2 shows a simplified representation of the first flexibility indicator.

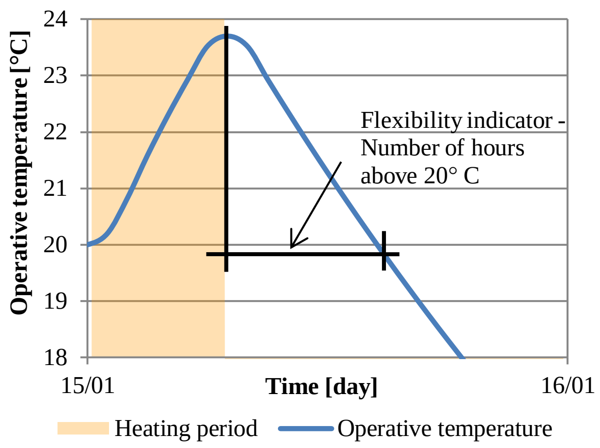


Figure 2: Representation of the flexibility indicator

The following formula (1) explains the concept adopted.

$$Ind_1 = min\ \{t \mid T_{op}(t) \geq 20\ °C\} \qquad (1)$$

where,

t is the time with a temperature above 20 °C after the heating cut-off,

T_{op} is the operative temperature in the room.

The indicator quantifies the time that the building can be independent of the heating supply, which corresponds to a period when the renewable energy is not available.

A second indicator is defined to evaluate the effect of implementing a dedicated heating program to take advantage of the flexibility potential in terms of power peaks. In fact, the new heating program created, considers raising the set point for a limited period and then switching-off the heating system. Usually, in buildings the heating set point is constant and the SH system controls the indoor temperature throughout the day. The consumption peak obtained when implementing a continuous control of the SH system is compared with the one obtained using the daily heating program presented in the previous paragraph. The indicator is defined as:

$$Ind_2 = \frac{P_{\max\_daily}}{P_{\max\_cont}} \qquad (2)$$

where,

P_{max_daily} is the consumption peak obtained with the daily heating program,

P_{max_cont} is the consumption peak obtained with the continuous control.

The flexibility indicator gives an evaluation of the impact that the implementation of energy flexibility has on the dimensioning of the heating system.

Results

Flexibility potential

The results of two reference rooms are plotted in the following graphs, Bedroom 1 and Kitchen-Living Room, since they have opposite orientation and they represent two “end-use” cases, with different occupancy schedules. In terms of thermal comfort, they represent the best and worst case respectively. The results shown refer to heating set point of 21 °C, since for higher set points, the same pattern is found and the difference is only in the magnitude of the indicator. It is important to mention that a basic case is created, and every time a single parameter is changed. The basic case considers narrow solar radiation, outdoor temperature at 0 °C and three occupants following the schedule for working people.

Figure 3 presents the flexibility indicator as number of hours above 20 °C. In this particular case, it quantifies the time that the building can be independent of the heating supply for the three different direct solar radiation cases.

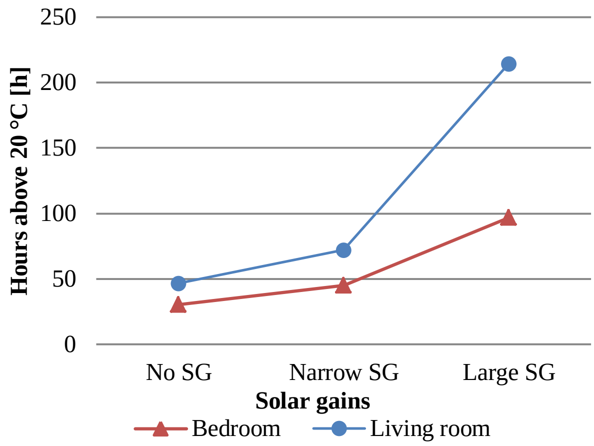


Figure 3: Flexibility indicator with different solar radiation

The impact of solar gains in the two rooms is presented with regards to the first indicator. This impact is higher in the living room than in the bedroom, mainly due to the orientation of the room that is facing south. Moreover, the glazing area is larger in the living room. It can be seen that the impact of the narrow solar gains compared to the absence of solar gains, increases the number of hours above 20 °C with a factor of 1.47 in the bedroom, while in the living room with 1.54. On the other hand, when considering the transition from narrow solar gains to large solar gains, the increasing factor is 2.14 in the bedroom and 2.96 in the living room. Therefore, it can be deduced that the indicator is affected more in case of large solar gains and that the living room is more sensitive to this change since it faces south.

In Figure 4, the results of the first flexibility indicator for different outdoor temperatures are presented. It is noticeable that the outdoor temperature influences the indicator with an exponential pattern. The results obtained with the outdoor temperature at -12 °C show that the temperature in the rooms drops rather faster and a lower flexibility can be achieved compared with the other temperatures tested. The decrease of the flexibility indicator, when the temperature at -12 °C is considered, is about 4.51 times for the bedroom and 3.93 times for the living room. On the other hand, when the temperature is changed to 5 °C, the indicator increased with a factor of 2.14 for the bedroom and 2.05 for the living room. Therefore, it is seen that the outdoor temperature is an important factor for the thermal comfort in the rooms as it was expected, and this is revealed from the flexibility indicator. In extreme outdoor conditions (-12 °C), the flexibility of the building is limited (i.e. 10 hours in the bedroom). As it is noticed also in the solar gains analysis, the living room always ensures a higher indicator due to the fact it faces south and it is more affected by the solar radiation.

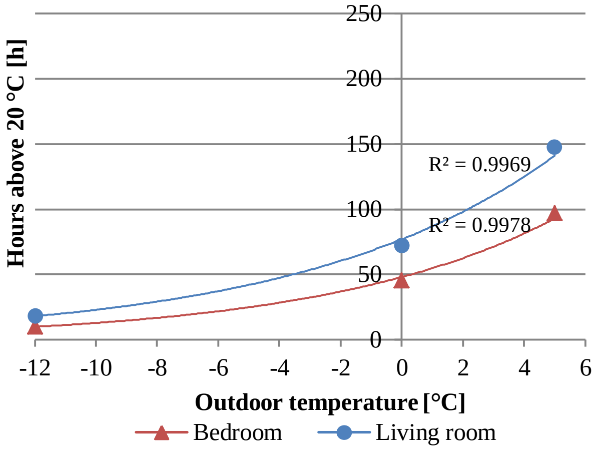


Figure 4: Flexibility indicator with different outdoor temperatures

Figure 5 shows the flexibility indicator results when different numbers of occupants are present in the apartment.

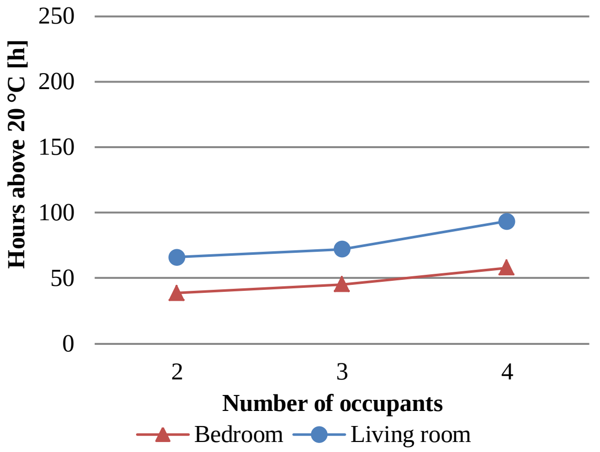


Figure 5: Flexibility indicator with different number of occupants

It can be seen that the increase in number of hours above 20 °C is almost linear for different number of occupants. The decrease of the flexibility indicator is about 1.18 times in the bedroom, when it is considered to decrease from three to two occupants, while it is about 1.09 in the living room. On the other hand, when four people are considered, compared to the case of three occupants, the flexibility indicator increased with a factor of 1.28 in the bedroom and 1.29 in the living room. It can be deduced that increasing the number of occupants by one unit can influence almost linearly the indicator. As it is seen in the previous cases, the living room guarantees a higher indicator due to the higher solar gains.

Figure 6 shows the results of the first flexibility indicator for the two different occupancy schedules. It is observed that for the schedule of retired people the flexibility indicator is higher. In particular, it increases with a factor 1.44 in the bedroom and approximately 1.26 in the living room. This is due to the time that the people spend in the apartment, which is assumed longer for retired people.

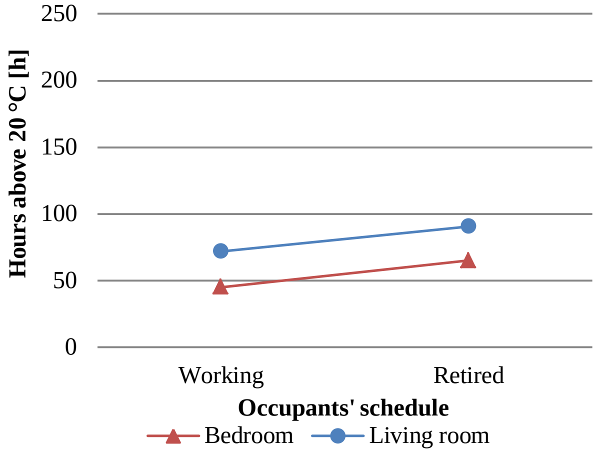


Figure 6: Flexibility indicator with different occupancy schedule

The following two analyses give an overview in case a different apartment in the same block is chosen as reference. In the first case, the building is rotated by 180°, in order to have the main façade oriented towards the north, while keeping the same allocation of the rooms. In the second one, the reference apartment is moved from the fourth floor to the ground floor and the top floor.

Figure 7 shows the flexibility indicator for two different orientation of the building.

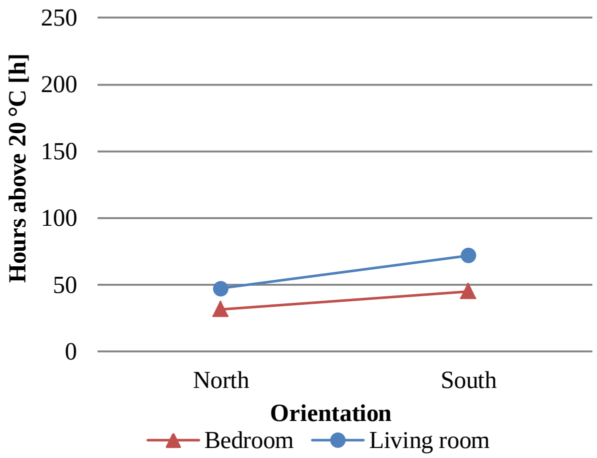


Figure 7: Flexibility indicator with different orientation of the building

When the main façade is oriented towards the south, the living room faces south and the bedroom faces north. The large glazing area of the first room guarantees the highest flexibility indicator, due to the high solar gains in the apartment. Consequently, the highest flexibility indicator is found in the living room. On the other hand, when the main façade faces north, the solar gains in the apartment are lower, resulting in a lower flexibility indicator. In this case, it is expected to get a higher indicator in the bedroom, since it is facing south. However, the room faces the internal yard, and the opposite side of the building creates shading thus reducing the solar gains in the room and attenuating the effect of being oriented to the south. Furthermore, higher internal gains in the living room still ensure a higher flexibility indicator in this room. This can be seen in Figure 7, where the indicator increases for the south orientation, 1.52 times higher than the north once, while in the bedroom the difference is lower, where the indicator is 1.43 times higher in the south oriented case.

Figure 8 represents the flexibility indicator when the apartment is moved at different floor levels.

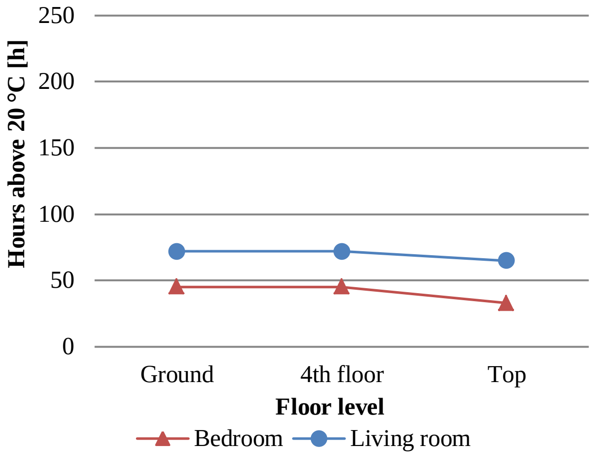


Figure 8: Flexibility indicator with different heights of the apartment

It is noticeable that for the cases at the ground floor and in the fourth floor, the results obtained are the same. Since the building has a heated basement, the losses in the ground floor are limited. However, the flexibility potential decreases 1.12 times in the living room and 1.37 times in the bedroom, when the apartment is located at the top floor. In this case, the apartment has higher heat losses since the exposed wall and roof area towards the ambient are bigger compared to the other cases.

Heating programs

In the second part of the study, the flexibility potential is investigated with particular attention to the heating system. Firstly, the flexibility potential is tested for different heating set points (20 °C – 24 °C) with the cut-off program for the basic case. Afterwards, the daily heating program is implemented and the possibility of implementing load shifting is investigated. Figure 9 shows the results of the flexibility indicator *Ind1* when the cut-off program is applied.

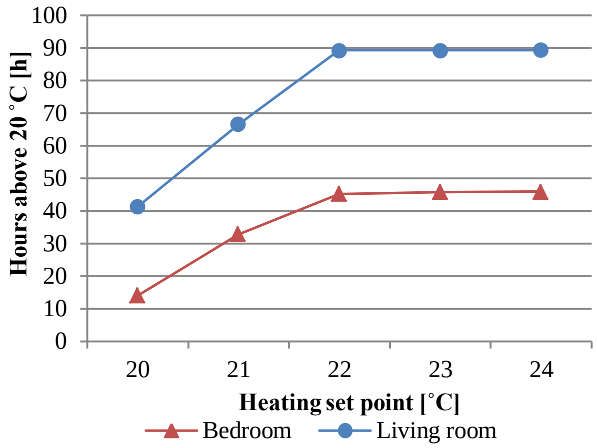


Figure 9: Flexibility indicator obtained with the cut-off program at different heating set points

It can be noticed that the heating set point used in the rooms during the heating period influences the flexibility indicator. In particular, for both the bedroom and the living room, the maximum amount of hours is reached with a set point of 22 °C, namely 89 h for the living room and 45 h for the bedroom. When higher set points are used (23 °C or 24 °C), the same results in terms of flexibility indicator are achieved, due to the thermal inertia of the floor heating system, namely that the temperature set point cannot be reached within the six hours of pre-heating. Thus, higher set points would not make any difference in the energy stored in the thermal mass. On the other hand, a set point lower than 22 °C, i.e. 20 °C and 21 °C, leads to a lower number of hours above the comfort limit. When the set point is increased from 20 °C to 22 °C in order to obtain the maximum result in terms of flexibility, the indicator increases with a factor of 3.22 for the bedroom and 2.16 for the living room.

Figure 10 shows the operative temperature pattern from January 14th until January 20th in the bedroom when the cut-off program is applied.

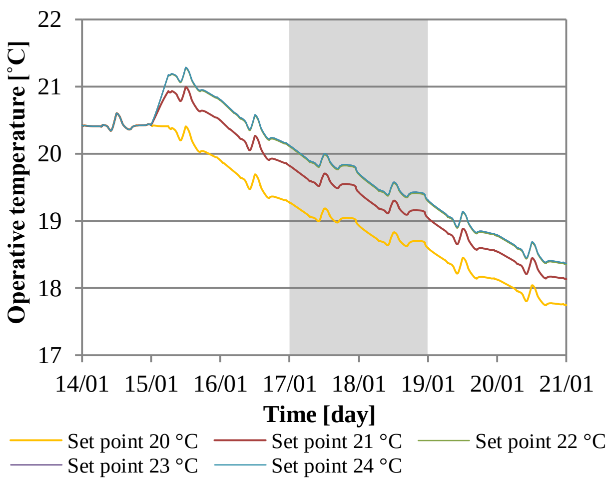


Figure 10: Temperature pattern in the bedroom for the cut-off program

As expected from the evaluation of the flexibility indicator, for set points 22 °C, 23 °C and 24 °C set points, the temperature pattern overlaps. It is possible to notice the influence of the solar radiation, which can be observed in the temperature peak in the middle of each day. The weekend is highlighted (grey area) and it can be seen that the operative temperature drop slightly decreases, due to the fact that the occupancy schedule considered that people spend more time at home. It is also noticed that the initial temperature in the beginning of the heating program is 20.4 °C. This is a bit higher than the set point (20 °C) for the period before the 15th of January. The reason is the proportional controller implemented in the model that introduces this error. It has a bandwidth of 2 °C and it stops the heating supply only when the operative temperature reaches the set point plus half of the bandwidth. The temperature pattern in the living room shows similar behaviour to the one in the bedroom with the only difference of higher temperatures, due to larger influence of the solar gains.

In the second step of the investigation, the heating program is applied on a daily basis. Figure 11 shows the temperature patterns in the bedroom at different heating set points when the daily heating program is applied. For the set points above 21 °C the six hours heating period guarantees acceptable thermal comfort in the rooms, thanks to the heat storage in the building mass. In fact, for all set points higher than 20 °C the temperature increases day by day. This means that during the heating period, the energy accumulated in the concrete slab increases day by day. Only for the set point at 20 °C, the temperature drops below the limit since the energy stored during the heating period is not sufficient to keep the temperature above the limit during the period with no heating. The grey area highlight the weekend, where the occupants spend more time in the apartment.

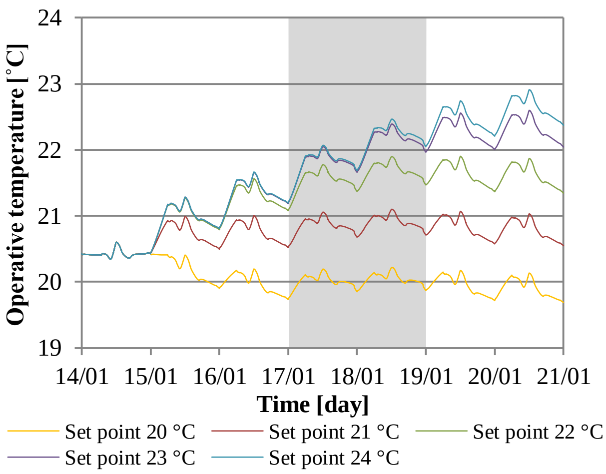


Figure 11: Temperature pattern in the bedroom for the daily heating program

The change of heating program is considered as a DSM solution and it is important to evaluate the effect of its implementation. A higher set point leads to higher power peak, compared to the continuous control of the heating system. Table 3 shows the consumption peaks for the cut-off program on the 15th of January.

Table 3: Consumption peaks for the cut-off program for different heating set points

Continuous control	
Heating set point [°C]	Peak consumption [W]
20	743
Cut-off program	
20	709
21	3785
22	4829
23	4877
24	4877

As it is expected from the operative temperature trends, the highest peak is obtained for the set points of 22 °C, 23 °C and 24 °C. This is due to the higher temperature difference between the operative temperature in the room and the set point temperature which influence the use of the proportional control. For this reason, in case of the 21 °C set point, the consumption is significantly lower, since the room temperature is closer to the set point. The highest peaks measured are slightly lower than 4900 W while for the set point of 21 °C it is approximately 3785 W. For the set point of 20 °C, the peak consumption is much lower, 743 W, but as previously shown, it provides lower flexibility. It can be concluded that, in order to enable the maximum energy storage and flexibility potential in the building mass with the increase of the heating set point, a higher power peak is required. Comparing the highest peak, (set point at 23 °C or 24 °C), with the one obtained with the continuous control of the SH system (set point 20 °C), the increase factor is approximately 6.56 times.

Regarding the daily heating program, in order to assess energy flexibility on a daily basis, the primary energy consumption of the heating system is evaluated. In this case, only the set point at 21 °C is considered, since it is the lowest one that guarantees an operative temperature above 20 °C in all the rooms. Table 4 shows the results in terms of primary energy consumption after one month.

Table 4: Primary energy consumption after one-month simulation

Primary Energy Consumption [kWh/m^2]	
Continuous control	2.89
Daily heating program	2.99

The primary energy consumption obtained with the continuous control is compared with the one obtained with the daily heating program. As it is noticed, the daily heating program leads to 3.5% higher energy consumption. Thus, in order to implement energy flexibility measures in buildings, an overconsumption of the SH system might need to be faced.

Discussion

The results obtained in this study highlighted the energy flexibility potential obtained in a newly constructed building in Denmark. The most updated construction technologies ensure lower energy consumption, lower heat losses and lower infiltration, leading to higher flexibility potential of new buildings, compared to buildings built following older legislations.

As explained also in the work presented by Le Dréau and Heiselberg (2016), which showed heating demand autonomy up to 48 h in a new passive house, new buildings can largely contribute to increase the flexibility of the energy system. In this study, the high amount of hours above 20° C (about 200 h) is found with a specific weather file, where the solar radiation and the outdoor temperature highly contribute to keep an acceptable indoor temperature. The higher flexibility potential, besides the differences in the weather condition, is due to the type of house considered in the two studies. In fact, a single family house, with higher external wall surface, has higher heat losses compared to an apartment.

During the investigation procedure, many assumptions had to be made in order to simplify the comprehension of the parameters changed. For this reason, the results need to be treated carefully according to the specific study case. However, they represent a starting point for further investigations, in which all assumptions can be sharpened in order to evolve it to a more generic model.

The specific heating program used during the investigation is based on the evaluation of consumption patterns found in the literature. According to (Fischer et al., 2016) in the daily heating demand, two peaks can be identified: the first one in the morning around 6 a.m. and the second one around 6 p.m. Since the building investigated is located in a district, where there are both residential and office buildings, synergies between the two types of buildings could be developed. The heating operator would supply the two types of buildings at different moments during the day, in order to reduce the peak consumption. Since the lowest consumption of the offices is during night, in this paper it is decided to implement a heating program that ensures the energy supply during night for the residential buildings. This solution could also be applied in case of electrical flexibility, since during the night the electricity price is usually lower due to the night tariffs regulations.

The results obtained from the investigations refer to a new low-energy building; therefore, using different type of buildings (different age of construction or not refurbished) to run the same simulations would lead to different flexibility potentials. In particular, with the perspective of extending the smart grid concept to a larger area, other types of buildings need to be studied to evaluate the impact and the relation of different constructions on the flexibility potential.

The main part of the investigations conducted is based on a simulation model that considered the weather conditions, i.e. outdoor temperature and solar radiation, constant throughout the period. This assumption was made in order to limit the influence of external parameters and it is possible to isolate the effect of each parameter. Definitely, variable weather conditions would have led to different results in terms of flexibility. However, the constant temperature and solar radiation are estimated based on the average conditions during the heating season. This makes the assumption consistent and the results more indicative.

The number of occupants and the schedule that regulate their presence are defined equally for each weekday and for each weekend. In reality, people behave in a random way and this could influence the results. In a further study, it would worth implementing a stochastic model. Moreover, to simplify the investigations, the occupied hours are not distinguished from the non-occupied ones when analysing the results and the minimum indoor temperature is set at 20 °C all times. Finally, the minimum temperature was set according to Category II of the Standard DS/EN 15251, while in reality, each occupant has a different level of acceptability and it could be interesting for further analysis to investigate if different set point temperatures in the rooms can influence the flexibility (i.e. lower set point than 20 °C in bedrooms).

Regarding the apartment's position, it emerged that a different orientation can affect its energy flexibility potential. In the same way, the floor level at which the apartment is located, can affect the indicator due to the different heat losses. Apartment and rooms oriented to the south can ensure higher flexibility as expected, while apartment located at the last floor or at the corners have a lower potential. Therefore, in order to take advantage of the flexibility potential in an apartment block, a dynamic system has to be considered, which evaluates the spatial differences and adapts the heating supply in order to ensure the desirable indoor temperature.

Conclusion

This study gives a first evaluation of how the flexibility potential depends on internal gains and weather conditions. To reach the aim, two flexibility indicators are defined and applied during the investigation. The study highlighted the high influence of solar radiation and outdoor temperature on the indicator, while showed a lower impact of the number of occupants and the occupant's schedule. The analyses about the apartment's position demonstrated that the orientation influences the flexibility indicator the most, while the floor level has a lower impact on it.

The south oriented rooms proved to be more sensitive to solar radiation, as the first indicator is found to be always higher compared to the rooms facing north. The highest impact on the flexibility indicator is given by the large solar radiation and the highest number is found at 215 hours in the living room. In this study case, the number of hours increased around four times compared to the lowest case obtained with no solar radiation. In the same way, the outdoor temperature is found to have high influence on the indoor comfort and therefore, on the possibility of using the flexibility of the building. In particular, for the extreme outdoor condition of -12 °C, the indicator showed the lowest result equal to 10 hours above 20 °C. Regarding the internal gains related to the number of occupants in the apartment and their schedules, it is noticed that increasing the number of occupants by one unit, the increase in hours is almost linear. From the study appears that solar radiation and outdoor temperature have high impact on the indoor comfort in buildings. Therefore, weather forecasts can be used as means to predict the flexibility potential of buildings and if integrated in a smart city network they can be used to achieve results in terms of energy consumption reduction, while maximizing the use of renewable energy.

Implementing flexibility with a dedicated heating program, such as the cut-off program, shows that higher power peaks have to be faced. In fact, the power peak found for the heating systems results 6 times higher than in case of continuous control. This result has to be treated carefully, since it is dependent on the heating control used, as well as it has to be taken into account that only the coldest month was considered. However, despite of the good results in terms of flexibility, changes in heating program need to be evaluated carefully, since due to the higher power peaks, the building's systems might have to be designed accordingly. Furthermore, the primary energy consumption of the heating system with the daily heating program is found to be 3.5 % higher than the continuous control. Therefore, implementing energy storage in the building mass with dedicated heating programs leads to a slight overconsumption, which can be justified if this ensures higher use of energy from RES.

Acknowledgements

This study is part of the project EnergyLab Nordhavn - New Urban Energy Infrastructures and the Danish research project CITIES (Centre for IT-Intelligent Energy Systems in cities).

References

Aduda, K.O., Labeodan, T., Zeiler, W., Boxem, G. and Zhao, Y. (2016). Demand side flexibility: Potentials and building performance implications. *Sustainable Cities and Society*, 22, 146–163.

Chen, D., Wang, X. and Ren, Z. (2012). Selection of climatic variables and time scales for future weather preparation in building heating and cooling energy predictions. *Energy and Buildings*, 51, 223–233.

Dansk Standard. (2007). DS/EN 15251 indoor environmental input parameters for design and assessment of energy performance of buildings addressing indoor air quality, thermal environment, lighting and acoustics. *Danish Standard*, 1st ed.

Firląg, S. and Murray, S. (2013). Impacts of airflows, internal heat and moisture gains on accuracy of modeling energy consumption and indoor parameters in passive building. *Energy and Buildings*, 64, 372–383.

Fischer, D., Wolf, T., Scherer, J. and Wille-Haussmann, B. (2016). A stochastic bottom-up model for space heating and domestic hot water load profiles for German households. *Energy and Buildings*, 124, 120–128.

Gellings, C. W. (1985). The concept of demand-side management for electric utilities. *Proceedings of the IEEE*, 73(10), 1468–1470.

Labeodan, T., Zeiler, W., Boxem, G. and Zhao, Y. (2015). Occupancy measurement in commercial office buildings for demand-driven control applications - A survey and detection system evaluation. *Energy and Buildings*, 93, 303–314.

Lazos, D., Sproul, A. B. and Kay, M. (2014). Optimisation of energy management in commercial buildings with weather forecasting inputs: A review. *Renewable and Sustainable Energy Reviews*, 39, 587–603.

Le Dréau, J. and Heiselberg, P. (2016). Energy flexibility of residential buildings using short term heat storage in the thermal mass. *Energy*, 111, 991–1002.

Ling, Z.G., Zhang, S.Y., Fu, X.Z. and Wang, Y. (2006). Investigation of floor heating with thermal storage, *Journal OSUT* 13(4), 1974–1975.

Lund, P.D., Lindgren, J., Mikkola, J. and Salpakari, J. (2015). Review of energy system flexibility measures to enable high levels of variable renewable electricity. *Renewable & Sustainable Energy Reviews*, Elsevier Ltd, 45, 785–807.

Marin, P., Saffari, M., de Gracia, A., Zhu, X., Farid, M. M., Cabeza, L. F. and Ushak, S. (2016). Energy savings due to the use of PCM for relocatable lightweight buildings passive heating and cooling in different weather conditions. *Energy and Buildings*, 129, 274–283.

Neupane, B., Bach, T. and Thiesson, B. (2014). Towards Flexibility Detection in Device-Level Energy Consumption. *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 8817, 1–16.

Rotger-Griful, S. and Hylsberg Jacobsen, R. (2015). Control of Smart Grid ResidentialBuildings with Demand Response. Chaos Modelling and Control Systems Design. (A. T. Azar & S. Vaidyanathan, Eds.). *Studies in Computational Intelligence* (Vol. 581), 133-161, Springer International Publishing.

Sikula, O., Plasek, J. and Hirs, J. (2012). Numerical simulation of the effect of heat gains in the heating season. *Energy Procedia*, 14, 906–912.

Werner, S. and Olsson Ingvarsson, L. (2008). Building mass used as short term heat storage. *In Proceedings of the 11th International Symposium on District Heating and Cooling*. Reykjavik, Iceland.

Xue, X., Wang, S., Yan, C. and Cui, B. (2015). A fast chiller power demand response control strategy for buildings connected to smart grid. *Applied Energy*, 137(Jan), 77–87.

Zheng, M., Meinrenken, C.J. and Lackner, K.S. (2015). Smart households: Dispatch strategies and economic analysis of distributed energy storage for residential peak saving. *Applied Energy*, 147, 246–257.

Paper IX: Gianniou, P., Heller, A., Nielsen, P. S., & Rode, C. (2016). Identification of parameters affecting the variability of energy use in residential buildings. In Proceedings of the *12th REHVA World Congress, Aalborg, Denmark.*

Identification of Parameters Affecting the Variability of Energy Use in Residential Buildings

Panagiota Gianniou[#1], Alfred Heller[#2], Per Sieverts Nielsen[*3], Carsten Rode[#4]

[#]*Department of Civil Engineering, Technical University of Denmark*

[1]pagian@byg.dtu.dk

[2]alfh @byg.dtu.dk

[4]car@byg.dtu.dk

[*]*Department of Management Engineering, Technical University of Denmark*

[3]pernn@dtu.dk

Abstract

Energy use of buildings varies significantly. When aggregating the demand profiles of a group of buildings, the variations of energy demand are critical to determine the aggregated load profile. Especially when dimensioning district energy systems, it is important to know the variability of energy demand that can guarantee the efficient operation of the system. For this reason, it is useful to distinguish the parameters that affect building energy performance the most and to estimate the magnitude of these variations on each parameter.

The aim of the present study is to identify the parameters that lead to the largest variations in energy performance of residential buildings in Denmark. A set of sensitivity analysis has been carried out using an extensive search algorithm. These sensitivity analyses were then applied for modelling a reference building representing Danish single-family houses of the 1940's. The study was able to determine the key variables that affect energy use in old Danish single-family houses using sensitivity analysis and proposes a methodology for parameter optimization. This analysis pointed out that the insulation in external walls and roof lead to the largest variations in space heating demand. Also, the infiltration rate and occupancy behavior play important role on space heating consumption. It was concluded that these findings highly depend on the specific case study and the characteristics of the buildings that are examined. If outdoor climate and location differ from the current case, a different set of parameters should be investigated upon its effect on building energy use.

Keywords – heat demand; variability; sensitivity analysis; building parameters

1. Introduction

Variations in energy use of buildings are considered significant issues for their integration with the overall energy grid. Especially in the framework of Energy Hubs [1] where integrated district energy systems are dominant, the accurate prediction of the operational energy use of a building or clusters of buildings is crucial. According to the findings of IEA EBC Annex 53 [2], these variations are mainly attributed to the following areas: building envelope and equipment, operation and maintenance, occupant behavior and indoor environmental conditions, as well as outdoor climate.

Uncertainty and sensitivity analyses are an integral part of the modelling process [3]. Uncertainty characterizes both passive building parameters (e.g. building structure) and active components (e.g. users, climate). Uncertainty analysis has proven to be the main approach to determine the effect of uncertainty on building energy performance. Moreover, sensitivity analysis is conducted to analyze variations in building parameters and their effect on energy use. Generally, the methodologies found in literature are classified into local and global sensitivity analyses (i.e. regression, variance-based, screening-based and meta-model based method). The main difference between these two methodologies is that the local sensitivity analysis emphasizes on the individual effect that each parameter has on the output, while the global analysis considers the interactions among the examined parameters [4]. However, the local sensitivity analysis is easily applied and often requires less simulation runs than the global one.

Previous studies have extensively used sensitivity analysis in building energy analysis such as [5], [6]. In particular, [7] showed quite early that variation of inhabitants' has a much higher effect on energy use than variation of outdoor climate using Monte Carlo technique. [8] identified the most important parameters affecting the total energy use of sustainable buildings to be lighting control and ventilation rate during winter. When heating demand was examined, the authors concluded that ventilation plays the most important role. In another study by [9], the most critical parameters on heating consumption were found to be thermostat set points and ventilation flow rate followed by the properties of windows. [10] concluded that the biggest factor in heating loads was fabric properties (insulation, glazing, infiltration) followed by floor area, heating set points and orientation, while occupancy behavior and lighting had a very low impact. Therefore, results from previous studies vary quite a lot depending on the reference case and on the purpose of analysis.

First, the present study aims at identifying the parameters that affect the variability of building energy use and specifically space heating demand focusing on residential districts. Second, it aims at estimating the changes in heating demand that the variations of these parameters lead to. In the case of district-level energy systems, the variability of building energy use makes it difficult to estimate the aggregate demand and then, dimension the systems accordingly. Thus, it is very important to determine the parameters that highly influence energy use in buildings when aggregated to specific districts or cities.

2. Methodology

A typical Danish single-family house of the 1940's was used as a reference building. Information about the building was collected by the TABULA Webtool [11] and [12]. The building model was created in IDA ICE using one single zone. The model properties are presented in Table 1. The building was connected to the local district heating system, while no mechanical ventilation or cooling system was installed in the building. The location of the building was set to Copenhagen and it was simulated using the Danish Design Reference Year (DRY) weather file [13].

The parametric analysis was made with the open source software MOBO [14]. MOBO is originally meant to handle multi-objective optimization problems with constraint functions. It provides the possibility to be coupled to building performance simulation tools, while selecting different algorithms to perform the calculations. For this study, no optimization was needed but a sensitivity analysis was carried out instead. An exhaustive search method was selected, namely Brute-force algorithm to conduct a local sensitivity analysis. Brute-force is a very simple straight-forward approach that can sample the whole solution space. For this reason, it is very computationally expensive and not very fast. However, it is very accurate and for the simple case examined in the present study the computation time did not exceed 15 minutes for each parameter.

First, the parameters were identified. It was decided that properties of building envelope and occupancy behavior would be investigated for their effect on heating demand. Outdoor climate was not studied hereby because the purpose of the analysis is to emphasize on residential buildings that are located in the same location (district, city) so they share the same climatic conditions. Moreover, no mechanical ventilation was installed in the building since it is very uncommon in older Danish single-family houses.

MOBO requires setting the initial and final value of each parameter to restrict the spectrum of calculations, as well as the steps. These are illustrated in Table 2. The steps were defined such that 50 simulations are executed in every case. The range of each parameter was very important because it determined the variations in heat demand. Thus, it was selected so that it is realistic for Danish single-family houses based on the TABULA database, as well as on Danish Statistics. More variables were examined, but the ones with the most significant effect are presented in the following.

Since the range of each parameter differs, a dimensionless impact factor had to be defined to indicate the changes in heating demand the variations of parameters caused. So, this was defined according to (1), where output is the space heating demand in kWh.

$$I = \frac{\max(output) - \min(output)}{average(output)} \tag{1}$$

Table 1. Properties of reference building

Properties	Value
Model floor area	90.9 m^2
Model volume	236.2 m^3
Model ground area	90.9 m^2
Model envelope area	282.6 m^2
Window/floor area	16%
Average U-value	0.82 W/K m^2
Number of occupants	2.9
Infiltration rate	0.4 ACH
Orientation	50 deg

Table 2. Examined parameters for sensitivity analysis

Parameters	Type	Min. value	Max. value
Orientation (deg)	Continuous	0	350
U-value of glazing (W/m^2K)	Continuous	1	3
Insulation thickness in ext. walls (m)	Continuous	0.001	0.32
Infiltration rate (ACH)	Continuous	0.1	1
Number of occupants	Continuous	0	6
Insulation thickness in roof (m)	Continuous	0.001	0.32
Insulation thickness in floor (m)	Continuous	0.001	0.35

3. Results and Discussion

The results of the parametric simulations in MOBO are illustrated in Figure 1. The blue curves in the graphs represent the variable that was changing while the red ones represent the respective annual space heating demand that was calculated for these changes after 50 simulations.

It can be observed that for many parameters the influence on space heating demand is almost linear. This means that the impact does not differ significantly depending on the parameter range. Orientation has a more complicated correlation to energy use, since solar gains change dramatically with the positioning and orientation of the building. Increase of insulation thickness of walls and roof, as well as occupancy behavior, result in a decrease of energy use and specifically space heating demand. In particular, the more these parameters increase, the lower the heating demand of the building is, as it was expected. The effect of floor insulation thickness on heating demand had a similar trend with the roof insulation and it is not presented in Figure 1 for brevity. Specifically, the curves of the roof and external wall insulation thickness show a hyperbolic behavior. It was also found that increase of infiltration rate results in a linear increase of the heating demand. Moreover, the U-values of the windows were investigated. It was observed that heating demand increased linearly with the increase of the U-value of the windows. It is evident from Figure 1 that the range of heating demand differs significantly depending on the investigated parameter.

Thereafter, the impact factor of all the afore-mentioned parameters on space heating demand was calculated for the reference building according to (1) representing their magnitude on the variability of heating demand. These results are presented in Table 3. The positive values indicate an influence to increase the heating demand, while the negative ones indicate a diminishing impact on the heating demand. It can be observed that the most influential parameters for the reference building, for the given ranges corresponding to Danish houses, were the insulation thickness of external wall with an impact factor of -1.021. The roof insulation followed with an impact factor of -0.242. Then the infiltration rate followed with an impact factor of 0.141. The influence of occupancy on heating demand was found to be -0.101. The impact factor for glazing U-values was calculated to be 0.058. Lastly, variations of floor insulation thickness and orientation were found to lead to the smallest changes in annual heating demand. Specifically, the impact factor of floor insulation was found to be significantly lower than the one of wall or roof insulation. This is attributed to the fact that the heat losses through the ground are lower for the floor, since the ground layers below the floor slab have thermal resistance and operate as a buffer for heating.

Figure 2 presents the distribution of annual space heating demands for each parameter. The reference line corresponds to the annual heating demand of the reference building which was found to be 14,135 kWh, having the properties as presented in Table 1. Many variations in heating demand are pronounced, the biggest factor being wall insulation thickness with a total range of 8,178 kWh, a mean heating demand of 8,011 kWh and a median heating demand of 7,232 kWh. Variations in roof insulation thickness result to lower changes in heating demand with a total range of 3,410 kWh, a mean value of 14,066 kWh and a median value of 13,759 kWh. Both wall and roof insulation thickness have five or four outliers, meaning that these lie an abnormal distance from the rest of the values. This was expected from Figure 1 due to the hyperbolic shape of the curve, and the low values for minimal insulation thickness. Infiltration and occupancy follow with a total range of 2,039 kWh and 1,423 kWh respectively. Lastly, the total range of heating demand for the variations in glazing U-value, floor insulation thickness and orientation are 844 kWh, 414 kWh and 178 kWh, respectively. This means that orientation and window U-values have minimal impact on the annual heating demand, with glazing U-values having greater influence.

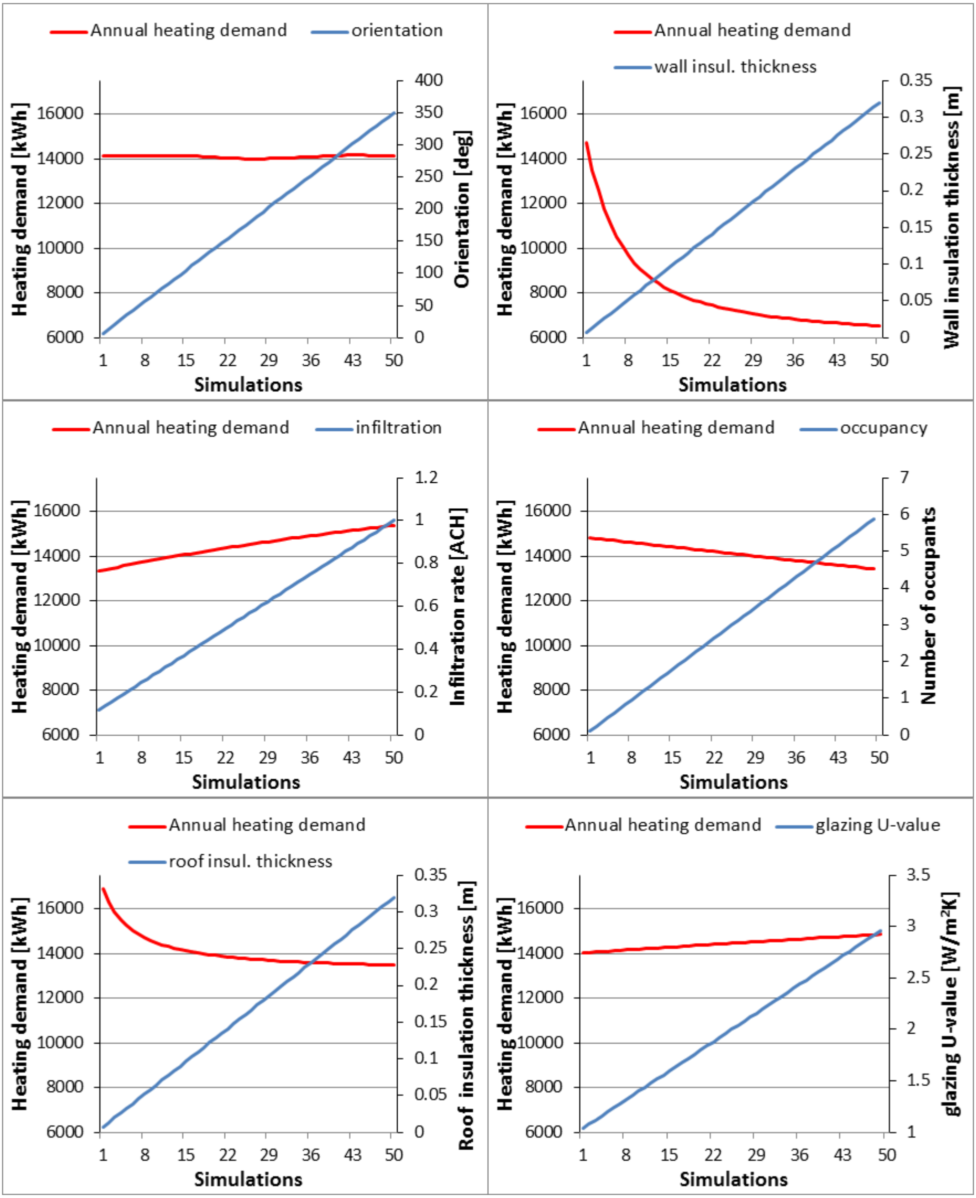


Figure 1. Annual space heating demand for each parameter after 50 simulations in MOBO

Table 3. Results of impact factor for each parameter

Parameter	*I*, Impact factor
Orientation	0.013
U-value of glazing	0.058
Insulation thickness in ext. walls	-1.021
Infiltration rate	0.141
Number of occupants	-0.101
Insulation thickness in roof	-0.242
Insulation thickness in floor	-0.030

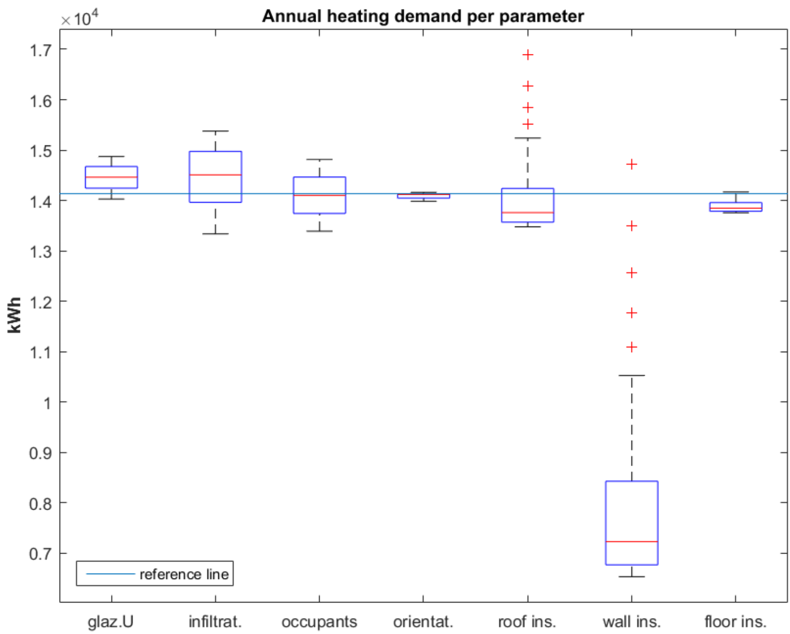


Figure 2. Distributions of annual space heating demand for each parameter. Box plots present the median (red line), quartiles (blue box) and extremes (whiskers). The red crosses indicate the outliers.

The structure of the reference building affects significantly the findings described before, so they cannot be generalized for every single-family house. It has to be pointed out that the reference building representing an old Danish house of the 1940's was poorly insulated with an average U-value of 0.82 W/m^{2}K. Specifically, the external walls and floor of this type of houses are very poorly insulated and only the roof has undergone some energy refurbishment concerning the addition of insulation [12], [15]. The poor insulation of the reference building determined the large impact of insulation thickness of walls on heating demand. At the same time, it affected the impact of secondary parameters such as orientation and occupancy, which were calculated to be significantly lower. If a very-well insulated building was examined instead, the impact factor of these parameters would have been much different. However, the purpose of this study was to propose a methodology for the estimation of magnitude of parameters affecting the variability of energy demand in buildings. In general, homogeneity characterizes most houses built in the same period in Denmark. Variability in occupancy was only accounted for through the number of occupants in this study. Nevertheless, stochasticity in occupant profiles may result in significant differences in similar buildings. So, probabilistic profiles accounting for stochasticity should be considered in the next step.

The weight of the building envelope properties (insulation thickness, infiltration) on the energy use is outlined in the present study. This is also due to the cold Danish climate that causes increased heat losses during the winter period and subsequently increased heating demand. If a warmer climate was applied, the results would be much different. The parameters that affect cooling demand would be more in focus in that case, such as windows g-value and ventilation. The interactions of parameters could be considered through implementing global sensitivity analyses. These interactions are even more important in cases where total energy use is examined because internal heat gains from solar radiation, occupants, appliances and natural ventilation interact highly with each other. Moreover, if the total Danish building stock was to be investigated, the combination of variations in parameters would be necessary to conduct sensitivity analysis.

4. Conclusions

A local sensitivity analysis was applied in the present study through an exhaustive search algorithm. The reference building representing an old Danish single-family house of the 1940’s was modelled in IDA ICE and parametric simulations were run in MOBO. The reported results of the sensitivity analysis indicated that the parameters resulting to the greatest variations in space heating demand were the insulation thickness of external walls and roof. The infiltration rate and number of occupants had significant although lower impacts on heating demand. These findings are highly dependent on the structure of the examined reference building, which in this case was poorly insulated. Different occupancy profiles were not examined in the study. Even if a district containing the same types of buildings would be studied, variations in occupant behavior may result in significant variations in total energy use. To be able to determine the total variations of heating demand in Danish houses, the combination of the aforementioned parameters is suggested to cover the whole range of Danish buildings' properties.

This study has aimed at proposing a simple methodology of estimating the impact of building parameters on energy use focusing on specific building typologies. Nevertheless, more sophisticated approaches should also be investigated and compared with regard to the accuracy of their results. The next step will be to apply this methodology to housing stock models considering any sources of uncertainty, such as heterogeneity and parameter uncertainty.

Acknowledgment

This study is part of the Danish research project CITIES (Centre for IT-Intelligent Energy Systems in cities).

References

[1] M. Geidl and G. Andersson. Optimal Power Flow of Multiple Energy Carriers. IEEE Transactions on Power Systems (2007) 22: 145 – 155.
[2] IEA EBC Annex 53. Final Report. Total energy use in buildings, November 2013.
[3] A. Saltelli, S. Tarantola, F. Campolongo. Sensitivity Analysis as an Ingredient of Modeling. Statistical Science (2000) 15: 377-395.
[4] W. Tian. A review of sensitivity analysis methods in building energy analysis. Renewable and Sustainable Energy Reviews (2013) 20: 411-419.
[5] T. L. Hemsath and K. A. Bandhosseini. Sensitivity analysis evaluating basic building geometry's effect on energy use. Renewable Energy (2015) 76: 526-538.
[6] Y. Sun. Sensitivity analysis of macro-parameters in the system design of net zero energy building. Energy and Buildings (2915) 86: 464-477.
[7] T. D. Pettersen. Variation of energy consumption in dwellings due to climate, building and inhabitants. Energy and Buildings (1994) 21:209-218.
[8] P. Heiselberg, H. Brohus , A. Hesselholt, H. Rasmussen, E. Seinre, S. Thomas. Application of sensitivity analysis in design of sustainable buildings. Renewable Energy (2009) 34: 2030–6.
[9] A. Ioannou, L.C.M. Itard. Energy performance and comfort in residential buildings: Sensitivity for building parameters and occupancy. Energy and Buildings (2015) 92: 216-233.
[10] R. Evins, K. Orehounig, V. Dorer. Variability between domestic buildings: the impact on energy use. Journal of building performance simulation (2015). http://dx.doi.org/10.1080/19401493.2015.1006526 .
[11] TABULA WebTool. http://webtool.building-typology.eu/#bm .
[12] S. Aggerholm. Cost-optimal levels of minimum energy performance requirements in the Danish Building regulations. Danish Building Research Institute (2013), SBI 2013:25.
[13] Danish Ministry of Climate, Energy and Building: 2001-2010 Danish Design Reference Year. Technical Report 13-19. Danish Meteorological Institute (2013), Copenhagen.
[14] M. Palonen, M. Hamdy, A. Hasan. MOBO a new software for multi-objective building performance optimization. Proceedings of BS2013, Chambery, France, August 26-28.
[15] J. Kragh. Energirenovering af etagebyggeriet. Videncenter for energibesparelser i bygninger (2010).

Paper X: Heller, A., Liu, X., & Gianniou, P. (2017). A Science Cloud for Smart Cities Research. *Energy Procedia 122*, 679-684.

Available online at www.sciencedirect.com

ScienceDirect

Energy Procedia 122 (2017) 679–684

Energy Procedia

www.elsevier.com/locate/procedia

CISBAT 2017 International Conference – Future Buildings & Districts – Energy Efficiency from Nano to Urban Scale, CISBAT 2017 6-8 September 2017, Lausanne, Switzerland

A Science Cloud for Smart Cities Research

Heller A.[a]*, Liu X.[b] and Gianniou P.[a]

[a] *Department for Civil Engineering, Brovej 1, Technical University of Denmark, 2800 Kgs.Lyngby, Denmark*
[b] *Department for Management Engineering, Produktionstorvet, Technical University of Denmark, 2800 Kgs.Lyngby, Denmark*

Abstract

Cities are densely populated and heavily equipped areas with a high level of service provision. Smart cities can use these conditions to achieve the goals of a smart society for their citizens. To facilitate such developments, the necessary IT-infrastructure has to be in place for supporting, amongst many other things, the whole lifecycle of big data management and analytics for research activities. At the Centre for IT-Intelligent Smart Energy for Cities, we have therefore been developing a flexible infrastructure, based on open sourcetechnologies. This paper presents this solution and its application in a city and building research.

Peer-review under responsibility of the scientific committee of the CISBAT 2017 International Conference – Future Buildings & Districts – Energy Efficiency from Nano to Urban Scale

Keywords: Infrastructure for smart cities; cloud computing; big data research; case examples, building application

1. Introduction

"Smart cities" is not a well-defined term, and its complexity is rather great [1]. The "smartness" of a city can stem from its citizens, organizations or technology. The last of these is the main characterization applied in this paper. The simplified idea is that the gathering of information from a city together with its intelligent handling to achieve smart decision-making and control is what constitutes a "smart city". The practical result is an innovative infrastructure.

* Corresponding author. Tel.: +45 4525 1861. *E-mail address:* alfh@byg.dtu.dk

1876-6102
Peer-review under responsibility of the scientific committee of the CISBAT 2017 International Conference – Future Buildings & Districts – Energy Efficiency from Nano to Urban Scale
10.1016/j.egypro.2017.07.369

Other definitions have been found in the literature, basically with the same vision, which apply some additional pieces to the common issue of enabling smart cities. In particular, geospatial representations have been widely applied to support urban planning, which enables information visualization on a map [2][3]. These examples show ways of integrating building information into geo-information, including all the aspects of buildings such as construction details (BIM), construction year, energy-related data, and much more [3].

This paper presents an operational example of a data management and computation infrastructure that can handle different aspects of a smart city for research. The proposed solution is different from other cloud solutions, in that it uses open source technologies and is able to handle open data and confidential data within the same infrastructure. All this is done in ways to provide researchers with high-level services of easy access.. In reality, a full-fledged city infrastructure would be much larger, but all central aspects are implemented and operational in the current setup. The goal is to develop the infrastructure that can handle the integration of building data into city systems, especially energy systems. This paper describes the infrastructure from the researchers' point of view and gives some examples of its use in the building and energy domain.

2. IT Infrastructure – An Implementation

The current infrastructure of the presented "cloud" is generic and can be applied to other subjects. However, most of the research involved in current activities is related to energy within cities. The objective of the infrastructure is to be able to do smart-city research, which is characterized by using a wide variety of data. Data can originate from sensors, legacy systems, operational databases, or other sources, with different formats and sizes. The heterogeneous data have to be handled in an efficient way by the resulting cloud. Meanwhile, generic routines will be developed to enable researchers to do the collecting work themselves and without the support from skilled-IT experts.

Adequate dealing with sensitive data is crucial. Some data are sensitive with respect to personal information. To ensure safe handling of sensitive data, the current infrastructure is designed to handle data within the system. The data are not allowed to leave the secure area, e.g., copied to personal computers and spread to elsewhere.

Open data is a highly prioritized requirement, and it is enabled via integration with international projects such as CKAN (http://ckan.org), a data publication infrastructure used by many cities in Europe to publish open data [4]. The European research repository at CERN, Zenodo (http://zenodo.org) is also supported and well-integrated. Export to Zenodo is designed into the solution by supporting the DataCite metadata (http://datacite.org), a standard that enables the description and citation of all kinds of research results, including research data.

To enable indexing and discovery of semi-sensitive and sensitive data, we support publishing metadata on the open data platforms such as CKAN and Zenodo, with a link to the data to get an authorized access. In that case that sensitive data have to be shared, anonymization will be applied.

All these services are provided to researchers within the current cloud infrastructure where the BigETL tool (see https://github.com/ xiufengliu/BigETL for more details) is used for the data transformations, such as solving missing values, removing duplicate data, and merging subsets of data into comprehensive datasets. In addition, the same package enables scheduling service for automating the transformations.

Origo is used to manage the Cloud. (https://www.origo.io). The Cloud has 18 physical servers, with 80 cores and 564 GB memory, and 4.2 TB node storage, with additional 1.2 TB network storage. Origo manages a pool of virtual machines (VMs) that can be scaled on the distributed infrastructure. The setup is composed of a core that is a centralized component managing the lifecycle of the VMs, a component for the identity and security management of users for safety of the data in the cloud, and a capacity management component that adjusts the placement of VMs based on a set of predefined policies. The default capacity scheduler implements a simple pairing policy and supports user-driven consolidation constraints.

To facilitate its use, the cloud provides Windows and Linux-based VM images with different pre-installed software packages. These include data science images with all the commonly used data analysis tools installed, such as R, Python, Pandas, Scikit-learn; and data management images with pre-installed different types of databases (e.g., PostgreSQL, MySQL, OpenTSDB, etc). Many other applications can be deployed within a VM. These settings can sufficiently satisfy different needs from our research.

The volume of data is a key value of the current infrastructure. A single fully equipped building can easily generate more than 10,000 data points ranging from temperature and movement sensors to system control sensors. Sampling rates can be almost real time, but 5- to 15-minute values are commonly applied. This means that the above-mentioned data points can easily generate 20-100 kB pr. hour. Scaling this up to a whole city, data volume and handling speed are decisive. With this in mind, the current infrastructure is designed to be scalable and flexible.

The data lifecycle describes the structure of the current infrastructure well. The data enters and is made ready to be served to the users. Then the data is used by researchers, perhaps transformed and prepared in certain ways. The intermediate and final results are stored and finally published.

Figure 1 shows the system, consisting of the following components:

1) Data ingestion layer that handles the many data sources
2) Data staging area
3) Data transformation layer
4) Data storage
5) Publication layer

Liu and Nielsen [5] present this workflow in detail, as implemented in the current setup.

In addition to this “normal data lifecycle”, there is a requirement for the archiving of research data in a university archive, in a national archive, or in Zenodo.

The infrastructure itself is analysed for efficiency, robustness and security by the system architect. More details can be found in Liu et al. [5], [6], [7] and [8].

3. Tooling for Researchers

The current cloud infrastructure aims at a high service level to researchers. In this section we elaborate on these services.

From a researcher’s perspective, the current infrastructure feels like a personal computer because of the integration in the preferred operating system in virtual machines. This way all the remote handling is hidden for the user. It is this aspect in particular, together with the computational efficiency that a cloud solution can offer, that avoids researchers copying data into their ‘unsafe’ local computers. This means data security is kept very high.

Virtualization is very important for a high service to the researchers. Therefore, a set of visualization tools are implemented.

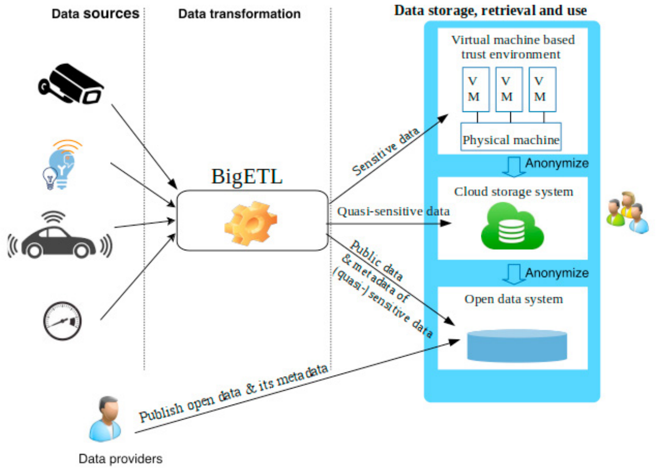


Figure 1 ICT system architecture

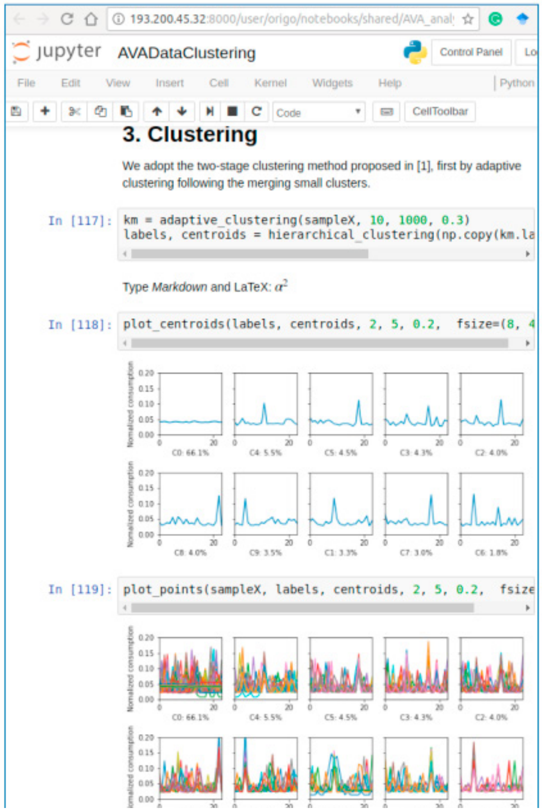


Figure 2 Data analysis by Jupyterhub

The researchers who use the current infrastructure come from a wide variety of subjects, so their traditions for methodologies and tools are also various. To enable cooperation between researchers from such different origins, a set of tools has been implemented aimed at making data management and analysis efficient, enabling reuse of work, and improving communication between researchers. Notebook tools have therefore been implemented:

JupyterHub (http://jupyter.org) enables the combination of documentation of the data handling and analytics with the coding of scripts in R and Python, two widely applied open source scripting languages where R is originated from the statistical domain, and Python is a more generic web scripting language. The scripts can be executed directly from within the notebook text, which makes it very convincing and efficient to develop the relevant data analytics routines.

In the current example, shown in Figure 2, the clustering analysis of the building dataset of a city is documented in the notebook; it enables running alternative scenarios in the same notebook environment. In such a notebook, the connection to the data, the handling of the data, the analytics and visualization are lined up for reproducibility and alternative examinations. In addition, such documents can directly be applied for educational purposes, giving students the description and their own dataset.

With this tool, the researchers can keep notes on their work in the form of codes, words and visualizations. These notes can be directly utilized for communication with peers and students, who are able to reproduce the work directly or transfer the codes to their own datasets from within a note copy. This makes it easy to share research in a dynamic and visualized form. An example is given in Figure 2, which shows documentation, codes, results and visualization in form of plots.

4. Case applications

At the research centre CITIES, the infrastructure presented here has been used for a number of smart-cities, smart grid(s), smart-buildings and smart components in the form of IoT applications. In the current paper, an example of such an application is given that can be relevant for professionals in building research and smart cities.

The city of Sønderborg in Denmark, Jutland has the ambition to be CO_2 neutral by 2029. This target drives a lot of research, development and innovation. A few years ago, the first monitoring and modelling approaches were carried out by Bacher et.al., developing grey box models for space-heating [9] and hot-water consumption [10]. They monitored and analysed a set of 16 buildings from the municipality. Before doing additional research, we first aimed at reproducing this earlier research [11]. For this purpose, we had to find data that was not publically available and had not been explicitly published. We moved the data onto the current data management system from where follow-up research can be carried out and new data can be found. This is an example of reusing, repurposing and reproducing research that is enabled and streamlined by a cloud infrastructure.

The data involved in the previous example was based on annual data. With the availability of e-meters, a follow-up study has been carried out since 2014, in which hourly values of the heat demand drawn from the district heating are being collected. 54 buildings from "Sonderborg Fjernvarme" are being monitored and the data collected in the science cloud by batch scripts. Figure 1 shows the principle of data management from collection to analysis. This case setup can be seen as a generic application of the science cloud infrastructure on typical energy and water meter data analysis for city area that can be repurposed for other cases and even applied to a collection of data between district and city cases. Hereby the infrastructure is applied in a way that differs from other implementation by the fact that it is on the one hand flexible to any application, on the other hand adapted to smart energy cities.

In the event of e.g. a loose connection to the source, warnings are sent to the relevant responsible person. In this way, losses of data are limited. Advanced automated procedures have been implemented to make data ready for use. In this case, the data is cleansed of errors and missing values, then stored in a database from where a number of predefined data packages are generated. These can be hourly, monthly or annual lists of values to be utilized by others. These data packages are then published through the channels mentioned elsewhere in this paper. This data is currently being used by the city of Sonderborg, for a series of student projects and other research. In the IEA EBC Annex 67 project, the flexibility that buildings offer to the energy system is a central aspect of the research. In this work, monitoring data and simulation data are applied to analyse these potentials. The work is still ongoing and the first publications are expected in the course of 2017. At the present time, a collection of time series data for typical Danish houses, based on the Tabula Webtool of typology for Denmark [12], has been developed and published through the current cloud infrastructure for open sharing. The published time series was generated synthetically using simulation models that will be made available parallel to the time series work to be published. This application of IT solutions is planned to be extended with the collection of other models and methodologies relevant for this research.

CITIES: Centre for IT-Intelligent Energy Systems for Cities, http://smart-cities-centre.org.

5. Conclusion

Smart City research is very interdisciplinary and involves huge amounts of data from various sources. The availability of powerful IT infrastructures that can handle this complexity and volume of data is vital. The current project demonstrates such an infrastructure developed using open source software only. This approach was preferred against existing large-scale infrastructures by well-known companies because of the flexibility that it gives. The drawback is that the size of the components is limited to the hardware available. We plan to overcome this problem by expanding the infrastructure with additional research projects that have the budget for this.

The solution is unique by the fact that the basic software architecture is open, flexible and extensible, and in our implementation simultaneously adjusted to the demand for energy analysis in a smart city context involving big data stemming from smart meters. This adjustment ensures secure handling of sensitive data in an environment that researchers from the relevant knowledge domain can handle without getting to be IT and data specialists.

The current infrastructure has been used for research projects, and this has shown its applicability and limitations with respect to smart-city and smart-building applications. Further development is ongoing and planned for the near future.

6. Acknowledgement

This research was supported by the CITIES project (NO. 1035-00027B) funded by Innovation Fund Denmark. The infrastructure components have been partly supported by the Danish e-Infrastructure Cooperation (DeIC) through the project "Science Cloud for Cities".

7. Literature

[1] Neirotti P, De Marco A, Cagliano AC, Mangano G, Scorrano F. Current trends in Smart City initiatives: Some stylised facts. Cities 2014;38:25–36. doi:10.1016/j.cities.2013.12.010.

[2] Kim SA, Shin D, Choe Y, Seibert T, Walz SP. Integrated energy monitoring and visualization system for Smart Green City development: Designing a spatial information integrated energy monitoring model in the context of massive data management on a web based platform. Autom Constr 2012;22:51–9. doi:10.1016/j.autcon.2011.07.004.

[3] CityZenith. CityZenith 2016. http://cityzenith.com.

[4] Open Data DK n.d. http://www.opendata.dk/ (accessed April 1, 2017).

[5] Liu X, Nielsen PS. A hybrid ICT-solution for smart meter data analytics. Energy 2016;115:1710–22. doi:10.1016/j.energy.2016.05.068.

[6] Liu X, Golab L, Golab W, Ilyas IF, Jin S. Smart Meter Data Analytics: Systems, Algorithms, and Benchmarking. ACM Trans Database Syst 2016;42:1–39. doi:10.1145/3004295.

[7] Liu X, Nielsen PS. Streamlining Smart Meter Data Analytics. Proc. 10th Conf. Sustain. Dev. Energy, Water Environ. Syst., 2015.

[8] Liu X, Heller A, Nielsen PS. CITIESData: a smart city data management framework. Knowl Inf Syst 2017:1–24. doi:10.1007/s10115-017-1051-3.

[9] Bacher P, Madsen H, Nielsen HA, Perers B. Short-term heat load forecasting for single family houses. Energy Build 2013;65:101–12. doi:10.1016/j.enbuild.2013.04.022.

[10] Bacher P, de Saint-Aubain PA, Christiansen LE, Madsen H. Non-parametric method for separating domestic hot water heating spikes and space heating. Energy Build 2016;130:107–12. doi:10.1016/j.enbuild.2016.08.037.

[11] Gianniou P, Heller A, Rode C. Building energy demand aggregation and simulation tools a Danish case study. Proceeding CISBAT 2015, 2015, p. 797–802. doi:10.5075/epfl-cisbat2015-797-802.

[12] EU project Tabula. Tabula Webtool for European Building Typology 2014. http://webtool.building-typology.eu/ (accessed June 18, 2014).

Paper XI: Liu, X., Nielsen, P. S., Heller, A., & Gianniou, P. (2017). SciCloud: A Scientific Cloud and Management Platform for Smart City Data. In Proceedings of *28th International Workshop on Database and Expert Systems Applications (DEXA)* (pp. 27-31).

2017 28th International Workshop on Database and Expert Systems Applications

SciCloud: A Scientific Cloud and Management Platform for Smart City Data

Xiufeng Liu and Per Sieverts Nielsen

Department of Management gineering
Technical University of Denmark
{xiuli, pernn}@dtu.dk

Alfred Heller and Panagiota Gianniou

Department of Civil Engineering
Technical University of Denmark
{alfh, pagian}@byg.dtu.dk

***Abstract*—The pervasive use of Internet of Things and smart meter technologies in smart cities increases the complexity of managing the data, due to their sizes, diversity, and privacy issues. This requires an innovate solution to process and manage the data effectively. This paper presents an elastic private scientific cloud, *SciCloud*, to tackle these grand challenges. SciCloud provides on-demand computing resource provisions, a scalable data management platform and an in-place data analytics environment to support the scientific research using smart city data.**

***Keywords*—Cloud, Platform, IoT, Smart city data.**

I. Introduction

Smart cities have been under development globally over the past decade, with a particular focus on the use of information and communication technologies (ICT) to manage urban infrastructure, including building, energy, transport and pollution monitoring [20]. Internet of Things (IoT) is widely deployed to make cities more green, safer and more efficient. Sensors and other smart devices are connected to urban infrastructure to obtain real-time information for decision-making purposes. IoT-based services will continue a substantial development, which is expected to reach 212 billion deployed entities globally by the end of 2020 [11]. In addition, it is clear that smart cities have been revolutionized by cloud-based ICT infrastructures to address the complex IoT services required for urban infrastructure. Cloud-based ICT infrastructures can integrate IoT technologies as needed, and can collect real-time data and the data from other sources, such as operational systems, legacy systems, and applications. According to the recent Forbes survey [21], more than 80 percent of today's organizations are using at least one cloud-based service in their businesses. A typical example is geospatial representation, which is showing an increasing popularity in recent years, such as in the application of urban planning and building modelling [7], [14]. This also illustrates the need for innovative solutions for building and geographic data integration, e.g., integrating the information of all aspects of buildings including construction details, related energy and more [7]. The data can be provided as a service to city governors, urban planners and citizens for decision making.

Moreover, the Big Data trend generates data increasingly complex and large, which makes analysis, archiving and sharing challenging. For example, in our current smart city work, a single building with indoor air-quality sensors can easily generate more than 10,000 data points per minute. In fact, smart cities have many use cases that produce large data sets, such as smart energy systems, weather monitoring, and transport systems. The sampling rate can be very fine granularity, e.g., per second or millisecond. As a result, the sizes of data are typically large, along with diverse types and formats. Therefore, it is difficult to orchestrate smart city data sets.

In this paper, we present a private cloud platform, *SciCloud*, for assisting our smart city research project [6]. SciCloud as an ICT infrastructure is designed to handle different aspects related to smart cities. In particular, it deals with the data streams originating from smart systems, such as the energy systems and IoT networks that our project emphasizes. In order to facilitate the use of SciCloud, we have developed a scalable data framework to simplify smart city data management, including data collection, cleaning, anonymization, and publishing. This paper describes the applicability of the Cloud and the framework from the perspective of research, and demonstrates two use cases in the field of smart energy and air quality. The Cloud and the data management platform can be applied to many other use cases, due to their flexibility.

II. Related Works

Cloud computing becomes increasingly popular today, due to its ability of providing unlimited resources. Customers can use cloud resources based on the pay-as-you-go business model, which only pays for what actually have been used, e.g, the number of cores, the amount of memory and space. There are a number of well-known public cloud platforms available, including Google Cloud, Amazon EC2, Microsoft Azure, Rackspace and AliCloud. Many organizations intend to set up their own private cloud to get better privacy protection and to support their business goals. The presented SciCloud is a private cloud platform that provides a secure environment for our research of smart cities.

Clouds need effective tools to manage their computing resources. There are several main-stream open source cloud management frameworks, including OpenNebula [19], OpenStack [23], Eucalyptus [22], Nimbus [13], Snooze [9] and Origo [25] (used in the SciCloud). These frameworks greatly facilitate the cloud management: allocating computing resources according to user demands, failure recovery, load balancing, system monitoring and others. In order to provide a secure cloud computing platform, in this paper we present a novel smart city data management framework in this private cloud environment, and create an in-place data analytic service to support our scientific research of smart cities.

Smart city data are often characterized with the Big data characteristics: high volume, high variety and high velocity,

2378-3915/17 $31.00 © 2017 IEEE
DOI 10.1109/DEXA.2017.22

which is difficult to manage. Some attempts have been made towards smart city data management. Examples include the SCOPE [28], which is a cloud-based smart city open platform and ecosystem; CiDAP [5], which is a real-time smart city data platform; and FIWARE [10], which is a framework of providing intelligent application development in the Future Internet. They focus primarily on infrastructure development, data collection, test bench deployment or applications/services-specific development, but less emphasis on data sensitivity management. The [1], [4], [12] explore smart city data management from the perspectives of data security and privacy, which involve data collection, transmission, processing and visualization services. In contrast, in this paper we present a complete solution for simplifying smart city data management, which includes data extraction, processing, storing, sharing and publishing. Especially we emphasize privacy and sensitivity management of smart city data, and propose a three-level sensitivity model for publishing and sharing data.

III. The ICT infrastructure - SciCloud

The current trend of data processing is increasingly transferring traditional centralized solutions to cloud environments. Within our Centre of IT-Intelligent Energy System in Cities [6], we have created a private cloud platform to support our research in smart cities. The purpose of implementing a private cloud is to integrate infrastructure, platform and resources; and to provide a consolidated infrastructure for scientific experiments, software development and provision of analytics services.

Figure 1 illustrates the architecture of the SciCloud. It consists of 18 physical servers, with 80 cores and 564 GB of memory, and 4.2 TB of node storage, plus 1.2 TB of network storage.

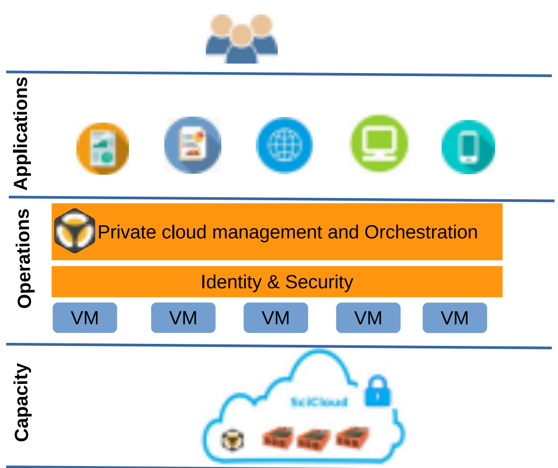


Fig. 1: The architecture of SciCloud

SciCloud is used to provide on-demand capacity for our research. SciCloud has integrated the Origo Virtual Infrastructure Engine [25] for resource management and to support a variety of demands of our researchers. Origo can efficiently manage virtual machines (VMs) that are scaled on a distributed infrastructure. Origo Virtual Infrastructure Engine provides the functionality to deploy, monitor and control VMs on a pool of distributed physical resources. Origo consists of the following three main components. The core is a centralized component that manages the lifecycle of a VM by performing basic VM operations, including deployment, monitoring, migration or termination. The core component also has a basic management and monitoring interface for the physical hosts. The second is the identity and security component. This component manages the security of user accounts and the safety of data in the cloud. The third is a capacity management component that adjusts the placement of VMs based on a set of predefined policies. The default capacity scheduler implements a simple pairing policy and supports user-driven integration constraints.

To support the use of this cloud, SciCloud offers Windows-based and Linux-based VM images with different pre-installed software packages. These include data science images with all the commonly used data analysis tools, such as R, Python, Pandas, Scikit-learn; and data management images with pre-installed different types of databases (e.g., PostgreSQL, MySQL, OpenTSDB, etc). Among others, many applications can be deployed in the VM. These settings can adequately satisfy the different needs of our research.

IV. Data Management Infrastructure

A. Smart city data management in the Cloud

Smart city data are collected from a variety of sources such as IoT devices, video surveillance systems, social networks, transport, government documents, or open data platforms, location-based services, and more. In addition, some data such as socio-economic data, contain sensitive personal related information such as social security number, name, age, home address and health, etc. Therefore, smart city data have the characteristics of big data, including big volume, high velocity and variety. These pose a grand difficulty in dealing with the data. In order to facilitate smart city data management, establishing a complete and flexible data management platform becomes essential. This is a key step between data sources and the applications of using the data. Although some studies have been done to study the big data platform of smart cities, most focus implementing a specific functional requirement and architectural design. At the same time, as there are many different tools and platforms with similar functionalities available in the big data community, we are often overwhelmed and confused by their features and capabilities. As a result, there is still a gap between a big data platform for smart cities and how they can be properly realized. We, therefore, present a cloud-based platform, called CITIESData [16], to manage smart city data, which is a trust framework that ensures that data can be properly shared, published, and used without compromising the data privacy.

The platform has the capability of processing the data with high diversity and complex, e.g., with different types, formats, meanings, and sizes; as well as handling the data with different quality issues, e.g., missing values and/or incorrect values. This platform aims at streamlining the whole data process: collecting, cleansing, storing, anonymizing, publishing and analyzing data. We have made a particular emphasis on data privacy and data quality. Figure 2 illustrates the system architecture which

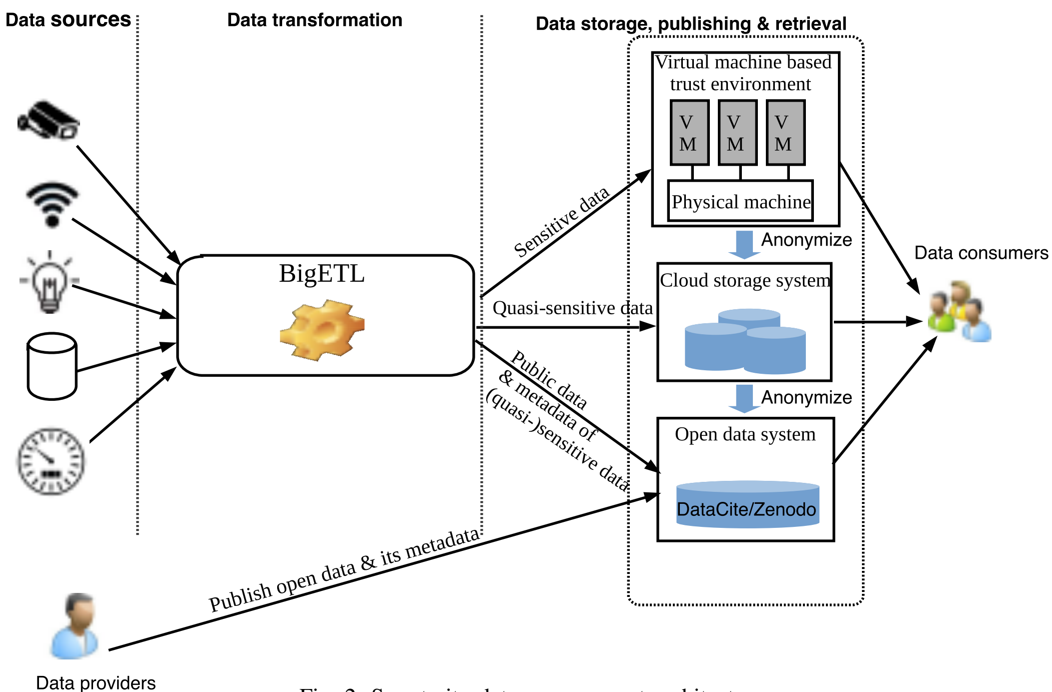


Fig. 2: Smart city data management architecture

comprises the phases including data extraction, transformation, anonymization, archiving and publishing.

This platform integrates a BigETL tool for data transformations (see https://github.com/xiufengliu/BigETL for more details), such as fixing missing values, removing duplicate data, and merging. The core feature of this system is on how to manage sensitive data for data publishing or sharing. This system classifies the data according to a three-level sensitivity model in terms of information disclosure. The data are classified into sensitive data, quasi-sensitive data, and public (open) data; and different strategies are used for the publishing. Sensitive data is shared with authorized users within a virtual machine based trust environment in the cloud - the data are not allowed to leave this environment when it is used. Quasi-sensitive data is shared via a cloud-based storage system, e.g., a private OwnCloud [26] that requires authorization; whereas public data is shared on an open data platform, namely Zenodo [29] or CKAN [8]. The sensitivity level can be mitigated by anonymization,e.g., sensitive data becomes non-sensitive after being anonymized. An open data platform itself is integrated with a data management system, such as CKAN, which allows data publishers to upload and publish data directly. An open data platform can also be restricted to publish metadata (i.e., the data of describing other data). This feature is useful for sharing information from (quasi-)sensitive data, i.e., only publishing the metadata while not the (quasi-)sensitive data itself. The benefit is that (quasi-)sensitive data can still be indexed and discovered through the open data platform even though the data themselves are not accessible. If a user needs to access (quasi-)sensitive data, (s)he has to link to a secure environment where user authorization is enforced. The open data portal is the single entrance to search the data available in all data repositories. This solution becomes a feasible way to maintain privacy and openness of smart city data.

In addition to the above "normal data life cycle", there is a demand for archiving research data, e.g., for a university archive and a national archive (see [16] for the details).

B. *In-place data analysis*

This data management system offers an additional setup for in-place analysis using data in the Cloud. This is done by adding a data analytics layer on top of CITIESData (see Figure 3). This layer has Jupyterhub and RStudio installed on the virtual machine, and both analytic tools can access data directly in the underlying CITIESData platform. As Jupyterhub and RStudio both offer the web-based interface for researchers to interact with the data, it means that they can do the analysis without copying the data out of the Cloud infrastructure. Another benefit is that the virtual machine has pre-installed all the necessary data analysis tools, software and packages. Researchers can be released from the tedious software installation, but focus on their analytics tasks. In addition, as virtual machines run on the Cloud, researchers can take full advantage of the computing power provided by the Cloud, for example, to handle a big data set which is usually not possible on a personal computer. The Jupyterhub and RStudio support multiple users, but each user has her/his own working environment. They can install additional software packages by themselves. Multiple users can share their work and work together, e.g., on the same notebook of Jupyter.

The data analysis platform itself is efficient, robust and secure (see [15], [17], [18] for more details). As mentioned earlier, releasing data is an important service, and legislation requires the secure handling of sensitive data. To ensure handling of sensitive data safely, we design this architecture to store the data within the system. The data never leaves the secure area, and hence no data is copied to personal computers that may be backed up and propagated. In addition, this platform uses OwnCloud to assist data security management. In OwnCloud, a role-based model is applied to manage fine-level data permissions. For each user, only the data that (s)he should access will be granted read permission. The data will be automatically synchronized to the analytics layer for users to use. When logging into Jupyterhub or RStudio, (s)he can find the data residing in the home directory.

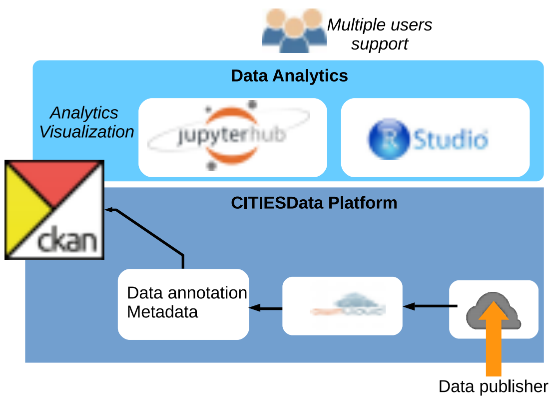


Fig. 3: In-place data analytics service

V. User Cases

In this section, we present two use cases as the examples of using the SciCloud platform.

A. *Smart meter data management*

A smart meter is an advanced meter that measures energy consumption at a regular time interval, typically every 15 minutes in electricity meters [15]. Smart meters communicate information back to the local utility for monitoring and billing purposes. The detailed energy usage information could lay bare the daily energy usage patterns of a household and even go so far as to enable deduction of what kind of device or appliance was in use at any given time. Besides, the unique ID of a smart meter is co-related to an individual or a household. This raises important privacy issues regarding the availability and processing of such data [27].

The data involved in our use case are the district heating data from 54 households in Sonderborg, Denmark. The meter readings are recorded every hour, and streamed into our system through the CITIESData platform. The heating consumption data is time-series data with two metrics, *volume* and *heat energy*, and the timestamp. The heating consumption data has some data quality issues, including row duplication and having missing values for some time series. Each time series has meter ID, which is associated with the household. The household data contains the sensitive information, including the name of customers and their addresses (road and building no.).

First, we address the data quality issues by cleansing the data, which involves fixing the missing values and outliers (anomaly high value above a set threshold value), transformation (extracting the date and hour from the timestamp), and removing the unnecessary values (e.g., the unit of the readings). This is done automatically by running a batch job in BigETL, which is triggered at a specified time.

Second, we address the data privacy problem by anonymizing the smart metering data so that information gleaned from it cannot easily be associated with an identified person. This is done by the following steps: 1) we replace the meter ID with a meaningless surrogate key, and decouple the time series from customer data; 2) we aggregate the time series, and provides daily and monthly readings to users; 3) we make use of our secure platform for ensuring the data safety. Data users are granted read permission in order to use the data. Data users can only do the online analytics within Jupyterhub or RStudio, while the data is not able to be distributed or downloaded onto personal computers.

Data quality and privacy protection are considered to be of prime importance to smart meter data management. This paper proposes a solution for streamlining smart meter data processing, and anonymizing high-frequency metering data through several strategies without compromising the use of the data. For these reasons, the SciCloud is well suitable for smart energy data management.

B. *Air quality IoT data management*

Air pollution is one of the most important factors that affects the health of people and the quality of life in smart cities. Over the past decades, the air quality in many global metropolitan cities has deteriorated significantly, especially in developing countries such as China and India. In the European countries, air pollution is not as prominent as the developing countries, but many governments have set their goals of reducing greenhouse gas emission, e.g, Danish government sets the goal of reducing greenhouse gas emission by 40% by 2020 and becoming a completely fossil-fuel free country by 2050. As a result, it is important to monitoring the air quality so as to provide the real-time information for the government and citizens for decision-making purposes. To support this, we have developed a cloud-based monitoring system [2], and deployed in Vejle, Denmark and Trondheim, Norway to monitor the air quality. We make use of the SciCloud in this project, and do the following:

- *Stream IoT data*: IoT sensors are installed in several places around the city, e.g., intersections of the roads, to monitor the air quality (including CO2 and CO level, particle sizes of PM1.0, 2.5 and 10) as well as weather condition (temperature, humidity, air pressure, wind speed). The resulting IoT data is streamed into the SciCloud in a real-time fashion. Besides, we also stream the traffic data at the places where the sensors are located into the SciCloud from a third-party traffic monitoring system.
- *Data storage*: To integrate multiple data sources, we deploy the time-series database, OpenTSDB [24], in the SciCloud to manage all the time series. All the

time series are stored in a uniform format, i.e., metric, timestamp, values, and a number of tags of labeling time series (e.g, the locations of the sensors). OpenTSDB is a distributed database, which can store and query data efficiently through its RESTful APIs. The data is provided as a service for application developers.

- *Air quality analytics*: There is a processing and data mining module, which is implemented as the Python program in Jupyter. The program reads the time series of air quality, weather condition and traffic, and studies impact of weather and traffic flow on the air quality by correlation. This model is done offline, but updates the analytic charts regularly through reading the data from OpenTSDB.
- *Monitoring dashboard*: Local city authorities are enabled to view air quality and the analytic results through a dashboard. The dashboard is implemented on the Apache Zeppelin [3] that is deployed on the SciCloud. The dashboard can be exported as an iFrame to be embedded into any websites, e.g., the government websites.

VI. Conclusion and Future Work

Smart cities necessitate management of the so-called Big Data. In this paper, we have presented a private cloud platform, SciCloud, for smart city data management. SciCloud provides the on-demand computing resources for researchers to experiment, develop prototypes and proof of concept. To facilitate the use of this cloud, we have developed a secured scalable smart data management system to streamline the processing of the data life cycle. In addition, we have proposed an in-place analytics environment for analyzing the data in the Cloud, and we verified its effectiveness by demonstrating two use cases on this Cloud platform.

In future work, we intend to further develop this SciCloud to support other research projects, and verify the Cloud by managing the data from more domains.

Acknowledgement

This research was supported by the CITIES project (NO. 1035-00027B) funded by Innovation Fund Denmark. The infrastructure components are partly supported by the Danish Electronic Infrastructure (DeIC) through the project "Science Cloud for Cities".

References

[1] A.G. Abbasi, Z.A. Khan, and Z. Pervez, "Towards Cloud Based Smart Cities Data Security and Privacy Management," Proc. of the 7th International Conference on Utility and Cloud Computing, pp. 806–811, 2014.

[2] D. Ahlers, P.A. Driscoll, K.F. Kraemer, F.V. Anthonisen, and J. Krogstie, "A Measurement-Driven Approach to Understand Urban Greenhouse Gas Emissions in Nordic Cities," NIK, 2016.

[3] Apache Zeppelin. Avail. at http://zeppelin.apache.org/ as of 2017-5-15.

[4] J. Bohli, D. Garcia, P. Langendorfer, M.V. Moreno, and A. Skarmeta, "SMARTIE project: Secure IoT data management for smart cities," Proc. of the International Conference on Recent Advances in Internet of Things (RIoT), pp. 1–6, 2015.

[5] B. Cheng, S. Longo, F. Cirillo, M. Bauer, and E. Kovacs, "Building a big data platform for smart cities: Experience and lessons from santander," In Big Data (BigData Congress), pp. 592–599, 2015.

[6] CITIES. Avail. at http://smart-cities-centre.org as of 2017-5-15.

[7] CityZenith. Avail. at http://cityzenith.com as of 2017-5-15.

[8] CKAN: The open source data portal software. Avail. at https://ckan.org as of 2017-5-15.

[9] E. Feller, C. Morin, and L. Rilling, "Snooze: A Scalable and Autonomic Virtual Machine Management Framework for Private Clouds," Proc. of the 12th IEEE/ACM International Symposium on Cluster, Cloud and Grid Computing, pp. 482–489, 2012.

[10] FIWARE: Open Source Platform. Avail. at http://www.fi-ware.org as of 2017-5-15.

[11] J. Gantz, and D. Reinsel, "The digital universe in 2020: Big data, bigger digital shadows, and biggest growth in the far east," IDC iView: IDC Analyze the Future, vol. 2007, pp. 1–16, 2012.

[12] N. Gruschka, and D. Gessner, "IoT-A: Internet of Things Architecture, Project Deliverable D4.2 - Concepts and Solutions for Privacy and Security in the Resolution Infrastructure," 2012.

[13] K. Keahey, T. Freeman, J. Lauret, and D. Olson. "Virtual workspaces for scientific applications," Journal of Physics: Conference Series, 78(1), 2007.

[14] S.A. Kim, D. Shin, Y. Choe, T. Seibert, and S.P. Walz, "Integrated energy monitoring and visualization system for Smart Green City development: Designing a spatial information integrated energy monitoring model in the context of massive data management on a web based platform," Autom Constr, 22:51–9, 2012.

[15] X. Liu, and P.S. Nielsen, "A hybrid ICT-solution for smart meter data analytics," Energy, 115(3):1710–1722, 2016.

[16] X. Liu, A. Heller, and P.S. Nielsen, "CITIESData: A Smart City Data Management Framework," Knowledge and Information Systems. 2017:1-24, 2017.

[17] X. Liu, L. Golab, W. Golab, I.F. Ilyas, and S. Jin, "Smart Meter Data Analytics: Systems, Algorithms, and Benchmarking," 42:139, 2016.

[18] X. Liu, and P.S. Nielsen, "Streamlining Smart Meter Data Analytics," Proc. of the 10th Conference on Sustainable Development of Energy, Water and Environment Systems, SDEWES2015.0558,1-14, 2015.

[19] D. Milojicic, I.M. Llorente, and R.S. Montero, "OpenNebula: A cloud management tool". IEEE Internet Computing, vol. 15, March 2011.

[20] P. Neirotti, A. De Marco, A.C. Cagliano, G. Mangano, and F. Scorrano, "Current trends in Smart City initiatives: Some stylised facts,". Cities 2014; 38:25–36.

[21] New Stats From The State Of Cloud Report. Avail. at https://www.forbes.com/sites/ benkepes/2015/03/04/new-stats-from-the-state-of-cloud-report/ as of 2017-5-15.

[22] D. Nurmi, R. Wolski, C. Grzegorczyk, G. Obertelli, S. Soman, L. Youseff, and D. Zagorodnov, "The Eucalyptus open-source cloud-computing system". Proc. of the 9th IEEE/ACM International Symposium on Cluster Computing and the Grid, 2009.

[23] OpenStack: Open source cloud computing software, Avail. at http://www.openstack.org as of 2017-5-15.

[24] OpenTSDB. Avail. at http://OpenTSDB.net as of 2017-05-15.

[25] Origo: cloud computing. Avail. at https://www.origo.io as of 2017-5-15.

[26] Owncloud. Avail. at http://owncloud.org as of 2017-05-15.

[27] E.L. Quinn, "Privacy and the New Energy Infrastructure," Social Science Research Network (SSRN), 2009.

[28] SCOPE: A Smart-city Cloud-based Open Platform and Ecosystem. Avail. at http://www.bu.edu/hic/research/scope/ as of 2017-5-15.

[29] Zenodo. Avail. at http://zenodo.org as of 2017-05-15.

Energy is one of the major drivers in smart cities. In this PhD thesis, enhanced modelling methodologies were introduced that provide building energy demand related insights in a high spatial and temporal resolution, which can help to evaluate energy policies and demand side management strategies. Moreover, the potential that buildings can have on the stabilization of the energy grid and the flexible operation of the energy system was determined -considering a smart energy cities context- both at building level and urban level.

DTU Civil Engineering
Technical University of Denmark

Brovej, bygning 118
2800 Kongens Lyngby

www.byg.dtu.dk

ISBN 9788778774859
ISSN 1601-2917